Multiple Memory Systems: A Neurophilosophical Analysis

Elizabeth Leigh Ennen
Department of Philosophy
McGill University, Montréal

August, 1995

Copy #1

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

ISBN 0-612-12363-4

TABLE OF CONTENTS

INTRODUCTION 1

CHAPTER ONE: TRADITIONAL COGNITIVISM 9

HISTORICAL BACKGROUND 9
Behaviourism 10
Physicalism 12
Functionalism 14
A Progress Report 18
THE REPRESENTATIONAL THEORY OF THE MIND 19
THE REPRESENTATIONAL THEORY OF INTELLIGENCE 30
The Scope of Cognitivism: Perceptual-Motor Skills and Tacit Knowledge 30
The Scope of Cognitivism à la Psychosemantics 38
Alleged Counter-Examples to 1A: [PA → (R - MP)] 42
Alleged Counter-Examples to 1B: [(R - MP) → PA] 45
FODOR'S DEFENSE OF THE RTM: THE ARGUMENTS FOR LOT 48

CHAPTER TWO: PHENOMENOLOGY AND BACKGROUND SKILLS 59

DREYFUS' ACCOUNT OF HEIDEGGERIAN COPING 67
Ontic Transcendence 69
Originary Transcendence 81
PHENOMENOLOGICAL CRITICISMS OF COGNITIVISM 86

CHAPTER THREE: BERGSON'S THEORY OF MULTIPLE MEMORY TYPES 96

BERGSON THEORY OF PERCEPTION 97
BERGSON'S THEORY OF MEMORY 104
Two Kinds of Memory: Representational Memory and Motor Memory 105
The Interaction of Representational Memory and Habit Memory 109
The Two Forms of Recognition 111
CATEGORIZATION AND THE GENERAL IDEA 117
CONTEMPORARY NEUROSCIENTIFIC ACCOUNTS OF PRACTICAL PERCEPTION 122

CHAPTER FOUR: HISTORY OF MULTIPLE MEMORY SYSTEMS RESEARCH 130

THE TOLMAN - HULL DEBATE 131
Hull's Theory of Learning 133
Tolman's Theory of Learning 138
HUMAN AMNESIA STUDIES PHASE ONE: HIPPOCAMPAL AMNESIA 153
HUMAN AMNESIA STUDIES PHASE TWO: THE SPARED CAPACITIES 162
CONCLUSION 169

CHAPTER FIVE: CONTEMPORARY THEORIES OF MULTIPLE MEMORY SYSTEMS 172

A FEW PRELIMINARY NOTES 175
AN ANIMAL MODEL OF HUMAN AMNESIA 179
Hirsh's Theory of Multiple Memory Systems 179
O'Keefe and Nadel's Theory of Multiple Memory Systems 185
MISHKIN'S THEORY OF MULTIPLE MEMORY SYSTEMS 199
The Representational Memory System 210
The Non-Representational Memory System 217
THE MULTIPLE MEMORY SYSTEMS THEORY OF COHEN AND EICHENBAUM 222
THE PERCEPTUAL-MOTOR SKILL SYSTEM: THE BASAL GANGLIA 233
CONCLUSION 239

CONCLUSION 242

ABSTRACT

Neuroscientific data may be usefully invoked in the arbitration of debates concerning the scope of representational theories of the mind. Contemporary cognitivists (e.g. Fodor) tend toward theoretical imperialism in that they argue that all types of intelligent behaviour, including perceptual-motor skills, can be explained within the framework of representationalism. Phenomenologists (e.g. Heidegger, Merleau-Ponty, and Dreyfus) argue that the scope of cognitivism is not as vast as its proponents suppose. They claim that perceptual-motor skills are non-representational and thus fall beyond the purview of cognitivism. I argue that this debate can be resolved in favour of the phenomenologists by citing the neuroscientific evidence for the claim that there are two distinct neural memory systems: (1) a hippocampal system which operates over neurally realized Fodorian representations and subserves rational thought and action and (2) a non-representational striatal system which subserves perceptual-motor skills.

Abrégé

Il peut être utile d'avoir recours à des données neuroscientifique pour abritrer les débats au sujet de la portée relative des théories représentationalistes de l'esprit. Les cognitivistes contemporains (Fodor, par exemple), penchent vers un impérialisme théorique, puisqu'ils argumentent que tous les genres de comportement intelligent, les habiletés perceptives-motrices incluses, peuvent être expliqués par le représentionalisme. Les phénoménologues (Heidegger, Merleau-Ponty et Dreyfus, par exemple) argumentent que la portée du cognitivisme n'est pas aussi étendue que ses préponderants ne le prétendent. Ils affirment que les habiletés perceptives-motrices sont non-représentationelles, et qu'elles sont donc en dehors du champs d'explication du cognitivisme. Je défend la thèse que ce débat peut être résolu en faveur des phénoménologues en faisant appel à l'évidence neuroscientifique à propos du fait qu'il y a deux systèmes de mémoires neuronale distincts: (1) un système basé au niveau de l'hippocampe qui opère au niveau des représentations Fodoriennes neuronales, et qui est subordonné à la pensé et l'action rationelles, et, (2) un système non-représentationnel constitué des corps striés du cerveau, subordonné aux habiletés perceptives-motrices.

Acknowledgments

I would like to thank the following people for their help: Nick Adamson, Brent Barraclough, Aki Caramanos, Emily Carson, Laura Duggan, Sue Dwyer, Richard Ennen, Angela Fotopoulos, Barbara Fultner, Michael Hallett, Iris Hardinge, Bill Hutchinson, Bilquis Khatoon, John Krotchko, Alison Laywine, Claudine Lefort, Calvin Lei, Marilyn Marshall, Cindy Mitch, Jessie Newman, Michael Petrides, Paul Pietroski, Matthew Shapiro, Ursula Staubli, Dale Turner, Ruel Tyson and Norm White.

Thomson House has been an important place of refuge during my years at McGill. I would like to thank all the kind people who work there. I feel particularly indebted to Jessie Newman, the resident chef. I hope she knows how much we appreciate her work.

Special thanks go to Any François for all her moral support and for her well-crafted translation of the abstract.

Special thanks also goes to my supervisors, David Davies and Charles Taylor. Professor Davies helped me navigate the terrain of contemporary cognitivism and spent endless hours reading and re-reading the text. His committed involvement in this project and his careful and extensive suggestions for improvement have been invaluable. Professor Taylor has influenced my way of thinking about philosophical issues and his seminars and reading courses have been important in the development of this project. I thank him for both his criticisms and his encouragement.

Natalie Oman's moral and intellectual support have kept me going these past two years. She believed in me and in this project even when I didn't and helped me in ways too numerous to list.

My deepest thanks go to my mother, Monica Ennen. Had it not been for her love and loyalty this dissertation would have remained unwritten.

To My Mother

Introduction

The cognitive revolution has fostered the renaissance of the mental representation. Behaviourism is no longer fashionable and internal states with representational content once again figure prominently in our explanations of intelligent behaviour. Descartes' theory of the mind has been up-dated and combined with the computer metaphor to yield the "cognitivist" view that intelligent behaviour is caused by the algorithmic manipulation of mental representations. This view, which is widely endorsed throughout the cognitive sciences, is the foundational principle of a vast number of successful research projects. The time has come, however, to re-evaluate cognitivism, to identify its blind spots and its excesses. The cognitive revolution has gone too far.

In this dissertation I will argue that some forms of intelligent behaviour are non-representational and that non-representational behaviour cannot be understood within a cognitivist framework. In today's cognitivist climate, non-representational behaviour is either overlooked altogether or misconstrued in representational terms. If we want a full account of intelligent behaviour, one which recognizes both representational and non-representational comportment, we must call into question the hegemony of cognitivism and articulate a more pluralistic approach to the study of human behaviour.

Philosophers and psychologists commonly assume that we face a choice between two alternative approaches to the study of intelligent behaviour: cognitivism and behaviourism. Given contemporary attitudes towards behaviourism, it is no wonder that cognitivism has become entrenched and that we have become preoccupied with representational mental states. Behaviourism's fall from grace has tainted the study of non-representational behaviour. Non-representational behaviour is, however, a legitimate object of study. What we need is a non-behaviouristic account of it.

The foundations for an alternative approach to non-representational behaviour were laid earlier this century by phenomenologists who took issue with Cartesian representationalism. Bergson, Merleau-Ponty, and Heidegger all direct our attention to the non-representational "background" skills that collectively constitute an agent's primary repertoire of intelligent behaviours. Representational cognition, it is argued, cannot be fully understood until its dependence on this "primary repertoire" is brought to light. Dreyfus and Taylor, among others, have argued that these phenomenological rejoinders to Descartes are relevant to current debates concerning the status and scope of cognitivism. The phenomenological tradition speaks to two issues of contemporary importance: (1) the nature of non-representational skills and (2) the relationship between representational and non-representational intelligent behaviour. It is phenomenology, not behaviourism, that will provide the appropriate framework for the study of non-representational comportment.

There are two immediate problems with the plan to invoke phenomenological principles in this inquiry, two problems which serendipitously admit of a common solution. The first problem is substantive. While phenomenology can "disclose" the phenomenon of non-representational behaviour and provide a rich description of it, a full account will require theoretical resources that lay outside the phenomenological tradition. The second problem is strategic. Cognitivist attitudes toward phenomenology range from disinterest to contempt. We want an account of non-representational behaviour that is intelligible and persuasive to cognitivists. Phenomenology alone will not do the trick.

The solution is to take a page from the game plan of cognitivists and opt for an inter-disciplinary approach. The ultimate goal is to have a full and detailed inter-disciplinary account of (1) representational cognition, (2) non-representational skills, and (3) the relationship between them. In this dissertation, I will focus on providing an inter-disciplinary account of non-representational skills. While I will offer a few suggestions concerning the potential relevance of certain research programs in computer science, I will concentrate on building bridges between phenomenology and neuroscience. The task at hand is thus to show that phenomenological insights into the nature of non-representational comportment are reinforced by contemporary neuroscientific research. I will attempt to circumvent the cognitivist's suspicion of phenomenology by showing that key phenomenological principles are being unwittingly rehearsed in more "reputable" fields of inquiry.

My strategy is to tell two seemingly unrelated tales and then show how they are deeply related. The first tale, the subject of Part One, concerns the general differences between cognitivist and phenomenological approaches to the study of intelligent behaviour. In Chapter One I review the historical motivations for cognitivism as well as its central principles, representationalism and computationalism. Fodor, a paradigmatic cognitivist, will serve as our guide to this material. In this portion of the thesis, I will focus exclusively on "classical" cognitivism. (The cognitivist community can be divided into two groups: classicists, who invoke conventional serial architecture in their computational models of the mind, and connectionists, who prefer to think of the mind as a neural net whose architecture resembles a PDP [parallel distributed processing] machine. The issue of connectionism will be addressed briefly in the conclusion.)

Classical cognitivism is based on the ingenious trick of coupling folk psychology with computationalism. The central principle of folk psychology is that intelligent behaviour is produced when agents reason sensibly from their beliefs and desires to efficacious courses of action. If Jean believes that it is going to rain, believes that an umbrella will protect her from the rain, and desires not to get wet, she will, all things being equal, elect to carry her umbrella with her. If we want to explain Jean's umbrella-carrying behaviour, all we need do is (1) assume that Jean is rational and (2) cite her relevant beliefs and desires and the logical relations among them. Cognitivism is said to provide a scientific foundation for folk psychology. Its *modus operandi* is to translate the agent's reasoning processes into computational processes. Mental processes are thus re-described as computational processes in which discrete symbols representing the content of individual beliefs and desires are algorithmically manipulated according to the rules of deductive logic.

The natural explicandum of cognitivism is rational thought and behaviour. When cognitivists turn their attention to perceptual-motor skills, they treat these skills as though they, too, were caused by the computational processing of mental representations. In Chapter One I argue for the claim that even from within the cognitivist world view, the representational theory of mind is an inappropriate theoretical framework for the study of perceptual-motor skills.

In Chapter Two I turn to the topic of phenomenology and argue that with respect to perceptual-motor skills, a Heideggerian approach is an impressive alternative to cognitivism. The phenomenological text with which I will work is the first few sections of Heidegger's *Being and Time* as they are interpreted by Dreyfus. (Dreyfus' reading of Heidegger is controversial in some circles. I shall take sidestep all issues concerning exegesis and take Dreyfus' interpretation of Heidegger at face value. What I am interested in is the position Dreyfus attributes to Heidegger. Whether or not Heidegger holds these views is irrelevant here.)

Heidegger distinguishes between two ways of being in the world. In the first instance, the agent is immersed in an ongoing flow of activities, guided by a pre-reflective understanding of the possibilities for action afforded by the environment. The comportment of the agent manifests a variety of practical perceptual-motor skills. The agent can navigate terrain, discern salient objects, and manipulate these objects effectively. The relationship between the agent and the world remains direct, unmediated by representational thought and instrumental reason. At some point in her development, the agent acquires the ability to forge a different type of relationship with the world. She develops the ability to distance herself from the pressing practical concerns of the moment at hand. She disengages from the world, represents features of the world to herself, thinks of both the past and the future, and reflects rationally on her options, goals, and plans. These two ways of being, engagement and disengagement, give rise to different sorts of intelligent behaviour. Engagement is associated with primordial perceptual-motor skills. Disengagement allows for the development of the high-order representational skills we associate with instrumental reasoning, theorizing, logical deduction, and planning.

The organizing principle that emerges from Part One is that there are two types of agent-world relationships (engagement and disengagement) which give rise to two distinct types of intelligent behaviour (skilled perceptual-motor comportment and rational action). We may speak in terms of a distinction between "low-order intelligent behaviour" (LOIB) and "high-order intelligent behaviour" (HOIB). The problem with cognitivism may now be formulated more precisely. Cognitivists tend to focus exclusively on HOIB. When they do turn their attention to LOIB, they apply the principles developed specifically for the explanation of HOIB to LOIB. This, however, will not work. Unlike HOIB, LOIB is not representational.

The second tale, which is the subject of Part Two of the thesis, concerns the study of multiple memory systems. I will use the neuroscientific data on multiple memory systems to show that while high-order cognitive skills are caused by internal processes that involve the manipulation of representations, perceptual-motor skills are caused by internal processes that do not involve the manipulation of representations. Contemporary neuroscientists have identified two functionally and anatomically distinct memory systems in the brain: a representational system centred in the hippocampus and a "skill" or "habit" system centred in the basal ganglia. (These anatomical terms will serve, for the moment, as convenient placeholders.) There are interesting similarities between cognitivist interpretations of HOIB and the neuroscientific interpretations of the hippocampal system, on the one hand, and phenomenological interpretations of LOIB and neuroscientific interpretations of the basal ganglia system, on the other. I will argue, in particular, that neuroscientific accounts of basal ganglia memory reinforce many of Dreyfus' controversial claims concerning the nature of background skills.

The hippocampal memory system is associated with tasks in which organisms must acquire and manipulate discrete representations of environmental states. The hippocampus assists in the consolidation of such representations. Once the representations are sufficiently "mature", they are stored in various cortical regions throughout the brain. The hippocampal system also serves as the central relay station for the manipulation of these mature representations.

The basal ganglia memory system is associated with tasks that require the deployment of perceptual-motor skills. Of particular interest here is the fact that some neuroscientists refer to basal ganglia memory as "non-representational" memory. The mechanism that underpins the development and execution of perceptual-motor skills involves the incremental modification of numerous perceptual-motor circuits, not the acquisition and manipulation of neurally realized representations. As these systems become increasingly fine-tuned, the organism responds more successfully to an increasing number of environmental demands.

In Chapter Three, I begin by discussing the phenomenological account of memory types offered by Bergson in his text *Matière et mémoire*. Bergson serves as an ideal bridge figure between phenomenology and neuroscience. As part of an extended argument against Cartesian representationalism, he draws a distinction between what he calls "representational memory" and "habit memory". Not only is this distinction closely related to Heidegger's distinction between representational thought and engaged coping, it is also striking similar to the neuroscientific distinction between hippocampal memory and basal ganglia memory. Contemporary neuroscientists describe Bergson's work on memory types as prescient.

Chapter Four is devoted to an examination of the historical neuroscientific literature on multiple memory systems. This chapter covers the period between 1930 and 1960 and focuses, in particular, on two historical events: (1) the Tolman-Hull debate and (2) Milner's early work on human amnesia.

Early signs of the transition from behaviourism to cognitivism are apparent in the extended debate between Tolman, a "cognitivist" neo-behaviourist, and Hull, a more traditional S-R theorist. Tolman broke rank with his fellow behaviourists in claiming that some high-order forms of intelligent behaviour require the ability to form and manipulate "mental" representations. Hull rejected this claim and remained loyal to the central principles of S-R theory. The Tolman-Hull debate is relevant here because it provides the necessary theoretical background for an understanding of contemporary work on multiple memory systems.

The second portion of the historical review will be devoted to an examination of Milner's ground-breaking work on the distinction between the lost and spared capacities of human amnesics. In the late 1950's and early 1960's, Milner conducted a number of studies on a patient known in the literature by his initials, H.M. At that time, the medial temporal lobe was poorly understood. It was known, however, that this region of the brain was often implicated in epileptic seizures. After H.M. had been diagnosed as suffering from debilitating seizures, it was decided that he would undergo an experimental operation in which the hippocampus would be removed bilaterally. To the surprise of his physicians and surgeons, H.M. emerged from the procedure with a case of severe amnesia. (His seizures were, however, brought under control.) Milner carefully catalogued H.M.'s lost and spared capacities and concluded that while he could not retain new information about the world, he could acquire new perceptual-motor skills. Milner's work thus inaugurated the new field of multiple memory systems research.

In Chapter Five I turn to the examination of four contemporary accounts of multiple memory systems, accounts proffered by (1) Richard Hirsh, (2) John O'Keefe and Lynn Nadel, (3) Mortimer Mishkin, and (4) Neal Cohen and Howard Eichenbaum. The first three of these theories are based on the principle that Hull's stimulus-response account of conditioning and other simple forms of habitual behaviour and Tolman's cognitivist account of more complex "representational" behaviour can be jointly accommodated via a theory of multiple memory systems. Cohen and Eichenbaum develop an account of multiple memory systems that is similar in many respects, though it is not framed in terms of the debate between Hullian behaviourism and Tolmanian cognitivism.

These theories of multiple memory systems constitute a demand for a more pluralistic approach to the study of intelligent behaviour. Most importantly, they provide strong foundations for the claim that skilled perceptual-motor behaviour is not caused in the same way as more highly developed "cognitive" behaviour. I will demonstrate that the data on multiple memory systems can be forcibly invoked in an argument against the Fodorian claim that perceptual-motor skills are caused by the manipulation of internal representations.

This thesis may be construed as a case study of sorts, a demonstration of the general principle that neuroscientific data are sometimes useful in the arbitration of competing philosophical theories. Phenomenologists and cognitivists offer inconsistent accounts of perceptual-motor skills. Cognitivists claim that skilled perceptual-motor behaviour is brought about via the processing of representations. Phenomenologists argue that perceptual-motor skills are not the product of internal representational processes. An examination of contemporary accounts of the memory systems in the brain indicates that the phenomenologists are right.

Chapter One: Traditional Cognitivism

In the first decade or so of its existence, cognitivism was an implicit creed that motivated and shaped the research programs of cognitive scientists in a number of fields. As a working hypothesis, cognitivism was so successful that its theoretical underpinnings went unexamined for many years. This situation was corrected in 1975, when Fodor presented a detailed version of the cognitivist manifesto in his book, *The Language of Thought.* [1] In this and more recent texts, Fodor discusses the central principles of cognitivism and codifies them in his Representational Theory of the Mind (RTM). The natural explicandum of Fodor's RTM is rational action. In some contexts, however, Fodor also discusses a more general theory of intelligent behaviour that is based on the RTM but that is intended to be applicable to a wider range of behaviours, including low-order perceptual-motor skills. I will refer to this theory as the Representational Theory of Intelligence (RTI). In this chapter I will review the historical motivations for cognitivism, discuss the central principles of both the RTM and the RTI, and argue that the RTI is an inappropriate theoretical tool for the study of low-order intelligent behaviour.

Historical Background

A critical puzzle facing philosophers and psychologists in the twentieth century can be described as follows: how are we to explain intelligent behaviour in a way consistent with the principles of ontological materialism. Four general solutions have been proposed: behaviourism, physicalism, functionalism, and cognitivism. (While cognitivism draws on several key features of both physicalism and functionalism, it represents a stronger set of claims.) To understand cognitivism is to understand why its proposed solution to the

---

1 Fodor, Jerry, *The Language of Thought,* Eds. Jerrold J. Katz, D. Terence Langendoen, and George A. Miller (Cambridge, Mass.: Harvard University Press, 1975).

puzzle is purportedly more efficacious than the solutions proposed by behaviourists, physicalists, and functionalists.

Behaviourism

The philosophical version of behaviourism, "logical behaviourism", makes a claim about the ontological nature of mental states. According to the logical behaviourist, mental states are particular types of behavioural dispositions. Mental causation is re-described in terms of the activation of such dispositions. The realization of a behavioural disposition is, on this view, all there is to mental causation. Fodor notes that "logical behaviourism provides a construal of mental causation and the glaring question is whether the construal it provides is adequately robust to do the jobs that need doing." [2]

Logical behaviourism is clearly counter-intuitive; common sense tells us that behaviour is (genuinely) caused by *mental* states. Logical behaviourism tells us that this cannot be, on pain of violating our commitment to materialism. On Fodor's view, behaviourism lacks the resources necessary to do battle with the common sense view. Behaviourism is simply not rich enough to explain the patterns that common sense perceives in intelligent behaviour. Consider the following example suggested by Fodor. If asked to explain John's aspirin consumption behaviour, common sense posits a chain of causally-related mental states and attributes them to John. For example, we might say that John has the following mental states: the belief that he has a headache, the desire to be rid of the headache, the belief that aspirin will provide relief, the belief that the benefits of aspirin consumption outweigh any harmful side effects, etc. On the common sense view, these mental states interact causally in a sequence of events we call mental processing and it is this mental processing that causes John's behaviour.

---

2 Fodor, Jerry, *Representations* (Cambridge, Mass.: MIT Press, 1981), p. 4.

Given this kind of story, in which a number of distinct mental states are seen to be implicated in the relevant mental processes, we can explain (and predict) John's behaviour in a number of related situations. John's aspirin consumption behaviour may vary over time; certain patterns in his behaviour may emerge. We may note, for example, that if John has a headache, he takes aspirin only if he has not eaten spicy food. On the common sense view, we can explain this pattern in John's behaviour by citing revisions and variations in John's mental states. In this case, we might suppose that John has modified his belief that the benefits of aspirin consumption outweigh its deleterious effects. His cost-benefit analysis might be altered by the discovery that aspirin, when taken on an irritated stomach, produces a pain far worse than the pain of a headache. We could explain John's non-consumption of aspirin by showing how the revised belief changes the causal relationships among his relevant beliefs and desires. The point is that these patterns in John's behaviour are easily explained via the invocation of causally efficacious mental states.

The logical behaviourist, in contrast, has an relatively impoverished set of explanatory tools at her disposal. Her account of John's behaviour is constrained; there can be no appeal to mental processing. John's behaviour must be described as being the result of a behavioural disposition - in this case, the disposition to produce headache behaviour. Here, in Fodor's words, is the logical positivist's proposed explanation.

> John was disposed to produce headache behaviours and being disposed to produce headache behaviours involves satisfying the hypothetical if there were aspirin around, one takes some, and there were aspirin around. [3]

While this explanation does not advert to mental processing, the gain in methodological asceticism is outweighed by the loss in explanatory and predictive power. Without an account of mental processing, the logical behaviourist cannot explain the patterns in John's aspirin consumption.

---

[3] Fodor, *Representations*, p. 4.

On Fodor's view, banishing mental predicates from psychological explanation is a grave error: there are too many cases in which mental processing is genuinely implicated in the etiology of intelligent behaviour. As he puts it, "these cases seem to be most glaringly the norm in reasoning and problem solving."[4] Note that on this point, Fodor's analysis is consistent with the claim that a cognitivist approach is most obviously suited for explanations of high-level or complex behaviour (e.g. ratiocination). In any case, Fodor summarizes the problem with behaviourism as follows.

> It seems perfectly obvious that what's needed to construe cognitive processes is precisely what behaviourists proposed to do without: causal sequences of mental episodes and a 'mental mechanics' to articulate the generalizations that such sequences instantiate. The problem was, and remains, to accommodate these methodological requirements within the ontological framework of materialism.[5]

The lesson learned from the downfall of behaviourism is that our desiderata include both materialism and genuine mental causation. Physicalism, functionalism, and cognitivism represent three distinct attempts to provide a theory of psychological explanation in which both materialism and mental causation are given their due.

Physicalism

The physicalist argues that we can have both materialism and mental causation if we can defend the view that mental events are identical to physical events. A mental event thus inherits all the causal properties of its physical instantiation. The problem of mental causation is thereby subsumed under the problem of physical causation and the ontological worry is put to rest.

---

4 Fodor, *Representations*, p. 6.

5 Fodor, *Representations*, p. 6.

Physicalism comes in two basic flavours: token physicalism and type physicalism. The token physicalist argues that any particular mental event is identical to some particular physical event. The belief that Paris is the capital of France is thus taken to be identical to some physical state, typically a physical state of the brain. The causal properties of the belief are thus the same as the causal properties that accrue to the corresponding neural state. Type physicalism is a stronger claim, for the type physicalist argues that mental state *types* are identical to physical state *types*. Thus, for example, since all tokens of the belief "Paris is the capital of France" belong to one mental state type, each neural state that instantiates the belief will belong to the same neural state type; the mental state type is identical to the neural state type.

Fodor rehearses the standard arguments against type physicalism. As he argues in his article "Special Sciences," the natural kinds (or "types") of a particular science are the entities over which the explanatory generalizations of a science range. [6] Since, for example, the explanatory generalizations of psychology range over beliefs and desires, beliefs and desires are natural kinds in psychology. Type physicalism requires that the natural kinds of psychology be coextensive with the natural kinds in neuroscience (or some other physical science). For each natural kind in psychology, there must be a legitimate neuroscientific kind that consists of the same members, suitably re-described in neuroscientific terms. While the natural kinds in psychology need not correspond to the "basic" kinds in neuroscience, they must at least correspond to kinds that have neuroscientific legitimacy. If the members of a psychological kind constitute a disjunctive set of entities at the neuroscientific level, then type physicalism fails. Consider, however, the example cited above concerning the belief that Paris is the capital of France. It seems wildly implausible to suppose that all the neural states that instantiate this belief would themselves form a natural kind *of any sort* at the neuroscientific level. Type physicalism

[6] Fodor, Jerry, "Special Sciences," *Readings in Philosophy of Psychology*, Vol. I, Ed. Ned Block (Cambridge, Mass.: Harvard University Press, 1980), pp. 120 - 33.

fails because the way in which we classify psychological states is incompatible with the way in which we classify neural states. The natural kinds of psychology are not coextensive with the natural kinds of neuroscience.

Token physicalism, on the other hand, puts no constraints on our principles of individuation. The psychological individuation of mental states is thus autonomous with respect to the individuation of physical states. Properties that do not define kinds at the physical level may count as causally relevant properties at the psychological level. Fodor (and just about everyone else) rejects type physicalism and endorses token physicalism. The problem with token physicalism, however, is that it does not say enough; having granted the requisite autonomy at the psychological level, it fails to specify just what we are to do with this freedom. Token physicalism raises the following question: if psychological individuation is not dependent upon neural individuation, how *are* we to individuate psychological states? Functionalism provides an answer to this question.

Functionalism

Functionalism, in the sense that concerns us here, is a thesis about the ontological nature of mental states. On the functionalist view, the mind is an input-output device and mental states are functional states of that device. That is to say, mental states are internal states which play certain functional roles within the system. Functionalism yields an account of mental state individuation, an account in which mental states are individuated in terms of their causal role at something other than the physical level. On this view, "psychological-state tokens (are)... assigned to psychological-state types *solely* by reference to the causal relations to proximal stimuli ('inputs'), to proximal responses ('outputs'), and to one another." [7] Functionalism is motivated by worries about type physicalism. As Fodor puts it, "we (are) driven to functionalism ... by the suspicion that

[7] Fodor, Jerry, "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade Mecum," *Mind* 94 (1984), p. 81.

there are empirical generalizations about mental states that cannot be formulated in the vocabulary of neurological or physical theories."[8] If the internal states of a system may be functionally individuated, the behaviour of the system will lend itself to functional explanation. In functional explanation, the comportment of the system is said to be caused by the causal interactions of its functionally-defined states. Functional individuation and functional explanation are related in the following way: we explain behaviour functionally by citing the internal state(s) which play certain functional roles.

All putative explanations of intelligent behaviour face the following test: they must explain intelligent processing in terms of unintelligent processing. Success depends on it, for if one set of intelligent processes is simply re-described in terms of another set of intelligent processes, the intelligence of the behaviour has clearly been left unexplained. This is the venerable problem of lurking homunculi.

The homunculus problem is a worry about infinite regresses in our explanations of intelligent behaviour. We start out with a bit of intelligent behaviour and we attribute its intelligence to some internal process. But if all we do is posit an internal processor (the homunculus) that is itself intelligent, we will then need to explain the intelligence of the homunculus. Positing further homunculi will not do; it can't be homunculi "all the way down." At some point, intelligent processes must be explained in terms of processes that are not themselves intelligent.

Ryle's version of the homunculus problem is instructive. On Ryle's account, intellectualism fails in its explanations of behaviour because it unwittingly posits an infinite regress of homuncular "rule-followers." The intellectualist "explains" intelligent behaviour by arguing that it is caused by an internal process involving the rule-governed manipulation of propositions. But what, asks Ryle, is performing the task of manipulating these propositions? What is performing the function of *applying* the rules to the propositions?

---

8 Fodor, *Representations*, p. 25.

As Ryle points out, the application of rules is a procedure which can itself be done intelligently or stupidly. We need a second-order rule to guide the application of the first-order rules. And third-order rules, and forth-order rules, and so on, infinitely. The only way to halt this regress would be to posit an intelligent homunculus. Since homunculi are against the rules, we are left with the regress and without an explanation of intelligent behaviour. It looks as though functionalism may be in the same sort of trouble.

The problem is that functions are cheap. Fodor makes the point by asking us to consider the following scenario. Humans have the capacity, at least on occasion, to provide true answers to questions. The functionalist might explain such a capacity by positing a "universal question-answering device." [9] When provided with input in the form of a question, the device supplies output in the form of a true response. It looks like a perfectly good functional story in that it confines itself to specifying a function from inputs to outputs. As Fodor notes, however, this kind of story is worse than question-begging. The problem is that there is no conceivable candidate for the mechanistic implementation of such a function. Only a homunculus would do the trick. If functionalism is to be taken seriously, there must be a principled way to bar the positing of such chimerical functions. Homunculi and pseudo-explanations must not be allowed.

The general goal is to explain intelligent process in terms of unintelligent processes. For the functionalist, this amounts to providing an account of psychological functions in terms of unintelligent mechanistic processes. This responsibility is the flip side of the functionalist's autonomy. Physicalism has no problem with mechanism; if psychological states are neural states, psychological explanation can avail itself of the wealth of causal mechanisms available at the neural level. In achieving its independence from the physical level, functionalism incurs the responsibility of finding its mechanisms elsewhere.

---

[9] Fodor, *Representations*, p. 12.

Here, making a grand entrance at long last, is the concept of computational mental states. The introduction of computational mental states is motivated by the desire to halt the regress and slay the homunculus by specifying the requisite unintelligent processes in mechanistic terms. Recall that cognitivism is the combination of representationalism and computationalism. As we will see, the account of computational mental processes that lies at the heart of cognitivism is imported from computational functionalism. It is important for our purposes, therefore, to get clear on the functionalist's notion of computationalism.

The functionalist aims to explain intelligent behaviour by positing psychological functions that are mechanistically realizable and thus causally efficacious. The challenge is met by arguing that psychological functions are *identical* to computational functions. On this view, the mind computes input-output relations in the same way that a computer computes input-output relations. Turing has shown that if a function is computable, it is computable on a universal Turing machine. A Turing machine is an abstract description of a input-output device. The machine itself is exhaustively described in terms of its input-output table. Given that Turing machines compute their input-output functions syntactically, and given that such functions are mechanistically realizable, if psychological functions can be identified with Turing machine functions, the mechanistic realizability of psychological functions is guaranteed. The possibility of specifying psychological functions in the language of Turing machine programs looks promising since, as Fodor points out, the mind is often conceived as some kind of symbol manipulation device. Fodor thus offers the following cheerful analysis.

> Suppose that, through some significant range of cases, mental processes are functionally specified by reference to operations on some kind of symbol-like objects. Then we know, at very least, that there exists for each such mental process

a corresponding Turing machine process, and that there exists for each such Turing machine process a mechanical realization. [10]

These days, Turing machines are not the only types of computers of interest to functionalists. Whatever type of computer is invoked, however, the basic theoretical point remains the same. Computational functionalists analyze mental causation in terms of the causal efficacy of the syntax of mental symbols. Computational functionalism thus constitutes one answer to the question of how we can account for mental causation within a materialist framework. On the functionalist view, mental states are individuated in virtue of their causal properties at the computational level.

A Progress Report

It is time to check the road map - to see where we have been so far and where we need to go. More importantly, the notion of progress implicit in the tale must now be made explicit. Behaviourism, physicalism, and functionalism are three distinct attempts to answer the following question: how can one give an account of mental causation that is consistent with the principle of ontological materialism. In some sense, all three theories are allegedly "successful" - each purportedly meets the challenge of providing a materialistic account of mental causation. Yet we construe physicalism as an improvement on behaviourism and functionalism as an improvement on physicalism. From whence this sense of progress?

Implicit in this tale of philosophical advance is the idea that we already have a good account of intelligent behaviour: folk psychology. The central assumption of folk psychology is that human agents are rational, that actions are caused by mental processes in which agents manipulate their beliefs and desires logically. If Pat desires a loaf of bread and believes that there are loaves of bread at the dépanneur, she will, all things being equal,

[10] Fodor, *Representations*, p. 14.

make her way to the corner store. This common sense view of intelligent behaviour is, on Fodor's account, the best account on offer. What is needed, says Fodor, is a scientific vindication of folk psychology.

Behaviourism fails because it denies that mental processes cause behaviour. Type physicalism fails because it does not preserve the generalizations of folk psychology. Token physicalism is a step in the right direction, for it allows us to acknowledge the legitimacy of individuating psychological states independently of their physical instantiations. It does not, however, provide a specific account of how psychological states are to be individuated. Functionalism provides such an account. On the functionalist view, mental states are individuated with respect to their computational causal properties. Folk psychology is not yet vindicated, however, because the functionalist claims that causal properties depend solely on syntactic properties; the semantic features of beliefs and desires, so important on the folk psychological view, are irrelevant to the functionalist's causal account of behaviour generation. Content must somehow be put back into the picture.

The Representational Theory of the Mind

The raison d'être of Fodor's Representational Theory of the Mind (RTM) of the mind is the vindication of folk psychology. Folk psychology is based on the principle that intelligent behaviour is caused by the manipulation of beliefs and desires (or "propositional attitudes" [11]). In order to vindicate folk psychology, we must show that there is a *scientific* psychology whose ontological commitments include a commitment to theoretical entities which have the same properties as the properties the folk ascribe to beliefs and desires. What properties do the folk attribute to propositional attitudes? The two most critical, as

[11] The term "propositional attitude" derives from the following sort of analysis. The belief that Ottawa is the capital of Canada may be thought of as an attitude of believing toward the proposition "Ottawa is the capital of Canada."

far as Fodor is concerned, are semantic evaluability and causal efficacy. Propositional attitudes have content and they cause behaviour. So in order to vindicate folk psychology, we need to find a scientific psychology that is committed to there being theoretical entities which are both semantically evaluable and causally efficacious. As it turns out, such a theory is readily available in the form of the RTM and the RTM just happens to be, on Fodor's view, the best scientific account of behaviour in town.

> There is already in the field a (more or less) empirical theory that is, in my view, reasonably construed as ontologically committed to the attitudes and that, again, in my view - is quite probably approximately true. If I'm right about this theory, it is a vindication of the attitudes. Since, moreover, it's the only thing of its kind around (it's the only proposal for a scientific belief / desire psychology that's in the field), defending the commonsense assumptions about the attitudes and defending this theory turn out to be much the same enterprise; extensively, as one might say. [12]

Fodor thus devotes himself to the task of "defending the commonsense assumptions about the attitudes," noting that this task is more or less equivalent to the task of defending the RTM (or "intentional state psychology"). The claim that must be defended, in either case, is the following: the internal states that figure in mental processes must be both semantically evaluable and causally efficacious. Showing that the propositional attitudes of the folk have content is relatively "easy". If anything is semantically evaluable, beliefs and desires are; propositional attitudes have content paradigmatically. Showing that the intentional states of intentional state psychology have causal efficacy is also relatively straightforward; the entities over which a science generalizes are causally efficacious and intentional state psychology generalizes over intentional states. The trick is to show how propositional attitudes can have causal

---

12 Fodor, Jerry, *Psychosemantics* (Cambridge, Mass.: MIT Press, 1987), p. 16.

efficacy, or, if you prefer, how intentional states can be semantically evaluable. The name of the game is to show how the same internal state can be *both* semantically evaluable and causally efficacious. Fodor accomplishes this feat by combining, in the RTM, a representational theory of the mind and a computational account of mental processing.

As we have seen, the computer metaphor figures in functionalist accounts of the mind. Functionalists recognized that if mental states were identified with computational states, a causal account of mental state processing could be generated via a causal account of syntactic processing. Fodor, in his RTM, gleans a second important insight from the computer metaphor. In a computer, syntactic symbols may be interpreted to have semantic content, and the rules of operations are designed in such a way that the syntactic manipulation of internal symbols respects the logical relationships among the propositions represented by these symbols. In other words, the machine operations of a computer are designed to be truth-preserving rules of inference.

> The machine is so designed that it will transform one symbol to another if and only if the symbols transformed stand in certain semantic relations; e.g. the relation that the premises bear to the conclusion in a valid argument. Such machines - computers, of course - just *are* environments in which the causal role of a symbol token is made to parallel the inferential role of the proposition it expresses. [13]

By identifying mental states and computational states, Fodor shows how mental states can be both semantically evaluable and causally efficacious AND how syntactic processing can be made to mirror semantic processing.

The RTM consists of two claims. The first is a claim about the nature of propositional attitudes. The second is a claim about the nature of mental processes. Fodor argues that an agent possesses a particular propositional attitude if and only if the agent bears a computational relation to a mental representation whose content is identical to the

[13] Fodor, "Fodor's Guide to Mental Representation," p. 93.

content of the propositional attitude. (So, for example, Jean has the belief that it is going to rain if and only if she bears the right sort of computational relation to a mental representation whose content is "it is going to rain.") Fodor formalizes this principle of representationalism in Claim 1 as follows:

> Claim 1 (the nature of propositional attitudes): For any organism O, and any attitude A toward the proposition P, there is a ('computational'/'functional') relation R and a mental representation MP such that MP means that P, and O has A iff O bears R to MP.[14]

Mental processing, on this view, amounts to the serial tokening of mental symbols endowed with both syntactic and semantic properties. Claim 2 of the RTM is the assertion that "mental processes are causal sequences of tokenings of mental representations."[15]

Given Fodor's agenda, the claim that mental states are computational states is clearly worth defending. He thus needs to provide us with a good reason for supposing that mental states are indeed computational states. At this point, Fodor simply directs our attention to the similarities between thought processes and computational processes. Both resemble arguments; both are intended to be truth-preserving. According to Fodor, "what you want to make thinking worth the while is that trains of thought should be generated by mechanisms that are generally truth-preserving so that a true inference (generally) suggests other inferences *that are also true*."[16] Computational mechanisms provide a model of the requisite "truth-preserving" mechanisms.

The RTM is concerned with mental representations of a very particular sort. The mental representations that figure in intentional states are, on Fodor's account, tokens of a "Language of Thought" (LOT). The LOT Hypothesis will prove to be a critical component in my argument against Fodor's representational account of perceptual-motor skills, so it is important to get clear on exactly what LOT entails. The first thing to note is that Fodor

---

14 Fodor, Jerry, *Psychosemantics* (Cambridge, Mass.: MIT Press, 1987), p. 17.

15 Fodor, *Psychosemantics*, p. 17.

16 Fodor, "Fodor's Guide to Mental Representation," p. 92.

makes life difficult by using the term "LOT" in two distinct ways, thus inviting a great deal of slippage in philosophical discussions about LOT's virtues and vices. We can avoid this sort of problem by distinguishing between what I will call the "weak LOT Hypothesis" and the "strong LOT Hypothesis".

In some contexts, Fodor uses the phrase "the LOT Hypothesis" as though it were synonymous with Claim 1 of the RTM. Fodor introduces Claim 1 by noting that "at the heart of the theory (the RTM) is the postulation of a language of thought: an infinite set of 'mental representations' which function both as the immediate objects of propositional attitudes and as the domains of mental processes." [17] We can construe propositional attitudes as relations to representations because the mental representations of the language of thought are both the objects of the attitudes and the syntactic symbols over which computational mental processes operate. I will refer to the claim that propositional attitudes are relations to representations as the "weak LOT Hypothesis".

A stronger version of the LOT Hypothesis, one which is of critical importance to the argument of this thesis, is discussed in detail in the appendix to *Psychosemantics*. In order to understand the strong version of LOT, we must get clear on the relationship between Fodor's RTM and "Intentional Realism". Intentional Realism, according to Fodor, is the view that "psychological explanations need to postulate a network of causally related intentional states." [18] The RTM thus qualifies as a version of Intentional Realism. But, as Fodor explains, the RTM goes beyond Intentional Realism in that it specifies that the mental representations constitutive of intentional states must have the following two features: (1) internal structure and (2) transportable parts. This claim about the nature of mental representations is the strong version of LOT. I will henceforth refer to representations that have internal structure and transportable parts as "Fodorian representations". We will look carefully at the requirements of strong LOT in a moment,

---

17 Fodor, *Psychosemantics*, pp. 16-17.

and investigate just what Fodor means by "internal structure" and "transportable parts". The point to emphasize here is that Intentional Realists may agree with the weak LOT Hypothesis and object vigorously to the strong LOT Hypothesis. In the following passage, Fodor discusses his claim that Intentional Realists will find nothing objectionable in the weak version of LOT. He refers to his interlocutor, the Intentional Realist, as "Aunty".

> What - over and above Intentional Realism - does the Language of Thought Hypothesis *claim*? Here, I think, the situation is reasonably clear. To begin with, LOT wants to construe propositional attitude tokens as relations to symbol tokens. According to standard formulations, to believe that *P* is to bear a certain relation to a token of a symbol which means that *P*. (It is generally assumed that tokens of the symbols in questions are neural objects, but this assumption won't be urgent in the present discussion.) Now, symbols have intentional contents and their tokens are physical in all known cases. And - qua physical - symbol tokens are the right sorts of things to exhibit causal roles. So there doesn't seem to be anything that LOT wants to claim *so far* that Aunty needs to feel uptight about. What, then, exactly is the issue? [19]

Aunty accepts the weak version of LOT in that she accepts the following sort of analysis of intentional states. To say that Laura believes that Montréal is south of Québec City is for Laura to have a computational relationship to a mental representation whose content is "Montréal is south of Québec City." The *content* of this mental representation has internal structure, i.e. the structure of the proposition in question. For Aunty, however, the mental representation itself does not have any internal structure. Fodor argues, in contrast, that the mental representation does have internal structure, a structure that reflects the structure of the propositional content of the representation. In this example,

---

[18] Fodor, *Psychosemantics*, p. 135.

the representation whose content is "Montréal is south of Québec City" may be said to have three internal parts: "Montréal", "Québec", and "is south of". The critical difference between Fodor and non-Fodorian Intentional Realists is that only Fodor posits mental representations *with internal structure.*

> It's here that LOT ventures beyond mere Intentional Realism, and it's here that Aunty proposes to get off the bus. LOT claims that *mental states* - and not just their propositional objects - *typically have constituent structure.* So far as I can see, this is the *only* real difference between LOT and the sorts of intentional Realism that even Aunty admits to be respectable. So a defense of LOT has to be an argument that believing and desiring are typically structured states. [20]

To be clear, intentional realists accept that intentional states are relations to mental representations. This sort of internal structure is not at issue. The disagreement between Fodor and Aunty concerns the issue of whether or not the mental representations that figure in intentional states have internal structure. Fodor says yes. Aunty says no. In order for Fodor to defend his version of representationalism, then, he must demonstrate the necessity of positing mental representations with constituent structure. We will carefully examine Fodor's defense of LOT in due course. For now, I will simply note that my argument against Fodor's account of perceptual-motor skills has the following form: Fodor's account of why we are justified in positing internally structured mental representations *does not apply in the case of perceptual-motor skills.*

Fodor takes his RTM to be "a framework for a science." [21] His agenda has been to formalize the assumptions at work in cognitive psychology in order to justify the claim that these assumptions constitute a legitimate working proposal with regard to the task of explaining intelligent behaviour. "My contentions have been modest: The program is far

---

19 Fodor, *Psychosemantics*, p. 135.

20 Fodor, *Psychosemantics*, p. 136.

from fully clear, but there's no obvious reason to believe that it is fundamentally confused; the program engages issues that are abstract by anybody's standards, but there is no obvious reason to deny that it's a program of empirical research." [22]

Fodor's RTM is intended to serve as a vindication of folk psychology. Folk psychology is (allegedly) the theory of intelligent behaviour that we all use to understand the rational behaviour of our cohorts. The RTM is supposed to show how instrumental, logical, or tactical reasoning could be instantiated in the mind. We can, for the sake of argument, grant Fodor his RTM. The issue here is whether or not the RTM provides an account of all types of intelligent behaviour. Even Fodor acknowledges that the scope of the RTM is limited but, as we will see, he has an optimistic view of the range of behaviours that will admit of a cognitivist-style analysis.

In the final chapter of *The Language of Thought*, "Conclusion: Scope and Limits," Fodor identifies the two features a mental state must have in order to qualify as an appropriate explicandum for the cognitivist. First, the mental state must be caused by another *mental* state. Sensations are thus excluded, as are beliefs that are "the consequence of brute incursion from the physiological level." [23] Eat enough oysters, says Fodor, and you may well come to have certain beliefs about being in pain. But these beliefs are not caused by a mental process and thus cannot be explained within the framework of cognitivism.

Secondly, a mental state must be caused by a mental state *in a certain way*, i.e. computationally. That is to say, there must be a rational relation between the mental state and its cause. Computational manipulations of mental states respect the logical or rational relations among the propositions represented by the mental states. If a mental state is

---

21 Fodor, *The Language of Thought*, p. 199.

22 Fodor, *The Language of Thought*, p. 200.

23 Fodor, *The Language of Thought*, p. 200.

caused non-computationally by another mental state, it falls beyond the purview of cognitivism. Consider the following example provided by Fodor.

A man wishes to remember to send a message to a friend, so he puts his watch on upside down as a reminder. Later, upon noticing the position of his watch, he remembers to send the message. The belief, "my watch is on upside down" causes the belief "I should send a message to my friend," but the causation is not computational. For a causal sequence to be computational, the manipulation of belief tokens must occur in a manner that respects the logical relationships among the propositions constitutive of the beliefs. In this case, "although the mental states are causally connected, they aren't connected by virtue of the *content*; compare the cause of the man who is reminded to send a message to his friend when he (a) hears and (b) understands an utterance token of the type 'send a message to your friend.'" [24] Fodor notes that we must await empirical verification of the claim that certain types of mental relations are of the "causal-but-not-computational" variety. He suggests, however, that such cases may be common and probably include instances where emotional states have causal influence on belief states. "If this hunch is right, then these are bona fide examples of causal relations between mental states which, nevertheless, fall outside the domain of (cognitive) psychological explanation." [25]

Under the heading of "causal-but-not-computational" mental processes, Fodor discusses certain mental activities that do not seem, on inspection, to be rule-governed. "Some of the most striking things that people do - "'creative' things like writing poems, discovering laws, or generically, having good ideas - don't *feel* like species of rule-governed processes." [26] As Fodor notes, appearances may be misleading; there may be implicit rule-governed procedures for creative processes that our psychological science has yet to discover. Then again, these sorts of mental processes may simply fall outside the

---

[24] Fodor, *The Language of Thought*, p. 203.

[25] Fodor, *The Language of Thought*, p. 203.

[26] Fodor, *The Language of Thought*, p. 201.

domain of cognitive psychology. "My main point ... is that the mere fact that creative mental processes are *mental* does not ensure that they have explanations in the language of psychology under *any* of their descriptions."[27] It is important to note that Fodor takes it to be an empirical question whether or not certain "suspicious" behaviours are indeed "rule-governed". Furthermore, it is important to stress that what Fodor means here by "rule-guided behaviour" is behaviour caused by a mental processes in which the agent's mental states are computationally related - i.e. "rationally" related. In *the Language of Thought*, Fodor allows for the possibility that many types of intelligent behaviour may indeed *not* have the right type of etiology. Intelligent behaviour not caused by computational mental processes is simply beyond the scope of cognitivism.

> There is no reason to believe that the kinds of mental phenomena which are thus excluded from the domain of theories of information flow are restricted to the occasional detritus of the mental life. On the contrary, some of the most systematic, and some of the most interesting, kinds of mental events may be among those about whose etiology cognitive psychologists can have nothing at all to say.[28]

On Fodor's view, the cognitivist has the theoretical tools necessary for the explanation of mental states only in cases in which the mental states are implicated in computationally defined mental processes. In other words, content matters, for such mental states must be causally interrelated in virtue of the logical relations among their propositional objects.

All of this counts as progress towards pluralism. In the old days, Ryle's task was to argue against that view that intelligent behaviour is intelligent just in virtue of its having been caused by the prior manipulation of mental representations. In those days, the cognitivist view was that intelligent behaviour just is behaviour with this sort of etiology.

---

27 Fodor, *The Language of Thought*, p. 202.

28 Fodor, *The Language of Thought*, p. 201.

Fodor's view is more ecumenical, for he takes it to be a matter of empirical investigation whether or not all the various types of behaviour we call intelligent are indeed caused by the computational processing of representations. The debate between cognitivists and non-cognitivists now takes a different form. The debate has shifted from a discussion of the whether or not all intelligent behaviour is caused by cognitivist-style mental processes to a discussion of *which types of intelligent behaviour* are caused by cognitivist-style mental processes.

On Fodor's view, cognitivism's domain is limited to the study of mental states that are computationally caused by other mental states. His discussion thus makes clear what kinds of *mental states* fall within the purview of cognitivism. He does not tell us here, however, what kinds of *intelligent behaviour* are appropriately studied by cognitivists. For a full understanding of the intended scope of cognitivism, we must turn to Fodor's discussion of an RTM-based theory of intelligence.

The Representational Theory of Intelligence

In *Psychosemantics*, Fodor discusses the relationship between the standard RTM and the representational theory of intelligence (RTI). The standard RTM borrows a "theory of rationality" from the computer metaphor. The mind, like a computer, instantiates a semantically evaluable syntactic system whose rules of operation are truth-preserving. "What RTM borrows from computers is, in the first instance, the recipe for mechanizing rationality: Use a syntactically driven machine to exploit parallelisms between the syntactic and semantic properties of symbols." [29] The "intelligent behaviour" version of the RTM goes one significant step further. "Some - but not all - versions of RTM borrow more than this; not just a theory of rationality but a theory of intelligence too." [30] It is *this* version of

---

[29] Fodor, *Psychosemantics*, p. 23.

[30] Fodor, *Psychosemantics*, p. 23.

the RTM that is critical to my project. It is important, therefore, to get clear on exactly what theory of intelligence is on offer. The limitations of cognitivism will then be more clearly visible.

The theory of intelligence that the RTI "borrows" from computationalism is based on the principles of functionalism. Given a piece of intelligent behaviour to explain, we begin by specifying a general function from input to output, from environmental and internal stimuli to the intelligent behaviour in question. This function is then broken down into a series of sub-functions, which are in turn broken down into another series of sub-functions, until we reach a level where each sub-function is elementary, mechanistically realizable. The RTI is a theory of intelligent behaviour which applies the traditional model of functional explanation to representational mental processes.

The Scope of Cognitivism: Perceptual-Motor Skills and Tacit Knowledge

The RTM is conceived as a theory which will provide a scientific justification for folk psychology. Since folk psychology serves as a means of explaining rational behaviour, the RTM is also concerned, first and foremost, with the explanation of rational thought and action. The RTI, on the other hand, may be applied to intelligent behaviour that does not involve conscious mental processing at all. It can be applied, that is, to low-order perceptual-motor skills. Fodor's favorite example of this type of competence is shoe-tying, a skill which he discusses both in *Psychosemantics* and in an early article titled, "The Appeal to Tacit Knowledge." [31] In *Psychosemantics* he defends the RTI against Dennettian criticisms. In "Tacit Knowledge" he defends the RTI against Rylean criticisms. Since shoe-tying is a reasonable example of the type of perceptual-motor skill at issue here, Fodor's discussion of it is extremely relevant.

---

[31] Fodor, Jerry, "The Appeal to Tacit Knowledge in Psychological Explanation," *Representations: Philosophical Essays on the Foundations of Cognitive Science* (Cambridge, Mass.: MIT Press, 1981), pp. 63-78.

In his article on tacit knowledge, Fodor sets out "to defend 'intellectualist' accounts of mental competences." [32] The article focuses on an intellectualist account of shoe-tying behaviour. Fodor begins with a sketch of an explanation worth quoting in full.

> There is a little man who lives in one's head. The little man keeps a library. When one acts upon the intention to tie one's shoes, the little man fetches down a volume entitled *Tying One's Shoes.* The volume says such things as "Take the left free end of the shoelace in the left hand. Cross the left free end of the shoelace over the right free end of the shoelace...," etc. When the little man reads the instruction "take the left free end of the shoelace in the left hand," he pushes a button on a control panel. The button is marked "take the left free end of a shoelace in the left hand." When depressed, it activates a series of wheels, cogs, levers, and hydraulic mechanisms. As a causal consequence of the functioning of these mechanisms, one's left hand comes to seize the appropriate end of the shoelace. Similarly, *mutatis mutandis*, for the rest of the instruction. The instructions end with the word "end". When the little man reads the word "end", he returns the book of instructions to his library. That is the way we tie our shoes. [33]

As Fodor quickly notes, the details of the explanation are surely wrong. Mechanical and hydraulic principles will have to be replaced by neurophysiological processes, for starters. It is unclear how many homunculi would be necessary and the task description for each remains to be specified. But, as Fodor says, these empirical details should be left to psychologists and neuroscientists. "What this paper will be most concerned with is the philosophical opinion that there is something *methodologically* wrong with the sort of account I have sketched." [34]

---

[32] Fodor, "Tacit Knowledge," p. 63.

[33] Fodor, "Tacit Knowledge," pp. 63-64.

[34] Fodor, "Tacit Knowledge," p. 65.

Fodor begins by addressing the Rylean complaint that this sort of explanation leads to an infinite regress. The Rylean argues that knowledge of a set of rules is insufficient for skilled performance. At a minimum, the agent must also "know how" to apply these rules. The ability to apply the rules cannot itself be identified with the possession of a set of rules for applying rules, on pain of an infinite regress. Fodor rightly argues that this criticism misses the mark, for the functional homunculi are ultimately discharged in mechanistic processes. Each homunculus in the hierarchy is given a rule to follow. At the lowest level of homunculi, we find rules that provide *elementary* instructions for *elementary* operations. "Each elementary instruction specifies an *elementary operation*, and an elementary operation is one which the normal nervous system can perform, but of which it cannot perform a proper part." [35] We cannot ask about *how* an elementary operation is performed. As Fodor notes, one may ask how to spell "add", but it makes no sense to ask how to spell "n".

> The nervous system carries out its complex operations in some way or another (i.e. by performing one or another sequence of elementary operations). But the nervous system performs elementary operations in no *way* at all; it simply performs them. If every operation of the nervous system is identical with some sequence of elementary operations, we get around proliferating men by constraining a completed psychological theory to be written in sequences of elementary instructions. [36]

The concept of elementary operations is crucial to Fodor's defense of the RTI, not only because it is useful in blocking a Rylean regress, but also because it is prominently featured in Fodor's account of how we justify particular functional explanations of particular actions. In order to get at Fodor's views on this issue we must, as he puts it, "vary the image" and turn from organic brains to computational machines. [37]

---

35 Fodor, "Tacit Knowledge," p. 66.

36 Fodor, "Tacit knowledge," p. 66.

37 Fodor, "Tacit Knowledge," p. 66.

The intellectualist argues that intelligent behaviour is produced by mental processes in which the agent consults and follows a number of instructional rules. The Rylean replies that this analysis does not work because the agent knows how to perform the actions without knowing the rules. When we tie our shoes, we do not consciously guide our movements by consulting an algorithm. Furthermore, in many cases, the agent cannot produce the rules in question even on reflection. Why then, should we suppose that such rules play a causal role in the production of the behaviour?

Fodor acknowledges that the agent does not consciously follow a set of rules. He grants that the agent may not even be able to produce the rules in question after careful thought. The agent may, on Fodor's account, know how to do X without knowing how X is done. Nevertheless, says Fodor, there are good grounds for asserting that the agent possesses these rules, in some sense, and that the behaviour is caused by the implementation of these rules. But where, we may ask, does this list of causally efficacious rules come from and why should we think that these rules are involved in the production of the behaviour?

We need to return to the issue of elementary operations. Fodor argues that if the behaviour can be simulated *optimally* in a computer then we can justifiably assert that when the behaviour is produced by an organic agent, the agent possesses and follows the rules embedded in the elementary operations of the computer. An optimal simulation of the behaviour is one in which we "specify a sequence of instructions which (a) convert the input to the box into the output from the box and (b) can be written in a way that mentions only operations that are elementary." [38] What is the relationship between the agent and these rules? Fodor argues that the agent possesses "tacit knowledge" of the rules.

> Now I want to say: if X is something an organism knows how to do but is unable to explain how to do, and if S is some sequence of operations, the specification of

---

[38] Fodor, "Tacit Knowledge," p. 75.

> which would constitute an answer to the question "How do you X?," and if an optimal simulation of the behaviour of the organism is by running through the sequence of operations specified by S, then the organism *tacitly knows* the answer to the question "How do you X?," and S is a formulation of the organism's tacit knowledge.[39]

Fodor notes that this analysis "leaves us in a slightly curious conceptual situation."[40] The problem is that it is discomforting to infer from facts about a machine simulation to the assertion that there is some epistemic relationship between an organism and some proposition or rule. Fodor's defense of this move is based on the principle of inference from like effects to like causes. If two behaviours B_1 and B_2 are of the same type, and B_1 is caused by an event of description D, we can, says Fodor, reasonably assume that B_2 is caused by an event e' that is accurately described by D as well.

> If machines and organisms can produce behaviours of the same type and if descriptions of machine computations in terms of the rules, instructions, etc. that they employ are true descriptions of the etiology of their output, then the principle that licenses inferences from like effects to like causes must license us to infer that the tacit knowledge of organisms is represented by the programs of the machines that simulate their behaviour.[41]

There are several problems with Fodor's analysis. The first concerns Fodor's claim that if a behaviour can be optimally simulated then we can infer that whenever that behaviour is produced, it is produced via the implementation of the elementary operations of the simulation. Even if we were to grant this claim for the sake of argument, we should still take notice of the fact that simulating perceptual-motor skills is hardly a *fait accompli*. It is clear in the history of AI that scientists have found the more basic exercises of

---

39 Fodor, "Tacit Knowledge," p. 75.

40 Fodor, "Tacit Knowledge," p. 75.

intelligence to be the most difficult to simulate. Computers could play some version of chess, in a fashion, before they could move a simple block. Perceptual-motor skills are notoriously difficult to simulate on classical machines. In the next chapter, we will review Dreyfus' discussion of why progress in AI research on basic perceptual-motor skills has been so slow and painstaking, and examine his arguments concerning the futility of attempting to model some aspects of basic intelligent comportment on classical machines. For the moment, however, it is enough to note that the inference that begins "if it can be simulated optimally in a computer..." may simply be counter-factual in a significant number of cases, particularly in the realm of perceptual-motor skills.

Let's assume, for the moment, that we can produce an optimal simulation of shoe-tying behaviour. Fodor's invocation of the principle of inference from like effects to like causes is problematic in its own right. In order to see why, we must first clarify Fodor's notion of an *optimal* simulation. If we take Fodor at his word, an optimal simulation is any simulation in which each of the posited functions is mechanistically realizable. On this reading of optimality, however, we could have multiple optimal simulations of the same behaviour in the absence of a criterion for choosing among them. This is inconsistent with Fodor's agenda, for as he notes, "the intellectualist is required to say ... that there are unconscious processes of learning, storing, and applying rules which in some sense 'go on' within the organisms and contribute to the etiology of its behaviour." [42] Fodor thus needs a stronger sense of optimality, one that can generate an account of which causal story is the *right* causal story. If we opt for a stronger reading of "optimal", one that does not lead to multiple (optimal) simulations of the same behaviour, we run into a second problem. If "optimal" means optimal in any traditional sense (e.g. most efficient), we can no longer infer from like effects to like causes. This sort of optimal simulation clearly provides an answer to the question of how a *machine* might produce the behaviour. There

---

41 Fodor, "Tacit Knowledge," p. 78.

42 Fodor, "Tacit Knowledge," p. 69.

are good reasons to suppose, however, that humans do not produce the behaviour in the same way. In his book *Microcognition*, Andy Clark stresses that the human brain is the product of evolution, not engineering. Evolution is a matter of "tinkering", not design. If what we want is an account of what actually goes on in the production of perceptual-motor skills, optimal simulations should be construed as a particularly unpromising source of inspiration.

> Gradualism requires that each structural step in the evolutionary process involve only a small adjustment to the previous state. Jacob compares evolution to a tinkerer who must use whatever is immediately at his disposal to achieve some goal. This case contrasts with that of an engineer who, within certain limits, decides on the appropriate materials and design, gathers them, and then carries out the project. The point, then, is that what the tinkerer produces is heavily dependent on his historical situation in a way in which the engineer's product is not.[43]

Optimal simulations may provide a model of the behaviour. There are good reasons to suppose, however, that optimal simulations do not provide a reliable picture of the actual etiology of the behaviour in question. We should ask whether the neural processes that subserve shoe-tying behaviour are aptly described in a computational terminology that is based on the machinations of a serial processing device, a device which processes representations. As I will discuss presently, we have good neuroscientific reasons for thinking that perceptual-motor skills are *not* caused by the processing of representations.

Finally, there is the question of why we might be motivated to apply the RTI to basic perceptual-motor skills in the first place. At the heart of the RTI is the RTM, and at the heart of the RTM is folk psychology. Whatever one's views are on the virtues and vices of the RTM, it is reasonably clear what motivates the theory. There are patterns in high-order behaviour that deserve explanation. One way of explaining the patterns is to

[43] Clark, Andy, *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing* (Cambridge, Mass.: MIT Press, 1990), p. 71.

assume that agents are more or less rational, that they possess a number of beliefs and desires, and that they are capable of acting upon an appreciation of the logical relations among beliefs and desires. The RTM formalizes these commonsense intuitions.

In the case of perceptual-motor skills, however, there is no analogous commonsense view that requires formalization or justification. The folk may claim that Pat is tying her shoes because she believes that untied laces are dangerous and she prefers not to live dangerously. Beyond these sorts of claims, the folk have little to say on the issue of shoe-tying behaviour production. The folk view gives us a partial description of the causal factors involved in shoe-tying; it sheds light on the agent's *reasons* for tying her shoes. It does not, however, illuminate the causal processes involved in the generation of the perceptual-motor acts that are constitutive of this skill.

Representationalism is well-motivated when our explicandum is rational thought. Even if we assume that the RTM is successful in explaining high-order behaviour, there are no obvious reasons for thinking that the representational framework that the RTI imports from the RTM is an appropriate framework for the study of perceptual-motor skills. In "The Appeal to Tacit Knowledge," Fodor assumes that since perceptual-motor skills are a form of intelligent behaviour, they must be caused by the manipulations of representations. The assumption is, however, unfounded.

Cognitive simulation was a popular research program in the 1960's. As Dreyfus points out in *What Computers Still Can't Do*, [44] the simulation paradigm is no longer very popular among computer scientists. Fodor himself no longer speaks in terms of computational simulations of behaviour. In *Psychosemantics*, Fodor presents an updated version of his views on perceptual-motor skills, in general, and shoe-tying behaviour, in particular, and it is to these more recent remarks that we now turn.

---

[44] Dreyfus, Hubert, *What Computers Still Can't Do* (Cambridge, Mass.: MIT Press, 1993).

The Scope of Cognitivism à la Psychosemantics

The RTM's principle of representationalism has been criticized by those who claim that the identification of "propositional attitudes" with "relations to mental representations" simply does not hold in a significant number of cases. These critics cite cases in which there are propositional attitude tokens but no corresponding mental representations, and cases in which there are relations to mental representations with no corresponding propositional attitudes. In the context of addressing the latter complaint, Fodor returns to the example of shoe-tying homunculi. Before turning to that specific example, I will discuss Fodor's general strategy for replying to his critics.

The alleged problem with the RTM is that its principle of representationalism is too strong. In its original form, the principle of representationalism requires both of the following:

1A. "For each tokening of a propositional attitude, there is a tokening of a corresponding relation between an organism and a mental representation."[45]

1B. "For each tokening of that relation, there is a corresponding tokening of a propositional attitude."[46]

1A states that if an agent has a particular propositional attitude (for example, a particular belief), then (i) there must be a mental representation tokened at the causal level whose propositional content is identical to the content of the belief and (ii) the agent must have the right sort of computational relation to the mental representation. We can represent claim 1A as follows:

$$\mathbf{1A}:\ (x)[PA^{i}_{m}(x)] \rightarrow (\exists y)[MR_{m}(y) \bullet R^{i}xy]$$

---

[45] Fodor, *Psychosemantics*, p. 20.

[46] Fodor, *Psychosemantics*, p. 20.

For all x, if x has a propositional attitude (PA) of kind i (e.g. believing) with content m, then there exists a y such that y is a mental representation (MR) with content m and x bears relation R' to y. For the sake of convenience, I will use the following shorthand label for this claim: [PA → (R - MR)]. We can represent claim 1B as follows:

$$1B: \ (x)(\exists y)[(MR_m(y) \bullet R'xy) \rightarrow PA'_m(x)]$$

For all x, if there exists a y such that y is a mental representation (MR) with content m and x bears a relation R' to y, then x has a propositional attitude (PA) of kind i with content m. For the sake of convenience I will use the same sort of shorthand for claim 1B: [(R - MR) → PA].

The problem is that, strictly speaking, both 1A and 1B are false. As Fodor puts it, "the equivalence fails in both direction."[47] Fodor is forced to acknowledge the difficulties with **1A** and **1B** because Dennett, among others, has provided convincing counter-examples to both. In this context, I am primarily interested in the alleged counter-examples to IB (for example, the shoe-tying case), in which we seem to have relations to representations but no propositional attitudes. Fodor's discussion of the IA cases is, however, directly relevant to my criticisms of Fodor's treatment of IB cases. Both types of counter-examples will, therefore, be examined.

Fodor's strategy is to argue that (1) the RTM need apply only in "core cases" of propositional attitude attribution and (2) none of the alleged counter-examples to the RTM is a core case. It is incumbent upon Fodor, therefore, to specify what counts as a core case and why. Fodor derives his account of what counts as a core case from the second claim of the RTM - the claim that "mental processes are causal sequences of transformations of mental representations."[48] From this it follows that if a propositional

[47] Fodor, *Psychosemantics*, p. 20.

[48] Fodor, *Psychosemantics*, p. 24.

attitude *is tokened as an episode in a mental process*, then an explicit representation of its content must be tokened at the causal level. Since the point is important, it is worth putting Fodor's wording on record here. We find in the text, four consecutive restatements of the same point:

> It follows that tokenings of attitudes *must* correspond to tokenings of mental representations when they - the attitude tokenings - are episodes in mental processes.
>
> If the intentional objects of such causally efficacious attitude tokenings are *not* explicitly represented, then RTM is simply false.
>
> I repeat for emphasis: If the occurrence of a thought is an episode in a mental process, then RTM is committed to the explicit representation of its content.
>
> The motto is therefore No Intentional Causation without Explicit Representation. [49]

He describes the "defense by appeal to core cases" as a reasonable defense in instances where science purportedly vindicates common sense views. We say, for example, that chemistry vindicates many common sense beliefs about water - i.e. the belief that water exists, that it is water flowing in the rivers, and that water can quench one's thirst. In picking out samples of water, and in attributing certain properties to water, the folk are, for the most part, quite right. Note, however, that when the chemist speaks about water, the chemist is speaking about "chemically pure" samples of water - pure H_2O. The samples of water picked out by the folk are rarely, if ever, samples of pure H_2O. The discrepancy can be explained by noting that chemists are interested in "core cases" - in pure samples. The folk know how to pick out samples of water well enough, even if they never pick out the core cases as identified by the chemist. "What chemistry does is reconstruct our commonsense categories *in what the theory itself identifies as core cases: chemically pure* water is H_2O." [50] Similarly, common sense tells us how to pick out propositional

---

[49] Fodor, *Psychosemantics*, pp. 24 - 25.

[50] Fodor, *Psychosemantics*, p. 21.

attitudes and how to attribute them to agents even though the propositional attitudes we do pick out may not conform to the specifications of the theorist.

How then, is the concept of a core case related to the task of vindicating folk psychology? We vindicate folk psychology by showing that folk psychological explanations line up with, for the most part, our scientific explanations, by showing that the beliefs and desires of folk psychology are, for the most part, relations to mental representation. There may well be cases in which the beliefs that figure in folk psychological explanation turn out not to be identical to relations to representations. In such cases, the contents of the beliefs are not explicitly represented in the mental processes that the scientist must posit to explain the behaviour in question. Furthermore, in some cases, the contents of the mental representations that are explicitly represented in the mental processes posited by the scientist are not tokened as the contents of the beliefs that figure in folk psychological explanations. All this is as it should be, says Fodor. What the RTM requires is that *in core cases*, "what common sense takes to be tokenings of propositional attitudes are indeed tokenings of a relation between an organism and a mental representation." [51] As long as the alleged counter-examples to the RTM are not among the core cases, the RTM is safe. In other words, there may well be instances of attitudes without representations, and representations without attitudes. These instances would be problematic for the RTM only if they turned out to be among the core cases for which RTM must assume responsibility. Fodor states that these "other cases - where you get either attitude tokenings without the relation or relation tokenings without the attitudes - the theory treats as derivative. This is all, I repeat, *exactly* what you'd expect from scientific precedent." [52]

---

[51] Fodor, *Psychosemantics*, p. 21.

[52] Fodor, *Psychosemantics*, p. 21.

Alleged Counter-Examples to 1A: [PA → (R - MP)]

Here is Dennett's proposed counter-example to 1A. Keep in mind that on Fodor's view, the case qualifies as a genuine counter-example to claim 1A, [PA → (R - MP)], but is consistent with the modified claim 1A*, [PA *is an episode in a mental process* → (R - MP)].

> In a recent conversation with the designer of a chess-playing program I heard the following criticism of a rival program: 'It thinks it should get its queen out early.' This usefully ascribes a propositional attitude to the program in a very useful and predictive way, for as the designer went on the say, one can usually count on chasing the queen around the board. But for all the many levels of explicit representation to be found in the program, nowhere is anything roughly synonymous with 'I should get my queen out early' explicitly tokened. The level of analysis to which the designer's remarks belongs describes features of the program that are, in an entirely innocent way, emergent properties of the computational processes that have 'engineering reality'. I see no reason to believe that the relation between belief-talk and psychological-process talk will be any more direct. [53]

Dennett's chess case highlights the fact that the folk attribute propositional attitudes to agents on the basis of perceived patterns in the agent's behaviour. Given that the computer's behaviour is consistent with the rule in question, it is sensible and "useful" for the folk to attribute to the computer a propositional attitude whose content is "get your queen out early." We might choose to say, as does Dennett, that the computer does indeed act upon the rule. "Dennett's point is that the program actually operates on this principle ('get your queen out'); but not in virtue of any tokening of any symbol that represents

[53] Dennett, Daniel C., "A Cure for the Common Code?," *Brainstorms: Philosophical Essays on Mind and Psychology* (Cambridge, Mass.: MIT Bradford Books, 1981), p. 107.

it."[54] Since no such rule is tokened in the computer's program, we have a violation of 1A: [PA → (R - MP)].

Dennett's point can also be put in the following way: just because an agent's behaviour is consistent with a rule, it does not follow that the rule is explicitly represented in the agent's mental processes. The rule may be simply "emergent" out of the computer's program. Thus on Dennett's view, the RTM errs in licensing an inference from rule-consistency to rule-representation, so to speak. While the unmodified RTM is vulnerable to such a charge, the modified RTM (which incorporates the core principle) is, on Fodor's view, safe.

The claim [PA → (R - MP)] is vulnerable because there is nothing in the formulation that ensures that the propositional attitude in question plays a causal role in the production of the behaviour. If the attitude does not play a causal role, we should not expect to find the content of the attitude represented at the syntactic level. Fair enough. What Fodor needs is a re-formulation of the claim 1A that specifies that the expectation of (R - MP) is warranted only when the PA in question plays a causal role. The formulation [PA is an episode in a mental process] is intended to make such a requirement explicit. To say that a propositional attitude is an episode in a mental process is just to say that the propositional attitude is an element in a "causal sequence of tokenings of mental representations."[55] On Fodor's view, Dennett has simply misunderstood the RTM.

> The allegation is that intentionalist methodology permits the inference from "x's behaviour complies with rule R" to "R is a rule that x explicitly represents." ... But in fact no such principle of inference is assumed. What warrants the hypothesis that R is explicitly represented is not mere behaviour in compliance with R; it's an

---

54 Fodor, *Psychosemantics*, p. 22.

55 Fodor, *Psychosemantics*, p. 17.

etiology according to which R figures as the content of one of the intentional states whose tokenings are causally responsible for x's behaviour. [56]

The RTM claims that if the propositional attitude whose content is "get the queen out early" is tokened in a mental process (i.e., if the propositional attitude is causally efficacious), *then* the content of the propositional attitude will be explicitly represented. [57] Fodor notes that in Dennett's example, the relevant propositional attitude is *not* tokened in the computer's mental processes. As a result, the example does not count as a core case and the RTM is off the hook.

> Thus, in Dennett's chess case, the rule 'get it out early' may or may not be expressed by a 'mental' (/program language) symbol. That depends on just how the machine works; specifically on whether *consulting* the rule is a step in the machine's operations. I take it that in the machine Dennett has in mind, it isn't; *entertaining the thought 'Better get the queen out early'* never constitutes an episode in the mental life of the machine. But then, the intentional content of this thought need *not* be explicitly represented consonant with 'no intentional causation without explicit representation. [58]

---

56 Fodor, *Psychosemantics*, p. 156.

57 A point of clarification is in order. The RTM requires that if a rule is the content of a causally efficacious propositional attitude, the rule must be explicitly represented in the computer's program. In other cases, rules need not be explicitly represented. As we have seen, rules that are emergent out of the computer's program will not be explicitly represented. Furthermore, hard-wired rules are not represented in the computer's program. The general principle to keep in mind is that "according to the RTM, programs - corresponding to the 'laws of thought' - *may* be explicitly represented; but 'data structures' - corresponding to the contents of thoughts - have to be." (Fodor, *Psychosemantics*, p. 25.)

58 Fodor, *Psychosemantics*, p. 25.

Alleged Counter-Examples to 1B: [(R - MP) → PA]

As has been noted, the RTM's analysis of shoe-tying behaviour involves the positing of many tiers of decreasingly intelligent homunculi. At the lowest level, we find homunculi that are discharged into elementary operations of the form (if A then do B), where A and B are, in theory, causally efficacious tokenings of mental representations. This form of explanation spells trouble for claim 1B: [(R - MP) → PA]. Here is the problem in Fodor's words.

> At the very top are states which may well correspond to propositional attitudes that common sense is prepared to acknowledge (knowing how to tie one's shoes, thinking about shoe tying). But at the bottom and middle levels there are bound to be lots of symbol-processing operations that correspond to nothing that *people* - as opposed to their nervous systems - ever do. [59]

Fodor has returned to the same issue that occupied him in his earlier article on tacit knowledge. In "Tacit Knowledge" Fodor was concerned with two questions: (1) how do we justify the claim that an agent's behaviour is caused by the rules that are generated in an optimal simulation of the behaviour? and (2) how do we describe the relationship between the agent and the rules so generated? Fodor replies, in effect, that we read the relevant rules off of optimal simulations of the agent's behaviour and then stipulate that the agent has "tacit knowledge" of these rules. As I suggested in my discussion of "Tacit Knowing," there are serious problems with Fodor's response. First, optimal simulations of perceptual-motor skills are difficult to come by. Second, even if we could generate an optimal simulation of a particular perceptual-motor skills, we cannot infer that the neural mechanisms responsible for the deployment of this skill in humans are in any way like the computational mechanisms responsible for the deployment of the skill in a computer. In

[59] Fodor, *Psychosemantics*, p. 24.

Psychosemantics, Fodor is interested in the same set of issues, though he no longer speaks in terms of optimal simulations. The general features of his analysis, however, have not changed. He still claims that such behaviour can be explained by invoking a form of hierarchical functionalism. The question, then, is the following: is Fodor's defense of hierarchical functionalism now more persuasive? Unfortunately, his comments are brief, ambiguous, and dissatisfying. Here, in its entirety, is Fodor's response to the claim that $[(R - MP) \rightarrow PA]$ is false.

> What about the ... examples where you have mental representation tokenings without attitude tokenings? Commonsense belief/desire explanations are vindicated if scientific psychology is ontologically committed to beliefs and desires. But it's not also required that the folk-psychological inventory of propositional attitudes should turn out to exhaust a natural kind. It would be astounding if it did; how could common sense know all that? What's important about RTM - what makes RTM a vindication of intuitive belief/desire psychology - isn't that it picks out a kind that is precisely coextensive with the propositional attitudes. It's that RTM shows how intentional states could have causal powers; precisely the aspect of commonsense intentional realism that seemed most perplexing from a metaphysical point of view.[60]

In this passage, Fodor elects to preserve the folk psychological usage of the term "propositional attitude". This clarifies an important issue. In order for an example to count as a core case, two conditions must be met. First, the case must involve the sorts of propositional attitudes generated by the folk and second, these propositional attitudes must be tokened in mental processes; they must, that is, be causally efficacious, Fodor's formulation of the core principle bears repeating here. The RTM requires that "what common sense takes to be tokenings of propositional attitudes" must be "tokenings of a

---

[60] Fodor, *Psychosemantics*, p. 26.

relationship between an organism and a mental representations."[61] In the shoe-tying example, the folk do not attribute propositional attitudes to the agent. As a result, shoe-tying does not count as a core case. Fodor is untroubled by this result. He emphasizes that we should *expect* folk psychology and the RTM to diverge. That is indeed the motivation for all his talk of non-core cases.

In 1A cases, the folk attribute propositional attitudes that turn out not to be causally efficacious. In 1B cases, the folk do not attribute propositional attitudes to the agent at all. What matters, for Fodor, is that we identify the right causal account of the behaviour in question. The RTM provides an explanation of *how* it is that the folk are right, when in fact they are. What matters is that the RTM "shows how intentional states could have causal powers."[62]

Fodor claims that since the shoe-tying case is not a core case, he need not worry about it. I think Fodor should be worried about perceptual-motor skills in general, and about shoe-tying in particular. He needs to provide an account of why we should explain perceptual-motor skills in terms of representational processes. So far, we have two reasons for thinking that a representational account of shoe-tying is unmotivated. First, a functionalist account of shoe-tying provides us with a set of rules with which human shoe-tying behaviour is consistent. But as Fodor himself emphasizes in his discussion of Dennett's chess-playing computer, rule-consistency does not entail rule-representation. A functionalist account of shoe-tying may provide a good causal story about shoe-tying production mechanisms in computers. We still have no reason to believe that computers and human beings tie shoes in the same way, no reasons to believe, that is, that the causal mechanisms implicated in functionalist accounts are the same causal mechanisms at work in the production of shoe-tying in humans. Second, the goal of the RTM theorist is the vindication of folk psychology. The view that folk explanations are worthy of vindication

---

61 Fodor, *Psychosemantics*, p. 21.

is prompted by the observation that the folk enjoy a great deal of predictive success. In the shoe-tying case, however, Fodor has lost his bearings; there are no folk explanations in need of vindication. We should ask, therefore, the following question: has Fodor provided a cogent account of why we should explain perceptual-motor skills in terms of internal processes that operate over representations?

Fodor's Defense of the RTM: Arguments for LOT

In the appendix to *Psychosemantics*, Fodor offers several arguments in defense of the RTM, arguments to the effect that in order to explain intelligent behaviour we *must* posit internally structured mental representations that are causally efficacious and semantically evaluable. I will argue that *none* of these arguments apply to skilled perceptual-motor behaviour. Fodor's interlocutor is the Intentional Realist, Aunty. Aunty wants to argue that the internal states causally responsible for the production of intelligent behaviour do not have internal structure. Fodor argues, in contrast, that intelligent behaviour is caused by the manipulation of internally structured mental representations. To be fair to Fodor, we should note up front that his agenda is to show that *at least some* types of intelligent behaviour are caused in this way. Since Aunty is claiming that *no* forms of intelligent behaviour are caused by the processing of internally structured representations, this is all he need do. My view is that *some* types of intelligent behaviour are *not* caused by the processing of Fodorian representations. Fodor's comments are thus not directed my way. On the other hand, some of the comments he makes in his arguments against Aunty can be usefully incorporated into my argument for the claim that there is no justification for applying the RTM to perceptual-motor skills.

Aunty accepts the following two claims: (1) beliefs have internal structure and (2) the internal states that are causally efficacious in the production of intelligent behaviour are

---

62 Fodor, *Psychosemantics*, p. 26.

semantically evaluable. She denies, however, that these internal states have internal structure. Fodor argues for the claim that these internal states, the ones Aunty admits are both causally efficacious and semantically evaluable, *are* internally structured. This state of affairs lends itself to a bit of terminological confusion, confusion which we might as well clear up sooner rather than later.

There are two ways of describing the debate between Fodor and Aunty. It doesn't really matter which we choose as long as we stick with it. (Fodor does not. Thus the confusion.) Either we say that Fodor and Aunty disagree about whether or not mental representations have constituent structure, or we say that since mental representations have constituent structure, Fodor and Aunty are fighting over whether or not there are mental representations. Consider the following passage in which Fodor gives us his account of one of Aunty's proposals. (The proposal, which is not itself of interest here, concerns the "right" way of construing mental processes.)

> Notice that this ... account, though it recognizes mental states with their intentional contents, does not recognize mental representations. Indeed, the point of the proposal is precisely to emphasize as live for Intentional Realists the option of postulating representational mental states and then crying halt.[63]

Aunty, on this view, accepts representational mental states but rejects mental representations. Elsewhere, Fodor speaks as though Aunty accepts the claim that there are mental representations and rejects only the claim that they have constituent structure. It strikes me as far too proprietary of Fodor to suggest that representational mental states are not mental representations unless they have constituent structure, so I will proceed as follow. Fodor and Aunty agree that there are internal states that are causally efficacious and semantically evaluable. Let's call them mental representations. The issue on the table, then, is whether or not mental representations have constituent structure.

---

[63] Fodor, *Psychosemantics*, pp. 144-145.

Fodor offers three arguments for the claim that mental representations have constituent structure: a methodological argument, an argument from "systematicity", and an argument based on an analysis of mental processes. The second and third arguments are so closely interrelated that I will discuss them together under the heading of "the argument from systematicity". I will show that neither the methodological argument nor the argument from systematicity applies to perceptual-motor skills.

Argument 1: The Methodological Argument

Here's the argument:

P1. If a behaviour is complex, its cause is complex.

P2. Some behaviour is complex.

C. In some cases, the cause of behaviour is complex.

An event, entity, or behaviour is "complex" just in case it has constituent structure. Fodor takes it to be an empirical fact that "behaviour does - very often, exhibit constituent structure." [64] He notes that complex behaviour includes, but is not limited to, linguistic behaviour. What interests me in this argument is the quantifier "some" in the second premise. We will, however, bracket this issue for the moment. Fodor defends the first premise by citing the *methodological* principle.

> Principle P: Suppose there is a kind of event c1 of which the normal effect is a kind of event e1; and a kind of event c2 of which the normal effect is a kind of event e2; and a kind of event c3 of which the normal effect is a complex event e1 and e2. Viz:
>
> $c1 \rightarrow e1$
>
> $c2 \rightarrow e2$

---

[64] Fodor, *Psychosemantics*, p. 143.

$c3 \rightarrow e1 \& e2$

Then, ceteris paribus, it is reasonable to infer that c3 is a complex event whose constituents include c1 and c2.[65]

If Fodor's argument goes through, Aunty is in trouble. If complex behaviour behaviour has complex causes, Aunty does not have the resources to explain complex behaviour. Fodor counts all this as a point for LOT, for while the mental representations of LOT are complex, the mental representations of Intentional Realism are not.

Let's assume that Fodor's argument goes through and return to the issue of the quantifier in the second premise - *some* behaviour is complex. Any comments Fodor makes about what kinds of behaviour are not complex would certainly be relevant here. Fodor notes that psychologists make a distinction between "segmented" behaviour and "synergisms". Segmented behaviours are complex; they have constituent structure. "Synergisms", on the other hand, "are cases where what appear to be behavioural elements are in fact 'fused' to one another, so that the whole business functions as a unit."[66] Fodor triumphs over Aunty, on his own view, because "it's empirically quite clear that not all behaving is synergistic."[67] That means that some behaviour is segmented (complex) and thus that the causes of some behaviours are complex. "If as a matter of fact, behaviour is often segmented, then principle P requires us to prefer the theory that the causes of behaviour are complex over the theory that they aren't, all else being equal."[68]

Fodor does not linger on the issue of synergistic behaviour, but he does pause long enough to provide one example of it: "when a well-practiced pianist plays a fluent

---

65 Fodor, *Psychosemantics*, p. 141.

66 Fodor, *Psychosemantics*, p. 143.

67 Fodor, *Psychosemantics*, p. 143.

68 Fodor, *Psychosemantics*, p. 143.

arpeggio." [69] The production of an arpeggio is clearly a perceptual-motor skill. Note that in the process of learning this skill, the pianist will most probably engage in behaviours that are complex. Principles concerning the proper positions of the arms and hands may be reviewed and may play a causal role in the early, deliberative performances of the arpeggio. The pianist may weigh the relative merits of different fingering sequences. The individual finger movements involved in the arpeggio may, for a time, be well-defined and distinct - not yet "fused", so to speak. But once the skill is acquired, all the movements of the fingers, hands, and arms do become "fused". Indeed, the performance breaks down entirely if the pianist analyzes and re-evaluates the individual movements that constitute the skilled playing of the arpeggio while playing the arpeggio. [70] The acquired perceptual-motor skill of arpeggio playing is synergistic.

It is a general feature of perceptual-motor skills that they begin as complex behaviours and are gradually transformed, with practice, into synergisms. We cannot give a causal account of a synergistic behaviour by providing a causal account of the earlier complex behaviour required for the acquisition of the skill. In his discussion of shoe-tying, Fodor makes just this mistake. He posits a tier of homunculi that are responsible for tasks such as "grasp the right lace" and "loop the right lace over the left lace." These procedures may well be involved when someone *first* learns how to tie her shoes. They are *not* involved in the production of skilled, synergistic shoe-tying behaviour.

Fodor argues for the claim that *some* behaviours are complex and thus have complex causes. What I want to show is that the RTM does not apply to perceptual-motor skills. Fodor helps my cause considerably by noting that the perceptual-motor skill of arpeggio playing is not complex. Since there are no compelling reasons to think that arpeggio playing is a peculiar perceptual-motor skill in this regard, we can conclude that

---

69 Fodor, *Psychosemantics*, p. 143.

the RTM does not apply to perceptual-motor skills and that Fodor's account of 1B cases (perceptual-motor skills such as shoe-tying) is misguided. He says only that the divergence of folk psychology and the RTM is not a cause for concern. What he should have said is that perceptual-motor skills fall beyond the purview of the RTM.

The Argument from Systematicity

Fodor notes that the standard argument invoked in defense of the claim that mental representations have internal structure is an argument that cites the productivity of propositional attitudes. Propositional attitudes are productive in that "there is a (potentially) infinite set of - for example - belief-state types, each with its distinctive intentional object and its distinct causal role." [71] Productivity is "immediately explicable" if we help ourselves to the assumption that mental representations have "combinatorial structure"; if we assume that mental representations are "somehow built up out of elements and that the intentional object and causal role of each state depends on what elements in contains and how they are put together." [72]

Fodor's view is that we need to posit LOT in order to account for the productivity of thought and language. The problem with the argument from productivity is, however, that it involves controversial idealizations. "The facts of mortality being what they are, not more that a finite part of any mental capacity ever gets exploited. So it requires idealization to secure the crucial premise that mental capacities really *are* productive." [73] If Aunty wants to block the conclusion that the LOT Hypothesis is justified because it provides an

---

70 For further discussion of this point, see Michael Polanyi's *Personal Knowledge: Towards a Post-Critical Philosophy* (Chicago: The University of Chicago Press, 1958), especially Chapter Four, "Skills".

71 Fodor, *Psychosemantics*, p. 147.

72 Fodor, *Psychosemantics*, p. 147.

73 Fodor, *Psychosemantics*, p. 148.

account of productivity, she can, as Fodor puts it, "simply refuse to idealize." [74] Fodor's response to this worry about idealization is to develop another argument for LOT, one which does not depend on idealization: the argument from systematicity.

Systematicity is a difficult term to define. It is typically used to refer to a property exhibited by linguistic capacities, a property in virtue of which the ability to understand or produce one sentence is "intrinsically connected" to the ability to understand or produce many other sentences. [75] We say that linguistic capacities are systematic because there is data showing, for example, that there are no speakers of English who understand the sentence "John loves Mary" but who fail to understand the sentence "Mary loves John." Systematicity is that feature of linguistic capacities in virtue of which anyone who understands the first sentence will also understand the second.

Fodor defends the claim that mental representations have constituent structure by arguing that cognitive capacities are systematic. Here's the argument.

P1. Cognitive capacities are systematic.

P2. If cognitive capacities are systematic, then mental representations have constituent structure.

C. Mental representations have constituent structure.

Fodor claims that the second premise requires very little defense. "I get the second claim for free for want of an alternative account." [76] The only way to explain the systematicity of cognitive capacities is, on Fodor's view, to assume that mental representations have constituent structure. With respect to the first premise, the claim that cognitive capacities are systematic, Fodor offers the following "fast argument."

P1 If linguistic capacities are systematic, then cognitive capacities are systematic.

---

74 Fodor, *Psychosemantics*, p. 148.

75 Fodor, *Psychosemantics*, p. 149.

P2. Linguistic capacities are systematic.

C. Cognitive capacities are systematic.

Fodor defends P1 by noting that "cognitive capacities must be *at least* as systematic as linguistic capacities, since the function of language is to express thought." [77] Fodor defends P2 by arguing that (1) if natural languages have constituent structure then linguistic capacities must be systematic, and (2) natural languages do have constituent structure.

The claim that "mental representations have constituent structure" thus depends on the claim that cognitive capacities are systematic. Fodor summarizes his argument as follows. "Linguistic capacities are systematic and that's because sentences have constituent structure. But cognitive capacities are systematic too, and that must be because *thoughts* have constituent structure. But if thoughts have constituent structure, then LOT is true. So I win and Aunty loses. Goody." [78]

Let's look more carefully at Fodor's claim that Intentional Realism alone does not have the resources to explain systematicity. It is in virtue of the systematicity of linguistic capacities that a speaker can transform the sentence "John loves Mary" into the sentence "Mary loves John." It is in virtue of the systematicity of cognitive capacities that an agent can perform the following inferential feat. If told that Kim is taller than Pierre, and that Laura is taller than Kim, the agent can infer that Laura is taller than Pierre. In order for the speaker to transform "John loves Mary" into "Mary loves John," and in order for the agent to engage in inferential reasoning, mental representations must have constituent structure AND mental processes must operate over the constituent parts of these mental representations. In the inference case, for example, the representation of Kim must be tokenable in both premises. The representation of Laura must be tokenable in both the second premise and in the conclusion. The semantic content of each representation must

---

[76] Fodor, *Psychosemantics*, p. 151.

[77] Fodor, *Psychosemantics*, p. 151.

remain constant or context-independent. Furthermore, mental processes that can *transport* these representation-parts must be available. On Fodor's view, if Aunty does not posit mental representations with constituent structure, she will have no account of mental processing. "As things stand now, the cost of not having a Language of Thought is not having a theory of thinking." [79] How does Aunty *purport* to explain such things as productivity and systematicity. On Fodor's view, Aunty takes refuge in the practice of citing "Unknown Neurological Mechanisms."

> You can imagine a story - vaguely Gibsonian in spirit - according to which cognitive capacity involves a sort of "tuning" in the brain. What happens, on this view, is that you have whatever experiences engaged such capacities, and the experiences have Unknown Neurological Effects (these Unknown Neurological Effects being mediated, it goes without saying, by the corresponding Unknown Neurological Mechanisms), and the upshot is that you come to have a very large - but finite - number of, as it were, *independent* mental dispositions. [80]

Aunty may be onto something, but in order to see just what it is, we need to put Fodor's argument from systematicity into a broader context. Fodor's argument for LOT is an argument to the best explanation. In order to explain the systematicity of cognitive capacities, we must posit internally structured mental representations. If a capacity is not systematic, however, there is no justification for positing internally structures mental representations. Such is the case with skilled perceptual-motor behaviours. Think in terms of the "set" of all behaviours that are skilled perceptual-motor behaviours. If each individual behaviour is synergistic, if each individual behaviour lacks constituent parts, the set of behaviours cannot exhibit a "combinatorial semantics", so to speak. In such a case, there are no grounds for claiming that the capacity that underlies the set of behaviours is

---

78 Fodor, *Psychosemantics*, p. 151.

79 Fodor, *Psychosemantics*, p. 147.

80 Fodor, *Psychosemantics*, p. 148.

systematic. And if the capacity underlying a certain kind of behaviour is not systematic, there is no justification for positing internally structured mental representations in explanations *of that kind of behaviour*. Perceptual skills are not complex; they do not have constituent structure. We therefore cannot assume that the capacities underlying perceptual-motor skills are systematic. There is therefore no justification for the claim that perceptual-motor skills are caused by the processing of internally structured mental representations. Fodor errs in including shoe-tying behavior and other perceptual-motor skills in the domain of phenomena that can be explained by appeal to the RTM.

Fodor is interested in cognitive behaviour, complex behaviour whose explanation requires the positing of internally structured mental representations. The virtues and vices of his explanation of cognitive behaviour are not at issue here. What is at issue here is the *alleged scope* of the RTM. My argument against Fodor is not an argument against the RTM per se, it is an argument against the claim that the RTM is applicable to perceptual-motor skills. My claim is that Fodor provides no support for his assertion that perceptual-motor skills are caused by the manipulation of Fodorian representations.

How might we explain the production of perceptual-motor skills? This is where Aunty's conception of brain "tuning" becomes relevant. Fodor dismisses all talk of Unknown Neurological Processes as unhelpful. Science continues apace, however, and these Unknown Neural Processes are not as unknown as they used to be. In fact, as we will soon see in more detail, neuroscientists make a distinction between "cognitive" memory systems and "noncognitive" memory systems. The cognitive system, which is implicated in high order "complex" thought, involves the manipulation of neurally realized representations with constituent structure. The non-cognitive system, which subserves perceptual-motor skills, involves the gradual tuning of performance systems throughout the brain. Fodor offers no cogent philosophical support for the claim that perceptual-motor skills are caused via the manipulation of internally structured mental representations. Neuroscientists, on the other hand, provide powerful support for the claim that perceptual-motor skills are subserved by neural processes that involve the gradual "tuning" of performance systems, *not* the manipulation of neurally realized Fodorian representations.

Arguing about the nature of perceptual-motor skills is not high on Fodor's list of priorities. In fact, it is fair to say that he doesn't much care who wins this game and his disinterest shows in his performance. The home team, on the other hand, thinks that this game is *very* important. All we ask, really, is that Fodor confine himself to the sort of phenomena that lend themselves to an RTM analysis. Perceptual-motor skills clearly do not.

Chapter Two: Phenomenology and Background Skills

The cognitivist account of agency is based on an analysis of the rational thought processes of the disengaged agent. On the cognitivist view, intelligent behaviour is caused by the computational processing of mental representations. Furthermore, the cognitivist assumes that this sort of internal processing is causally responsible for both rational action and skilled perceptual-motor behaviour. In the last chapter, I argued that the application of a representational theory of the mind to perceptual-motor skills is problematic, even from within the world view of cognitivism. In this chapter, I turn from cognitivism to phenomenology, from Fodor to Dreyfus. Dreyfus' account of agency, which he calls Heideggerian "Skill Pragmatism", will provide an additional set of reasons for calling into question the assumption that perceptual-motor skills are caused by representational mental processes.

Cognitivists and phenomenologists speak to each other across a formidable divide, one that is not easily negotiated. I am interested in one particular area of disagreement: the question of whether or not perceptual-motor skills are the product of representational mental processes. The cognitivist says yes. The phenomenologist says no. This dispute is difficult to adjudicate, as cognitivist and phenomenologists have differing views on what counts as successful philosophical discourse. In this chapter, I will outline the skill-pragmatic view of intelligent behaviour by (1) discussing Taylor's account of the differences between cognitivist and neo-Heideggerian conceptions of agency and (2) discussing Dreyfus' interpretation of Heideggerian background skills. I will then distinguish between those aspects of skill-pragmatism that are most likely to seem compelling *to a cognitivist*, and those aspects for which the cognitivist is likely to offer important objections. In Chapters Four and Five, I will argue that the skill-pragmatist can respond to these objections by citing recent neuroscientific work on multiple memory systems.

Taylor identifies two different conceptions of the self that figure in contemporary philosophy: the disengaged self of cognitivism and the engaged self of neo-Heideggerian phenomenology. [1] Among the many differences between these two views of the self, one looms particularly large. The neo-Heideggerian, but not the cognitivist, focuses on what Taylor calls the "significance feature" of agency. The term "significance feature" will require a great deal of unpacking, most of which will be done only when we turn to Dreyfus' work on Heidegger. For the moment, I will use it more or less as a place-holder. The point to emphasize here is that *the* critical difference between cognitivists and phenomenologists is that while cognitivists systematically eliminate the significance feature from their accounts of intelligent behaviour, phenomenologists take it as their starting point.

According to Taylor, the concept of the disengaged self is a byproduct of the scientific revolution of the seventeenth century. The triumph of the natural sciences led to the valuation of detachment, objectivity, and logical reasoning. These virtues were embraced, not only as properties of well-conducted science, but as features of general human agency at its finest. The scientist, detached and rational, became a model for our conception of the self. A critical figure in the development of the ideal of the disengaged self was Descartes, who urges us to seek clarity of thought through detachment from the physical world. Detachment creates the distance required for the acquisition of clear mental representations. Furthermore, it fosters the sort of objectivity required for the rational or logical processing of these representations. "The liberation through objectification wrought by the cosmological revolution of the seventeenth century has become for many

---

1 Taylor, Charles, *Human Agency and Language: Philosophical Papers*, Vol. I (Cambridge: Cambridge University Press, 1985).

the model of the agent's relation to the world, and hence sets the very definition of what it is to be an agent." [2]

Natural science, in keeping with its objectivity requirement, must identify and set aside the "subjective" features of its explicanda. Thus were secondary properties, for instance, banished from the domain of the natural science in the seventeenth century. Such properties were deemed to be unsuitable for scientific investigation because they depend upon the subjective experience of an observer. "They concerned purely the significance of things for us, not the way the things were." [3] As Taylor notes, the decision to identify and bracket "subjective" properties paved the way for great scientific advances. He states that the "eschewal of anthropocentric properties was undoubtedly one of the bases of the spectacular progress of natural science in the last three centuries." [4]

In the twentieth century, the methodology of the natural sciences has been imported into our study of human behaviour. The success conditions of the inquiry are clear; we must eschew all references to anthropocentric qualities in the interest of providing a properly scientific account of the agent's behaviour. In philosophy, the movement toward objectivism has produced a distinctly modern conception of the human agent - one in which the agent's mind causes behaviour via the computational manipulation of representations in the context of a significance-free world.

Taylor notes that this conception of agency has a distinct *moral* appeal. It portrays the agent as an autonomous being that exerts responsible control over its life via the formulation and implementation of rational plans for the achievement of well-specified and explicitly represented goals. "The ideal of disengagement defines a certain typically

---

2 Taylor, *Human Agency and Language: Philosophical Papers*, Vol. I, p. 5.

3 Taylor, Charles, "The Concept of a Person," *Human Agency and Language: Philosophical Papers*, Vol. I (Cambridge: Cambridge University Press, 1985), p. 106.

4 Taylor, "The Concept of a Person," p. 106.

modern notion of freedom, ... its own peculiar notion of human dignity." [5] Taylor argues that an allegiance to these moral virtues underpins the commitment to the disengaged construal of the self. The construal should thus be afforded a certain degree of respect.

> The disengaged identity is far from being simply wrong and misguided and besides, we are all too deeply imbued with it to repudiate it. The kind of critique we need is one that can free it of its illusory pretensions to define the totality of our lives as agents, without attempting the futile and ultimately self-destructive task of rejecting it altogether. [6]

The sort of critique Taylor has in mind involves demonstrating that the "significance feature", for all its incompatibility with objectivism, cannot be simply omitted. Neo-Heideggerians argue that while some forms of intelligent behaviour can be understood in terms of the disengaged self, certain forms of skillful comportment require that we think in terms of the engaged self. The agent is thus said to be capable of two very different types of relationships to the world: disengagement and engagement. Engagement is the more primordial of the two; indeed it is a necessary precondition of disengagement. For that reason, an accurate account of agency must not only acknowledge the phenomenon of engagement, it must accord it a place of honour. If we begin with the practices of the engaged agent, the "significance feature" comes back into view.

The "significance feature" is a property of the relationship between an embodied agent and the world it inhabits. The skill-pragmatist claims that "significance" relationships arise in the everyday practices of what Heidegger calls engaged coping. *Au fond*, these practices consist of a nexus of perceptual-motor skills that manifest a certain know-how, a preliminary grasp of the possibilities for action afforded by the environment. This "grasp" of the world is practical, non-representational. A brief discussion of one example of

---

5 Taylor, *Human Agency and Language: Philosophical Papers*, Vol. I, p. 5.

6 Taylor, *Human Agency and Language: Philosophical Papers*, Vol. I, p. 7.

engaged coping will shed light on the relationship between engaged coping and significance. The oft-used example of Heideggerian hammering is useful in this context.

Let us imagine that we want to understand the behaviour of an agent who is immersed in the project of constructing, say, a wooden box, and who needs to drive in a number of nails. For such an agent, a nearby hammer will "show up" in a particularly salient way. In describing the hammer, we could begin by listing its so-called "objective properties" - its size, weight, and morphology. A moment's reflection reveals, however, that such properties are relevant to the agent only insofar as they are constitutive of more significant anthropomorphic qualities: of-the-right-shape-for-this-job, light-enough-for-me-to-manipulate, heavy-enough-to-use-in-driving-in-these-nails, etc. These qualities belong to the hammer in virtue of its relationship to the agent. They are a product of the agent's "directedness", her "thrownness" into a particular project, and the "serviceability" of the hammer with respect to this project.

A first pass understanding of the significance feature can be had, then, by reflecting on the relationship between embodied agents and their practical dealings with useful objects in the environment. An agent's understanding of the world is grounded in her *dealings* with it, not in her internal representations of it. The way in which these objects appear to the agent is critical; we cannot understand the agent's comportment if we bracket out the object's anthropomorphic qualities. The phenomenologist invokes the image of the engagement in order to help us envision the agent in something other than Cartesian terms. The agent is not detached. The subject-object dichotomy does not apply. Most importantly, the relationship between agent and world is direct; it is not mediated by the interpolation of internal representations. Significance is not a property which can be added onto an internal representation. It is a feature of an irreducibly dyadic relationship between the agent and the world.

> We can draw a neat line between my *picture* of an object and that object, but not between my *dealing* with the object and that object. It may make sense to ask one to focus on what one *believes* about something, say a football, even in the absence of that thing; but when it comes to *playing* football, the corresponding suggestion would be absurd. The actions involved in the game cannot be done without the object; they include the object. Take it away and we have something quite different - people miming a match on the stage, perhaps. The notion that our understanding of the world is grounded in our dealings with it is equivalent to the thesis that this understanding is not ultimately based on representations at all, in the sense of depictions that are separately identifiable from what they are of. [7]

The self is an active agent whose grasp of the world cannot be exhaustively described by itemizing its internal representations. While the knowledge of the disengaged agent may be portrayed in "knowing-that" terms, the understanding of the engaged agent is a form of active and embodied knowing-how. "But this puts the role of the body in a new light. Our body is not just the executant of the goal we frame. Our understanding itself is embodied. That is, our bodily know-how, and the way we act and move, can encode components of our understanding of self and world." [8] As has been mentioned, the skill-pragmatist acknowledges that agents have both engaged and disengaged relationships to the world. With respect to disengagement, the phenomenologist accepts the importance of representational processing. The caveat, however, is that representational processing is always dependent on the agent's prior engagement with the world; it is dependent on the agent's background skills. If we focus exclusively on high-order representational thought, as the cognitivist tends to do, these underlying background

[7] Taylor, Charles, "Overcoming Epistemology," *After Philosophy: End or Transformation*, Ed. Kenneth Baynes, James Bohman, and Thomas McCarthy (Cambridge, Mass.: MIT University Press, 1987), p. 477.

[8] Taylor, Charles, "To Follow a Rule...," Cahiers d'épistémologie, Cahier no. 9019 (Montréal: Université du Québec à Montréal, 1990), p. 6.

practices are lost from view and the significance feature is overlooked. This is why, on the skill-pragmatic view, it is absolutely crucial to enlarge the scope of inquiry, to include non-representational perceptual-motor skills in our list of intelligent behaviours.

> There is all the difference in the world between a creature with and one without the significance feature. It is not just a detachable feature that action has in some medium of internal representation, but is essential to action itself. The supposedly secondary, dispensable character of the significance feature disappears once we cease to dwell on that small range of actions which have plausible machine analogues. [9]

The computer model of intelligent behaviour seems plausible only if we confine our inquiry to high-order representational thought. The problem with the machine paradigm is that it does not have the resources to acknowledge the significance feature of agency. Whatever similarities there may be between machines and human beings, it is clear that machines do not enjoy significance relationships with the world. Computers do not become "immersed" in engaged coping. The computer is a paradigm of disinterest; objects and events do not *matter* to a computer. "Once you see the importance of the significance feature, it is evident that computing machines can at best go some of the way to explaining human computation, let alone intelligent, adaptive performance generally." [10] The burden of the phenomenological argument against cognitivism, then, is to show that the significance feature is an essential element of at least some forms of intelligent behaviour. The phenomenologist wants to counter the claim that the significance feature is a "merely" subjective property that must be put aside for the sake of objectivity.

> The crucial distinction to understand the contrast between us and machines is not mental/physical, or inner/outer, but possessing /not possessing the significance

---

[9] Taylor, Charles, "Cognitive Psychology," *Human Agency and Language: Philosophical Papers*, Vol. I (Cambridge: Cambridge University Press, 1985), p. 200.

[10] Taylor, "Cognitive Psychology," p. 204.

> feature. Once we understand this, we can see that this feature cannot be marginalized as though it concerned merely the way things *appear* to us, as though it were a feature merely of an inner medium of representation. On the contrary, it plays an absolutely crucial role in explaining what we do, and hence defines the kind of creatures we are. [11]

In his article, "Overcoming Epistemology," Taylor discusses Heidegger's contributions to our understanding of the phenomenon of significance. On Taylor's view, "the tremendous contribution of Heidegger ... consists in having focused the issue properly." [12] Heidegger appreciates the fact that cognitivists *knowingly* omit the significance feature from their accounts of intelligent thought and behaviour. It won't do, then, simply to point out that the significance feature has been overlooked. A case must be made for the claim that intelligent behaviour cannot be understood in the absence of an account of the significance feature. Heidegger takes his cue from the transcendental arguments of Kant and argues that not only is significance a critical feature of intelligent behaviour, it is the *sine qua non* of all experience. With respect to the cognitivist's "epistemological construal" of agency, a construal in which a disengaged subject confronts an objective world, Taylor notes the following:

> We argue for the inadequacy of the epistemological construal, and the necessity of a new conception, from what we show to be the indispensable conditions of there being anything like experience or awareness of the world in the first place. [13]

Heidegger's phenomenological depiction of being-in-the-world is a description of the conditions necessary for experience. Once being-in-the-world is described in a phenomenologically persuasive way, we will see that representational theories of the mind err (1) in positing mental representations in contexts where there are none, and (2) in

---

11 Taylor, "Cognitive Psychology," p. 201.

12 Taylor, "Overcoming Epistemology," p. 476.

overlooking the dependence of representational cognition on non-representational background skills. Taylor states that Heidegger shows that "(the) conditions of our forming disengaged representations of reality is that we be already coping with the world, dealing with things in it, at grips with them." [14]

Dreyfus' Account of Heideggerian Coping

Dreyfus offers a detailed account of Heidegger's "transcendental arguments" against cognitivism in his recent book, *Being-in-the-World: A Commentary on Heidegger's* Being and Time, *Division I.* [15] One of Dreyfus' critical claims is that Heidegger seeks to defeat cognitivism via an implausibility thesis: once the activities of the engaged agent are accurately described, the implausibility of cognitivism will be made manifest.

The implausibility thesis is directly related to the issue of representationalism. For the cognitivist, mental representations are implicated in intelligent behaviour in a number of ways. Behaviour caused by instrumental reasoning is directed toward an explicitly represented goal. The content of causally efficacious beliefs and desires is explicitly represented in mental processes. If behaviour is rule-governed, the rules are explicitly represented. Finally, if objects in the external world are implicated in an agent's behaviour, the "objective" properties of these objects are captured in an explicit mental representation. Heidegger shows, however, that we can describe many types of intelligent behaviour in non-representational terms. As Dreyfus notes, such a demonstration does not constitute a refutation of cognitivism, but it does make clear the fact that cognitivism is not our only option. It is *not* the only game in town.

The task at hand is to sketch the basic principles of Heidegger's analysis of being-in-the-world. I will start by drawing a simplified map of Heidegger's ontology. There are

---

13 Taylor, "Overcoming Epistemology," p. 473.

14 Taylor, "Overcoming Epistemology," p. 473.

15 Dreyfus, Hubert, *Being-in-the-World: A Commentary on Heidegger's* Being and Time, *Division I* (Cambridge, Mass.: MIT Press, 1991).

three phenomena at issue here: representational cognition, local engaged coping, and being-in-the-world. For my purposes, local engaged coping is the most important of the three and will thus receive the most attention. In order to understand local engaged coping, however, we must understand its relationship to both representational cognition and being-in-the-world. The first thing to note is that all three phenomena are forms of what Heidegger calls transcendence. Transcendence, cast in the most general of terms, is a relationship that obtains between agents and objects in the world. Representational cognition and local engaged coping are forms of "ontic transcendence" while being-in-the-world *is* "originary transcendence". Heidegger introduces the term "transcendence" in order to disabuse us of the cognitivist tendency to construe all agent-object relationships in terms of mental-state intentionality. As we shall see, only representational cognition may be understood as a form of mental-state intentionality. Local engaged coping and Being-in-the-world, on the other hand, are non-representational forms of engagement which are more primordial than mental-state intentionality. Heidegger's goal is to show that representational cognition is dependent on local engaged coping, and that *both* are dependent on being-in-the-world.

Ontic Transcendence	Representational Cognition Local Engaged Coping
Originary Transcendence	Being-in-the-world

Figure 2.1: Two Kinds of Transcendence

Ontic Transcendence

In order to understand ontic transcendence we must understand originary transcendence, and vice versa. Since we have to start somewhere, and because originary transcendence is a notoriously difficult concept to grasp, I will begin with a sketch of ontic transcendence.

Ontic transcendence is the way in which Dasein relates to objects in the world. (Dasein itself is a particular "way of being" which accrues to individual human agents as they are socialized into a culture, into a way of life or a set of practices. We will examine the concept of Dasein more carefully when we discuss originary transcendence. For now, we may think of Dasein as a way of being that is equivalent to "being with other beings" [i.e. being with objects in the world, being with other agents]. Following Heidegger, I will adopt the convenience of speaking of agents who exhibits Dasein as "Dasein".)

Dasein interacts with objects in the world in a number of different ways, each of which constitutes a distinct type of "comportment". A comportment, for Heidegger, is a way of being with things. The way in which we "are" with things depends on the nature of the objects we encounter. The way in which Dasein relates to an object depends, that is, on the "way of being" of the object encountered. At the ontic level, there are two kinds of comportment, cognition and engaged coping, which correspond to two kinds of objects, occurrent objects and available objects.

Ontic Transcendence	Comportment: Cognition	Object's Way of Being: Occurrent
	Comportment: Coping	Object's Way of Being: Available
Originary Transcendence	Being-in-the-world	

Figure 2.2: Two Types of Comportment

Dasein and Available "Objects"

In everyday coping, Dasein's activity exhibits a certain flow, a directedness. This flow is not undifferentiated; it can be analyzed into a series of "projects" or endeavors, each of which is purposive. As a first pass, we may think of each of these projects as a task which is oriented towards a goal. A critical feature of engaged coping is, however, that it has no "goal" in any traditional sense. Each task is directed toward an endpoint, a "towards-which". The agent does not have the "towards-which" in mind; it is not explicitly represented. As Dreyfus points out, "phenomenological examination confirms that in a wide variety of situations human beings relate to the world in an organized purposive manner without the constant accompaniment of representational states that specify what the action is aimed at accomplishing. This is evident in skilled activity such as playing the piano or skiing, habitual activity such as driving to the office or brushing one's teeth." [16]

The point of introducing the concept of a towards-which is to help us see that an activity can be purposeful even though the agent does not possess an explicit representation of the activity's purpose. A particularly clear example of purposive activity of this sort can be found in the comportment of athletes. Dreyfus cites Larry Bird's description of the purposive act of passing a basketball to a teammate. "(A lot of the) things I do on the court are just reactions to situations ... I don't think about some of the things I'm trying to do ... A lot of the times I've passed the basketball and not realized I've passed it until a moment or so later." [17] Elsewhere Dreyfus discusses the dynamics of a tennis swing. One of the "goals" in playing tennis well is to hold the racket correctly when returning the ball to the opponent. A beginner may well rehearse certain guidelines to herself, but the accomplished player does not. As Dreyfus states, "I cannot represent how I should turn my racket since I do not know what I do when I return the ball. I may have once been told to hold my racket perpendicular to the court, and I may have succeeded in doing so, but now experience has sculpted my swing to the situation in a far more subtle and appropriate way than I could have achieved as a beginner." [18] Such skillful coping activities do not require a mental representation of their goals. The agent does not pause to deliberate. Rather, she exhibits a

---

16 Dreyfus, *A Commentary*, p. 93.

17 Dreyfus, *A Commentary*, p. 93.

18 Dreyfus, "The Hermeneutic Approach to Intentionality," p. 4.

form of pre-theoretical know-how. Experience has left her with the ability to move forward into an activity without "thinking" about what she is doing. The same analysis may be applied to many different types of purposive activities, skillful perceptual-motor coping as well as intellectual coping. The point is that, at least in some cases, we should cease to think of purposive behaviour as being directed toward an explicit goal.

> We can make sense of our own comportment, or the comportment of others, in terms of such directedness towards long-range and proximal ends. But this should not mislead us into postulating mental intentions in actions, since there is no evidence that this division into intelligible subsets of activity need be in the mind of the person who is absorbed in the activity any more than an athlete experiencing flow is purposefully trying to achieve a basket or a touchdown. The "towards-which" is Heidegger's nonintentionalistic term for the end points we use in making sense of a flow of directed activity. [19]

Given the nature of engaged coping, we must re-evaluate our practices of judging when an activity is successful. Since there are no explicit goals in engaged coping, success cannot be defined in terms of the attainment of a specific goal. The solution is to define success in terms of Dasein's directedness. Dasein's very nature is to be involved in ongoing activity; it is natural for Dasein to "press into possibilities". [20] The activity of Dasein has "non-representational success conditions" that are defined in terms of Dasein's directedness. [21] Dreyfus notes that "either the activity presses transparently into the future, or else it runs into trouble. ... Skilled activity can thus succeed or fail even though the skilled performer does not represent its success conditions." [22]

---

19 Dreyfus, *A Commentary*, p. 94.

20 Dreyfus, *A Commentary*, p. 240.

21 Dreyfus, "The Hermeneutic Approach to Intentionality," p. 6.

22 Dreyfus, "The Hermeneutic Approach to Intentionality," p. 6.

In engaged coping, Dasein is directed towards a toward-which. Dasein's stance is not therefore one of detachment or disinterest. As a result, objects in the world *matter* to Dasein. They matter insofar as they can help or hinder Dasein's activity. Proximal objects which appear to be "useful" with respect to a "towards-which" are said to be "available" or "ready-to-hand". The set of available objects in the environment is called "equipment". Heidegger states that "we shall call these entities which we encounter in our concern 'equipment'." [23] Dasein's concern reveals the utility of equipment; it reveals that "equipment is essentially something-in-order-to'." [24] Equipment is thus defined in terms of its relationship to the directedness of Dasein's purposive activity.

> In dealings such as this, where something is put to use, our concern subordinates itself to the "in-order-to" which is constitutive for the equipment we are employing at the time; the less we just stare at the hammer-Thing, and the more we seize hold of it and use it, the more primordial does our relationship to it become, and the more unveiledly is it encountered as that which it is - as equipment. [25]

It should be noted that, for Heidegger, there is no such thing as a single isolated piece of equipment. Useful objects exist in a matrix of equipment. In Dreyfus' words, "an 'item' of equipment is what it is only insofar as it refers to other equipment and so fits in a certain way into an 'equipmental whole'." [26] In a workshop, for instance, tables, hammers, nails, saws, benches, etc., are all interrelated; the function of each item is defined in relation to the functions of all the other items. Heidegger states that "what determines a

---

[23] Heidegger, Martin, *Being and Time*, Trans. John Macquarrie and Edward Robinson (New York: Harper and Row, 1962), p. 97.

[24] Heidegger, *Being and Time*, p. 97.

[25] Heidegger, *Being and Time*, p. 98.

[26] Dreyfus, *A Commentary*, p. 62.

piece of equipment as an individual is in each instance its equipmental character and equipmental nexus."[27]

What then, *is*, a piece of equipment? On Heidegger's view, the "in-order-to" quality of a piece of equipment is essential to the equipment. If we are limited to a list of the hammer's objective features, we will never arrive at a full understanding of the hammer. The hammer must be understood *as* equipment, and equipment must be understood in terms of its directionality, its involvement. Dreyfus notes that "equipment cannot be made intelligible in terms of objective substance plus use-predicates. Since equipment is in no way derivative ... we can say that equipment in use is equipment as it is in itself."[28] That a hammer is "heavy-enough-for-this-particular-Dasein-in-this-particular-task" is part of the hammer's relationship to Dasein. While this property belongs to the hammer, it cannot be captured except in an account that makes reference to both the hammer and Dasein. In order to understand the hammer, Dasein must use it, not represent it.

Our question can now be stated with greater precision: what is the relationship between Dasein and available equipment. This relationship cannot be described in terms of mental-state intentionality, for the way Dasein understands equipment is to use it. Furthermore, the relationship cannot be described in terms of the subject-object dichotomy. Both Dasein and equipment can be characterized only in terms of the communal tasks in which both are deeply involved. Dasein's "toward-which" and equipment's "in-order-to" are directionality vectors that become interlaced, so to speak, and create an irreducibly dyadic relationship between the agent and the world.

When all goes well, when the practical activity directed at a towards-which flows unimpeded, equipment becomes transparent. This means, among other things, that the

---

[27] Heidegger, Martin, *The Basic Problems in Phenomenology*, Trans. Albert Hofstadter, Revised ed. (Bloomington: Indiana University Press, 1982), p. 292.

[28] Dreyfus, *A Commentary*, p. 66.

agent becomes sufficiently absorbed in the task that she becomes unaware of the equipment she is using. Her attention is directed at the activity itself. The cane of the blind person is not itself an object of awareness, but a tool the agent uses for navigating the world. The hammer is not grasped as an objective object, but simply as a piece of equipment with a salient "in-order-to" quality. The agent, in both cases, is not attending to the equipment, but to the task, and the equipment itself becomes "unobtrusive."

> The peculiarity of what is primarily available is that, in its availableness, it must, as it were, withdraw in order to be available quite authentically. That with which our everyday dealings primarily dwell is not the tools themselves. On the contrary, that with which we concern ourselves primarily is the task - that which is to be done at the time.[29]

When practical activity is unimpeded, Dasein itself becomes transparent. Dasein is, in virtue of its involvement in an activity, "open" to the world. Goals are not explicitly represented. Equipment is used in a pre-theoretical manner. Dasein's involvement in the practical realm precludes the intrusion of reflection. There is no "room" for instrumental reasoning, for representational thought. With experience, Dasein accumulates a form of practical wisdom, a "know-how" that enables it to respond to an every-changing set of circumstances in a pre-reflective way. This know-how is accompanied by a special form of practical perception, which Heidegger calls "circumspection". Circumspection, in keeping with the transparency of engaged Dasein, is non-reflective and non-representational. It is that component of Dasein's practical comportment which grasps the utility of objects in the environment. Circumspection "discovers" the serviceability of ready-to-hand equipment. In other words, circumspection brings to light the relationship between Dasein's current "toward-which" and the "in-order-to's" of nearby equipment. The relationship is already there; circumspection makes it manifest. When the agent grasps

---

29 Heidegger, *Being and Time*, p. 99.

the hammer, she does not pause to inventory the hammer's objective properties and calculate their utility. She "sees" at a glance that the hammer will help her move forward into her activity.

> The view in which the equipmental nexus stands at first, completely unobtrusive and unthought, is the view and sight of practical *circumspection*, of our practical everyday orientation. "Unthought" means that it is not thematically apprehended for deliberate thinking about things; instead, in circumspection we find our bearings in regard to them. [30]

In unimpeded practical activity, Dasein becomes immersed in the task, involved in the task to such a degree that Dasein's "toward-which" and the equipment's "in-order-to" correspond so closely that they disappear together into the task itself. Both Dasein and equipment have, as a way of being, involvement. The agent is aware of the involvement, but not of herself, not of the equipment. She is completely absorbed in the task. Dasein is not a locus of subjective self-awareness but a "being-with" other beings, an involvement with available equipment. Dasein *is* this involvement, and the involvement comprises both the "agent" and the equipment of the "world".

> Self and world belong together in the single entity, the Dasein. Self and World are not two entities, like subject and object, or like I and Thou, but self and world are the basic determinations of the Dasein itself in the unity of the structure of being-in-the-world. [31]

Dreyfus emphasizes that the "workshop" analysis applies to a great deal of our practical lives, to "dressing, working, getting around, telling, eating, etc." [32] He notes that Dewey appreciated the difference between the "know-how" manifest in practical

---

30 Heidegger, *The Basic Problems in Phenomenology*, p. 163.

31 Heidegger, *The Basic Problems in Phenomenology*, p. 297.

32 Dreyfus, *A Commentary*, p. 67.

comportment and the sort of knowledge involved in conscious reflection. He cites the following passage from Dewey's *Human Nature*:

> We may ... be said to know by means of our habits ... We walk and read aloud, we get off and on street cars, we dress and undress, and do a thousand useful acts without thinking of them. We know something, namely, how to do them ... If we choose to call (this) knowledge ... then other things also called knowledge, knowledge of and *about* things, knowledge *that* things are thus and so, knowledge that involves reflection and conscious appreciation, remains of a different sort.[33]

Dasein's know-how does not consist in the possession of a set of rules for behaviour. And while "know-how" is "unthinking", it is neither innate nor mechanistic. The ability to manipulate equipment effectively is an acquired skill. Comportment is adaptable; it is sensitive to the changing details of the unfolding situation. Dasein's know-how is both non-representational and non-reflective, but it is a manifestation of intelligence all the same. Dreyfus states that "the description of the skilled use of equipment enables Heidegger to introduce both a new kind of intentionality (absorbed coping) which is not that of a mind with content directed towards objects, and a new sort of entity encountered (transparent equipment) which is not a determinate, isolable substance."[34]

Dasein and Occurrent Objects

We can now turn our attention to a different kind of intentionality, the kind of intentionality at issue in the relationship between Dasein and occurrent inner-worldly entities. It is in this context that Heidegger acknowledges the subject-object dichotomy. In its relation to occurrent entities, Dasein exhibits a different mode of being; it becomes a subject with mental states. When in this mode, Dasein can relate to entities in the world as

---

[33] John Dewey, *Human Nature and Conduct. An Introduction to Social Psychology* (London: George Allen and Unwin, 1922), p. 178. Cited in Dreyfus, *A Commentary*, p. 67.

objects. Heidegger cautions, however, that the basic relationship between subjects and objects is typically misunderstood. Dreyfus states that "we shall see that there are subjects and objects but that the tradition has introduced them too early in the analysis and, moreover, has misunderstood them so as to give them a foundational function they cannot perform." [35]

The type of intentionality that obtains between Dasein and occurrent entities is mental-state intentionality. This type of intentionality comes into play when the flow of Dasein's ongoing activity is impeded or interrupted. In such cases, Dasein must switch from an unreflective mode to a mode of deliberate attention. Objects go from being "available" to being "occurrent". Dreyfus states that "once ongoing activity is held up, new modes of encountering emerge and new ways of being encountered are revealed." [36] Heidegger illustrates these new modes of being in his discussion of "breakdown". A breakdown occurs when equipment becomes useless, when the tool at hand is unsuited for the task or becomes ineffective or defective. Breakdown is just one particularly clear example of a situation in which Dasein must shift to a more deliberative way of being. Dreyfus states that "although he concentrates on the special case of breakdown, Heidegger's basic point should be that mental content arises whenever the situation requires deliberate attention." [37]

There are two types of breakdown that warrant attention here: temporary breakdown and total breakdown. In temporary breakdown, objects are revealed to be unavailable, not-ready-to-hand. In total breakdown, objects are revealed to be merely present-to-hand, occurrent. The following chart indicates how these two types of breakdown figure in Heidegger's ontological map.

---

[34] Dreyfus, *A Commentary*, p. 69.

[35] Dreyfus, *A Commentary*, p. 69.

[36] Dreyfus, *A Commentary*, p. 70.

[37] Dreyfus, *A Commentary*, p. 70.

Ontic Transcendence		
Cognition / Theoretical Reflection	Total Breakdown	Object: Occurrent
Engaged Coping: Involved Deliberation	Temporary Breakdown	Object: Unavailable
Engaged Coping: Transparent Coping	No Breakdown	Object: Available

Figure 2.3: Two Kinds of Breakdown

In transparent coping, there is a natural fit between Dasein's "towards-which" and the "in-order-to" of equipment. This "assignment" of a "towards-which" and an "in-order-to" remains in the background. Circumspection grasps this assignment, but only tacitly. In cases of temporary breakdown, however, this assignment is made explicit. The agent becomes aware of it. Heidegger states that "*when an assignment has been disturbed* - when something becomes unusable for some purpose - then the assignment becomes explicit." [38] If the hammer breaks, Dasein becomes aware of the utility of the intact hammer, just in virtue of the fact that the hammer is no longer useful, no longer available. Furthermore, Dasein becomes aware of its own thwarted "towards-which", the directedness of the activity of hammering. When an assignment becomes explicit, the way of being of both Dasein and equipment changes. They become, in some sense, subject and object. It should be emphasized, however, that Heidegger's conception of subject and object is idiosyncratic.

When a temporary breakdown is sufficiently severe, a subject capable of deliberation emerges. The subject must pause, consider options, calculate strategies, plan reflectively. In some cases, the subject must envision plans that involve equipment that is not present. In reflecting on objects that are unavailable, either because they are defective or missing, Dasein possesses "mental content". We may speak of Dasein's "mental representations" in this context, so long as we keep in mind the differences between

[38] Heidegger, *Being and Time*, p. 105.

Dasein-as-subject and the traditional Cartesian subject. For Dasein, mental representations are not "purely mental"; they are dependent on the world in the following sense: representations are grounded in the agent's engaged coping. Representations figure in, and are defined by, the agent's towards-which, the agent's ongoing coping. If my hammer breaks, I may consider the possibility of borrowing Pat's hammer but conclude that her hammer is too heavy for me to use well. I may decide to use a rock to pound in the nail, as the rock can be manipulated in the right way and will not be too heavy. My understanding of these possibilities is a form of know-how that has access to the potential availability of other pieces of equipment. My "representation" of alternative tools is not a self-contained mental entity. Dreyfus states that "even when people have 'mental representations,' i.e., mental content, such as beliefs and desires, and make plans, and follow rules, etc., they do so against a background of involved activity." [39]

In temporary breakdown, objects that were formerly available become unavailable. Dasein becomes aware of the object in a new way. Certain functional attributions of the hammer come to light - the hammer is too heavy or too light, it is cracked or misshapen. These characteristics are defined in terms of Dasein's activities. Heidegger states that "anything available to us is, at the worst, appropriate for some purposes and inappropriate for others; and its 'properties' are, as it were, still bound up in these ways in which it is appropriate or inappropriate." [40] Equipment is thus no longer transparent; it's "situational characteristics" are revealed.

> Just as temporary breakdown reveals something like what the tradition has thought of as a 'subject,' it also reveals something like an 'object,' and just as the 'subject' revealed is not the isolable self-sufficient mind that the tradition assumed, but is

---

[39] Dreyfus, *A Commentary*, p. 74.

[40] Heidegger, *Being and Time*, p. 115.

involved in the world, so the 'object' revealed is not an isolable, self-sufficient, substance, but is defined by its failure to be available.[41]

On some occasions, practical activity comes to a standstill; there is a total breakdown. Work is not merely interrupted, but brought to a halt. The relationship between Dasein and the world undergoes a dramatic transformation: Dasein can adopt the state of theoretical reflection. Heidegger states that "if knowing is to be possible as a way of determining the nature of the occurrent by observing it, then there must first be a *deficiency* in our having to do with the world."[42] Theory thus presupposes practice, for it is when the practical attitude is thwarted that the theoretical attitude takes over. Dasein perceives objects differently; they are no longer merely unavailable. The object is "decontextualized"; it becomes merely "occurrent" or "present-to-hand". As Dreyfus states, "Once our work is permanently interrupted, we can either stare helplessly at the remaining objects or take a new detached theoretical stance towards things and try to explain their underlying significance."[43]

Occurrent entities are the objects left over when the involvement relations between Dasein and the world are severed. In the theoretical stance, Dasein is attuned to the "objective" properties of occurrent objects. The difference between the "aspects" of available objects and the "properties" of occurrent objects may seem, at first glance, to be subtle. When Dasein is engaged, the hammer's heaviness is revealed in a practical context. Once in the theoretical mode, Dasein appreciates the heaviness of the hammer in a new way. Heidegger states that "in the 'physical' assertion that 'the hammer is heavy' we overlook ... the tool-character of the entity we encounter."[44] Now the property of heaviness is decontextualized; the hammer has a certain weight, and the weight is

---

[41] Dreyfus, *A Commentary*, p. 76.

[42] Heidegger, *Being and Time*, p. 88.

[43] Dreyfus, *A Commentary*, p. 79.

[44] Heidegger, *Being and Time*, p. 413

independent of Dasein's concern. The new relation between Dasein and inner-worldly entities paves the way for natural science. Dreyfus notes that "once characteristics are no longer related to one another in a concrete, everyday, meaningful way, as aspects of things in a particular context, the isolated properties that remain can be quantified over and related by scientific covering laws and thus be taken as evidence for theoretic entities."[45] Dasein's way of being toward occurrent entities, which has its roots in the breakdown of practical coping, paves the way for the theoretical stance of the scientist.

Originary Transcendence

Local engaged coping is the phenomenon of primary interest here. In order to understand local coping more fully, however, we must understand its relationship with a more primordial form of coping: originary transcendence. I will begin by discussing the concept of originary transcendence in the context of the workshop, and then turn to the issue of how and why originary transcendence is equivalent to Dasein's primordial way of being: being-in-the-world.

Originary transcendence is the *sine qua non* of experience. When a particular Dasein is in a "workshop", ontic coping reveals the "availableness" of particular pieces of equipment. Before this is possible, however, Dasein must first be familiar with the entire nexus of equipment and of its own relationship to this "totality" of tools. When an item of equipment is available, its way of being is involvement. In dealing with an available tool, Dasein's way of being is involvement. But before Dasein can "discover" any particular relationship of involvement between itself and a particular piece of equipment, it must already grasp the deeper matrix of involvements in which both Dasein and the tool are implicated. Heidegger states that "as the Being of something available, an involvement is itself discovered only on the basis of the prior discovery of a totality of involvements."[46]

---

[45] Dreyfus, *A Commentary*, p. 81.

[46] Heidegger, *Being and Time*, p. 118.

Ontic circumspection reveals the relationship between pieces of equipment and Dasein's directedness; it "discovers" equipment. Originary circumspection reveals the network of relationships that define the involvement of both Dasein and equipment; it "discloses" the world. Ontic transcendence is a way of being with entities in the world. Originary transcendence is a practical understanding of this way of being.

The totality of involvements which must be grasped prior to any particular involvement is itself the "world". Dasein is so constituted that it always has a pre-thematic grasp of the matrix of involvements that constitute the world. This matrix is made explicit, however, in times of breakdown. Breakdown thus reveals both ontic transcendence and originary transcendence. Ontic transcendence is revealed in that a particular assignment of a "towards-which" and an "in-order-to" is brought to light. On a deeper, ontological level, breakdown reveals the totality of relationships that obtain between Dasein and the network of available tools, a totality which Dasein has already understood pre-thematically via originary circumspection.

> When equipment cannot be used, this implies that the constitutive assignment of the 'in-order-to' to a 'towards-this' has been disturbed. ... When an assignment to some particular 'towards-this' has been circumspectively aroused, we catch sight of the towards-which' itself, and along with it everything connected with the work - the whole 'workshop' as that wherein concern always dwells. The context of equipment is lit up, not as something never seen before, but as a totality constantly sighted beforehand in circumspection. With this totality, however, the world announces itself.[47]

Originary transcendence is a mode of being of Dasein in virtue of which anything can be discovered in the world at all. Once the world is "disclosed", equipment can be "discovered". Originary circumspection is Dasein's natural familiarity with the way of

[47] Heidegger, *Being and Time*, p. 105.

being of the world. In order for Dasein to be involved in the world, Dasein must already be familiar with it. Before individual objects can show up in particular ways for Dasein, the conditions must be right for "showing up" in general. Originary transcendence discloses the world by allowing inner worldly entities to show up at all. Dreyfus states that "disclosing as letting something be involved is originary transcendence." [48]

Dasein, in virtue of its circumspective throwness into the world, creates a "clearing", a region in which objects and activities show up as significant, as mattering. Significance just is the very structure of the world; the nexus of involvements. Dasein and the world arrive together, and immediately constitute a web of involvements that provide a background against which objects and actions matter. Dreyfus states that "significance is the background upon which entities can now make sense and activities can have a point." [49]

The term "world" must be examined with more care. There are two basic uses of "world" which correspond to two ways an entity can be "in" a world. An object can be located spatially inside a world. All physical objects are in the physical world. The chair is in the room which is in a house which is in Montréal, which is in the world. Certain types of entities, those which have Dasein as a way of being, can be in the world in a second way; they can be involved in it. There are two sense of involvements which must be distinguished here. The first sense of involvement is ontical; actors are in the world of theatre, bankers are in the world of finance, philosophers are in the world of academia. The second sense of involvement is more primordial; it is ontological. It is the way of being of all types of ontic involvement, the general structure of involvement. The worldiness of the world, as Heidegger calls it, is this primordial structure of involvement.

Dasein's way of being is to be-in-the-world, to be circumspectively absorbed in engaged coping. Dasein is involvement. We may imagine, for a moment, that this

---

48 Dreyfus, *A Commentary*, p. 106.

49 Dreyfus, *A Commentary*, p. 97.

involvement has a subjective pole that coincides with the towards-which of Dasein, and an objective pole which corresponds to the in-order-to of equipment. In the end, however, there is no distinction between Dasein and the world. Dasein is being-there, and the there of its being is the world. Dasein is not a locus of subjectivity but a centre of significance, of practical involvement in a way of life. Dasein itself is transcendence. Heidegger states that "Dasein does not sort of exist and then occasionally achieve a crossing over outside itself, but existence originally means to cross over. Dasein itself is the passage across." [50]

Dasein is a type of being for whom its own being is an issue: Dasein is self-interpreting. The self-interpretation of Dasein is not, however, a set of beliefs, but a set of practices, a way of life. Dasein, in virtue of its inherent thrownness, naturally falls into relationships with other entities in the world. The throwness of Dasein is manifest in the myriad "towards-whichs" of everyday life. These "toward-whichs" are themselves directed toward a more fundamental directionality vector - a "for-the-sake-of-which". A Dasein's for-the-sake-of-which is something like a life goal or life role, but neither of these terms is adequate. The for-the-sake-of-which of a Dasein is its self-understanding, an understanding that itself consists of a set of practices. Being a teacher, intellectual, farmer - these are some of the options Dasein confronts, some of the possible for-the-sake-of-whichs in light of which Dasein may choose to understand itself. The for-the-sake-of-which of a Dasein is its primary way of being, a way of being in light of which its daily towards-whichs make sense.

It is part of Dasein's way of being to have a pre-reflective understanding of the world, a familiarity with significance that makes possible the activity of encountering objects that matter. Dasein's understanding of the world, of the practices that constitute it, is a form of self-understanding. This self-understanding is self-interpretation. The interpreter is not carried back into the self but remains open to the world. It is Dasein's

---

50 Heidegger, Martin, *The Metaphysical Foundations of Logic*. Trans. Michael Heim (Bloomington: Indiana University Press, 1984), p. 164.

way of being to understand itself in terms of its practices. Dasein's way of being is to be with other entities, dealing with them. Dasein's self-interpretation is thus an understanding of its own involvement in the world. Dreyfus states that "the stand Dasein takes on itself, its existence, is not some inner thought or experience: it is the way Dasein acts. ... Dasein takes a stand on itself through its involvement with things and people." [51]

In summary, ontic transcendence may take the form of either mental-state intentionality or engaged coping. Both types of ontic transcendence depend on originary transcendence, which is itself a more primordial form of engaged coping. While originary transcendence is an understanding of the world, it does not consist in a set of beliefs; it consists in a set of practices in terms of which Dasein grasps both itself and the world.

With respect to the task of understanding perceptual-motor skills, the most relevant form of transcendence is engaged coping *at the ontic level*. Heidegger's analysis of the perceptual-motor skills involved in ontic engaged coping is important for two reasons. (1) He distinguishes ontic coping from mental-state intentionality. Most importantly, he demonstrates that in the case of ontic coping, no mental representations are implicated. (2) Heidegger demonstrates that the significance feature of the relationship between Dasein and equipment is an ineliminable component of engaged coping. Coping is simply unintelligible in a "significance-free" context. This means, among other things, that a representational account of background skills will always be unsatisfactory.

Phenomenological Criticisms of Cognitivism

Heidegger demonstrates that there are three types of transcendence: ontic cognition, ontic engaged coping, and originary being-in-the-world. If Heidegger's description of transcendence is accurate, only ontic cognition could be described within the framework of a representational theory of the mind. To the degree that ontic cognition is dependent on

---

[51] Dreyfus, *A Commentary*, p. 61.

both ontic and originary coping, even it cannot be *fully* captured within the framework of representationalism. Since computationalism depends on representationalism, it looks as though computational models of intelligent behaviour will be, at best, incomplete. They will be of use in modelling only those elements of ontic cognition which can be theoretically isolated from engaged coping.

In the early 1970's Dreyfus articulated a critique of cognitivism based on Heidegger's phenomenological analysis of human comportment. He analyzed certain failures in the research programs of cognitive scientists. He begins with an analysis of the patterns of failure in AI research. Workers in AI were finding that high-level rational thought was relatively easy to model in comparison with basic perceptual-motor skills. Furthermore, AI theorists confronted two problems which are clearly related to the significance feature: (1) the commonsense knowledge problem and (2) the frame problem. The commonsense knowledge problem concerns the difficulty of programming background understanding and background skills into computers. The frame problem concerns the difficulty of programming a computer in such a way that it can "comport" itself appropriately in "situations", an ability which depends on the capacity to detect the salient items in a situation or the relevant information in a data bank. In more recent discussions of the issues, both in his commentary on Heidegger and in his book, *What Computers Still Can't Do*, [52] Dreyfus concludes that the branch of cognitivism known as AI has failed in precisely the ways one might have predicted, given Heidegger's depiction of Dasein and the world.

> Having to program computers keeps one honest. There is no room for the armchair rationalist's speculations. Thus AI research has called the Cartesian cognitivist's bluff. ... Actual difficulties in AI - its inability to make progress with what is called the commonsense knowledge problem, on the one hand, and its inability to define

[52] Dreyfus, Hubert, *What Computers Still Can't Do* (Cambridge, Mass.: MIT Press, 1993).

the current situation, sometimes called the frame problem, on the other - suggest that Heidegger is right. it looks as though one cannot build up the phenomenon of the world out of meaningless elements.[53]

I want to conclude this discussion of Dreyfus by reviewing and clarifying two issues, one epistemological and one ontological, on which Heideggerians and cognitivists disagree. The first issue concerns the claim that all knowledge can be formalized and thus adequately represented in a computer. The second issue concerns the claim that the world consists of a totality of occurrent objects which can be described in objective terms and thus represented in a decontextualized fashion. In so doing, we will have the opportunity to examine some of Dreyfus' more specific comments on the nature of the commonsense knowledge problem and the frame problem.

One of Heidegger's goals is to undermine the tradition of privileging disinterested knowledge. Disinterested knowing, which is based on the framework of the subject-object dichotomy, is taken to be the paradigmatic form of intentionality. As a result, other forms of transcendence are overlooked or misunderstood. Heidegger states with regret that "every act of directing oneself towards something receives the characteristic of knowing." [54] Heidegger subverts this tradition in arguing that practice precedes theory, that engaged coping is prior to representational cognition, and that the most basic form of knowledge is embodied knowing-how, not theoretical knowing-that.

Commonsense knowledge is not propositional; it does not consist in a set of implicit beliefs or tacit rules. This is why, argues Dreyfus, AI researchers have encountered roadblocks in their attempt to model human understanding in computational terms. By 1979, AI theorists were aware of the difficulties of modelling "background" knowledge computationally and of providing the computer with some tenable means of

---

53 Dreyfus, *A Commentary*, p. 119.

54 Heidegger, *The Metaphysical Foundations of Logic*, p. 134.

detecting salience. Proposed solutions to these problems took a number of forms, including Minsky's frames, Schanks' scripts, and Winograd's micro-worlds. In each case, the goal was to supply the computer with background "information" about particular situations. It was hoped that once the computer was equipped with a sufficient amount of "information", it would be capable of detecting salience, retrieving relevant information from its own data banks quickly, and "comporting" itself more appropriately in a given "context". Unfortunately, as Dreyfus points out, these proposed solutions have not worked and on his view, these failures are not temporary setbacks for AI but signs of fundamental theoretical flaws. Know-how cannot be captured in propositional form; it cannot be explicitly represented. That is because it is a way of "dealing" with the world, not a way of representing the world. Dreyfus notes that "since our familiarity does not consist in a vast body of rules and facts, but rather consists of dispositions to respond to situations in appropriate ways, there is no body of commonsense knowledge to formalize. The task is not infinite but hopelessly misguided." [55]

The epistemological assumption - that all knowledge is formalizable - is untenable. Its ontological corollary - objectivism - is equally problematic. On the objectivist view, the world is composed of a collection of entities which admit of a completely "objective" description. [56] The world itself must be objective in order for it to supply the sort of information that can be captured in discrete representations. Dreyfus notes that "the data with which the computer must operate ... must be discrete, explicit, and determinate, otherwise, it will not be the sort of information which can be given to the computer so as to be processed by rule." [57] Since the mind is construed as a computer, the same goes for

---

[55] Dreyfus, *A Commentary*, pp. 117 - 118.

[56] See George Lakoff, *Women, Fire, and Dangerous Things*, particularly Chapter 11, "The Objectivist Paradigm," for a detailed discussion of objectivism, its relation to cognitivism, and the problems inherent in it. (Chicago: University of Chicago Press, 1987).

[57] Dreyfus, *What Computers Still Can't Do*, p. 206.

the information that constitutes the "input" to the mind. Is the world so structured as to lend itself to objectivist descriptions and explicit representations?

Dreyfus argues, following Heidegger, that cognitivists have confused the physical universe and the world of human involvement. The universe, but not the *world*, may be represented objectively. Human intelligent behaviour occurs in the world of involvement, so no matter how convenient it may be to invoke the physical universe in one's computational theory of intelligence, in the end, the "significance" world must be acknowledged. We misunderstand the world altogether if we view it as a collection of occurrent properties and occurrent objects. We misunderstand inner-worldly entities if we strip them of significance, convert them from equipment to occurrent objects.

Dreyfus, like Taylor, attributes the widespread allegiance to objectivism to a reverence for the physical sciences and a commitment to an ethics based on instrumental rationality. "The goal of the philosophical tradition embedded in our culture is to eliminate uncertainty; moral, intellectual, and practical." [58] Interpreting the world of involvement in terms of the physical universe, passing over the significance feature, creates an illusory sense of control and clarity. The relationship between the scientist and the physical universe is illicitly *read back* into the relationship between the human agent and the world of practice.

Merleau-Ponty discusses this "read-back" phenomenon, as Taylor calls it, in his discussion of a patient who, due to brain injury, is locked into the scientific stance, the stance of disinterested knowledge. Merleau-Ponty's point, as we shall see, is that this stance of disengagement may be suitable for science, but it is abnormal and pathological in everyday comportment.

Schneider, the patient in question, confronts a world very much like the objective world posited in cognitivism. When Schneider is shown common, everyday items, he

---

58 Dreyfus, *What Computers Still Can't Do*, p. 211.

cannot interact with them naturally. When shown a fountain pen, for example, Schneider recites a list of the pen's objective properties. He notes that it is long, shiny, and pointed. He then infers from these objective properties to possible uses of the item; he notes, for example, that the object is probably an instrument of some type, perhaps a pencil or a pen. As Merleau-Ponty notes, such neuropsychological cases are evidence for the claim that the stance of disinterest is derivative; it is a diminishment of the agent's natural powers.

> This procedure contrasts well with, and by so doing throws into relief, the spontaneous method of normal perception, that kind of living system of meanings which make the concrete essence of the object immediately recognizable. ... It is this familiarity and communication with the object that is here interrupted. In the normal subject, the object "speaks" and is significant.[59]

Merleau-Ponty speaks of the patient's injury in terms of the severing of the "arcs of intentionality" that normally wed the agent and the world. Scientists may seek to sever these arcs of transcendence in order to achieve some degree of objectivity. Schneider's case, and others like it show, however, that disinterest is not the normal state of affairs. In everyday comportment, transcendence is a necessary precondition for normal, purposive activity. For Schneider, the destruction of his primordial way of being-with entities in the world has left him in a world without meaning.

> The world in its entirety no longer suggests any meaning to him and conversely the meanings which occur to him are not embodied any longer in the given world. We shall say, in a word, that the world no longer had any *physiognomy* for him.[60]

If we begin with the sorts of objects that figure in Schneider's world, or in the world of natural science, we will never arrive at an understanding of the world of significance. While you can arrive at an occurrent object by taking equipment and

---

[59] Merleau-Ponty, Maurice, *Phenomenology of Perception*, Trans. Colin Smith (London: Routledge & Kegan Paul, 1962), p. 131.

subtracting its significance, you cannot arrive at equipment by starting with occurrent objects and "adding" significance. The stance of disinterest diminishes the world. The world is not the totality of occurrent entities; it is the matrix of involvements that unite Dasein with the where of its being.

If we are to understand perceptual-motor skills correctly, we must investigate them in the context of the world of significance, the world of involvement. We must keep in mind that not all purposive activities have represented goals, that not all skillful comportment is a matter of rule-following. Fodor and his fellow cognitivists pass over the phenomenon of the world altogether and, as a result, fail to appreciate the *non-representational* character of embodied know-how. In the case of perceptual-motor skills, the Heideggerian agenda is to demonstrate in a phenomenologically persuasive way that such skills are non-representational. While Dreyfus has been successful in persuading a number of cognitivists of the merits of a Heideggerian analysis of engaged coping, there are still those who remain unconvinced; cognitivism remains the default view among North American philosophers of mind and psychologists. By way of conclusion, I will offer a few reasons why a cognitivist might find Dreyfus' analysis unpersuasive and offer an alternative strategy for persuading these skeptics.

The first issue to sort out is the issue of goal-directed behaviour. Dreyfus argues that much of everyday life is spent engaging in behaviour that is purposeful but not goal-directed. If it is not goal-directed, says Dreyfus, then the agent is clearly not acting upon an explicit representation of a goal. Dreyfus supports his claim that everyday coping is not goal-directed by means of a phenomenological analysis in which what is at issue is whether or not the agent has a goal "in mind", whether or not the agent has conscious access to an explicit goal. If the agent has no goal in mind, then there is no goal to be explicitly represented.

---

[60] Merleau-Ponty, *Phenomenology of Perception*, p. 132.

This portion of Dreyfus' discussion is problematic - *if the intended audience is composed of cognitivists.* Recall that in Fodor's shoe-tying example, we have a case in which there are all sorts of explicit representations to which the agent has no conscious access. Fodor handles these cases by invoking the core case principle. He suggests that the representational theory of the mind is the best theory on hand, and that it is not undermined by non-core cases in which there are representations but no propositional attitudes. It is possible, on Fodor's account, for an explicit representation (e.g. of a goal) to be implicated in the computational processes causally responsible for the production of intelligent behaviour even though the agent does not have conscious access to the representation. If what Dreyfus has done is show that in purposive behaviour, the agent has no goal "in mind", it does not follow, says Fodor, there is no representations of a goal *at the computational level.* The Dreyfussian might reply that the cognitivist has misunderstood the charge: it is not simply that the agent has no goal in mind, but that there is no goal at all, anywhere. The behaviour is not goal-directed. The cognitivist has misunderstood the nature of her own explicandum. But if we want to convince a true cognitivist that a piece of behaviour is not goal-directed, there is only one surefire way to do it: identify the causal mechanism involved in the production of the behaviour and show that the mechanism does not involve the manipulation of a representation whose content is a propositional description of the goal in question.

Consider the following example: Jean is in her car driving to Quebec City. How might the cognitivist explain Jean's behaviour. The first thing to note is that the cognitivist can tell a story about both *why* Jean in her car and *what* it is she is doing while there. So, for example, the cognitivist, following the folk, might say that Jean is in her car because she desires to go to Québec City and believes that driving is the best way to get there, all things considered. These beliefs and desires are cashed out in causally efficacious mental representations at the computational level. The phenomenologist acknowledges that cognition involves the manipulation of mental representations, so while there might be disagreement over why we would want to call folk psychology a theory, or why we would need to speak in terms of computational mechanisms, there is agreement on the issue of whether or not mental representations might be involved.

Now consider the cognitivist's explanation of Jean's actual driving behaviour. Extrapolating from Fodor's analysis of shoe-tying, we would first break down the task of driving into a series of goal-oriented tasks, e.g. at the corner, the goal is to turn the steering wheel 90° to the right. We then articulate a function whose input is the intention to achieve the goal and whose output is the performance of the behaviour that satisfies the goal. This function is then broken down into a number of sub-functions, which are in turn broken down into sub-functions, etc., until we reach a level where the sub-functions are themselves mechanistically realizable. Each of these sub-functions constitutes a rule which must be, on Fodor's account, explicitly represented in the agent's mental processes.

The cognitivist points out that whether or not the agent's first-order goal is explicitly represented is irrelevant to the claim that, au fond, the causally efficacious mechanism responsible for the behaviour does indeed involve the manipulation of mental representations, keeping in mind that mental processes are equivalent to computational processes. The phenomenologist might, however, ask the following question: from whence the conviction that there are *any* representations involved at all if the behaviour is itself not goal-directed at the psychological or phenomenological level? What if there is no first-order function from an intention to pursue a goal to the behaviour that fulfills the goal? What if the functionalist analysis simply doesn't get off the ground?

By the time the cognitivist turns her attention to skilled perceptual-motor behaviour (e.g. shoe-tying), the representational theory of the mind is already firmly in place. Since it works so well with respect to rational behaviour, the cognitivist assumes that she has found the holy grail - *the* mechanism of the mind - representational mental processing. Given that all intelligent behaviour is caused by mental processes, that mental processes involves the manipulation of mental representations, and that skilled perceptual-motor behaviour is a form of intelligent behaviour, it follows that skilled perceptual-motor behaviour is caused by the manipulation of mental representations. Mental processing is equivalent to computational processing. To say that a representation is tokened in an agent's mental processes is *not* to say that the agent is aware of the content of the representation. It is to say that the representation is implicated in the causal mechanism responsible for the production of the agent's behaviour. If we want to speak to the cognitivist *on her terms*, we must show that the causal mechanisms responsible for the production of skilled perceptual-motor behaviours do not involve the manipulation of representations.

The second issue to unravel concerns the Dreyfussian claim that engaged coping cannot be replicated computationally because know-how cannot be captured in propositional/representational terms. The abilities to detect salience, appreciate significance, manipulate equipment - they are all the result of a certain history of embodied dealings with the world that cannot be duplicated in a computational device. The cognitivist may concede that AI theorists have been thwarted in their attempts to model perceptual-motor skills. The cognitivist may even concede that Dreyfus is right to argue that these failures in AI research are symptoms of a deep problem. The cognitivst might elect to retreat to the following position: perhaps some, but not all, of the causally efficacious mental processes that subserve perceptual-motor skills are representational. If even one of the necessary processes responsible for such skills is non-representational, it would make sense that successful computational models of these skills have not been forthcoming. If we want to show that *none* of the processes that subserve perceptual-motor skills is representational, in the Fodorian sense, we will need an additional argument.

Dreyfus' account of Heideggerian coping, coupled with his account of the failures of AI theorists to model engaged coping computationally, constitutes a strong alternative to cognitivism - or, to be more precise, to cognitivist accounts of background skills. There are documented cases in which cognitivists exposed to Dreyfus have undergone a "Gestalt shift" of sorts. (Terry Winograd is a prime example.) Indeed, it would take something like a Gestalt shift to disabuse cognitivists of their commitment to the claim that intelligent behaviour is caused by the processing of mental representations. (We must remember that there are "mental" representations and then there are mental representations. For the cognitivist, x counts as a mental representations if it is tokened in the agent's computational processes. It matters not whether the representation is tokened in the agent's "mental" processes, in the good old-fashioned sense of the term.) If we want to persuade a few more cognitivists of the legitimacy of Heidegger's non-representational account of skilled comportment, we will have to argue with them on their own terms. That involves focusing on the issue of causal mechanisms. In the second part of the thesis, I will attempt to beat the cognitivist at her own game by showing that the mechanisms involved in skilled-perceptual-motor behaviour are not representational.

Chapter Three: Bergson's Theory of Multiple Memory Types

In their article, "Cognitive Neuroscience Analysis of Memory: A Historical Perspective," Polster, Nadel, and Schacter note that one of the first serious examinations of multiple forms of memory can be found in the work of Henri Bergson. [1] In his 1911 text, *Matière et mémoire*, Bergson makes a distinction between what he calls "representational memory" and "habit memory". [2] His work is relevant here because Bergson serves well as a bridge figure between phenomenology and neuroscience. On the one hand, his distinction between representational memory and habit memory resonates with Heidegger's distinction between ontic representational thought and ontic engaged coping. On the other, Polster et al. note that Bergson's classification and description of memory types will be "familiar to contemporary students" of neuroscience. [3] Bergson's work also sets an appropriate tone for the second half of this dissertation. He claims that the study of memory is a central issue in philosophy and that phenomenological and neuroscientific inquiries into the nature of memory can be mutually supportive. In this chapter, which serves as a preface to my examination of the neuroscientific literature on multiple memory systems, I will outline Bergson's theory of memory types in such a way that the shared philosophical views of Bergson and Heidegger are brought into relief.

Bergson's central agenda in *Matière et mémoire* is to offer an account of ontological dualism that corrects the deficiencies of Cartesian dualism. On Bergson's view, Descartes errs in positing a subject that is disembodied and disengaged from the material world. The Bergsonian subject is, in contrast, an embodied and engaged agent. Bergson's discussion

---

1. [1] Polster, Michael R., Lynn Nadel, and Daniel L. Schacter, "Cognitive Neuroscience Analyses of Memory: A Historical Perspective," *Journal of Cognitive Neuroscience* 3.2 (1991), p. 106.

[2] Bergson, Henri, *Matière et mémoire* (Paris: Presses Universitaires de France, (1911). Page references here are to the English translation prepared by N. M. Paul and W. S. Palmer, *Matter and Memory*, (New York: Zone Books, 1991).

of memory types arises in the context of his account of dualism - his account of mind, matter, and the relationship between the two. By way of introduction to his work, it should be noted that for Bergson, mind is equivalent to pure memory and matter is equivalent to pure perception. While actual memory and perception are distinct from pure memory and perception, we can gain an understanding of the interaction of mind and matter by studying the interaction of actual memory and actual perception. It is in this context, a study of actual memory and its relationship to actual perception, that Bergson makes a distinction between two kinds of memory: representational memory and habit memory. Before turning to his discussion of these two types of memory, I will outline the general principles of his ontology.

Bergson Theory of Perception

Bergson's account of perception begins with the rejection of two traditional views: realism and idealism. According to Bergson, the realist begins with the external world as described by science and sees perception as a tool for discovering the truths of such a world. The idealist begins with the findings of perception and views metaphysics as an attempt to describe the contents of our perception. Both science and metaphysics aim to provide a representation of the world as a whole. Both are dedicated to the generation of purely disinterested knowledge. It is their common preoccupation with knowledge that leads both the idealist and the realist to misunderstand perception.

> If we now look closely at these two doctrines, we shall discover in them a common postulate, which we may formulate thus: perception has a wholly speculative interest; it is pure knowledge. [4]

---

3 Polster, et. al., "Cognitive Neuroscience Analysis of Memory," p. 106.

4 Bergson, *Matter and Memory*, p. 28.

Perception, on Bergson's account, is not an instrument of knowledge but an instrument of action. That perception is devoted to the generation of action may be seen by reviewing the evolution of perception in the phylogenetic scale. The most basic form of perception is tactile. The simplest organism discovers the environment through literal contact, and this contact serves to prepare the organism for appropriate actions. Actions are "appropriate" to the degree that they service the basic survival needs of the organism. As we move up the phylogenetic scale, perception becomes increasingly complex. It does not, however, lose its utilitarian function. Visual perception and tactile perception differ only in degree. In both cases, perception serves to reveal possible opportunities for action.

> And no more in the higher centres of the cortex than in the spinal cord do the nervous elements work with a view to knowledge; they do but indicate a number of possible actions at once, or organize one of them. [5]

When realists and idealists overlook the basic utilitarian function of perception, they obscure the true nature of matter. The disinterested perception of realism delivers the world-in-itself. The disinterested perception of the idealist reveals the idea of the world. The world of both the realist and the idealist is devoid of all practical significance. According to Bergson, the true nature of the world is revealed only when we replace disinterested perception with practical perception. Between nature *en soi* and the idea of nature, there is nature as it appears to the practical perception of embodied human agents. Practical perception reveals the "world-for-us". For Bergson, the appearance of nature is matter itself. Matter is a seamless continuity of what Bergson calls "images" - the manifestation of salient objects in the world.

> Matter, in our view, is an aggregate of "images". And by "image" we mean a certain existence which is more than that which the idealist calls a *representation*,

---

[5] Bergson, *Matter and Memory*, p. 31.

but less that which the realist calls a *thing* - an existence placed halfway between the *thing* and the "representation". [6]

Perception operates on a plane midway between idealism and realism, between metaphysics and science. It is incapable of contacting the world in its entirety, for it is interested only in that portion of the world that is actually manifest to the perceiver. Nature always outruns the images that appear to the perceiver. It is not that these images give lie to nature, but that these images portray only those elements of nature that are of interest to a perceiver in action. Images do not distort nature, they distill it. They represent a diminution of the whole to its salient parts.

Perception discerns those features of the world that are salient with respect to the agenda of the body. Embodied organisms are endowed with needs and desires. The fulfillment of such needs and desires requires that the body be prompted into action. Perception is charged with the task of highlighting those objects in the world that provide an occasion for action. "I call matter the aggregate of images, and perception of matter these same images referred to the eventual actions of one particular image, my body." [7] Perception consists of a process that may be negatively described as the diminution of nature to "presence" and positively defined as the discernment of the salient.

> Perception ... consists in detaching, from the totality of objects, the possible action of my body upon them. Perception appears, then, only as a choice. It creates nothing; its office, on the contrary, is to eliminate from the totality of images all those on which I can have no hold, and then, from each of those which I retain, all that does not concern the need of the image which I call my body. [8]

---

6 Bergson, *Matter and Memory*, p. 9.

7 Bergson, *Matter and Memory*, p. 22.

8 Bergson, *Matter and Memory*, p. 229.

The notion of the "image" is critical to Bergson's claim that perception is material. Since matter is a collection of images and perception is simply a subset of salient images, perception is itself material. By describing the relationship between perception and matter in this way, Bergson avoids the task of explaining how some "inner" perceptual process is related to an "external" world. Perception is an irreducibly dialogical process that unites embodied agents with the material world in which they reside. Bergson argues against two alternative views of perception: the neural view, in which perception is localized in the brain, and the representational view, in which perception is localized in the mind. We will consider Bergson's rejection of both the neural and representational views in turn.

The error committed by those who proffer a neural view of perception is the mistake of overlooking perception's dialogical nature. On Bergson's account, the brain does indeed subserve some portion of perception - the portion that is generated by the perceiver. But a neural account fails to acknowledge the indissoluble bond between the perceiver and the perceived; it fails to recognize the material continuity between the act of perception and the object perceived. Bergson was conversant with the neuroscience of his day; he was aware of the basic neural processes involved in vision. He insists, however, that the neuroscientific view of vision is but a useful *façon de parler*. Science does not undermine the phenomenological claim that perception is a dyadic relationship between the body of the perceiver and the objects perceived.

Bergson illustrates this claim by asking us to consider a simple case of perception in which an organism is said to perceive a luminous point "P" in space. We may choose to assume the stance of natural science and describe the process whereby nerve impulses travel from the retina to cortical and subcortical visual centres. We may describe the creation of a "representation" of P in these visual centres. But we should not confuse such a scientific description with an authentic phenomenological account. In describing the neural formation of an internal representation we are, according to Bergson, "merely bow(ing) to the exigencies of the scientific method; we in no way describe the real process." [9] In actual perception, the brain is not disengaged from the external world. The two form a continuous circuit, and perception, rather than carrying the agent inward, delivers the agent directly into the world.

> There is not, in fact, an unextended image which forms itself in consciousness and then projects itself into P. The truth is that the point (P), the rays which it emits, the retina and the nervous elements affected form a single whole; that the luminous point P is part of the whole; and that it is really in P, and not elsewhere, that the image of P is formed and perceived. [10]

To claim that perception is reducible to neural processing is to be seduced by "the fiction of an isolated material object." [11] This seduction leads to the disengagement of the brain from the material world of which it is a part. Bergson argues that "it makes no sense to conceive the nervous system as living apart from the organism that nourishes it, from the atmosphere that in which the organism breathes, from the earth which the atmosphere envelops." [12]

The error of the representational view of perception is more profound. To locate perception in the mind is to misconstrue the very nature of perception, for perception is material, not mental. Perception concerns itself with matter - with the images of the world as they concern the body. But these images are material, not mental. Mental representations, by definition, fall beyond the purview of perception, as it is defined by Bergson. On Bergson's account, perception is dedicated to the preparation and execution of actions, not the processing of mental representations. "Our body is an instrument of action, and action alone. In no degree, in no sense, under no aspect, does it serve to prepare, far less to explain, a representation." [13]

---

9 Bergson, *Matter and Memory*, p. 42.

10 Bergson, *Matter and Memory*, pp. 42-43.

11 Bergson, *Matter and Memory*, p. 24.

12 Bergson, *Matter and Memory*, p. 24.

13 Bergson, *Matter and Memory*, p. 223.

The common error of associating perception with mental representations can be traced to the habit of thinking of perception and memory as continuous. Perception is said to provide a representation which fades with time into a faint image stored in memory. Perception and memory, on this account, differ only in the intensity of their representations. Bergson argues, on the contrary, that perception and memory differ in kind, not degree. This is one of the central claims of *Matter and Memory*: Bergson returns to it repeatedly in order to stress that with the distinction of perception and memory comes the distinction between matter and mind. To collapse memory and perception into a continuous phenomenon is to forego the possibility of an adequate ontology.

Philosophers fail to distinguish perception and memory because they overlook the phenomenon of pure perception. Pure perception is the grasp of the possibilities for action afforded by the environment. It is the immediate material contact between the perceiver and the perceived. It is unmediated by representations. This pure perception is, however, a theoretical abstraction. When we turn from theory to phenomenology, we find that perception is always infused with memory. Memory, the stored narrative of personal experience, resides in consciousness; it is immaterial. When memory is added to perception, it imbues perception with the individual and subjective features that colour the world for a particular observer. Philosophers pass over the phenomenon of pure perception and take as their object of inquiry this mixture of perception and memory. As a result, they attribute to perception characteristics that should be ascribed to memory, and memory alone.

> The individual accidents (of memory) are merely grafted on to this impersonal perception ... (and) because philosophers have overlooked it, because they have not distinguished it from that which memory adds or subtracts from it, they have

> taken perception as a whole for a kind of interior and subjective vision, which would differ from memory only by its greater intensity. [14]

Once perception is construed as subjective, it does seem to differ from memory only in degree. But this is to overlook the fundamental difference between perception and memory. Pure memory is an immaterial collection of images from the past. Perception, on the contrary, abides in the present and is always turned toward the future, toward incipient action.

> The actuality of our present perception thus lies in its activity, in the movements which prolong it, and not in its greater intensity; the past is only idea, the present is ideo-motor. [15]

Perception is traditionally explained as the processing of mental or neural representations. When the perceiver's relationship to the world is perceived as indirect, the mediation of a representation is required. A different picture of perception emerges when the relationship between the agent and the world is characterized as direct and immediate. It is no longer necessary to assign to perception the task of bridging the divide between a representation and the object. Perception need not solve the difficult problems raised by positing a disengaged agent. It serves, instead, as the very medium of engagement.

> Restore, on the contrary, the true character of perception; recognize in pure perception a system of nascent acts which plunges roots deep into the real; and at once perception is seen to be radically different from recollection, the reality of things is no more constructed or reconstructed, but touched, penetrated, lived. [16]

---

14 Bergson, *Matter and Memory*, p. 34.

15 Bergson, *Matter and Memory*, p. 68.

16 Bergson, *Matter and Memory*, p. 69.

Bergson's Theory of Memory

Pure perception exists only in theory. If perception were always "pure", in this sense, the role of memory would be reduced to the filing away of "an uninterrupted series of instantaneous visions, which would be part of things, rather than ourselves."[17] In order to arrive at a more accurate picture of perception and memory, a more thorough account of their symbiosis is needed.

> The action of memory goes further than this superficial glance would suggest. The moment has come to reinstate memory in perception, to correct in this way the element of exaggeration in our conclusions, and so determine with more precision the point of contact between consciousness and things, between the body and the spirit.[18]

For Bergson, the phenomenon of memory is at the heart of ontology. The resolution of the mind-body problem lies in seeing memory as the point of contact between consciousness and the material world. The philosopher is urged to take up the study of memory, and to make use of the wealth of scientific data available on the subject.

> No one ... will deny that, among all the facts capable of throwing light on the psychophysiological relation, those which concern memory, whether in its normal or pathological state, hold a privileged position. Not only is the evidence here extremely abundant..., but nowhere else have anatomy, physiology, and psychology been able to lend each other such valuable aid. Anyone who approaches ... the classical problem of the relations of soul and body, will soon see this problem as centering on the subject of memory.[19]

---

17 Bergson, *Matter and Memory*, p. 65.

18 Bergson, *Matter and Memory*, p. 65.

19 Bergson, *Matter and Memory*, p. 13.

Traditional dualism posits a stable and rigid distinction between *res cogitans* and *res extensa*. The very stability of this distinction proves, however, to be the fatal flaw of dualism. Having posited a rigid ontological divide between mind and matter, the dualist is incapable of providing an adequate account of the interaction of consciousness and the brain, on one hand, and the subject and object, on the other. Bergson's goal is the articulation of an ontology in which mind and matter are described in such a way that their interaction is not merely possible, but inevitable.

Two Kinds of Memory: Representational Memory and Motor Memory

Bergson initially defines mind as memory and matter as perception. As we move closer to the line of demarcation between mind and matter, we see that memory, rather than remaining neatly on the immaterial side of the divide, slips across the line, and moves almost imperceptibly into the realm of the material. Our ontological map must therefore be modified.

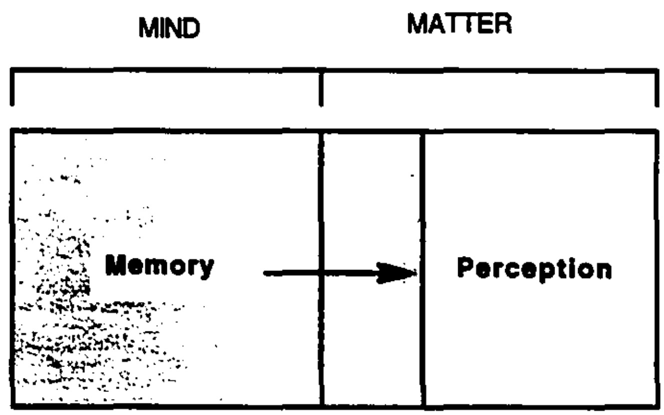

Figure 3.1: Bergson's Ontological Map

This new map reflects that fact that an agent's experience is retained in two distinct ways: (1) her acquired physical skills are stored in perceptual-motor pathways and (2) her personal narrative is stored as a sequence of representations. "The past survives under two distinct forms: first in motor mechanisms; secondly in independent recollections." [20] According to Bergson, the body is devoted exclusively to the preparation and execution of movements. So while the body can subserve "motor" memory, it cannot be the substrate of "representational" memory. "In the form of motor contrivances, and of motor contrivances only, it (the body) can story up the action of the past. Whence it results that past images, properly so-called, must be otherwise preserved." [21] The images of representational memory exist only in immaterial form, in consciousness.

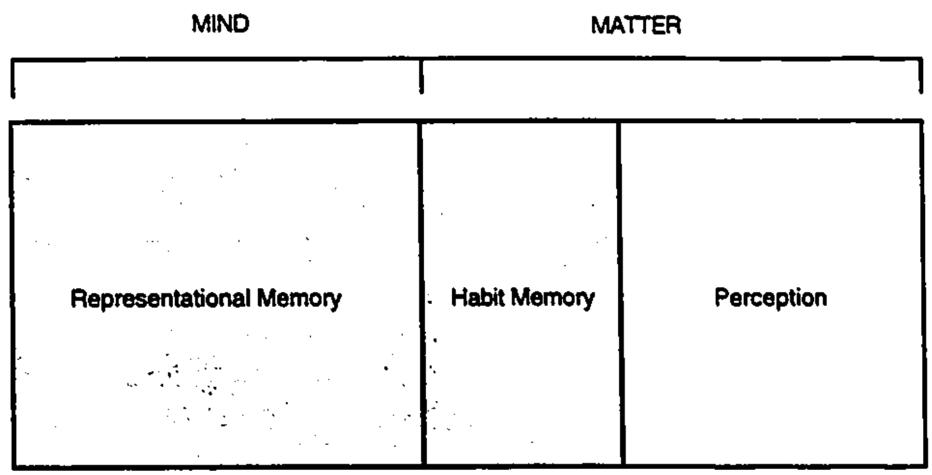

Figure 3.2: Representational Memory and Habit Memory

The differences between representational memory and motor memory may be elucidated by studying an example proposed by Bergson. Consider the case of a student who is faced with the task of committing a text to memory. The student adopts the strategy of reading the text aloud repeatedly until she can reproduce it without error. Once the process is complete, we may distinguish between the type of memory involved in the recollection of particular rehearsals of the text (the memory of a "reading"), and the memory that subserves the student's new found capacity to produce the text on demand (the memory of the "lesson".) One might suppose that these two types of memory differ

[20] Bergson, *Matter and Memory*, pp. 77-78.

[21] Bergson, *Matter and Memory*, pp. 77-78.

only in degree, that the memory of the lesson consists in a summation of the student's memories of particular readings, particular rehearsals. Bergson's objective here, however, is to argue that the memory of the lesson differs *in kind* from the memory of a particular readings.

The memory of the lesson is a capacity for action. It involves the acquisition of certain motoric abilities. On Bergson's view, this type of memory "has *all* the marks of a habit." [22]

> Like a habit, it is acquired by the repetition of the same effort. Like a habit, it demands first a decomposition and then a recomposition of the whole action. Lastly like every habitual bodily exercise, it is stored up in a mechanisms which is set in motion as a whole by an initial impulse, in a closed system of automatic movements which succeed each other in the same order. [23]

The memory of a reading is altogether different; it has "*none* of the marks of a habit." [24] The memory the student has of a particular rehearsal of the text is an immaterial image that resides in consciousness; it is a representation. "The memory of a given reading is a representation, and only a representation." [25] This type of memory is distinctive, in part, because it is "disinterested". "(Representational memory) records, in the form of memory-images, all the events of our daily life as they occur in time; it neglects no detail; it leaves to each fact, to each gesture, its place and date. Regardless of utility or of practical application, it stores up the past by the mere necessity of its own nature." [26]

Representational memory is immaterial and disinterested. Habit memory, on the other hand, is corporeal and practical. Each memory-image in representational memory is,

---

22 Bergson, *Matter and Memory*, p. 80.

23 Bergson, *Matter and Memory*, p. 80.

24 Bergson, *Matter and Memory*, p. 80.

25 Bergson, *Matter and Memory*, p. 80.

26 Bergson, *Matter and Memory*, p. 81.

in a sense, dated. It is formed at a certain point in time and fits into the narrative of a lifetime at a certain place. Future experience may, of course, affect or distort a memory image. But that which is distorted is a memory image which carries a certain date. Habit memory, on Bergson's view, resides in the present in the following sense: it is of the essence of habit memory to prepare actions that unfold in the here and now. It matters not when the capacity for a particular habitual action is acquired. "In fact, the lesson once learned bears upon it no mark which betrays its origin and classes it in the past; it is part of my present, exactly like my habit of walking or of writing; it is live and acted, rather than represented.[27] Habit memory is dispositional memory.

Habit memory works in close conjunction with perception. When perceptions are prolonged into actions, these actions gradually modify the motor pathways in the agent's nervous system. If an action becomes habitual, these motor pathways become increasingly fine-tuned. The movements constitutive of a habit "modify the organisms and create in the body new dispositions toward action."[28] Over time, the agent acquires an increasingly diverse repertoire of adaptive habits. The agent comes to possess two different types of experience; she acquires a representational narrative of events and a non-representational set of skilled perceptual-motor habits. "Thus is gradually formed an experience of an entirely different order, which accumulates within the body, a series of ready mechanisms, with reactions to external stimuli ever more numerous and more varied and answers ready prepared to an every growing number of possible solicitations."[29] Habit memory differs in kind from representational memory because "it (habit memory) no longer *represents* our past to us, it *acts* it."[30]

---

[27] Bergson, *Matter and Memory*, p. 81.

[28] Bergson, *Matter and Memory*, p. 81.

[29] Bergson, *Matter and Memory*, pp. 81-82. (Modified translation.)

[30] Bergson, *Matter and Memory*, p. 82.

The Interaction of Representational Memory and Habit Memory

Having distinguished representational memory and habit memory, Bergson next turns to the task of discussing their interaction. An agent equipped only with representational memory would be reduced to living in the past, in a world of deracinated "dreams". An agent who possessed only motor memory would live as a mindless automaton, unable to invoke the non-motoric past. In situations in which the agent confronts several alternative avenues of action at once, she would be unable to make an intelligent choice. Representational memory and habit memory must somehow join forces.

In the normal course of events, habit memory, drawn into the future, overpowers representational memory, which remains inert in the past. This claim is a variation on the observation that perception, as it is prolonged into adaptive action, inhibits the influence of the past and propels the agent into the future. Both habit memory and perception issue in action and, as a result, they tend to suppress representational memory, which is concerned, in the first instance, only with the preservation of the past. How, then does representational memory service the agent in the present?

The primary function of representational memory is the task of allowing the past to inform choices made in the present. As we have seen, Bergson sees the range of perceptual phenomena as a continuum that begins with tactile contact and ends with complex vision. This continuum may now be described in terms of the complexity of the reaction demanded of the organism. In all perception, the objects of the world reflect back to the agent various invitations for action. These invitations may be pressing or remote, varying with the distance between the perceiver and the object perceived. When the distance is quite small, the reaction of the organism is automatic.

> The more immediate the reaction is compelled to be, the more must perception resemble a mere contact: and the complete process of perception and of reaction can

then hardly be distinguished from a mechanical impulsion followed by a necessary movement. [31]

After the perceiver backs away from an object, the intensity of the object's influence diminishes. More objects, at varying distances from the perceiver, come into view. As the visual field is populated with an increasing number of objects, the number of possible actions available to the agent is multiplied. A "zone of indetermination" envelops the agent. No longer is action immediate and automatic. The purity of perception is lost because the perceiver cannot maintain an intimate relations with all the objects that appear in the scene. "To perceive all the influences of all the points of all bodies would be to descend to the condition of a material body." [32] When the actions of the agent become indeterminate, consciousness, in the form of memory, must intercede. Our actions are, for the most part, neither automatic nor pre-determined. Choices must be made, and these choices are guided by the lessons of past experience. "This indetermination of acts to be accomplished requires, then, if it is not to be confounded with pure caprice, the preservation of the image received." [33]

Representational memory thus seizes indeterminacy as an opportunity to influence the choices of the perceiver. At the moment when the agent falters, when she pauses to regard her options, the present loses some measure of its power. The agent assumes an attitude of momentary suspension. Both perception and habit memory are temporarily immobilized. The agent turns toward her past to await instruction. An image is selected from the inventory of representational memory, an image that seems likely to resonate with the suspended perception. This image is no longer inert, for it enters the perception and

---

31 Bergson, *Matter and Memory*, p. 32.

32 Bergson, *Matter and Memory*, p. 49.

33 Bergson, *Matter and Memory*, p. 65.

extends itself into an appropriate movement. "A memory becomes actual ... only by borrowing the body of some perception into which it slips." [34]

Representational memory thus sends forth a memory-image which can insert itself into a present perception. This memory-image represents an intermediate stage between pure perception and pure memory. The memory-image straddles the ontological divide between mind and matter. It is, in part, an image, for it reproduces past perception; it realizes a representation of pure memory. It is also, in part, a motor phenomenon, for it is the reification of the prolongation of a representation into nascent action. The memory-image may be pictured as located on the line between representation memory and habit memory.

The Two Forms of Recognition

The nature of the memory-image may be clarified by examining, in greater detail, the process in which images from the past come into contact with percepts in the present - the process of recognition. Bergson defines recognition as "the concrete process by which we grasp the past in the present." [35] As one might expect, Bergson's account of recognition diverges sharply from traditional accounts.

Recognition is typically described as the association of a perception with a memory. This melding of a memory and a perception is said to occur either psychologically, via the association of ideas, or physiologically, via the mechanistic process in which the neural representation of a percept triggers the activity of the neural circuits which subserve the memory. Bergson rejects both models. The psychological account reduces perception to an internal representation that may be associated with a memory. The physiological account errs in localizing the memory-representations in the brain, where they may physically interact with a percept. Recognition cannot be localized in the mind, for

[34] Bergson, *Matter and Memory*, p. 67.

perception is a material phenomenon. Nor can it be localized in the brain: representational memory is immaterial. Thus recognition cannot be, argues Bergson, a simple meeting of perception and memory.

Bergson acknowledges that recognition involves, *in some sense*, the joining of a memory and a percept. What he cannot accept, however, is that the union occurs in either the mind in isolation or the brain in isolation. Given that Bergson defines memory as immaterial and perception as material, recognition must involve a mechanism for crossing the ontological divide between mind and matter. He has provided a preliminary sketch of this mechanism in his discussion of the insertion of a memory-image into a perception. Recognition, in outline, is the process in which a memory-image is embodied in a perception and protracted into an action. We must now examine this process in greater detail.

Recognition crosses the ontological barrier between mind and matter because the memory-image is, itself, partly representational and partly motor. In order to analyze these two aspects of the memory-image, we must distinguish between representational recognition, based on representational memory, and motor recognition, based on habit memory. The critical distinction between the two is that representational recognition is "attentive" while motor recognition is "inattentive". Representational recognition begins with an attentive mind and works its way toward the object. Motor recognition requires no "attention", for it is pre-reflective. It begins with the movements of the body and works its way back toward the mind.

Inattentive motor recognition is a process in which a familiar object prompts the performance of the sequence of movements that constitute a habit. There are two classic examples of motor recognition - the "mindless" traversal of familiar routes and the "automatic" action involved in using a familiar object. When an agent first encounters a

---

[35] Bergson, *Matter and Memory*, p. 90.

network of streets to navigate on foot, her judgment is guided by a tentative and naive perception. Each moment in time is static, for the revelations of the next moment are not yet known. With each traversal of the route, perception loses some degree of its naiveté. As the route becomes familiar, each moment becomes linked to the next. Habit memory has come to the aid of perception. There is an "incipient consciousness" of the motor movements required to complete the course. Just as we move around familiar rooms avoiding the furniture without attending to particular chairs and tables, we navigate familiar routes by means of this "incipient consciousness". The initial naive perception has evolved into a perception that has some amorphous motor awareness of what is to be done next.

> Now if the latter perceptions differs from the first perception in the fact that they guide the body toward the appropriate mechanical reaction, and if, on the other hand, those renewed perceptions appear to the mind under that special aspect which characterizes familiar or recognized perception, must we not assume that the consciousness of a well-regulated motor accompaniment of an organized motor-reaction, is here is foundation of the sense of familiarity. [36]

When the subject comes to know the terrain thoroughly, she moves along it "automatically", as though motor memory has taken over from perception the task of guiding her. The organization of the route is mirrored in the organization of her movements. At this stage, the subject develops a full bodily awareness of the organization implicit in her own actions. In this awareness of the internal organization of motor plans lies the origin of recognition. "At the basis of recognition there would thus be a phenomenon of motor order." [37]

The type of recognition involved in navigation is closely related to a second form of motor familiarity - the skillful use of familiar objects. Here again the basis of recognition is

---

[36] Bergson, *Matter and Memory*, p. 93.

[37] Bergson, *Matter and Memory*, p. 93.

not an intellectual or physical association of memory and percept, but a motor schema that is learned through repetition over time. Just as in the case of the route, the subject acquires an awareness of the organization of the movements elicited by perception. As the habit of interacting with the object is developed, the agent experiences the latent organization of her movements.

> The habit of using the object has, then, resulted in organizing together movements and perception; the consciousness of these movements, which follow perception after the manner of a reflex, must be here also at the bottom of recognition. [38]

As we have seen, representational memory comes to the aid of a faltering perception. It can also comes to the rescue of a faltering motor mechanism. At times there is slippage between perception and the movement it spawns. The movement ceases to adapt itself well to the task of finding one's way or using an object. Perhaps a hand slips or a tool breaks. In these cases, the mind selects certain memory images and brings them to bear on the present situation. In such cases, recognition "implies an effort of the mind which seeks in the past, in order to apply them to the present, those representations which are best able to enter into the present situation." [39] The mind awaits a breakdown in the patterns of motor movements laid down by habit. A memory-image then slips into the perception. "This memory merely awaits the occurrence of a riff between the actual impression and its corresponding movement to slip in its image." [40]

There are vast numbers of memory-images available to the subject. The subject must, in some sense, choose which image will come forth to be embodied in movement. Here again Bergson rejects an intellectualized model. According to Bergson, it is the movement in which the subject is engaged that prepares this choice, which makes the "initial cut" as it were. Amongst the remaining memory-image candidates, certain will feel

38 Bergson, *Matter and Memory*, p. 94.

39 Bergson, *Matter and Memory*, p. 78.

less resistance in the face of the organizations of ongoing movements and current perceptions. These memory images will then find their way into the actions of the subject.

> By the very constitution of our nervous systems, we are beings in whom present impressions find their way to appropriate movements: it so happens that former images can just as well be prolonged in these movements, they take advantage of the opportunity to slip into actual perception and get themselves adopted by it. [41]

The opportunity to inform current perception is controlled by the current movements of the agent. When the agent is immersed in the activity, her bearing is toward the present and the future, and her absorption is too strong for past memory images to achieve potency. When there is a breakdown, an interruption, a pause, or a mismatch between perception and movement, the past experience surfaces in a 'non-motor' form, in the embodiment of memory images.

Motor recognition, based on habit memory, proceeds non-reflectively; it is "inattentive". Recognition based on representational memory is, however, attentive. Bergson describes his analysis of attentive recognition as "the essential turning point of our discussion." His analysis of attentive recognition clarifies the nature of the procedure in which the mind actually comes into contact with matter - the procedure in which an immaterial representation slips into a nascent action prepared by the brain.

Attention is typically defined as a heightened sense of awareness, as an attitude of the mind in which its resources are highly focused. Bergson prefers to see attention as an attitude of the body, one in which the body turns from the present in order to be receptive to the lessons of the past. Attention consists of a two part process. The negative work of attention occurs when the mind gives up its pursuit of the useful, ceases for a moment its forward movement into future action, and allows the body an instant of quiescence. There

---

40 Bergson, *Matter and Memory*, p. 95.

41 Bergson, *Matter and Memory*, pp. 95-96.

is an inhibition of movement; action is arrested. The body is in a state of suspended animation. Attention can then perform its positive work, seizing this moment of calm as an opportunity for memory to come forth and be realized.

The positive work of attentive recognition must be examined with care, for it is the central phenomenon of Bergson's ontology. Conventional accounts of attention liken it to a searchlight which plays over the details of a percept and thus reveals an increasing number of details. In keeping with his metaphor of perception as a dialogue between the body and the object, Bergson likens attention to the activities of a telegraph clerk who receives a transmission and sends it back to its source for verification. The message is sent back and forth, between the subject and the object, until the full accuracy of the text is ensured and all its meanings made clear.

Attention is directed by memory, for it is memory that makes an educated guess as to the meaning of the transmission and sends it through the brain and out to the object for verification. In the case of perception, memory selects an image that resembles the initial percept and projects this image back onto the object. "The 'analysis' of attention is actually a series of syntheses, our memory chooses, one after another, various analogous images which it launches in the direction of the new perception." [42] If the object "accepts" the memory-image, it sends back to the body a new, enriched percept. Memory does not select an image from its inventory in a random fashion. When the body is suspended in a certain receptive attitude, this attitude already contains the movements involved in tracing the outlines of the object. The choice made by memory is constrained by this outline.

> This choice is not made at random. What suggests the hypothesis, what presides even from afar, over the choice, it the movement of imitation, which continues the

---

[42] Bergson, *Matter and Memory*, p. 102.

> perception, and provides for the perception and for the image a common framework. [43]

The function of memory in attention is to repeat the imitative movements of perception and to add to them only those images which resonate with these movements.

Categorization and the General Idea

Bergson draws upon his theory of memory to provide an account of categorization - the process in which individual objects in the world are grouped together by kind. Here again the fundamental distinction between representational memory and habit memory provides the framework for his analysis. According to Bergson, the agent comes to grasp the nature of the general idea, or "category", as a result of the interplay between representational memory and habit memory.

Bergson describes the fictitious agent who possesses only representational memory as a dreamer who lives entirely in the realm of ideas. Such an agent sees each idea as an isolated entity and discerns only the differences which separate one idea from the next. The fictitious agent who possesses only motor memory is the ultimate creature of habit, an automaton who sees objects only in terms of their practical resemblances. Representational memory discerns difference; habit memory experiences similarity. In order to construct a genuine categorization of the world, the two processes must be combined.

According to Bergson, traditional accounts of "the general idea" oscillate between the agent-as-dreamer and the agent-as-automaton in an ineffective attempt to genuinely unite the processes of discerning difference and discerning similarity. The result is a vicious circle. Two accounts of the general idea manifest this vicious circularity: nominalism and conceptualism. Consider the process of constructing a general idea of the rose. The nominalist begins with the extension of the concept rose, with an unlimited number of

---

[43] Bergson, *Matter and Memory*, p. 102.

individual roses. She extracts from this group of objects a list of communal features and assigns to this list the name "rose". Taken collectively, these features constitute a criterion of admission for the category. The list of features represents the intension of the concept. The nominalist thus moves from extension to intension, from difference to similarity. But the process cannot begin until the agent has before her a collection of roses, until, that is, she already possesses an account of their similarities.

The conceptualist begins at the other side of the circle. She enumerates the properties of the world, distinguishing each from the others. Having identified the properties of "having a stem", "having petals", "having a certain morphology", etc., she then converts each *property* into a category, one under which individual objects may be subsumed. Categorization of the world then proceeds by noting certain patterns that emerge when the objects are classified according to their properties. In this case, the movement is from intension to extension. But the circle has not been broken. The classification of distinct groups of objects here relies on a prior capacity to distinguish differences among properties. We cannot explain the ability to discern difference by assuming that very same ability.

Both the nominalist and the conceptualist are trapped. "In order to generalize ... we have to extract similarity, but in order to disengage similarity usefully, we must already know how to generalize." [44] Bergson proposes to break out of the circle by highlighting the role of the body in generalization. The nominalist and the conceptualist fail to exit the circle, on Bergson's account, because they view generalization as an act of the intellect; they overlook its origin in the body.

The origin of the general idea is to be found in the mechanics of habit memory and motor recognition. Motor recognition occurs when an object provokes the set of actions that constitute a habit. The very same habitual response may be triggered by a variety of

---

44 Bergson, *Matter and Memory*, p. 160.

objects. A number of different objects may cause the same effect in the body. These objects are similar in virtue of the fact that they produce the same actions in the agent. The similarity that unites these objects is not a theoretical construct, but a brute physical force.

> This similarity acts objectively like a force and provokes actions that are identical in virtue of a purely physical law which requires that the same general effects should follow the same profound causes. [45]

The perception of similarity is not an act of conscious discernment. Perception, by its very nature, detaches from nature only those objects which are of interest to it. These objects are salient precisely because they provoke adaptive actions. It seems reasonable to suppose that actions which are adaptive will fall naturally into certain "practical" categories - eating, drinking, avoiding obstacles, and the like. There is no need for abstract generalization at this level, for similitude is experienced, not thought. The perception of similitude is a process in which "beings seize from the surroundings that ... which interests them, practically, without needing any effort of abstraction, simply because the rest of their surroundings take no hold upon them; this similarity of reaction following actions superficially different is the germ which the human consciousness develops into general ideas." [46]

Habit recognition provides the foundation for the perception of similarity. When the mind reflects on this process, it generates the concept of a category. This process requires the intervention of representational memory, a memory which "grafts distinctions upon resemblances which have been spontaneously abstracted." [47] In this process the mind "disengage(s) from the habit of resemblance the clear idea of generality." [48]

---

[45] Bergson, *Matter and Memory*, p. 159.

[46] Bergson, *Matter and Memory*, p. 159-60.

[47] Bergson, *Matter and Memory*, p. 160.

[48] Bergson, *Matter and Memory*, p. 160.

The categorization of the world thus requires two processes: habit memory discerns individuals and representational memory constructs images of these individuals which can be analyzed and classified according to their differences. Categorization involves a constant movement between the plane of action and the plane of thought.

> This idea of generality was, in the beginning, only our consciousness of a likeness of attitude in a diversity of situations; it was habit itself, mounting from the sphere of movement to that of thought.[49]

The movement from action to thought, and from thought to action, creates a productive tension between the subject and the world. This tension takes the form of an "attention to life" that draws the agent from the world of mere ideas to the world of reality. If the tension between memory and action is heightened, as in the case of attentive recognition, deeper levels of reality are revealed.

Bergson's account of the distinction between representational memory and habit memory resonates well with Heidegger's distinction between ontic representational thought and ontic engaged coping. In ontic coping, as in habit memory and inattentive recognition, the agent enjoys a close, non-representational relationship with salient objects in the environment. The role of perception in engaged coping, which Heidegger captures in his description of "circumspection", is similar to the role Bergsonian perception plays in the acquisition and development of skilled motor habits. In the case of ontic representational thought, as in the case of representational memory, the mind of the agent becomes disengaged from the pressing demands of the practical realm. Both Bergson and Heidegger argue that the realm of practice precedes the realm of representational thought, and that it is in times of practical breakdown that mental representations are summoned forth.

---

[49] Bergson, *Matter and Memory*, pp. 160-61.

One of Bergson's most important contributions is his account of habit, which foreshadows Merleau-Ponty's discussion of the same issue. The concept of habit is critical for those philosophers who wish to move out from under the shadow of the Cartesian tradition. In the concept of habit, we find a new form of understanding, one in which the body plays a central role. Merleau-Ponty states that "the acquisition of habit as a rearrangement and renewal of the corporal schema presents great difficulties to traditional philosophies, which are always inclined to conceive synthesis as intellectual synthesis." [50] In a passage which reads like a synthesis of Bergson's account of habit and Heidegger's account of engaged coping, Merleau-Ponty states that "a movement is learned when the body has understood it, that is, when it has incorporated it into its 'world', and to move one's body is to aim at things through it; it is to allow oneself to respond to their call, which is made upon it independently of any representation." [51] Perhaps the most delightfully apt comment Merleau-Ponty offers on the subject of habit concerns the relationship between habit and significance. "The acquisition of a habit is indeed the grasping of a significance ..., the motor grasping of a motor significance." [52]

Bergson's work is relevant here not only because it has Heideggerian overtones, but because it serves as a link between phenomenology and neuroscience. As we will see, Bergson's description of the differences between habit memory and representational memory is strikingly similar to contemporary accounts of the differences between basal ganglia memory ("skill" or "habit" memory) and hippocampal memory ("representational" or "declarative" memory). We should thus keep in mind Bergson's claims about the nature of both habit memory and representational memory.

Habit memory is a motor phenomenon closely linked with perception. Practical perception discerns salient or significant objects and triggers the unfolding of adaptive

---

[50] Merleau-Ponty, *Phenomenology of Perception*, p. 142.

[51] Merleau-Ponty, *Phenomenology of Perception*, p. 139.

[52] Merleau-Ponty, *Phenomenology of Perception*, p. 143.

actions. The similitude of objects which share the same practical significance is experienced via the repetition of these habitual actions. Over time, the motor mechanisms involved in such habitual actions are fine-tuned, and the agent gradually acquires a set of well-rehearsed dispositions for adaptive actions. Motor recognition and habit memory are inattentive and do not require conscious effort or representational thought. They are symptoms of engagement.

The activation of representational memory, on the other hand, is a symptom of disengagement. Bergson claims that representational memory is disinterested; it records the episodes of a lifetime indifferently and stores them in narrative order. Unlike habit memory, which is sensitive to the similarities among objects, representational memory excels at making distinctions, at discerning differences among its stored representations. It provides for the possibility of making conscious choices informed by lessons of the past.

Bergson's depiction of these two types of memory is prescient. His account of perception, however, requires revision. The problem with Bergson's account of perception is that it tells only half the story. He focuses exclusively on action-oriented perception and insists that perception never involves the acquisition of visual representations. Recent work on the visual system indicates that the nervous system contains two visual pathways, one of which is devoted to action-oriented perception, one of which is devoted to representational perception. These contemporary accounts of the two visual systems confirm that Bergson's distinction between action-oriented systems and representation-oriented systems is important. These accounts make clear, however, that the distinction applies not only to memory systems, but to perceptual systems as well.

Contemporary Neuroscientific Accounts of Practical Perception

It is common to think of perception as a means of obtaining visual representations of the world. Bergson's discussion of the practical and non-representational nature of pure perception serves as a needed correction to the tendency to understand all forms of perception in representational terms. In order to understand the contemporary significance of Bergson's theory of perception, it will be necessary to review a few basic principles of neuroscience and to examine some recent work on vision that has a direct bearing on Bergson's concept of "pure" perception.

The brain is conventionally divided into four lobes: the frontal lobe, the parietal lobe, the occipital lobe, and the temporal lobe. Within these lobes we find the four primary sensory cortices: the visual cortex (in the occipital lobe), the auditory cortex (in the temporal lobe), the somatosensory cortex (in the parietal lobe), and the motor cortex (in the frontal lobe). Near the centre of the brain is an important sub-cortical structure, the thalamus, which serves as a central relay station.

In the early 1980's, Leslie Ungerleider and Mortimer Mishkin conducted extensive studies on the visual system. It was already known at that time that visual information is received by the retina and passed along through the thalamus to the primary visual cortex in the occipital lobe. Mishkin discovered that visual information leaving the primary visual cortex is transmitted along *two* distinct pathways: a dorsal pathway which runs from the occipital lobe up to the parietal lobe, and a ventral pathway which runs from the occipital lobe down into the temporal lobe.

Mishkin investigated the functions of the two pathways by studying the deficits that were produced when each was selectively "disconnected" from the primary visual cortex. He found that monkeys who suffered damage to the occipito-temporal pathway were impaired in tasks of object recognition while monkeys who suffered damage to the occipito-parietal pathway were impaired on tasks which require that the monkey be able to detect an object's location. As a result of Mishkin's work, the ventral visual stream became known as the "what" pathway and the dorsal visual stream became known as the "where" pathway.

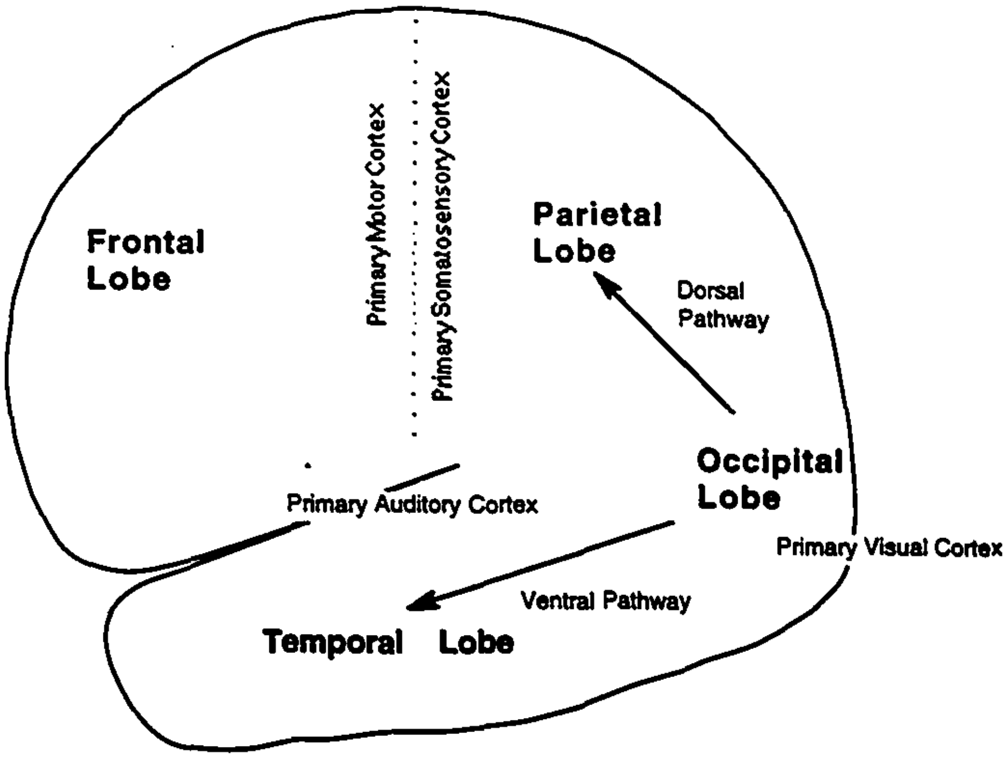

Figure 3.3: Two Cortical Visual Pathways

In 1992, Goodale argued that Mishkin had mis-characterized the function of the occipito-parietal pathway. While Mishkin emphasizes the type of input received by the two visual pathways, Goodale stresses the type of output required of them. He maintains that both visual streams receive information about the size, shape, and orientation of objects, but that they put this information to use in different ways. On Goodale's view, the ventral stream is indeed specialized, as Mishkin suggests, for object recognition. Goodale argues, however, that the dorsal pathway is specialized, not for spatial vision, but for the type of vision required for skillful perceptual-motor behaviours: accurate reachings, accurate orientation of the hand and limb for grasping objects, and object manipulation. The ability to detect the location of objects is just one element of this more general function of action-oriented perception.

> We propose that the ventral stream of projections from the striate cortex (the primary visual cortex) to the inferotemporal cortex plays the major role in the perceptual identification of objects, while the dorsal stream projecting from the striate cortex to the posterior parietal region mediates the required sensorimotor transformation for visually guided actions directed at such objects.[53]

Goodale reviews and re-evaluates the neuropsychological studies cited as supporting evidence for Mishkin's "what - where" theory of the two perceptual streams. Patients with damage to the ventral pathway often suffer from visual agnosia (a deficit in object recognition) and prosopagnosia (a deficit in face recognition) even though they are still able, as Goodale puts it, to "navigate through the everyday world ... with considerable skill."[54] Patients with damage to the dorsal pathway, on the other hand, often suffer from optic ataxia; they are "unable to reach accurately towards visual targets that they have no difficulty recognizing."[55] This pattern of deficits is consistent with Mishkin's claim that the ventral pathway subserves spatial vision. Goodale suggests, however, that we should look more carefully at the deficits that result from damage to the dorsal visual stream.

Goodale discusses the case of a patient who suffers from bilateral damage to the parietal lobe and who is, as a result, impaired in the task of picking up objects. The deficit cannot be exhaustively described in terms of an inability to detect the location of objects. The patient is unable, for example, to move fingers and thumbs in an appropriate way during reaching; the distance between thumb and forefinger bears no relation to the size of the object.

> Such studies suggest that damage to the parietal lobe can impair the ability of patients to use information about the size, shape, and orientation of an object to

---

53 Goodale, Melvyn A., and A. David Milner, "Separate Visual Pathways for Perception and Action," *Trends in Neurosciences* 15.1 (1992), p. 20.

54 Goodale, "Separate Visual Pathways," p. 21.

55 Goodale, "Separate Visual Pathways," p. 21.

control the hand and fingers during a grasping movement, even though this same information can still be used to identify and describe the object.[56]

Furthermore, a closer examination of subjects who suffer from damage to the ventral visual stream reveals that they are unimpaired on tasks of accurate grasping, even though they cannot recognize the objects for which they are reaching. Goodale discusses the case of patient D.F. who, "despite her profound inability to recognize the size, shape, and orientation of visual objects," shows "strikingly accurate guidance of hand and finger movements directed at the very same objects."[57] Even more striking is D.F.'s performance on the following set of tasks. If asked to distinguish between two rectangular blocks that differ only in size, D.F. is unable to comply. When asked to indicate the width of the blocks by using her thumb and forefinger, her performance is erratic and inaccurate. When, however, she is asked to pick up one of the blocks, "the aperture between her index finer and thumb changed systematically with the width of the object, just as in normal subjects."[58]

A similar pattern of abilities and deficits appears in tasks having to do with the spatial orientation. Asked to insert slabs into slots of various orientations, she performs well. If asked simply to indicate the right target orientation by holding the slab at the proper angle, her performance is severely impaired. Goodale concludes that "these disparate neuropsychological observations lead us to propose that the visual projections system to the human parietal cortex provides action-relevant information about the structural characteristics and orientation of objects, and not just their position."[59]

One way to investigate the function of neurons in a particular region of the brain is to determine what types of stimuli trigger a strong firing response. Goodale argues that

56 Goodale, "Separate Visual Pathways," p. 21.

57 Goodale, "Separate Visual Pathways," p. 22.

58 Goodale, "Separate Visual Pathways," p. 22.

59 Goodale, "Separate Visual Pathways," p. 22.

physiological studies of this sort, conducted on monkeys, support his interpretation of the function of the two visual pathways. He argues that neurons in the parietal lobe implicated in the occipito-parietal visual pathway are preferentially responsive to self-directed motions of the eyes and limbs. Many of these neurons fire at a rate that is dependent on the direction of the animal's gaze. Others are active during visual pursuit and fixation. Some parietal neurons are particularly sensitive to precisely those features of an object that would determine the appropriate hand posture required for reaching and grasping. While there are motion-sensitive cells in both the parietal and temporal visual pathways, it is only within the parietal pathway that cells "appear to be well-suited for the visual monitoring of limb position during reaching behaviour." [60] Motion-sensitive neurons in the occipito-temporal pathway, on the other hand, "have been reported not to respond to such self-produced visual motion." [61] Goodale also reminds us that the parietal visual pathway sends its output to regions of the frontal lobe that are known to be involved in "ocular control, reaching movements of the limb, and grasping actions of the hands and fingers." [62] He concludes that "the parietal cortex is strategically placed to serve a mediating role in the visual guidance and integration of prehensile and other skilled actions." [63] Monkeys with parietal lobe lesions exhibit the same pattern of impairments as their human counterparts; they have difficulty orienting and shaping their hands for grasping tasks and do not reach accurately for objects. They are not, however, impaired in tasks of object recognition.

Neurons in the occipito-temporal pathway have a difference response profile. They are "strikingly sensitive to form, pattern, and colour." [64] It is of particular interest in this content that neurons in this pathway show uniform response rates over a "wide range of

---

60 Goodale, "Separate Visual Pathways," p. 23.

61 Goodale, "Separate Visual Pathways," p. 23.

62 Goodale, "Separate Visual Pathways," p. 22.

63 Goodale, "Separate Visual Pathways," p. 22.

64 Goodale, "Separate Visual Pathways," p. 23.

size ... and viewpoint transformations of the object."[65] As Goodale notes, "such cells, far from providing the momentary information necessary for guiding action, specifically ignore such changing details."[66] The response patterns of these inferotemporal neurons are consistent with the theory that the ventral visual pathway is important for object recognition, for tasks in which the animal must extract size and shape constancies from visual stimuli so that the stimuli may be recognized in the future from different angles. Goodale suggests that we therefore think of the ventral stream as "object-oriented" and that we think of the dorsal stream as "action-oriented" or "viewer-oriented."[67] While size constancy is a property that is relevant to both pathways, in the case of the dorsal stream, size-constancy is combined with egocentric information concerning the position of the object and, if it is moving, the direction of the object relative to the observer.

Goodale's work has been well-received by Mishkin, who recently described the dorsal visual pathway as follows: "The dorsally directed streams ... appear to be critical for spatial perception and orientation as well as for the spatial guidance of motor responses."[68]

In the concluding section of his article, Goodale offers a few preliminary comments on the relationship between the two visual pathways and consciousness. He hypothesizes that information processed in the dorsal pathway is not available to consciousness and cites studies in which subjects reach appropriately for moving targets even though they could not report afterward whether or not the target stimulus had in fact moved. The ventral visual stream, on the other hand, is associated with conscious visual experience. Although the jury is still out on these questions, Goodale ventures that "it is feasible to maintain the hypothesis that a *necessary* condition for conscious visual experience is that the ventral

---

65 Goodale, "Separate Visual Pathways," p. 23.

66 Goodale, "Separate Visual Pathways," p. 23.

67 Goodale, "Separate Visual Pathways," p. 23.

68 Mishkin, Mortimer, "Cerebral Memory Circuits," *Exploring Brain Functions: Models in Neuroscience*, Ed. T.A. Poggio and D.A. Glaser (New York: John Wiley and Sons, 1993) pp. 113-125.

system be activated."[69] Because the two systems, dorsal and ventral, are often active at the same time, it is possible to have a conscious visual experience of skillful perceptual-motor comportment. On the other hand, it is possible for the dorsal stream to carry on its work without interacting with the neural mechanisms responsible for consciousness.

Bergson's account of pure perception, and of the relationship between perception and skillful movement, is in accord with Goodale's description of the occipito-parietal visual stream. Bergson errs, however, in arguing that perception never serves in the preparation of visual representations, as data on the occipito-temporal pathway indicate. Given the philosophical attention lavished on representational perception, however, Bergson theory of perception provides an important sense of balance. The same may be said of Bergson's account of memory. Bergson concludes, "Now, it is no doubt possible to conceive, as an ideal limit, a memory and a perception that are disinterested; but, in fact, it is toward action that memory and perception are turned; it is action that the body prepares."[70] We can now turn to the neuroscientific literature on multiple memory systems to investigate the extent to which Bergson's distinction between representational memory and habit memory was a harbinger of future scientific theories.

---

69 Goodale, "Separate Visual Pathways," p. 24.

70 Bergson, *Matter and Memory*, pp. 227 - 228.

Chapter Four: History of Multiple Memory Systems Research

In 1953, H.M. woke up from his bilateral hippocampectomy with a severe case of multi-modal amnesia. The hippocampus, a small cortical structure named for its alleged resemblance to the sea horse, is located in the medial portion of the temporal lobe. Prior to H.M.'s operation, it was thought that the hippocampus was involved primarily in the processing of olfactory information. As a result of H.M.'s profound misfortune, scientists realized that they had underestimated the importance of the hippocampus. Brenda Milner, a neuropsychologist at the Montreal Neurological Institute, studied H.M. extensively and catalogued his lost and spared mnemonic capacities. She noted that while H.M. could not retain new information about the world or about the events of his own life, he was still capable of acquiring, refining, maintaining, and deploying perceptual-motor skills. Milner concluded that there must be at least two memory systems in the brain, a hippocampal system devoted to the processing of representational memory and a nonhippocampal system devoted to the processing of perceptual-motor skills. As a result of Milner's groundbreaking work, neuroscientists have engaged in an extensive research program, one dedicated to the mapping and analysis of the brain's multiple memory systems.

Mishkin, one of the most active researchers in the area of multiple memory systems, offers an account of the difference between hippocampal memory and nonhippocampal memory that is of particular interest here. He argues that the data on multiple memory systems should prompt us to reopen the debate between cognitivists and behaviourists. Mishkin cites a particular instantiation of this debate, conducted in the 1930's and 1940's in the *Psychological Review*, between Tolman (a *cognitivist* neo-behaviourist) and Hull (a more traditional S-R theorist). Mishkin then proposes the following olive branch hypothesis: cognitivism à la Tolman and behaviourism à la Hull may be reconciled, for they consutute two compatible account of *different* phenomena.

Cognitivists are interested in the representational processes associated with the hippocampal system while behaviourists are interested in the types of learning and memory associated with the nonhippocampal system, the "habit" or "skill" system, as Mishkin calls it.

The next chapter will be devoted to an examination of contemporary theories of multiple memory systems, including the theory of Mishkin. In this chapter, I will lay the foundation for a study of this contemporary work by reviewing two historical events: (1) the Tolman - Hull debate and (2) Milner's early work on hippocampal amnesia.

The Tolman - Hull Debate

Before turning the details of the Tolman-Hull debate, a few preliminary words are in order. In Chapter One we reviewed the basic principles of philosophical cognitivism and philosophical behaviourism. We now turn to psychological versions of these two theories. In psychology, cognitivists argue that human behaviour cannot be adequately explained without reference to mental events or processes. Behaviourists counter that such references render an explanation unscientific. Cognitivism and behaviourism are thus typically viewed as inconsistent *methodologies* and the cognitivism - behaviourism debate is construed as a debate which centres on the propriety of including mental phenomena in our scientific explanations of behaviour. When the debate between cognitivists and behaviourists is cast in these terms - as a debate about the proper role of mental predicates in scientific explanations - there is no logical hope of reconciling the two positions. Either mental predicates are permissible in scientific explanations of behaviour, or they are not. No new facts about neural memory systems could possibly shed light on this methodological debate. In order for the "olive branch" hypothesis to be plausible, cognitivism and behaviourism must be construed as distinct *substantive* accounts of the nature of learning and memory. Only then is it conceivable that cognitivist-style memory and behaviourist-style memory could be jointly accommodated by means of a theory of multiple memory systems. This is precisely what neuroscientists like Mishkin have in mind.

The psychological version of the cognitivism - behaviourism debate is relevant to my argument concerning the scope of philosophical cognitivism. The reasons for which it is relevant will be made clear, however, only after we have reviewed the Tolman - Hull debate. For the moment, it will suffice to say that both Mishkin and Hull argue that perceptual-motor skills are subserved by neural processes that are non-representational.

The second preliminary point to be made concerns the relationship between memory and learning. Until now, I have spoken almost exclusively about memory. For Tolman and Hull, however, the critical issue is the nature of learning. It is not necessary, in this context, to worry about the conceptual relationship between learning and memory. For our purposes, it will be enough to think of learning as the process in which memories (whatever they turn out to be) are acquired. Tolmanian learning is thus part of the process in which Tolmanian memories are acquired, and likewise for Hullian learning. Throughout my discussion of the Tolman - Hull debate, I will speak primarily about learning, though I mean for this discussion to be directly relevant to the questions on the table concerning the nature of memory.

Both Tolman and Hull ascribe to the view that psychologists must eschew all references to unobservable phenomena in their explanations of behaviour. Their substantive accounts of behaviour are, however, quite different. Tolman is classified as a cognitive behaviourist because he argues that phenomena like consciousness and "mental" representations can and should be included in our explanations of behaviour, provided that these phenomena are re-defined in operational terms. On Tolman's view, animals learn by constructing representational cognitive maps of their environment. Hull, on the other hand, is a strict S-R theorist who endeavors to explain all learning in terms of mathematical relationships between stimuli and responses. As Kendler writes in his history of modern psychology, "the two theories, although committed to methodological behaviourism and constructed according to a common methodological blueprint, appeared to be in complete opposition and thus generated a lively theoretical controversy that influenced the subsequent course of psychology." [1]

Hull's Theory of Learning

Hull's early career is characterized by an ambivalence toward behaviourism and an interest in a variety of "cognitive" phenomena. In 1926 he was sufficiently interested in Gestalt psychology to invite Koffka over from Germany to spend a year with him in residence at the University of Wisconsin. Many years later, in an autobiographical sketch, Hull comments on his impressions of Koffka.

> While I found myself in general agreement with his criticisms of behaviourism, I came to the conclusion not that the Gestalt view was sound but rather that Watson had not made out as clear a case for behaviourism as the facts warranted. Instead of converting me to *Gestalttheorie*, the result was a belated conversion to a kind of neobehaviourism - a behaviourism mainly concerned with the determinism of the quantitative laws of behaviour and their deductive systematization. [2]

Because he is a behaviourist, Hull considers the task of determining the "quantitative laws of behaviour" to be equivalent to the task of rigorously describing the relationships between stimuli and responses. In classical behaviourism, S-R relationships are to be explained without appeal to any intervening processes, be they mental or neural. Hull, however, does not fully accept this type of "black box" psychology. He agrees that

---

[1] Kendler, H., *Historical Foundations of Modern Psychology* (Philadelphia: Temple University Press, 1987), p. 271.

[2] Hull, Clark, "Autobiography," In E. G. Boring, H. S. Langfeld, H. Werner, and R. M. Yerkes (Eds.), *A History of Psychology in Autobiography*, (Worcester, Mass.: Clark University Press, 1952). Cited in Amsel, A., & Rashotte, M. E., *Mechanisms of Adaptive Behaviour: Clark L. Hull's Theoretical Papers, with Commentary* (New York: Columbia University Press, 1964), p. 154.

mental predicates should be banned from scientific explanation on the grounds that they are unobservable. Hull's view on the role of neural predicates represents, however, a serious departure from classical behaviourism. Stimuli and responses can be defined most precisely in terms of their underlying neural substrates which are not, in the final analysis, unobservable. On Hull's view, therefore, not only should neural predicates be taken out of the black box, they should be placed at the centre of any respectable theory of learning.

Hull starts with the phenomenon of conditioning, a process in which particular responses come to be associated with specific stimuli. He uses his study of conditioning as the foundation for an elaborate axiomatic theory of learning. His goal is not particularly modest - his stated intention is to articulate an extensive deductive system with the resources and power to explain *all* forms of mammalian behaviour. He claims that history teaches us that the top-down approach of philosophers has been a failure. Starting with conscious experience, on Hull's account, is a doomed approach. We must instead begin with the basics, with the simplest relationship between a stimulus and a response, the "habit". He writes, "I shall invert the whole historical system. I shall start with action - habit - and proceed to deduce all the rest, including conscious experience, from action, i.e. habit." [3]

Hull's concept of habit is set forth most clearly in his book, *Principles of Behaviour*. Hull's use of the term "habit" is somewhat idiosyncratic. He notes that in common usage, the term "habit" typically refers to "a well-worn mode of action." [4] For Hull, however, the term "habit" is meant to pick out "a persistent state of the organism," [5] one which results from a past history of the strengthening of particular stimulus-response

---

[3] Ammons, R. B., "Psychology of the Scientist: IV. Passages from the 'Idea Books' of Clark L. Hull," *Perceptual and Motor Skills*, 15 (1962), pp. 807 - 82. Cited in Amsel, A., & Rashotte, M. E., *Mechanisms of Adaptive Behaviour: Clark L. Hull's Theoretical Papers, with Commentary* (New York: Columbia University Press, 1964), p. 6.

[4] Hull, C. L., *Principles of Behaviour* (New Haven: Yale University Press), footnote, p. 102.

[5] Hull, *Principles of Behaviour*, footnote, p. 102.

associations via reinforcement. A habit is the neural connection between the afferent neural impulse triggered by the stimulus and the efferent neural impulse that generates a motor response. When Hull refers to S-R associations, he is talking about the internal, physical connections set up between incoming sensory messages and outgoing motor messages. A habit, for Hull, consists of a strengthened connection between these two types of neural messages. "The process of habit formation consists of the physiological summation of a series of discrete increments, each increment resulting from a distinct receptor - effector conjunction ... closely associated with a reinforcing state of affairs." [6]

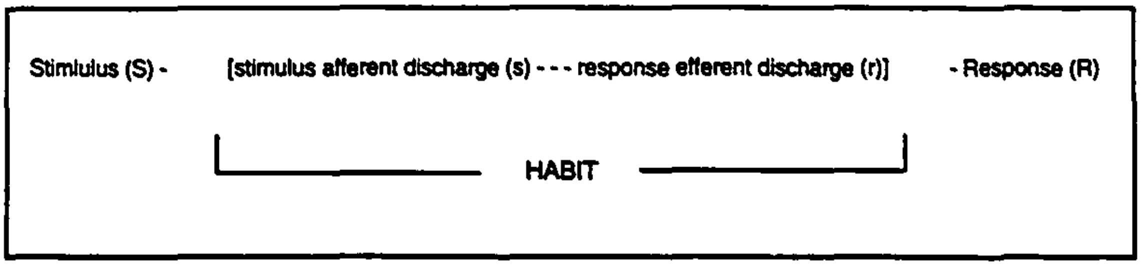

Figure 4.1: Hull's Conception of Habit

The important feature of a habit, the feature that plays a causal role in behaviour, is the strength of the neural association between the stimulus afferent discharge and the response efferent discharge, the "s - - - r" association. In Hull's terminology, this phenomenon is called the habit strength and is symbolized "sHr". Because the stimulus (S) is closely related to the neural message it incites (s), and because the motor efferent discharge (r) is closely related to the actual response (R) of the organism, habit strength may also be defined in terms of the actual stimulus and the actual response. In this case, habit strength is symbolized as "$S^{H}R$".

Habit strength ($S^{H}R$) is a logical construct which reflects the number of times the S-R connection has been reinforced in the past. Habit strength manifests itself in a variety of ways: the magnitude of the reaction, the latency of the reaction, and the percentage of

6 Hull, *Principles of Behaviour*, p. 102.

"correct" responses elicited by the stimulus in question. Use of the term "habit strength" is a symbolic convenience. As Hull writes, "the chief advantage to be expected of the logical construct *habit strength* arises from economies in thought, i.e. in symbol manipulation." [7] The bulk of Hull's learning theory consists of the derivation of precise mathematical formulae in which the various features of habit strength are clarified.

Hull thus begins with the concept of habit strength. It is necessary, however, to introduce several modifications to this concept in order to arrive at a more accurate and refined account of the relationship between stimuli and responses. I will discuss a few of these modifications in order to suggest the general flavour of Hull's explanatory program. This first important modification to the concept of habit strength derives from the observation that stimuli that are similar to, but not identical with, the reinforced stimulus are also somewhat efficacious in eliciting the reinforced response. Hull calls this phenomenon "stimulus generalization" and attempts to quantify the relationship between the strength of the response and the degree to which the new stimulus is related to the original stimulus.

When stimulus generalization is taken into account, the resulting habit strength becomes, in Hull's terminology, the "effective habit strength". The effective habit strength, in combination with the organism's relevant drives, yields the "reaction potential". Reactions do not always go forward, however, as there are a variety of inhibitory influences at work. If these inhibitory influences are taken into account, we arrive at the actual "effective reaction potential". This construct is the centrepiece of Hull's theory of learning. "The pivotal theoretical construct of the present system is that of the effective reaction potential." [8]

---

[7] Hull, *Principles of Behaviour*, p. 111. Hull describes habit strength both as a logical construct and as an empirical phenomenon. A sorting out of this problem would take us too far afield. We may think of the habit strength as an empirical phenomenon whose features serve as variables in Hull's theorems.

[8] Hull, *Principles of Behaviour*, footnote, p. 342.

Hull is committed to a mechanistic account of behaviour. All the same, he is forced to acknowledge one essential difference between organisms and machines. Organisms, but not machines, exhibit a great deal of variability in their reactions to stimuli. Hull compares a human calculator and a mechanical calculating device. The latter is far more reliable in its reactions, and anomalies in its performance can be traced to mechanical malfunctions. The same cannot be said of the human.

> While first rate calculating machines sometimes get out of order and make errors, ordinary inorganic machines under the same external conditions show, in general, much less variability in behaviour than do organisms. Indeed variability, inconsistency, and specific unpredictability of behaviour have long been recognized as the chief molar distinctions between organisms and inorganic machines.[9]

Hull attributes this behavioural variability to the complexity of the neurophysiological underpinnings of the "afferent discharge - efferent discharge" connection. The variability in responses to a single stimulus can be explained in terms of an oscillation function which "modifies the intensity of every muscle contraction involved in every coordinated reaction."[10] According to Hull, the principle of oscillation makes a scientific explanation of behaviour more difficult but not impossible. Hull claims that the variability of behaviour has been an obstacle to progress in the social sciences. "It may be said that the principle of behavioural oscillation is to a large extent responsible for the relatively backward condition of the social, as compared with the physical, sciences."[11] Hull argues, however, that oscillation itself can be studied and quantified. When oscillation is taken into account, we arrive at what Hull calls the "momentary effective reaction potential."

---

9 Hull, *Principles of Behaviour*, p. 304.

10 Hull, *Principles of Behaviour*, p. 318.

11 Hull, *Principles of Behaviour*, footnote, p. 317.

Hull's treatment of the issue of oscillation bears witness to his allegiance to the principles of mechanistic explanation. He sees all behaviour as explicable in terms of the neural association of stimuli and responses. All behaviour can be viewed as the result of the interplay of basic drives and S-R learning. Experience occurs when certain reactions are preferentially reinforced because they lead to drive reduction. In order to explain behaviour, we must, on Hull's account, formalize and quantify the various elements of the reinforced S-R associations.

Tolman's Theory of Learning

0He accepts the view that mental phenomena are out of place in scientific explanations of behaviour. On the other hand, Tolman seeks to preserves mentalist discourse by defining terms such as "consciousness" and "mental representation" operationally. The result is an unstable combination of faux behaviourism and stymied cognitivism. Tolman's work is relevant here because he serves well as a foil for Hull and because his account of "representational" memory lines up well with contemporary accounts of hippocampal memory.

Tolman begins by asking us to notice that there are two distinct ways of construing behaviour. He points out that behaviourists typically construe behaviour mechanistically, as a series of movements or motions made by the animal. "Molecular" behaviour of this sort requires a physiological explanation. Tolman points out that if all behaviour were molecular in this sense, psychology would be left without a true *raison d'être*. Psychology does have an important role to play, according to Tolman, precisely because all behaviour is not molecular; some behaviour is molar. "Molar" behaviour consists of patterns of goal-directed activities. "Behaviour for me is not as it is for many, probably most, behaviourists primarily a matter of mere muscle contraction and gland secretion, of mere 'motions.'" [12] Goal-directed activity is irreducible to physical movements; it exhibits emergent properties whose analysis is to be the proper subject matter of a scientific psychology. Purposive action can be studied as an independent phenomenon, without seeking recourse in a reductive physiology. "I conceive behaviour ... as presenting a new and unique set of descriptive properties all its own, - new properties which, as such, can be described and known, irrespective of whatever muscular or glandular activities underlay them." [13]

Tolman's argument against the S-R theorist may be summarized as follows:

P1. Most behaviour is purposeful.

P2. If a behaviour is purposeful, it is cognitive.

P3. If a behaviour is cognitive, it is not reducible to S-R associations.

C. Therefore, most behaviour is not reducible to S-R associations.

Purposes and cognitions are the molar characteristics of behaviour that require a psychological, not physiological, analysis. This definition of the agenda of psychology represents a departure from classical behaviourism. Tolman was confident that a mechanistic approach to behaviour was ill-founded but he sometimes mentions his discomfort with his apparent transgression of the rule to confine his research to observable phenomena. [14] As a result, he endeavors to redefine 'purpose' and 'cognition' in operational terms that would be more palatable to his peers.

---

12 Tolman, E. C., "A Behaviouristic Theory of Ideas," *Behaviour and Psychological Man: Essays in Motivation and Learning* (Berkeley: University of California Press, 1966), p. 49. (Originally published in *Psychological Review* 33 [1926], pp. 352-369.)

13 Tolman, "A Behaviouristic Theory of Ideas," p. 49.

14 See, for example, Tolman, E.C., "A Behaviourist's Definition of Consciousness," *Behaviour and Psychological Man: Essays in Motivation and Learning* (Berkeley: University of California Press, 1966), p. 63. (Originally published in *Psychological Review* 34 [1927], pp. 433 - 439.)

Tolman argues, therefore, that purpose is immanent in observable behaviour. In running a maze, a rat persists until it locates the reward. Once the reward is obtained, the rat ceases its search. According to Tolman, purpose is manifest in the rat's persistence. Conversely, a rat will attempt to move away from noxious stimuli or negative goals. "Whenever, in merely describing a behaviour, it is found necessary to include a statement of something *toward-which* or *from-which* the behaviour is directed, there we have purpose." [15] Purpose is manifest in pursuit and in avoidance. The ultimate goals in animal behaviour are the pursuit of pleasure and the avoidance of pain. The behaviour of the rat in the maze is explicable, on Tolman's account, only once we acknowledge the goal or "*telos*" of the rat. "What, now, are purposes? ... They are persistences to or from. But to or from what? We will answer quite dogmatically... that they are persistences to or from ... states of bodily quiescence or of bodily disturbance." [16] Tolman's operational definition of purpose is the persistence an animal exhibits in pursuing pleasure or avoiding pain. The rat runs until it finds food and then it ceases to run. This persistence is itself observable and is thus an appropriate element of an empirical explanation. "Purpose, adequately conceived, it will be held, is itself but an empirical aspect of behaviour." [17] Having analyzed purposive behaviour in this way, Tolman takes it as obvious that most behaviour is indeed purposeful.

According to Tolman, if a behaviour is purposeful, it is cognitive. Any action that counts as a pursuit of a goal manifests certain hypotheses about the nature of the environment. "Every behaviour act, in going off and being what it is, expresses, implies, certain specific characteristics in the environment." [18] Having a purpose renders certain

---

15 Tolman, E. C., "Behaviourism and Purpose," *Behaviour and Psychological Man: Essays in Motivation and Learning* (Berkeley: University of California Press, 1966), p. 35. (Originally published in *Journal of Philosophy* 22 [1925], pp. 36-41.)

16 Tolman, "A Behaviouristic Theory of Ideas," p. 52.

17 Tolman, "Behaviourism and Purpose," p. 33.

18 Tolman, "A Behaviourist's Definition of Consciousness," p. 64.

features of the environment salient. The environment serves as a collection of objects that aid or hinder the animal in its pursuit of a goal. "Behaviour is driven by organic needs, and in going off it postulates that the environmental characters and relations are such that it will prove an appropriate behaviour for satisfying those needs." [19]

Behaviour is a transaction between purposeful organisms and the environment. An action is the means by which the organism negotiates between its agenda and the opportunities afforded by its current environment. An action is thus specified as a means to a particular goal; it is "a specific pattern of commerce-, intercourse-, engagement-, community- with such and such intervening means-objects, as the way to get thus to or from (a goal or noxious stimuli)." [20] An action is thus a mutual exchange between a purposeful agent and the salient elements of its world. Organisms do not, on Tolman's account, perceive the world in terms of static bits of sense data. The environment is perceived in terms of meaningful Gestalt patterns.

> The environment as so envisaged is thus naught but a very field or tissue of means-end relations. It is a *means-end-field* in which the various component objects and situations appear ineluctably in their rôles of possible, or impossible, good, or bad, better or worse, *means* to, or from, such and such other objects or situations. [21]

In engaging in purposeful behaviour, an animal manifests three types of cognitive postulations: (1) postulations of discrimination features, (2) postulations of manipulation features and (3) postulations of units of behaviour that capitalize on the opportunities afforded by these discrimination and manipulation features. For example, a human faced with a chair detects certain of its physical properties (e.g. size and shape), appreciates the

---

19 Tolman, "A Behaviourist's Definition of Consciousness," p. 64.

20 Tolman, E.C., *Purposive Behaviour in Animals and Men* (New York: Century, 1932), p. 11.

21 Tolman, E. C., "Gestalt and Sign-gestalt," *Behaviour and Psychological Man: Essays in Motivation and Learning* (Berkeley: University of California Press, 1966), p. 86. (Originally published in *Psychological Review* 40 [1933], pp. 391 - 411.)

various ways in which the chair may be manipulated (e.g., one may sit or stand on it, use it for firewood or as a weapon, etc.) and postulates specific behaviour units (e.g. "sitting on the chair"). Depending on an animal's constitution and its agenda, a given object will present different features, different possibilities for manipulation, different invitations for action. "We must know that while a thing the size of a chair will, for a man, present manipulation-features such as to-be-sat-on-ness, to stand-on-ness...; for a rat, it will present quite a different set of manipulation features, those, say, of to build-a-nest-in-ness or to-hide-behind-ness." [22]

Discrimination features and manipulation features may be combined and re-combined to yield increasingly complex types of behaviour units. These behaviour units constitute a network of possible animal - environmental interactions. In order to explain any given piece of behaviour, the scientist must, on Tolman's account, specify the animal's purpose in engaging in the behaviour. Behaviour units, by their very nature, cannot be defined without reference to the animal's agenda. On Tolman's view, S-R theory is fundamentally flawed because it misconstrues the basic relationship between an organism and the environment. The relationship between the animal and the environment is not merely mechanistic; it is a relationship created by the purposes of the organism and the perceived usefulness of the environment with respect to these purposes. This relationship cannot be reduced to a series of stimulations that issue in atomized responses.

> Behaviour as a type of commerce with the environment can take place only in a whole organism. It does not take place in specific sensory and motor segments, which are insulated and each by itself.[23]

Tolman proposes an explanatory framework which is meant to replace S-R explanations of behaviour. He argues that behaviour is a function of three types of variables: (1) environmental variables (e.g. the nature of the goal-object, the types of motor

---

[22] Tolman, "A Behaviouristic Theory of Ideas," p. 56.

[23] Tolman, *Purposive Behaviour in Animals and Men*, p. 18.

responses required to approach or avoid it, etc.); (2) individual differences variables (e.g. heredity, age, previous training, physiological condition, etc.); and (3) intervening variables. Environmental variables and individual differences variables are straightforward - but the issue of intervening variables requires some explanation.

Behaviour = *f* (Intervening Variables + Environment + Individual Differences)

An intervening variable is a theoretical construct that bridges the gap between the dependent variable (in this case, behaviour) and various independent variables (e.g. heredity, physical constitution, training, etc.). Intervening variables are intended to provide a theoretical "short-hand" for capturing important patterns in the relationship between the dependent variable and some number of independent variables. Tolman justifies his use of intervening variables by noting that each intervening variable remains closely tied to both the dependent variable, behaviour, and the relevant independent variables.

We have already seen two examples of intervening variables: purposes and cognitions. We can now put these examples in a broader context. According to Tolman, there are three general types of intervening variables: (1) "Immanent Determiners of Behaviour", (2) Capacities, and (3) Behaviour Adjustments. Cognitions and purposes are the immanent determiners of behaviour; they are immanent in particular behaviour-acts. Capacities are "the endowments of the individual or the species which result from ... innate endowment and past training." [24] An animal's capacities are its abilities to make appropriate cognitive postulations.

This brings us to the third and final type of intervening variable: the behaviour-adjustment. The term "behaviour-adjustment" is meant to be the operational or behaviouristic term for the conscious evaluation of potential courses of action. "Behaviour-adjustments constitute our behaviouristic substitution for, or definition of,

---

[24] Tolman, *Purposive Behaviour in Animals and Men*, p. 439.

what the mentalists would call conscious awareness and ideas. They are unique organic events which may on certain occasions occur in an organism as a substitute, or surrogate, for actual behaviour." [25] An animal is typically immersed in its ongoing goal-directed behaviour. If, however, it can "pause" for a moment, so to speak, it will be able to "consider" what to do next. This moment of reflection alters the course of the animal's behaviour. "(Behaviour-adjustments) function to produce some sort of modifications or improvements in what were the organism's initially aroused immanent determinants, such that his final behaviour, corresponding to these new modified immanent determinants, is different from what it otherwise would have been." [26]

One way of thinking about "behaviour adjustments" is in terms of the animal's ability to respond differentially to distinct environmental cues. Consider a simple maze in which a rat must, at some point, enter either a green door or a red door. The relative right-left positions of the red and green doors are changed pseudo-randomly over trials, but the green door always leads to food, while the red door never does. When the rat runs its initials trials on the maze, its response to red and green is "undifferentiated". At some point, after a certain critical number of trials, the rat appears to "catch on" to the connection between the green door and the food. As Tolman puts it, some "internal change" occurs in the rat, such that its pattern of responses at the choice point is altered. Now the rat responds differentially; it chose green every time. The change from undifferentiated responses to differentiated responses is critical. "The moment of this switch is the moment of consciousness." [27]

This moment of consciousness may be fleeting, for once the animal has differentiated between the two types of cues and realized their import, the act of differentiation may become automatic. In other words, cognitive differentiation, in and of

---

[25] Tolman, *Purposive Behaviour in Animals and Men*, p. 20.

[26] Tolman, *Purposive Behaviour in Animals and Men*, p. 20.

[27] Tolman, "A Behaviourist's Definition of Consciousness," p. 65.

itself, is not a sign of consciousness. It is the *change* from undifferentiated responses to differentiated responses that signals the involvement of consciousness. Once the rat has associated "green door" with "food", his "correct" behaviour at the choice point may become quite automatic. "Acts which imply more cognitive differentiation may be just as automatic as ones which imply less cognitive differentiation. It is only the switch-over when it occurs in a given moment of stimulation that defines consciousness." [28]

Tolman's discussion of the behaviour adjustment provides the occasion for his introduction of the concept of representational cognition. According to Tolman, the achievement of moving from undifferentiated to differentiated responses requires the ability to "represent" the future in the present. The rat represents itself choosing the red door to no avail. It then "pictures" itself choosing the green door and finding food. Because the rat can represent the future accurately, it can adjust its behaviour so that its purposes and demands are most likely to be fulfilled. Tolman writes, "To make an adjustment to an act is to achieve a *representation* (based, of course, upon what has happened upon previous occasions when this earlier act or similar ones have actually been performed) of the probable stimulus results to be expected from the act." [29] The overt sign of a representation is a sudden and abrupt change in the pattern of responses made by the animal. This change will be manifest in the animal's learning curve. "Whenever there is a sudden drop in the learning curve, there is consciousness. For only by representation of its results (through memory or imagination) could acts hitherto infrequent become thus suddenly and consistently frequent." [30] Tolman's theory of behaviour is cognitivist in the following sense: he claims that if we wish to explain the patterns observable in intelligent behaviour, we must posit internal representational states.

---

28 Tolman, "A Behaviourist's Definition of Consciousness," p. 65.

29 Tolman, "A Behaviourist's Definition of Consciousness," p. 65.

30 Tolman, "A Behaviourist's Definition of Consciousness," p. 60.

The rat alters its behaviour by representing to itself actions from the past. The representation of future actions is a more complex procedure involving "foresight". Tolman cites Kohler's studies on tool use in apes as an example of foresight learning. In one experiment, the ape succeeds in obtaining a food reward only if he "realizes" that the act of extending his arm in conjunction with the act of holding a nearby stick, will allow him to reach the food. As Tolman describes it, "the new insight arises, in other words, out of an ability to achieve in representation the results of compound acts, only the component parts of which have previously been performed." [31] Foresight learning is, in some sense, more "cognitive" than purposive behaviour that does not involve foresight. This is because foresight learning manifests what we now take to be the trademark feature of cognitive functioning - the manipulation of "mental" representations. "Foresight learning involves free play among representations: - the ability to add them together (and also, in some instances, to subtract them) and thus to create new representations." [32]

Toward the end of his career, Tolman devoted himself to the development of a specific theory of cognitive learning - his theory of cognitive mapping. The notion of cognitive mapping is foreshadowed by his earlier discussion of foresight learning, insofar as foresight learning involves the ability to work with more than one representation at a time. Tolman later argues that animals create full-fledged representational "maps" of their environment - maps which consists of numerous representations of the features and locations of various objects in the environs. Tolman modifies his argument against the S-R theorist to take the principles of cognitive mapping into account. In the case of purposive learning in general, S-R theory fails because it does not acknowledge the complex relationship between a goal-driven animal and the environment. In the case of cognitive

---

31 Tolman, E.C., "Habit Formation and Higher Mental Processes in Animals, Part 2," *Psychological Review* 25 (1928), p. 50.

32 Tolman, "Habit Formation and Higher Mental Processes in Animals, Part 2," p. 50

map learning, S-R theory fails because it underestimates the complexity of what is going on "inside" the animal.

Tolman's paper, "Cognitive Maps in Rats and Men," begins with a general description of the behaviour of trained rats in baited mazes. If a rat is given the opportunity to wander through a maze repeatedly, and if there is a food reward at a constant and specific place on the maze, then the amount of time it takes the rat to reach the reward will decrease with the number of times the rat is exposed to the maze. Tolman notes that there are two competing interpretations of the rat's behaviour: S-R theory and "field" theory. The S-R theorist maintains that the highly trained rat's unerring and direct approach to the food site is the result of acquired S-R bonds.

> The rat's central nervous system, according to this view, may be likened to a complicated telephone switchboard. There are the incoming calls from the sense organs and there are the outgoing messages to muscles. ... Learning ... consists in the respective strengthening and weakening of various of these connections: those connections which result in the animal's going down the true path become relatively more open to the passage of nervous impulses, whereas those which lead him into the blinds become relatively less open. [33]

Field theorists, on the other hand, argue that the rat's successful behaviour should be attributed to the acquisition of a "cognitive map" of the environment. Field theorists, and he counts himself among them, claim that the rat's behaviour is too sophisticated to be attributed solely to passively experienced neural associations between stimuli and responses. On this view, the mastery of the maze requires that the animal be actively engaged in the pursuit of spatial information about the environment. The animal must, that is, form a cognitive map of its world.

---

[33] Tolman, E. C., " Cognitive Maps in Rats and Men," *Behaviour and Psychological Man: Essays in Motivation and Learning* (Berkeley: University of California Press, 1966), pp. 242-43. (Originally published in *Psychological Review* 55.3 [1948], pp. 189-

> We assert that the central office is far more like an old-fashioned telephone exchange. The stimuli, which are allowed in, are not connected by just simple one-to-one switches to the outgoing response. Rather, the incoming impulses are usually worked over and elaborated in the central control room into a tentative, cognitive-like map of the environment. And it is this tentative map, indicating relationships, which finally determines what responses, if any, the animal will finally release. [34]

Having identified the two alternative hypotheses about the rat's behaviour, S-R theory and field theory, Tolman proceeds to describe a number of experiments that provide support for the field theorists. I will describe two of these experiments here. Tolman argues that (1) latent learning and (2) place learning cannot be described in terms of the acquisition of S-R associations. In latent learning, animals acquire information about a stimulus even when the stimulus is not itself associated with a reward. In place learning, animals acquire spatial information that is flexible and map-like, i.e. it is not information about particular pathways but about the overall spatial structure of the environment.

The first experiment, performed by Spence and Lippitt, tests the capacity of rats to engage in latent learning. [35] In this experiment, rats who are satiated with respect to both food and water are placed on a forced-choice Y-maze which is food-baited on the left and water-baited on the right. The rats are enticed to run the maze via the reward of being returned to a "living cage" in which they can interact with other rats. After a week of training, the rats are divided into two groups: a food-deprived group and a water-deprived group. When these rats are returned to the maze, the hungry rats more frequently choose

---

208.)

[34] Tolman, "Cognitive Maps in Rats and Men," pp. 244 - 45.

[35] Spence, K. W., G. Bergman, and R. Lippitt, "A Study of Simple Learning Under Irrelevant Motivational-Reward Conditions," *Journal of Experimental Psychology* 40.5 (1950), pp. 539 - 551.

the food-baited arm while the thirsty rats more frequently choose the water-baited arm. Tolman interprets these results as evidence in favour of the field theory.

> In short, they had acquired a cognitive map to the effect that food was to the left and water to the right, although during the acquisition of this map, they had not exhibited any stimulus-response propensities to go more to the side which later became the appropriate goal.[36]

Tolman decided to investigate the spatial characteristics of rat learning in more detail. To that end, he devised his now famous "starburst" experiment. In phase one of this study, a group of rats is trained to run directly from the starting box to the goal box in an maze with the structure indicated here on the left.

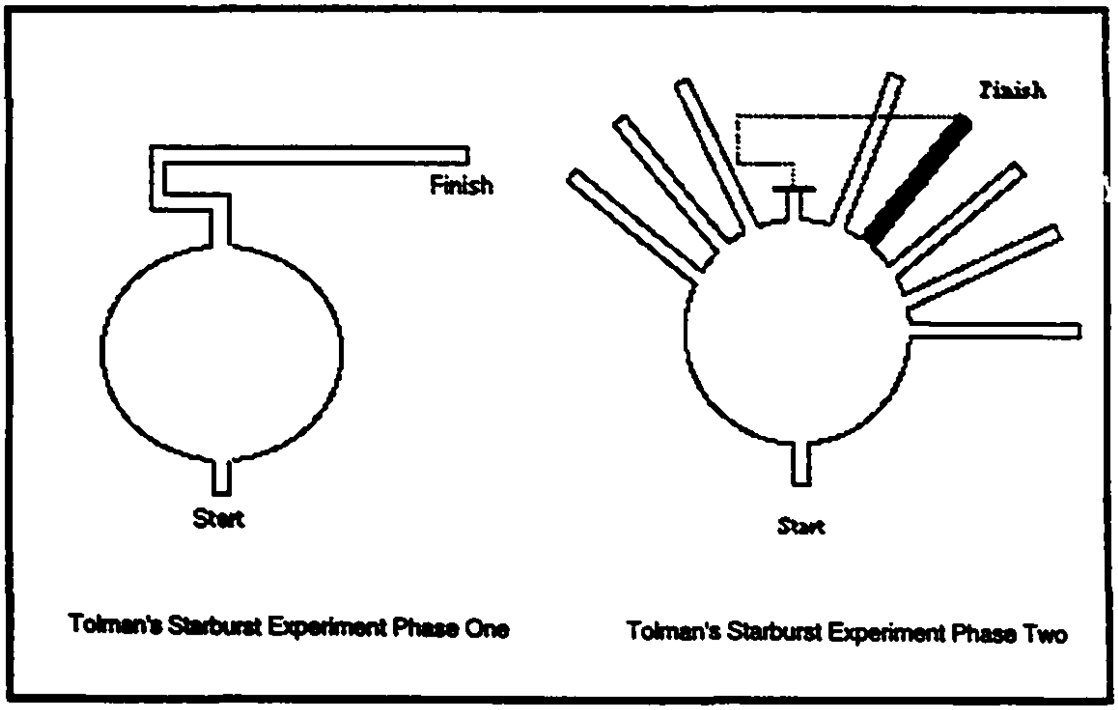

Figure 4.2: Tolman's Place Learning Experiment

In phase two, the maze is modified: a sunburst configuration of alleys is added to the maze and the original passageway is blocked. When rats are first introduced to this

[36] Tolman, " Cognitive Maps in Rats and Men," pp. 250.

new maze, they run to the original passageway only to discover that it has been blocked. The rats then essay a few of the starburst arms, traveling only a few inches into each arm. Each rat ultimately chooses one arm and travels its entire length. There is a marked tendency for the rats to choose the arm that ends at a point quite near the site of the original food reward box. Tolman writes, "As a result of their original training, the rats had, it would seem, acquired not only a strip map to the effect that the original specifically trained-on path led to food, but rather, a wider comprehensive map to the effect that food was located in such and such a direction in the room." [37]

It is interesting that Tolman notes that these rats have acquired two different methods for reaching the goal box. The first method is the rat's acquisition of a specific route between the start point of the maze and the goal box. It is because the rats have learned a particular route that they first attempt to repeat their steps on the starburst maze. Only when this option is no longer available do the rats invoke a second strategy - cognitive mapping. In this case, the rats appear to select a pathway based on a conception of where the food is located in space. Tolman notes that certain conditions will favour the use of route strategies while other conditions are more hospitable for cognitive map strategies. Route strategies or "strip" maps are favoured, according to Tolman, in cases of brain damage, inadequate environmental cues for the construction of a cognitive map, or over-training on a particular route.

In one of his final papers, "There is More Than One Kind of Learning," [38] Tolman seeks a truce with his more conventional behaviourist peers. He notes that while there are several important types of cognitive learning (e.g. drive discrimination, goal cathexis, and mapping), there is one type of learning that may not be cognitive at all - the learning of simple motor patterns. In discussing motor pattern acquisition, Tolman notes that we may

---

[37] Tolman, " Cognitive Maps in Rats and Men," p. 258.

[38] Tolman, E.C., "There is More than One Kind of Learning," *Psychological Review* 56 (1949), pp. 144-155.

think of goal-directed movements as involving combinations of motor patterns. These constitutive sequences of movements must themselves be learned at some point, in order to make purposive actions possible. Tolman then grants that the principles of simple conditioning (as analyzed by his contemporary Guthrie) may govern motor pattern acquisition.

> We do build up, I believe, many motor patterns (the old name was sensory-motor skills) which we carry around with us equipment for behaving in new situations. And, whereas, I do not think we as yet know much about the laws for the learning, unlearning and forgetting such motor patterns, I am willing to accept, for the present, Guthrie's notions concerning their learning and unlearning. [39]

In this article, Tolman suggests that competing psychological theories can be accommodated within a framework of pluralism and offers this prescription for détente. "I wish to suggest that our familiar theoretical disputes about learning may *perhaps* (I emphasize 'perhaps') be resolved, if we can agree that there are really a number of different kinds of learning. For then it may turn out that the theory and laws appropriate to one kind may well be different from the those appropriate to other kinds." [40]

Tolman's early work on cognitive learning emphasizes the notion of purposive behaviour. We may think of purposive behaviour as being cognitive in a relatively "weak" sense; it is cognitive insofar as is *manifests* cognitive postulations about the environment. Cognitive mapping behaviour is cognitive in a much stronger sense; it is cognitive because it involves the acquisition and manipulation of relational representations. To review, we may say that there are three types of learning in Tolman's classificatory schemata: (1) motor pattern learning which is not itself purposeful and which manifests no cognitive

[39] Tolman, "There is More than One Kind of Learning," p. 154. In summarizing the article Tolman writes, "As to the laws for the acquisition of motor patterns, per se, I suggested that Guthrie's principle of simple conditioning may perhaps be correct." Ibid., pp. 154-55.

[40] Tolman, "There is More than One Kind of Learning," p. 144.

postulations about the environment, (2) weak cognitive learning which is goal directed and does manifest simple egocentric postulations about the environment, and (3) strong cognitive learning which is not goal-directed in an immediate sense and which requires that the animal construct a non-egocentric [41] representational map of its environment.

When neuroscientists speak of Hullian learning, they are referring to the acquisition and maintenance of S-R connections at the neural level. When they speak of Tolmanian learning, the situation is not as clear. For the most part, the term "Tolmanian learning" has come to refer to cognitive learning in the strong sense - cognitive map learning. In some cases, however, scientists use the term "Tolmanian learning" to refer to cognitive learning in the weak sense - the learning associated with any type of purposeful behaviour.

While both Tolman and Hull are behaviourists in some sense, they offer decidedly different substantive accounts of learning. Hull proffers a neurally based S-R theory and argues that *all* behaviour can be explained in terms of the notion of habit strength. Tolman counters that intelligent behaviour cannot be explained without positing representational internal states. As has been mentioned, this debate provides the theoretical framework for Mishkin's account of multiple memory systems. It is also worth noting that Tolman's theory of cognitive mapping has been incorporated into one the most important theories of multiple memory systems in non-human animals. O'Keefe and Nadel, in their monograph of 1978, *The Hippocampus as a Cognitive Map*, argue that the hippocampus is specialized for the processing of representational spatial information, information that is coded into Tolmanian cognitive maps. In the next chapter, we will look more carefully at O'Keefe and Nadel's theory of hippocampal mapping. For the moment, however I will continue my

---

[41] Egocentric spatial information involves a reference to the position of the animal. (e.g. you should turn right when you reach the landmark). Non-egocentric spatial information does not involve a reference to the position of the organism (e.g. the red door is ten centimeters to the east of the green door).

account of the history of multiple memory systems research by reviewing Milner's early work on hippocampal amnesia.

Human Amnesia Studies Phase One: Hippocampal Amnesia

In the early 1950's scientists viewed the temporal lobes as a locus of visual processing. Milner, a neuropsychologist who specializes in the temporal lobes, details the evidence for such a view in her 1954 article, "Intellectual Function of the Temporal Lobes." [42] By the time this article was published, Milner had already begun an extensive new study which led her to modify her position. On her new view, the lateral portions of the temporal lobe are indeed devoted to visual processing, but the medial portions are devoted to the processing of memories. Milner reached this conclusion while conducting post-surgical cognitive studies on the temporal lobe patients of Scoville and Penfield. Here I will recount the cases of four particularly important patients, two patients of Scoville, H.M. and M.B., and two patients of Penfield, F.C. and P.B.

Case #1: M.B. Between 1950 and 1957, Scoville performed thirty partial temporal lobe resections on "severely deteriorated" psychiatric patients who had not responded to other forms of treatment. [43] In one case, Scoville removed an unusually extensive amount of temporal lobe tissue. "In one case only in this psychotic group all tissue mesial [44] to the temporal horns for a distance of at least 8 cm. posterior to the temporal tip was removed, a removal which presumably included the anterior two-thirds of the hippocampus complex bilaterally." [45] This radical operation was carried out on December 18, 1952 on

---

42 Milner, Brenda, "Intellectual Function of the Temporal Lobes," *Psychological Bulletin* 51.1 (1954), pp. 42-62.

43 Scoville, William, and Brenda Milner, "Loss of Recent Memory after Bilateral Hippocampal Lesions," *J. Neurol. Neurosurg. Psychiat.* 20 (1957), p. 11.

44 "Mesial" is a synonym for "medial".

45 Scoville and Milner, "Loss of Recent Memory," p. 11.

a 55-year old female psychotic known as M.B.[46] While others who had received similar but less radical operations suffered some degree of memory impairment, M.B. emerged from her operation with a severe case of "global" or multi-modal amnesia.

<u>Case #2: H.M.</u> On September 1 of 1953, Scoville performed the same operation on a young epileptic, H.M., in the hopes of reducing the frequency and severity of his seizures. H.M.'s seizures had proven to be pharmacologically intractable and the "frankly experimental" surgery was undertaken as a last resort.[47] Though there was no discernible focal area of epileptogenic activity in H.M.'s brain, the temporal lobes, particularly the hippocampus, had been identified as a particularly active area for epileptics in general. Scoville thus reasoned that the removal of the hippocampus and surrounding tissue might bring some relief to H.M., who was "totally incapacitated" by his seizures.[48] The surgery proved to be successful with respect to H.M.'s seizures, which were, for the most part, brought under control. H.M.'s memory was, however, profoundly disturbed by the operation. After the operation, H.M., like M.B., was severely amnesic.

<u>Case #3: F.C.</u> Between 1953 and 1958, neuropsychologists, neurosurgeons, and neurologists at the Montreal Neurological Institute tested over one hundred patients both before and after unilateral temporal lobe surgery undertaken to provide some relief from epileptic seizures. In two cases, this surgery resulted in a severe memory impairment. The first of these patients, F.C., received a unilateral left temporal lobectomy on October 21, 1952. While F.C.'s epilepsy was ameliorated by the surgery, he emerged from the operation with a severe case of amnesia.

---

[46] Because some "clinical" success had been achieved via the undercutting of the orbital frontal lobes in severely psychotic patients, and because the orbital frontal lobes are highly interconnected with the medial temporal lobes, it was thought that the removal of the medial temporal lobe would also prove to be clinically beneficial.

[47] Scoville and Milner, "Loss of Recent Memory," p. 11.

[48] Scoville and Milner, "Loss of Recent Memory," p. 11.

Case #4: P.B. The second patient, a 41-year old civil engineer, received a partial left temporal lobectomy on August 14, 1946. The purpose of the operation was the removal of a localized region of focal epileptogenic activity. In this surgery, the hippocampus was not removed. When the surgery proved to be insufficiently therapeutic, P.B. underwent complete unilateral left temporal lobe removal (including the hippocampus). After the second (but not the first) operation, P.B. developed a case of severe anterograde amnesia.

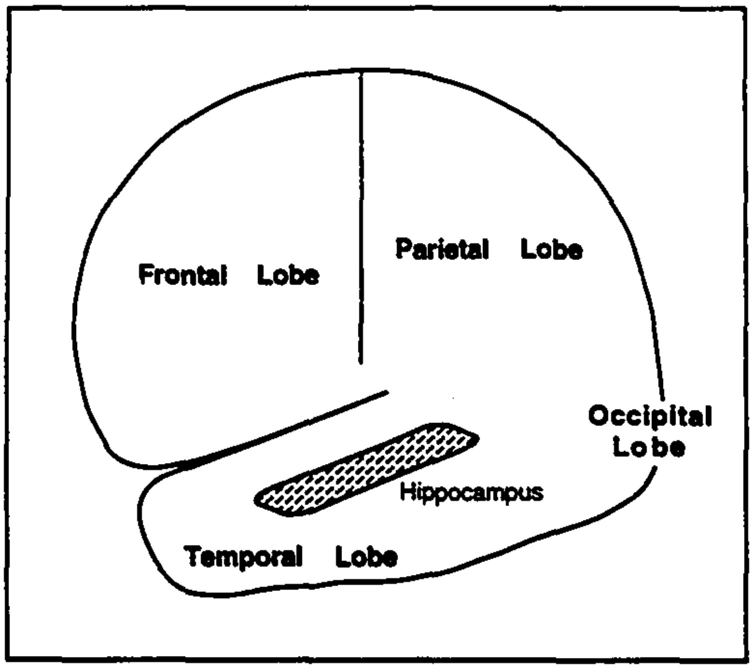

Figure 4.3: Location of the Hippocampus in the Medial Temporal Lobe

All four patients experienced an unusually severe form of amnesia post-operatively. It was as though, for these four individuals, experience left no memory trace. Descriptions of H.M.'s post-operative behaviour convey the severity of his memory loss. Scoville and Milner describe H.M.'s condition: "After the operation this young man could no longer recognize the hospital staff nor find his way to the bathroom, and he seemed to recall nothing of the day-to-day events of his hospital life." [49]

---

[49] Scoville and Milner, "Loss of Recent Memory," p. 14.

The memory deficit experienced by these individuals persisted over time, and when Milner interviewed H.M. 19 months after his surgery, he claimed that it was still 1953 and that he was still 29 years old. It was as though no time had passed for him in the interval. Milner writes that "this patient appears to have a complete loss of memory for events subsequent to bilateral temporal-lobe resection 19 months ago." [50] His family reported that H.M. could read the same magazines repeatedly without recognizing the material. As Scoville and Milner describe it, these patients "appear to forget the incidents of their daily life as fast as they occur." [51] Milner examined M.B. in 1955 and found that "she showed a global loss of memory similar to that of H.M." [52] The same type of amnesia also afflicted the unilateral patients, F.C. and P.B. Milner describes their deficit as follows. "It is an amnesia for ordinary day-to-day events, and on formal testing it is found to affect all kinds of test material, including stories, drawings, and new word-associations. It is a defect which is not specific to one sense modality and which also cuts across the distinction between verbal and nonverbal material." [53]

In studying the thirty temporal lobe patients of Scoville, Milner discovered that the degree of memory impairment correlated well with the extent of the hippocampal damage. Full blown amnesia occurred only when *significant* portions of *both* hippocampi were removed. Milner and Scoville concluded that "bilateral medial temporal-lobe resection in man results in a persistent impairment of recent memory whenever the removal is carried out far enough posteriorly to damage portions of the anterior hippocampus and hippocampal gyrus." [54]

---

[50] Scoville and Milner, "Loss of Recent Memory," p. 17.

[51] Scoville and Milner, "Loss of Recent Memory," p. 15.

[52] Scoville and Milner, "Loss of Recent Memory," p. 17.

[53] Penfield, Wilder, and Brenda Milner, "Memory Defect in Bilateral Hippocampal Lesions," *Psychiatric Research Reports* 11 (1959), p. 46.

[54] Scoville and Milner, "Loss of Recent Memory," p. 21.

While all of the Scoville patients received bilateral temporal lobe resections, Penfield's patients received unilateral temporal lobe removals. Most of Penfield's patients did *not* develop severe cases of amnesia. Of all the patients who underwent unilateral temporal lobectomy, only two, F.C. and P.B., suffered from a severe post-operative memory deficit. Milner concluded, therefore, that in the cases of F.C. and P.B. there must have been pre-existing damage to the contralateral hippocampal region. (This was eventually confirmed on autopsy, at least in the case of P.B.) Milner argued that a necessary condition for global amnesia is *bilateral* damage to the hippocampal complex. "To account for it (the profound memory loss in F.C. and P.B.) we have assumed that, in addition to the known epileptogenic lesion of the left hippocampal zone, there must have been a second and pre-operatively unsuspected destructive lesion of the opposite (right) hippocampal zone at the time of birth, so that when the surgeon removed the left hippocampal area the patient was functionally deprived of that area on both sides." [55]

If this anatomical analysis is correct, all four patients suffered from the effects of bilateral hippocampal loss. The similarity of their amnesic symptoms can thus traced to the fact that they have suffered from the same pathology. Milner writes that "the memory loss in the cases of bilateral surgical destruction was more striking and perhaps more persistent than in Dr. Penfield's two cases of unilateral temporal lobectomy reported above. Nevertheless, the difference is only one of degree and there can be little doubt that we are dealing with the same syndrome." [56]

We can now turn to the task of examining this "same syndrome" in more detail, noting in particular the differences between the cognitive capacities that have been lost and those which have been retained. The first thing to note is that the general intelligence of these four individual was not damaged by the operation. In fact, H.M.'s I.Q. scores improved post-operatively, perhaps due to the fact that he was no longer as fatigued by

---

[55] Penfield and Milner, "Memory Defect," pp. 46-47.

seizures.[57] M.B.'s I.Q. was not measured before the surgery, but her "verbal intelligence" after the surgery appeared to be "normal". In describing their post-operative studies of F.C., Penfield and Milner write that "the most striking feature of these follow-up examinations has been the contrast between the patient's good general intelligence and his loss of recent memory."[58] In describing P.B. post-operatively, Penfield and Milner note that "the most interesting feature of this postoperative examination was the contrast between the patient's memory defect and his other intellectual functions, which remained at their high preoperative level."[59] They note, for example, that P.B. suffers from "a selective disturbance of recent memory, with no corresponding impairment of attention, concentration, or reasoning ability."[60] The capacity to acquire and retain new information can thus be disassociated from the cognitive capacities that constitute general intelligence.

The memory loss associated with bilateral hippocampal damage was termed "global" amnesia because of its striking severity. All sensory modalities are affected and patients are impaired with respect to both verbal and non-verbal stimuli. As severe as the memory deficit is, however, the term "global" amnesia proved to be a misnomer. Milner and others described three types of mnemonic capacities that are spared in severe cases of amnesia: (1) the capacity to retrieve some memories formed pre-operatively (2) the capacity to remember post-operative information for a short time in the absence of distractions, and (3) the capacity to engage in skills acquired pre-operatively.

Bilateral hippocampal amnesia is described as severe anterograde memory loss accompanied by partial, temporally-graded retrograde memory loss. Patients are

---

[56] Penfield and Milner, "Memory Defect," pp. 47-48.

[57] Milner, Brenda, "Further Analysis of the Hippocampal Amnesic Syndrome: 14-Year Follow-Up Study of H.M.," *Neuropsychologia* 6 (1968), p. 219.

[58] Penfield, Wilder, and Brenda Milner, "Memory Deficit Produced by Bilateral Lesions in the Hippocampal Zone," *A.M.A. Archives of Neurology and Psychiatry* 79 (1958), p. 481.

[59] Penfield and Milner, p. 486.

[60] Penfield and Milner, p. 486.

profoundly impaired in the task of laying in new memories. Memories acquired just prior to the operation are lost as well. But early memories, and memories acquired a sufficient amount of time before the operation, are intact. This temporally-graded amnesia varies in severity from patient to patient. H.M.'s retrograde amnesia, for example, extends for about three years prior to his operation. F.C.'s extends about for about four years, and P.B.'s extends for the relatively short period of three months.[61] Of M.B., Scoville and Milner report simply that "her conversation centred around her early life and she was unable to give any information about the years of her hospital stay."[62]

Based on these findings, Milner argues that the hippocampus must be involved in the process of memory consolidation. On her view, new memories are established in a process whereby the hippocampus communicates with cortical tissue. The memory "exists" in some sense, in the cortical tissue, but it must be nurtured, for some period of time, by hippocampal-cortical interaction. Once the memory is well-consolidated, it achieves its independence from the hippocampus and comes to reside solely in cortical tissue.

This theory of the role of the hippocampus in memory consolidation explains the fact that hippocampal patients cannot create new memories but can access memories that have already been well consolidated. Furthermore, it explains the temporally-graded nature of retrograde amnesia. If a memory formed prior to the operation is still in the process of being consolidated, the loss of the hippocampus will result in the loss of the memory.

> But it seems logical to suppose ... that eventually ... the cortical linkage becomes autonomous, and recall is no longer dependent upon the simultaneous activity of hippocampal cells. On this view, then, removal of the hippocampal region will leave these well-established associations between cell assemblies essentially intact

---

[61] Penfield and Milner, "Memory Defect," p. 46.

[62] Scoville and Milner, "Loss of Recent Memory," pp. 17-18.

but will make it very difficult for new associations to be built up, and those not yet independent of hippocampal activity will be lost.[63]

Amnesic patients also have the ability to remember small amounts of information for short periods of time, if their attention is not diverted. H.M., for example, could remember the number 548 for up to fifteen minutes, if he was left undisturbed. Milner notes that H.M. accomplishes this task "by continuously working out elaborate mnemonic devices."[64] As soon as he is distracted, however, the number, as well as his mnemonic devices, are no longer available for recall. The engineer, P.B., also retained a relatively good capacity for short-term retention. He was able, for instance, to repeat nine-digit numbers forward and seven-digit numbers backward. Milner concludes that the capacity for attention is not compromised by bilateral hippocampal resection. She writes, "the evidence from our patients is that this attentional process is also independent of the hippocampal system. Any proposition, number series, or word association to which the patient can give attention at one time is available for recall provided attention has not been diverted elsewhere before this recall is required. In other words, the essential condition for recall is that the trace be kept ceaselessly active."[65] To put the point another way, the hippocampus is required for the process of converting the objects of attention into stable memories. "The fact that in our patients with bilateral hippocampal lesions forgetting appears to take place the instant their attention is diverted shows that this consolidation process involves hippocampal and not merely cortical cells."[66]

Finally, these patients retained the ability to perform skills. This is particularly striking in the cases of F.C., a glove cutter, and P.B., a civil engineer. Both were able to return to work, though P.B.'s memory deficits necessitated a demotion from manager to

---

63 Penfield and Milner, "Memory Defect," p.52.

64 Penfield and Milner, "Memory Defect," p. 49.

65 Penfield and Milner, "Memory Defect," p. 51.

66 Penfield and Milner, "Memory Defect," p. 52.

draftsman. "Both of these amnesic patients have continued to earn their living, one as a glove cutter, the other as a draftsman; and their professional skills are well-maintained." [67] The skills retained by amnesics include the ability to use language as proficiently as they did pre-operatively. (Vocabulary acquired a sufficient amount of time prior to the operation is retained.) When the glove-cutter was studied in 1959, seven years after his operation, Penfield and Milner note that "the post-operative memory deficit has persisted... but the patient has retained his old skills. Furthermore, formal tests show his problem-solving ability to be unimpaired. He does his work as well as ever and seems to enjoy it. Nor has be forgotten how to behave socially." [68] P.B., the draftsman, was no longer able to handle the complex memory tasks required of an administrator, but was able to produce complicated blueprints. Penfield and Milner describe his post-operative employment as follows. "He cannot learn the names of new business associates, and if, for example, he is interrupted while telephoning, he will forget completely the substance of the telephone conversation. Yet he is still able to prepare very complicated blueprints, though tending to work more slowly than before." [69]

Milner attributes the sparing of skills and the sparing of early memories to the same general physiological principle. Both are attributed to the fact that hippocampal resection does not affect long-lasting cortical changes elsewhere in the brain. Milner notes that "these patients with hippocampal lesions retain their professional knowledge and skills, their understanding and use of language, and their ability to recall early experiences" and that "this evidence of the continuing effects of past learning implies of course enduring changes in the brain." [70] She argues that it is reasonable to assume that "such changes primarily involve the cerebral cortex; certainly these long-established habits do not depend

---

67 Milner, Brenda, "Psychological Defects Produced by Temporal Lobe Excision," *Res. Pub. Assoc. Nerv. Ment. Dis.* 36 (1958), pp. 244-257.

68 Penfield and Milner, "Memory Deficit," p. 480.

69 Penfield and Milner, "Memory Deficit," p. 487.

70 Penfield and Milner, "Memory Defect," p. 51.

upon the hippocampus for their maintenance and reactivation. They imply rather a changed relationship between cortical cells." [71]

Human Amnesia Studies Phase Two: The Spared Capacities

During the 1950's scientists were agreed that amnesics retained previously acquired skills. During the 1960's and 1970's, Milner and her colleague Suzanne Corkin investigated the issue of whether or not amnesics could acquire new skills post-operatively. In the fall of 1960, Milner spent three days conducting additional studies on H.M. She tested him on two new tasks: a visual maze task and a mirror-drawing task. [72] While H.M. was unable to perform well on the visual maze task, he was able to acquire and retain the skills necessary to perform well on the mirror drawing task.

The visual maze is composed of a ten-by-ten array of bolt heads. H.M. is provided with a stylus which he is instructed to place on the "start" bolt-head. The goal is to move the stylus, in a "stepping-stone" fashion, from one bolt-head to the next until the "finish" bolt-head is reached. There is only one correct sequence of moves through the maze, a sequence which involves 28 steps. While there are no visible indications of the correct pathway, incorrect moves elicit the clicking noise of an error-counting device and signal to H.M. that he should pursue another path.

Milner conducted three days of testing on the maze. On each day, H.M. was tested on three separate occasions. On each occasion, he attempted the maze 25 times. Milner hypothesized that H.M.'s performance would improve slightly within each series of 25 attempts but that his performance level would drop back to baseline at the beginning of each series. As it turned out, H.M. showed no learning whatsoever over the three days of

---

71 Penfield and Milner, "Memory Defect," p. 51.

72 Milner, Brenda, "Les troubles de la mémoire accompagnant des lésions hippocampique bilatérales," In P. Passouant (Ed.), *Physiologie de l'hippocampe* (Paris: Centre Nationale de la Recherche Scientifique, 1962).

testing. On the second and third days of the test, H.M. had no memory of having attempted the maze on the previous day.

In the mirror-drawing task, H.M. was asked to draw a five-point star by "keeping within the lines" of an outline provided, using only a mirror for visual guidance. Normal subjects experience some difficulty on first exposure to this task, but master it quickly with practice. Success is defined in terms of the decreasing amount of time required to complete the task. Milner found that H.M.'s performance improved over trials, and that his new-found skill was retained over the testing period of three days. On his first right-handed attempt, H.M. took over 250 seconds to trace the star; by day three, he consistently completed the task in under 100 seconds.

> Ces résultats diffèrent nettement de ceux du test du labyrinthe et suggèrent que l'acquisition d'aptitudes visuo-motrices peut se révéler être indépendante du système hippocampique. On connait mal encore les autres types d'apprentissage qui peuvent être acquis indépendamment du système hippocampique. [73]

In 1962, Corkin conducted a study of the performance of various types of neurological patients on a short, 10-choice tactual maze. Her primary focus was on the comparative performance of right and left frontal lobe patients, right and left temporal lobe patients, and bilateral hippocampal patients. The three hippocampal patients in her study were H.M., F.C., and P.B.

The maze was hidden from view from the subjects, who were asked to trace the route from start to finish using a stylus. If a subject entered a blind allay, a warning bell was rung. The test was conducted over a two-day period. Subjects were asked to perform the task in blocks of ten trials until they reached the criterion of three runs without error or,

---

[73] Milner, "Les troubles de la mémoire," p. 270. ("These results differ completely from those of the labyrinthe test and suggest that the acquisition of perceptual-motor skills can be shown to be independent of the hippocampal system. We still have a poor understanding of other types of learning that might be acquired independently of the hippocampal system.")

in severely impaired subjects who could not master the task, until 50 trials had been completed.

Of all the patient types who attempted the task, the bilateral hippocampal patients were the most severely impaired. (Left frontals and left temporals performed normally, and right frontals and right temporals were impaired, but not as much as the bilateral hippocampals.) F.C. and P.B. did, however, improve significantly over trials, though they did not reach the criterion of three consecutive errorless runs. H.M.'s error score suggested, however, that he was not learning the maze at all. Even after Corkin increased the upper limit on the number of trials from 50 to 80, H.M. showed no evidence of learning. On the other hand, the amount of time required for H.M. to complete the task decreased steadily. Corkin interprets this finding to mean that H.M. acquired some degree of "proprioceptive-motor skill." [74] In describing the graph which plots the time taken to complete the task against the number of trials, Corkin notes that "the steady slope of the curve suggests that the patient acquired the proprioceptive-motor skills involved in the task despite being unable to retain the correct sequence of turns." [75]

Corkin decided to investigate the phenomenon of H.M.'s spared capacities in greater detail. In 1968, she published an article titled, "Acquisition of Motor Skill after Bilateral Medial Temporal-Lobe Excision, " in which she discusses H.M.'s performance on three additional tests of motor skill. [76] Corkin introduces the article by stating that there are good reasons to suppose that motor memories and non-motor memories are subserved by two different neural substrates. First, she points to a number of commonsense distinctions between what she calls "rote cognitive learning" and motor memory. "Complex motor skills (such as skating, swimming, pronouncing a foreign language,

---

[74] Corkin, Suzanne, "Tactually-guided maze learning in man: Effects of unilateral cortical excision and bilateral hippocampal lesions," *Neuropsychologia* 3 (1965), p 348.

[75] Corkin, "Tactually-guided maze learning in man," p. 348

[76] Corkin, Suzanne, "Acquisition of motor skill after bilateral medial temporal lobe excision," *Neuropsychologia* 6 (1968), pp. 255-265.

dancing and piano playing) are usually acquired most efficiently in childhood, are extremely difficult to erase through disuse, and show little relationship to the level of general intellectual function." [77]

Secondly, she cites the data collected by Milner and herself that suggest that H.M. has retained the capacity to acquire motor skills. She cites, in particular, the fact that H.M. was able to acquire Milner's mirror-tracing skill and the fact that his time scores on her own 10-choice tactual maze improved over time. Corkin describes Milner's appraisal of these findings. "On the basis of these two findings, it was hypothesized that other motor skills could also be acquired by patients with bilateral lesions of the medial structures." [78] Corkin proposes to follow through on Milner's hypothesis by studying H.M.'s performance on three new motor-learning tasks: "Rotary Pursuit, Bimanual Tracking, and Tapping." [79]

In the rotary pursuit task, the subject is asked to maintain contact between a hand-held stylus and a metal target that rotates, at various speeds, on a circular disk. Success is measured in terms of the amount of time, in a given trial, during which contact is maintained. The bimanual tracking task is similar, except that both hands are used to maintain contact with two wavering lines that are rotated on a drum placed before the individual. In the tapping task, subjects are asked to tap certain numbered segments of a round stationary disk in a particular sequence. Success here is measured in terms of the lack of errors (tapping in the wrong sequence) and in terms of the speed with which the task is performed. (In the bimanual version, each hand must operate independently to tap out the sequence on two different disks, whose internal segments have been numbered in a different order.)

While H.M.'s overall performance on these three tasks was somewhat impaired compared to controls, he did exhibit the capacity to learn the tasks and improve his

---

77 Corkin, "Acquisition of motor skill," p. 255.

78 Corkin, "Acquisition of motor skill," p. 256.

performance over trials.[80] Furthermore, when he returned to these tasks after intervals of various lengths, his skill was preserved in the interim. Corkin notes that "these results imply that motor learning involving visually- and proprioceptively-guided movements can be mediated to some extent by brain structures which are still intact in H.M."[81]

In 1968, Milner published an article titled, "Further Analysis of the Hippocampal Amnesic syndrome: 14-Year Follow-Up Study of H.M," in which she "attempts to delineate certain residual learning capacities of H.M."[82] In this article, Milner reviews H.M.'s amnesic symptoms and discusses his performance in a variety of cognitive tasks.

Milner notes that H.M.'s severe amnesia persists over the years. He cannot recognize people he has met since the operation, even those who have spent a good deal of time with him. While he can accurately report his own birth date, he underestimates his own age and "can only make wild guesses as to the date."[83] His only way of determining the time of the year is to make inferences from the weather. H.M. describes his own state of mind by saying that it is "like waking from a dream."[84] According to Milner, "his experience seems to be that of a person who is just becoming aware of his surroundings without fully comprehending the situation, because he does not remember what went before."[85]

H.M.'s high level general intelligence is unchanged and he continues to use language well. Milner notes that H.M.'s extensive vocabulary is consistent with this high

---

79 Corkin, "Acquisition of motor skill," p. 256.

80 The difference between the performance of H.M. and the performance of normals is attributed by Corking to non-mnemonic difficulties such as slow reaction times and a generally low level of spontaneous activity, both of which are common general consequences of brain damage.

81 Corkin, "Acquisition of motor skill," p. 262.

82 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 215.

83 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 216.

84 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 217.

85 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 217.

level of general intelligence. He can appreciate jokes, even those which turn on subtle semantic ambiguities. [86]

In the fourteen years since the operation, H.M. has experienced a few subtle changes in his anterograde memory capacities. He can sometimes remember events which are emotionally laden for him. He can report, for example, that J.F.K. was assassinated. [87] His retrograde memory loss is somewhat ameliorated in that vague memories for the three years prior to the surgery occasionally surface. Milner notes that H.M.'s perceptual capacities are intact. He can, for instance, perform the relatively difficult task of identifying the approximate age and gender of "poorly-defined silhouettes of faces." [88] His perceptual memory, however, is still severely impaired. If presented with a photograph of a face and then briefly distracted, he will be unable to recognize it two minutes later.

Milner discusses new evidence that suggests that H.M. may be capable of some types of perceptual learning. She cites a study conducted by Warrington and Weiskrantz in which amnesics showed normal "priming effects" for both pictorial and verbal material. [89] "Priming effects" are changes in behaviour that are attributed to the fact that an individual has retained information about a stimulus even though she does not have conscious access to the information. For example, if a subject is shown a list of words, this will prime her responses on a word stem completion task. (If she has just been shown the word "cyclone" she will tend to complete the word stem CYC____ with "cyclone" and

---

86 H.M.'s vocabulary was acquired pre-operatively. A study conducted in the 1980's confirms that H.M. cannot acquire new vocabulary post-operatively. (See Gabrieli, J.D., N. J. Cohen, and S. Corkin, "The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection," *Brain and Cognition* 7 (1988): 157 - 177.

87 We can assume that H.M.'s retention of this information is made possible either by the sparing of a nonhippocampal memory system that benefits from repeated exposure to emotionally laden material or by the partial sparing of his hippocampal system.

88 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 231.

89 Warrington, E. K., & Weiskrantz, L., "A new method of testing long-term retention with special reference to amnesic patients," *Nature* 217 (1968), pp. 972-974, and Warrington, E. K., & Weiskrantz, L., "The amnesic syndrome: consolidation or retrieval," *Nature*, 228 (1970), pp. 628-630.

not, for example, "cycle".) Milner tested H.M. on a perceptual priming test that involves the identification of degraded line drawing of objects before and after priming. (The subject is primed by exposure to the complete, detailed line drawing, which is easily identified.) While H.M. did not benefit from priming as much as normals, his performance did manifest a significant priming effect. Some perceptual information was in fact being retained, though H.M. did not have conscious access to it.[90]

Milner also reviews the literature on H.M.'s performances on both visual and tactual mazes. In all cases, the number of choice points involved exceeded H.M.'s immediate memory span. Milner decided to investigate H.M.'s ability to perform on mazes with a drastically reduced number of choice points to see if he could learn the shortened mazes, and retain the learning over time. Milner found that with extensive training, H.M. could learn very short visual and tactual mazes. H.M. learned these mazes very slowly and Milner was unconvinced that he would remember what he learned. "Acquisition of these short sequences was extremely slow, and it is not known how far his capacity for partial recall would have extended beyond the intervals employed in these experiments."[91]

After surveying H.M.'s lost and spared capacities, Milner states that hippocampal amnesia is first and foremost a deficit in the capacity to convert experience into long-term memories. "All these observations support the view that the essential difficulty of these patients is not in primary registration, or short-term memory as such, but in some secondary process by which the normal subject achieves transition to long-term storage of information."[92] At the same time, the deficit, even in the case of H.M., is not absolute. Valent memories may be retained and H.M.'s performance on the priming tests suggests

---

90 The issue of priming is important, but I will not pursue it further here. For a recent non-technical overview of the literature, see Daniel Schacter, "Implicit Memory: A New Frontier for Cognitive Neuroscience," *The Cognitive Neurosciences*, Ed. M. Gazzaniga (Cambridge, Mass.: MIT Press. 1995) pp. 815-24.

91 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 232.

92 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 232.

that he is capable of some type of perceptual learning. Furthermore, the maze studies indicate that with intensive training, H.M. can learn a small amount of information very slowly, though it is not clear whether or not he would be able to retain the information for long.

The most clear-cut example of a retained mnemonic capacity is H.M.'s ability to perform well in tasks of motor learning. Milner notes that, "the relative sparing of motor learning is now clearly established." [93] In fact, Milner urges caution in attributing other spared capacities to amnesics.

> Among the residual capacities tested, those involving the learning of motor skills were probably the least affected and it is questionable whether an even more drastic removal in the same region would have disturbed them seriously. In contrast, the very slight learning on other tasks, only demonstrable with intensive training, may well reflect the incompleteness of the lesion, which permits some recovery of function over time. [94]

Milner concludes, therefore, that the primary spared capacity in hippocampal amnesia is the ability to acquire motor skills.

Conclusion

The distinction between high-order cognitive behaviour and low-order perceptual-motor skills arises in a number of diverse contexts. As we have seen, the distinction is critical to the neo-Heideggerian project of articulating the limitations of cognitivism. A serious problem facing the neo-Heideggerian is the cognitivist's commitment to the primacy of causal explanation. Let the phenomenological account of the distinctiveness of perceptual-motor skills be as persuasive as you like, the cognitivist may simply sidestep it

---

93 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 232.

94 Milner, "Further Analysis of the Hippocampal Amnesic Syndrome," p. 232.

altogether, on the grounds that phenomenology is irrelevant to science. Dreyfus' successful and sustained attacks on the computer model of the mind raise serious questions about the scope of cognitivism. We can supplement Dreyfus' work by providing a causal account of perceptual-motor skills that makes clear the fact that high-order intelligent behaviour and low-order intelligent behaviour are indeed caused in two distinct ways. We can show that while high-order intelligent behaviour involves, at the neuroscientific level, the manipulation of neurally realized representations, low-order intelligent behaviour does not.

In this chapter, I have outlined the differences between Tolmanian cognitivism and Hullian behaviourism. The critical difference between the two is that Tolmanian cognitivism allows for the positing of full-fledged internal representations, while Hullian behaviourism confines itself to the positing of S-R associations at the neural level. In the next chapter, I will explain Mishkin's reasons for asserting that Tolmanian cognitivism is consistent with the mechanisms of the memory systems centred in the medial temporal lobe (an area which includes the hippocampus) and that Hullian behaviourism is consistent with the mechanisms of a second, nonhippocampal memory system.

Mishkin's work is based on Milner's studies of patients who suffer from bilateral medial temporal lobe damage as well on studies of amnesia in non-human animals. The most important element of Milner's work, for our purposes, is her documentation of the fact that H.M. retains the capacity to acquire the very sorts of perceptual-motor skills that Bergson describes in terms of habit memory. We can thus begin to connect a few important dots. Heideggerian engaged coping consists, in part, of the capacity to comport oneself skillfully in the everyday world. Both Heidegger and Dreyfus argue that engaged coping is non-representational, and their arguments are based on a phenomenological analysis of the everyday world. Bergson appears to describe the same type of phenomena in his account of habit memory. Neuroscientists cite Bergson's account of habit memory as a prescient description of nonhippocampal memory. Milner provides us with a set of neuropsychological reasons for supposing that habit memory is a distinct neuroanatomical memory system. If we can show that this nonhippocampal memory system is non-representational, we can collate the work of Heidegger, Dreyfus, Bergson, Milner, and Mishkin and provide a strong foundation for the claim that a cognitivist approach to perceptual-motor skills is misguided.

Chapter Five: Contemporary Theories of Multiple Memory Systems

Fodor's Representational Theory of the Mind (RTM) purportedly codifies our commonsense intuitions about the etiology of intelligent behaviour. Both Fodor and the folk claim that rational behaviour is caused by the logical "processing" of beliefs and desires. When it comes to perceptual-motor skills, however, Fodor and the folk part ways. For the folk, rational behaviour and perceptual-motor skills are quite different phenomena: there is no obvious reason for assuming that if rational behaviour is caused by the manipulation of representations, skilled perceptual-motor behaviour must be as well. Fodor, on the other hand, is guided by the conviction that intentional state explanations are the best causal accounts of intelligent behaviour on offer. Since perceptual-motor skills are a species of intelligent behaviour, they, too, must be explained in intentional terms, within the framework of the RTM. When Fodor attempts a representational account of perceptual-motor skills, he ends up having to posit a variety of mental representations whose contents are utterly foreign to the folk, but consoles himself with the thought that science does not always respect folk wisdom. As it turns out, Fodor should have been more attentive to the fact that the folk fall silent when the behaviour in question is a perceptual-motor skill. Fodor is wrong about perceptual-motor skills, but it has been difficult to argue against him in the absence of an alternate *scientific* account of perceptual-motor skills. Now we have one.

The critical issue here is the scope of the RTM. In order to determine the scope of the RTM, we need an account of the circumstances under which it is correct for cognitive scientists to invoke an intentional state explanation of particular piece of intelligent behaviour. Fodor's account is particularly rigorous. He is a Strong Realist about intentional states. Intentional explanations for him are not merely a useful *façon de parler.* On Fodor's view, a scientist correctly invokes an intentional state explanation of a particular behaviour only if the behaviour is genuinely caused by the manipulation of internal representations. Others, notably Dennett, argue that intentional state explanations are appropriate when they are predictively useful. Consider the infamous example of Dennett's thermostat. As far as Dennett is concerned, we may explain the behaviour of the thermostat in intentional terms if we like. An intentional explanation of thermostat behaviour is not, however, particularly useful; there is little gain in predictive power involved. [1] On Fodor's view, on the other hand, it would be *wrong* to explain the behaviour of the thermostat in intentional terms. We know how thermostats work and our best causal account of their behaviour makes no reference to intentional states.

Dennett argues that the adoption of the "intentional stance" is a useful strategy in that it allows us to predict behaviour accurately in a wide range of cases. For Fodor, however, intentional explanations are "correct" only if the behaviour is genuinely caused in the right way, i.e., if it is caused by the processing of representations. Recall Fodor's analysis of Dennett's chess-playing computer. With respect to the machine's "mental processes," Fodor claims that "either they are causal sequences of explicit representations, or the representational theory of chess playing is simply false of the machine." [2] We rule out a potential candidate for the RTM by showing that its behaviour is not caused by the manipulation of representations.

Here's the question of the hour. What is the best causal account of skilled perceptual-motor behaviour on offer today? Fodor proposes that we analyze the etiology of perceptual-motor skills within the framework of his RTM. As we have seen in Chapter One, there are good reasons for thinking that this account is flawed. It is based on the dubious technique of simulating human behaviour computationally and assuming that the simulation mirrors the genuine causal processes responsible for generating the behaviour in

---

[1] Dennett, Daniel, *The Intentional Stance* (Cambridge, Mass.: MIT Press, 1987).

[2] Fodor, Jerry, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass.: MIT Press, 1987), p. 25.

humans. On Fodor's more recent account, it is based on the equally dubious technique of generating a functionalist account of the behaviour, identifying the rules constitutive of this functionalist account, and attributing to the behaver a set of mental representations whose contents are the rules in question. In either case, we arrive at a set of rules with which the behaviour is consistent. On Fodor's own account, rule-consistency is not enough; rule-consistency does not entail rule-representation. The RTM counts as our best causal account of perceptual-motor skills only if we assume that, in general, our best causal account of intelligent behaviour is an account couched in terms of an intentional idiom. There is no other reason for supposing that perceptual-motor skills are caused by the manipulation of internal representational states. Up until now, I have emphasized the claim that there are no good reasons for thinking that Fodor's analysis is correct. In this chapter, I will go one step further and argue that there is one very good reason for thinking that his analysis is incorrect: Fodor's account is incompatible with our best neuroscientific account of the etiology of perceptual-motor skills. Our best causal account of skilled perceptual-motor behaviour rests on the account proffered by neuroscientists, an account which indicates that perceptual-motor skills are not caused by processes involving the manipulation of neurally realized Fodorian representations.

As we have seen in Chapter Two, the representations that figure in Fodor's RTM have two critically important properties: internal structure and transportable parts. These properties are said to explain the systematicity and generativity of rational thought and rational behaviour. The key principle to keep in mind is that if a representation has internal structure and transportable parts, each of its parts can figure flexibly in any number of composite representations and can do so without losing its own integrity. This type of representational system can support inferential reasoning; individual representations and representation-parts can serve as transportable elements in the premises and conclusion of arguments. Furthermore, generativity is accounted for in that representations can be combined in novel ways.

In contemporary neuroscientific theories of multiple memory systems, the critical difference between hippocampal memory and nonhippocampal memory is that hippocampal memory involves the processing of neurally realized Fodorian representations and nonhippocampal memory *does not*. This is crucial for my project for, as I have shown in Chapter Four, the production of skilled perceptual-motor behaviour is associated with a nonhippocampal memory system.

The argument I have been developing comprises two premises. First, skilled perceptual-motor behaviour is dependent on the mechanisms of a nonhippocampal memory system. As we have seen in Chapter Four, patients with severe amnesia caused by the destruction of their hippocampal memory system are still capable of acquiring, refining, and demonstrating perceptual-motor skills. (At the end of this chapter, I will present some additional evidence for this claim, though I take the claim to have been well defended in Chapter Four. If H.M. retains the capacity to acquire perceptual-motor skills in the absence of his hippocampi, perceptual-motor skill acquisition cannot be dependent on the hippocampal system and must be dependent on some other, nonhippocampal, memory system.) Second, nonhippocampal memory systems are not plausibly viewed as systems that operate over Fodorian representations. This chapter will be devoted to demonstrating that this second premise is true. Taken together, these two premises support the conclusion that skilled perceptual-motor behaviours are not caused by the processing of Fodorian representations.

A Few Preliminary Notes

In this chapter, I review four accounts of the distinction between hippocampal memory and nonhippocampal memory, the accounts of (1) Richard Hirsh, (2) John O'Keefe and Lynn Nadel, (3) Mortimer Mishkin, and (4) Neal Cohen and Howard Eichenbaum.

The going will be easier if we address a few potential sources of confusion, both conceptual and terminological, sooner rather than later. To begin with, there are more than two memory systems in the brain. The jury is still out on the exact number. One of the memory systems, the hippocampal system, is relatively well-defined, functionally and anatomically. The remaining systems are typically grouped together under the heading of "nonhippocampal memory". This should be born in mind when considering the various ways in which the "two" systems have been named: "declarative and procedural memory", "representational and dispositional memory", "fact memory and skill memory", to name but a few.[3] As Daniel Schacter points out in his recent book, *Memory Systems 1994*, only the first term in each of these couplets refers to a specific neural system.[4] The second term in each couplet is a catch-all term used to describe the general properties of some of the nonhippocampal systems. As Schacter states, "because of our present lack of information about the vast terra incognita that we call procedural memory, its most adequate description at the present is probably by exclusion."[5] This way of classifying memory systems has its advantages, since the same *general* principles of operation do appear to govern the processing of most nonhippocampal systems, a fact made more perspicuous when nonhippocampal systems are grouped together. At a very general level of description, then, the operating principles of the hippocampal system can be distinguished from the operating principles of the nonhippocampal systems.

While the exact number and neuroanatomical substrates of the nonhippocampal systems are not known, we can identify a number of these systems in very general terms. A system for conditioned emotional responses is centred in the amygdala, a small, sub-cortical structure located in the medial temporal lobe just anterior to the hippocampus. The

---

[3] Squire, Larry R., *Memory and Brain* (Oxford: Oxford University Press, 1987), p. 169.

[4] Schacter, Daniel L., and Endel Tulving, Eds., "What are the Memory Systems of 1994?," *Memory Systems 1994* (Cambridge, Mass.: MIT Press, 1994), pp. 22-23.

[5] Schacter and Tulving, "What are the Memory Systems of 1994?" p. 27.

cerebellum, a large cortical appendage located in the rear of the brain, is responsible for "classical eyelid conditioning (and other skeletal muscular) conditioning." [6] The basal ganglia, a collection of sub-cortical nuclei in the centre of the brain, are responsible for both perceptual-motor skills and certain simple cognitive skills that are acquired through practice over time. Nonhippocampal memory systems thus support a wide range of mnemonic functions, ranging from the acquisition of simple conditioned responses to the acquisition of simple "cognitive" habits. Several of the theories I will examine liken nonhippocampal memory to S-R memory. While this may be disquieting to philosophers who consider behaviourism a failed paradigm, we should remember that we are interested in the general operating principles that govern nonhippocampal systems. In the end, we can elect to reject the S-R terminology altogether, or use it only when speaking of the more basic forms of nonhippocampal memory. That is to say, nothing in my review of this literature should be taken to mean that we are going to find ourselves committed to a behaviouristic account of perceptual-motor skills.

When the four accounts of multiple memory systems are summarized, we will see that in each case, the emphasis is primarily on the hippocampal system. This is due, at least in part, to the fact that much more research has been done on hippocampal memory than on nonhippocampal memory. Furthermore, the general strategy employed with respect to nonhippocampal memory is to summarize the very general principles of operations that govern all or most of the nonhippocampal systems, as diverse as they may be. This means that we will have to glean our information about the memory system that supports perceptual-motor skills in a somewhat indirect fashion, by collating the claims made by these scientists concerning the general principles of nonhippocampal memory.

The most important task at hand is to make clear the similarities between hippocampal processing and Fodorian processing. The neuroscientists on our list agree that

---

[6] Cohen, Neal, and Howard Eichenbaum *Memory, Amnesia, and the Hippocampal System*, (Cambridge, Mass.: MIT Press, 1993).

the hippocampal memory system is dedicated to the processing of neural representations which have both of the principle features of Fodorian mental representations; these neural representations have both internal structure and transportable parts. Neuroscientists argue that *the* critical difference between hippocampal processing and nonhippocampal processing is that *only* hippocampal processing involves the manipulation of these sorts of representations. Since perceptual-motor skills are subserved by a nonhippocampal system, we can infer that perceptual-motor skills are not caused by the manipulation of Fodorian-style neural representations. My argument to the effect that Fodor's account of perceptual-motor skills is misguided is based on an account of the memory systems which shows that (1) hippocampal memory systems operate over neurally realized Fodorian representations and (2) nonhippocampal systems do not. My focus will therefore be on the various accounts of hippocampal memory on offer and on the claims made by each of the neuroscientists to the effect that nonhippocampal memory does not involve the same sort of representational processes that figure in hippocampal processing. At the end of the chapter, I will offer a few comments on the nature of the memory system that is dedicated to the acquisition and implementation of perceptual-motor skills.

One final preliminary note is in order. Care must be taken when comparing memory in animals and in humans. Some neuroscientists are particularly cautious about collating data and theories across species. Schacter claims that "we are probably better off developing separate classifications of memory systems for different species." [7] Others, like Eichenbaum, insist on the importance of bringing together the studies done on both animals and humans. "Analyses at the cognitive and neuropsychological levels can be performed appropriately in human subjects, but detailed circuit analyses can be accomplished only in animals." [8] Eichenbaum's theory is particularly rich because he does attempt to collate data

---

[7] Schacter and Tulving, "What are the Memory Systems of 1994?," p. 31.

[8] Eichenbaum, Howard, "The Hippocampal System and Declarative Memory in Humans and Animals: Experimental Analysis and Historical Origins," *Memory Systems*

from both animals and humans. Having noted that caution is warranted, I will avail myself of the relevant studies in both the human and animal literature.

An Animal Model of Human Amnesia

In the 1970's scientists sought to provide an animal model of human amnesia. The basic method was straightforward: the hippocampi of animals (primarily rats) were lesioned or pharmacologically incapacitated, and subsequent mnemonic deficits were carefully tested and catalogued. Interest was focused primarily on ascertaining the function of the hippocampus, though spared mnemonic capacities were duly noted and described. At this time, the Tolman - Hull debate provided a theoretical framework for the understanding of the lost and spared capacities of amnesic animals. Two important theories of animal hippocampal function emerged during this time, Richard Hirsh's theory of hippocampal contextual retrieval and O'Keefe and Nadel's theory of hippocampal cognitive mapping.

Hirsh's Theory of Multiple Memory Systems

By the early 1970's hippocampal damage had been clearly related to mnemonic deficits in humans, but the relationship between the hippocampus and memory in animals was still obscure. Attempts to model human amnesia in animals were initially thwarted by the fact that hippocampal ablates retain a number of significant mnemonic capacities and appear to suffer from a confusing variety of non-mnemonic deficits. In his paper, "The Hippocampus and Contextual Retrieval of Information from Memory: a Theory," Hirsh describes the situation as follows.

> The early clinical reports (from human amnesics) gave tremendous impetus to the study of the effects of hippocampal ablations in animals. Hippocampally ablated animals were observed to learn in a wide variety of tasks without difficulty.

1994, Eds. Daniel Schacter and Endel Tulving (Cambridge, Mass.: MIT Press, 1994), p.

> Further research revealed a bewildering variety of abnormalities in situations which most analysts regard as involving motivational, attentional, or motoric rather than mnemonic processes. Thus, hippocampal function has not been discussed in mnemonic terms by experimenters working with animals. [9]

Hirsh argues that the "bewildering" data on hippocampal deficits in animals can and should be re-interpreted. On his view, the hippocampus is responsible for a process he calls "contextual retrieval", a process that is explicitly mnemonic (and which will be explained in detail below). [10] The deficits catalogued in the early animal experiments that appear to be non-mnemonic (i.e. those related to motivation, attention, or movement) should be construed, according to Hirsh, as secondary consequences of damage to the contextual retrieval system. Furthermore, the retained mnemonic capacities of hippocampally ablated animals should be attributed, on his account, to a secondary, nonhippocampal memory system. "In the absence of contextual retrieval due to hippocampal ablation, an alternate process analogous to that advanced in early versions of S-R learning is assumed to operate, preventing the appearance of learning deficits in hippocampal lesioned animals in some tasks." [11]

Hirsh thus argues that there are two distinct types of memory in the rat: hippocampal memory that subserves contextual retrieval and nonhippocampal memory that subserves S-R habit formation. The critical difference between these two types of memory is the location of the memory storage with respect to what Hirsh calls the "performance line". He defines the performance line as "a system mediating the series of events or processes initiated by the overtly observable stimulus and resulting in the occurrence of the overtly observable response. It is considered to exist in real time and real space, and

---

150.

[9] Hirsh, Richard, "The Hippocampus and Contextual Retrieval of Information from Memory: A Theory," *Behavioral Biology* 12 (1974), pp. 421-444.

[10] Hirsh, "The Hippocampus and Contextual Retrieval," p. 422.

ultimately to be physiologically observable." [12] The role of the hippocampus, according to Hirsh, is to retrieve the appropriate information from an independent memory storage area and place it on the performance line. The hippocampus uses contextual clues to determine what information is relevant to the task at hand, to determine which pieces of information should be taken out of storage, so to speak, and made available on the performance line. One advantage of such a system is that it can avail itself of information that is stored more securely off line, "free from interference by information processing being carried out on the performance line." [13]

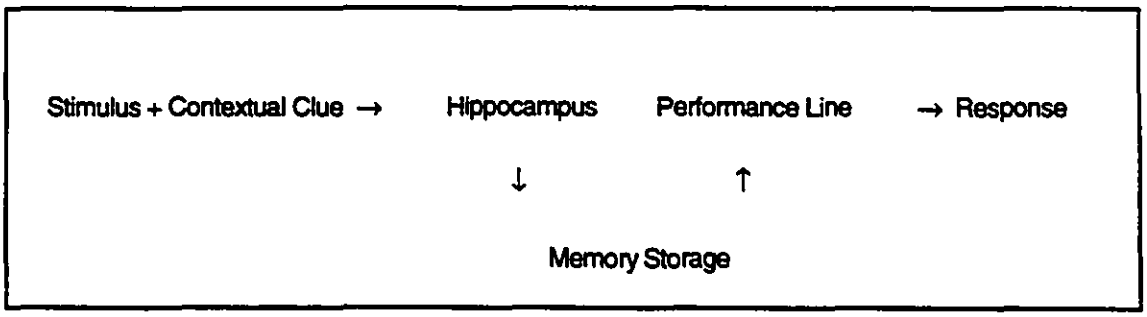

Figure 5.1: Hirsh's Theory of Contextual Retrieval

If the hippocampus is incapacitated, the animal must operate without the benefit of the contextual retrieval system; it must rely on more basic S-R mechanisms. In S-R behaviour, memory is stored directly on the performance line. The animal does not "access" other memory storage areas. The stimulus is thus said to "cause" the response directly. The response is conditioned by past experience, and in this sense only is "memory" stored on the performance line. Habit memory is just the continually altered disposition of the animal to respond in particular ways to particular stimuli. All changes to the system are made directly to the performance line itself. "It suffices the say that the

---

[11] Hirsh, "The Hippocampus and Contextual Retrieval," p. 422

[12] Hirsh, "The Hippocampus and Contextual Retrieval," p. 422.

[13] Hirsh, "The Hippocampus and Contextual Retrieval," p. 423.

occurrence of the stimulus causes the occurrence of the response. Contextual retrieval has no role in such a system."[14]

Stimulus → Performance Line → Response

Figure 5.2: Hirsh's Conception of Nonhippocampal Memory

An example will help to clarify the dynamics of contextual retrieval. Hirsh cites an experiment conducted in 1946 by Kendler which involves, on Hirsh's analysis, the use of motivational states as contextual cues. 15 16 Rats who are both hungry and thirsty are trained on a T-maze which is food-baited on one arm and water-baited on the other. The location of the water and the food remains constant across trials. After training, the rats were tested on the maze when they were either hungry or thirsty. The rats went reliably to the goal location appropriate to their motivational state. (The experiment is quite similar to the Spence and Lippitt experiment on latent place learning, except that now the rats are not sated during training trials.) According to Hirsh, each motivational state prompts the hippocampal system to retrieve information about specific goal locations and to place this information on the performance line.

A more recent experiment conducted by Hsiao and Isaacson in 1971 provides evidence for the claim that the hippocampus is critical for tasks involving such motivational contextual clues.[17] Once again, a T-maze is baited with food at one goal location and with water at the other. Rats were run on the maze when they were either hungry or thirsty. While normal rats learned the task of going to the appropriate goal location, "hippocampal

---

14 Hirsh, "The Hippocampus and Contextual Retrieval," p. 423.

15 Hirsh, "The Hippocampus and Contextual Retrieval," p. 424.

16 Kendler, H., "The Influence of Simultaneous Hunger and Thirst Drives upon the Learning of Two Opposed Spatial Responses of the White Rat," *Journal of Experimental Psychology* 36 (1946), pp. 212 - 220.

ablation caused severe learning deficits." [18] On Hirsh's view, hippocampal animals are impaired at this task because it cannot be solved using simple S - R mechanisms; it requires the use of the contextual retrieval system.

The key to contextual retrieval is that the animal must be able to appreciate the relationship between two distinct pieces of "information". In the above example, animals were required to associate a motivational cue with information about reward locations. Motivational states are not, however, the only type of contextual cue. Sometimes the contextual cue is the presence of what is called a "conditional" cue. In a conditional discrimination task, a particular stimulus is rewarded only if it is accompanied by a particular conditional cue. "In conditional discrimination the immediate stimulus never changes except for the presence of a particular cue. In the presence of one cue, one response is correct while in the presence of another cue a conflicting response it correct. The conditional discrimination cue operates as a contextual factor determining which information will be applied to the control of behaviour. Hippocampally ablated animals had deficits relative to normal controls in acquiring a conditional discrimination (Kimble 1963) [19] as the learning of one response interfered with the learning of the conflicting one (Isaacson and Kimble, 1972). [20] No deficit was observed when the same cues were used in a simple simultaneous discrimination, one involving no conditional rule." [21]

One way of clarifying the nature of hippocampal function is to study animals who have undergone complete bilateral hippocampal removal. Hirsh compares the "style" of learning in normals and hippocampal ablates. Learning in normals appears to occur in

---

[17] Hsiao, S., and R. L. Isaacson, "Learning of Food and Water Position by Hippocampus Damaged Rats," *Physiol. Behav.* 6 (1971), pp. 81 - 83.

[18] Hirsh, "The Hippocampus and Contextual Retrieval," p. 430.

[19] Kimble, D. P., "The Effects of Bilateral Hippocampal Lesions in Rats," *Journal of Comparative Physiological Psychology* 56 (1963), pp. 273 - 283.

[20] Isaacson, R. L., and D. P. Kimble, "Lesions of the Limbic System: Their Effects upon Hypotheses and Frustration," *Behavioural Biology* 7 (1972), pp. 767 - 795.

[21] Hirsh, "The Hippocampus and Contextual Retrieval," p. 429.

discrete steps. It is as though the normal animal learns, at least in some contexts, by acquiring particular bits of information about the world that allow it to make better choices. Intact animals are able, for example, to glean useful, behaviour-influencing information from single trials. This information can be added to memory-storage for use on future occurrences of the same problem solving task. This approach to learning requires an intact contextual retrieval mechanism, for the animal must be able to search its memory for solutions to problems as they arise. When the "solution" provided by memory proves to be successful, it is retained; when unsuccessful, it may be discarded. This process produces the appearance of "discrete" learning. "Depending on the results of that trial, the information remains in control or is replaced. Discrete changes in performance thus occur during learning." [22]

The learning of hippocampal ablates, on the other hand, is slow and gradual. On Hirsh's account, contextual retrieval is no longer available as a search mechanism - all influential experience must be stored directly on the performance line. Learning occurs via the reinforcement of successful responses to stimuli. At any given moment, there are a number of stimuli in the animal's sensorium. When reinforcement of a response occurs, it effects all the stimuli in the sensorium. There is no built-in mechanism for ensuring that reinforcement acts upon only those stimuli that are related to the animal's response. Over time, however, only the stimuli related to the response will receive consistent reinforcement, while the stimuli "only randomly related to reinforcement" will not. [23] Such stimuli will receive reinforcement only when they co-occur with response-related stimuli. Responses that are related to the stimuli are reinforced more consistently; behaviour is thus gradually shaped. Hirsh calls this learning by averaging. "In the absence of contextual retrieval due to hippocampal ablation, search procedures are impossible. Learning must

22 Hirsh, "The Hippocampus and Contextual Retrieval," p. 434.

23 Hirsh, "The Hippocampus and Contextual Retrieval," p. 434.

rely upon averaging and thus changes in the behaviour of hippocampally ablated animals are much less discrete than those of controls." [24]

Hirsh views the theory of multiple memory systems as a means for achieving some degree of peace between behaviourism and cognitivism. Hippocampal learning involves the retrieval of discrete pieces of information about the world. As such, it is consistent with the general principles of representationalism. Nonhippocampal learning involves the gradual acquisition and modification of dispositions to respond to particular stimuli in particular ways. It is consistent, on Hirsh's account, with the basic principles of S-R theory.

> In the absence of the hippocampus, associative retrieval operates. Behaviour is completely controlled by external stimuli and learning is a matter of habit formation. Readers familiar with learning theory will realize that the behaviour of normal animals is treated in neo-Tolmanian framework, while that of hippocampally-ablated animals is held to be everything for which early S-R theorists could have wished. [25]

O'Keefe and Nadel's Theory of Multiple Memory Systems

Hirsh's theory was not the only important new theory of hippocampal function to emerge in 1974. In their book, *The Hippocampus as a Cognitive Map* (which was sent to press in 1974), O'Keefe and Nadel argue that the hippocampus is responsible for *spatial memory*. [26] It is difficult to overestimate the importance and influence of this text. It received more attention than did Hirsh's piece on contextual retrieval and signaled the public emergence of a new, more "cognitive" approach to learning and memory within the

---

[24] Hirsh, 'The Hippocampus and Contextual Retrieval," p. 434.

[25] Hirsh, "The Hippocampus and Contextual Retrieval," p. 439.

[26] O'Keefe, J., and L. Nadel, *The Hippocampus as a Cognitive Map* (Oxford: Clarendon Press, 1978).

neuroscience community. One commentator on the book notes that "it represents a substantial forward step toward realizing that the *behaviour* of organisms is not directly caused by stimuli (environmental energies that fire receptor cells) ... The book represents an ambitious effort to spell out in neural terms what in menta :tic psychology is called *cognition*." [27]

Like Hirsh, O'Keefe and Nadel argue that there are multiple memory systems in the brain, that one of these systems is, in some sense, a "Tolmanian" cognition system, that this system is centred in the hippocampus, and that nonhippocampal memory systems can be understood in non-cognitive terms. Nadel has recently described the effect of the Tolman - Hull debate on the early theories of animal memory systems. "Roughly speaking, the notion that the hippocampus is responsible for the kind of cognitive learning Tolman emphasized (especially place learning), while the rest of the brain is responsible for the kind of noncognitive learning Hull emphasized, has been central to most of the early dichotomies in the field." [28]

One of the critical differences between the theory of hippocampal function proposed by O'Keefe and Nadel and the theory proposed by Hirsh, is that while Hirsh views the hippocampus as responsible for mnemonic processing in which *context* is critical, O'Keefe and Nadel argue that the hippocampus is more specialized, that it is responsible only for those mnemonic processes in which *spatial context* is important. In commenting on Hirsh's theory of contextual retrieval, O'Keefe and Nadel comment that "this plausible suggestion runs into one major stumbling block: putting all contextual functions into the hippocampus leaves little for the remainder of the brain to do. One of us has considered the

---

[27] Garth Thomas, "Cognition, Memory, and the Hippocampus," published as a commentary on O'Keefe, John, and Lynn Nadel, "Précis of O'Keefe and Nadel's *The Hippocampus as a Cognitive Map*," *The Behavioural and Brain Sciences* 2 (1979), pp. 515 - 16.

[28] Nadel, Lynn, "Multiple Memory Systems: What and Why, an Update," Memory Systems 1994, Eds. Daniel Schacter and Endel Tulving (Cambridge, Mass.: MIT Press, 1994), p. 44.

problem of context at length...., and concluded that the hippocampus plays a role in spatial, but not other, context effects." [29]

The question of whether or not the hippocampus is specialized *specifically* for spatial memory is one of the most significant and ongoing debates within the literature on multiple memory systems. The literature cannot be understood without an appreciation of the importance of this debate. Ever since they espoused the view in 1974, O'Keefe and Nadel have engaged in ongoing debates with other scientists in the field, many of whom claim that the hippocampus is responsible for the processes of *relational* memories, and that spatial memories are simply one particularly important type of relational memory. It should be noted and stressed, however, that O'Keefe and Nadel argue that the hippocampus is specialized for spatial memory only *in non-human animals*. They accept the claim that the human hippocampus is involved in more general types of relational memory processing. There are important and philosophically interesting reasons for supposing that the hippocampus plays a more general role in human mnemonic processing than in animal mnemonic processing, reasons that we will examine in due course. For the most part, however, we can sidestep the debate about the specificity of the hippocampus' role and concentrate on discovering why scientists claim that the hippocampus is involved in *relational* mnemonic processes, be they spatial or non-spatial.

The debate between Tolman and Hull may be interpreted in terms of the different types of spatial learning each attributes to animals. At issue is the question of whether rats acquire "cognitive maps" or learn "responses". Tolman argues that animals acquire non-egocentric cognitive maps of their environment while Hull claims that animals learn about simpler, egocentric spatial relationships involved in making responses to stimuli (e.g. turn right at the choice point). While Tolmanian rats engage in "place learning", Hullian rats engage in "response learning". Endel Tulving, an important theorist in the multiple

---

29 O'Keefe, John, and Lynn Nadel, "Précis of O'Keefe and Nadel's *The Hippocampus as a Cognitive Map*," *The Behavioural and Brain Sciences* 2 (1979), p. 525.

memory systems field, once joked that "place-learning organisms, guided by cognitive maps in their heads, successfully negotiated obstacle courses to food at Berkeley, while their response-learning counterparts, propelled by habits and drives, performed similar feats at Yale." [30] O'Keefe and Nadel argue that both Tolman and Hull can be accommodated, as there are two distinct spatial memory systems in the brain, a hippocampal "locale" system responsible for the acquisition and use of non-egocentric spatial maps of the environment and a nonhippocampal "taxon system" that tracks the egocentric spatial information involved in making responses to particular stimuli.

The term "locale system" is meant to convey the view that the hippocampus is specialized for spatial memory. The term "taxon system" requires a bit more explanation. As O'Keefe and Nadel stress, their primary purpose in discussing taxon systems is to provide an account of what hippocampal processing *is not*. The concept of a taxon system is thus defined negatively; taxon systems are neural memory systems that are not dependent on the hippocampus. One of the taxon systems is devoted to tracking egocentric spatial information. Others are responsible for a wide variety of mnemonic functions (perceptual memory, S - R associative memory, simple conditioning memory, skill memory, etc.) "It must be stressed that the 'taxon systems' embrace the vast majority of the central nervous system, including both sensory and motor systems, and the generalizations put forward here can scarcely be expected to apply at all levels of this set of systems." [31] As O'Keefe and Nadel point out, there is no one functional description that could adequately and accurately represent such a varied collection of memory systems. On the other hand, O'Keefe and Nadel are able to identify a few basic principles of operation which distinguish nonhippocampal mnemonic processing from hippocampal mnemonic processing, principles which are of interest here. As far as the name is concerned,

---

[30] Tulving, E., and S.A. Madigan, "Memory and Verbal Learning," *Annual Review of Psychology* 21 (1970), pp. 437-84. Cited in O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, footnote, p. 73.

[31] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 97.

O'Keefe and Nadel state that "the term 'taxon' was chosen to denote the fact that processing within the nonhippocampal systems was based on the taxonomic principles of category inclusion and generalization." [32] While some of the taxon systems described by O'Keefe and Nadel are clearly taxonomic in this sense, others are not obviously governed by the principles of "category inclusion and generalization." For the moment, we should not let this worry us. I will attempt, at the end of this discussion, to clarify this issue. Unless otherwise noted, I will use the term "taxon system" to refer to the nonhippocampal memory system devoted to the negotiation of egocentric space.

The locale system and the taxon system in animals are both dedicated to spatial navigation, but while the locale system is based on *map* navigation, the taxon system is based on *route* navigation. The differences between map navigation and route navigation may be explained by comparing two different types of navigational instruction we might offer a lost traveler. If a map is handy, we help the traveler by identifying two points on the map, the "you are here" point and the destination point. Equipped with this information and the map, the traveler can then select any one of a number of pathways that can get her from her current location to her desired location. Note that if a certain road is closed, the traveler has all the information necessary for selecting an effective detour route. Furthermore, the map provides the traveler with freedom and flexibility. If she wishes to change her destination, she has the information required to do so. The map has a high information content; each location is spatially related to every other location. As a result maps show "great resistance to degradation"; [33] the map remains useful even if portions of it become unreadable.

The taxon system is responsible for two different types of route strategies: those that emphasize the "stimulus" portion of the instruction and those that emphasize the "response" portion of the instruction. Stimulus-oriented route instructions focus attention

[32] Nadel, "Multiple Memory Systems: What and Why, an Update," p. 43.

on particular landmarks that can serve as *guidance* objects for the animals (e.g. go until you get to the cathedral). This type of route strategy is thus called a *guidance* strategy. Response-oriented route instructions focus attention on what is required of the animal and typically involve the notions of *orientation* and *directions* (e.g. walk straight two blocks, turn right at the corner).

Route strategies are a more basic form of navigation. They are useful only for going from one specific location to one specific destination. Route instructions are easy to follow; the destination is set and there is only one known pathway for getting there. On the other hand, route instructions are very inflexible. As a result of this inflexibility, route instructions are also very vulnerable. They can be rendered useless by even small amounts of physical damage to the instructions themselves or by small changes in the environment. While maps allow the traveler to choose from among a number of different pathways, routes are constrictive. "(Routes) are inflexible, must be used in the correct sequence, and only rarely allow freedom of choice to the traveler." [34] There is one more significant difference between a route and a map. While the existence of a map says nothing about the desirability of locations within the map, in a route, the endpoint is always construed as a goal. The existence of a route implies the value of an endpoint. "Routes *imply goals which imply motivations.*" [35]

So much for our general description of the differences between map strategies and route strategies. We must now see how these two sorts of navigational strategies are instantiated in animal memory systems.

A map may be defined as "the representation of a set of connected places which are systematically related to each other by a group of spatial transformation rules." [36] So far,

---

[33] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 88.

[34] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 83.

[35] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 83.

[36] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 86.

we have been thinking in terms of a physical map, one printed on a piece of paper. We may also think of maps in both psychological terms and neural terms. A psychological map consists of a number of "place representations". In a neural map, a place may be defined "in terms of the activation of a specific array of hippocampal neurons," i.e. in terms of a neurally realized representation.[37]

Studies conducted in the early 1970's by O'Keefe and Dostrovsky indicated that cells in the hippocampus are particularly sensitive to spatial stimuli. Electrodes were implanted in a number of rats who were then allowed to move freely within a certain circumscribed environment, an environment which was replete with various visual landmarks. The activity of hippocampal cells was then monitored as the rats explored their surroundings. Certain of these hippocampal cells fired "solely or maximally when the rat was *situated in a particular part of the testing platform facing a particular direction.*"[38] For obvious reasons, these neurons became known as "place cells". O'Keefe and Dostrovsky note that "these findings suggest that the hippocampus provides the rest of the brain with a spatial reference map. The activity of cells in such a map would specify the direction in which a rat was pointing relative to environmental land marks and the occurrence of particular tactile, visual, etc. stimuli whilst facing in that orientation."[39] Other researchers found hippocampal "mismatch" cells, neurons that fire maximally when the rat encounters a location in which a stimulus is "missing", i.e. a location in which a previously encountered stimuli is no longer in its usual location. Certain cells have firing patterns which appear to be linked to the animals "'orienting and approach behaviour".[40] Nadel

---

37 O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 93.

38 O'Keefe, J., and J. Dostrovsky, "The Hippocampus as a Spatial Map: Preliminary Evidence from Unit Activity in the Freely Moving Rat," *Brain Research* 34 (1971), p. 172.

39 O'Keefe and Dostrovsky, "The Hippocampus as a Spatial Map," p. 174.

40 Nadel, Lynn and John O'Keefe, "The Hippocampus in Pieces and Patches," *Essays on the Nervous system: A Fetschrift for Professor J. Z. Young*, Eds. R. Bellairs and E.G. Gray (Oxford: Clarendon Press, 1974), p. 382.

and O'Keefe note that such place cells, mismatch cells, and approach and orientation cells "could provide the raw material for a mapping system." [41]

When a rat is placed in a novel environment, it begins with "a tabula rasa of potential place representations." [42] Place learning can be divided into a series of discrete episodes in which the rat acquires a new piece of information about the spatial relations among objects in the environment. For the sake of simplicity, consider the information that might be acquired by a rat who remains in a fixed position. At time t_1, the rat may learn that objects 1 and 2 are located in specific positions, ten centimeters apart, with object 1 to the left of object 2. At t_2, the rat may turn in place and notice that object 3 is located five centimeters to the right of object 2. Note that the rat can now surmise the spatial relation between objects 1 and 3, even though objects 1 and 3 were never visually experienced at the same time. "Though the cartoons of the environment are initially generated by experience directly related to an organism (and are thus egocentrically inspired), the matrix of relationships constituted in a cartoon provides information about the environment which is not tied to the organism. The extreme power of mapping derives from this fact. Every representation within a cartoon stands in some relation to every other, whether or not they have ever been experienced in spatial or temporal contiguity." [43] When rats engage in exploratory behaviour, their calculations become more complex. The same basic principles do, however, apply. "Some interaction between sensory data and movement feedback is required to identify the procession of places as an animal moves through its environment." [44]

Several properties of this sort of spatial learning must now be described. First, the learning occurs in specific contexts. The animal *now* perceives that two objects are located

[41] Nadel and O'Keefe, "The Hippocampus in Pieces and Patches," p. 382.

[42] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 94.

[43] O'Keefe and Nadel, "The Hippocampus in Pieces and Patches," p. 382.

[44] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 94.

in particular positions relative to one another. That information is acquired all at once, in an instance. The location of these objects is noted on the map. Subsequent exposure to these two objects yields no additional information about the locations of these objects. (Similarly, if a subsequent exposure to the same location reveals that an object is missing, mismatch cells fire and the object is deleted, in an instant, from the map.) Information acquisition in the locale system is episodic: it occurs very rapidly in an all-or-none fashion.

Secondly, the underlying motivation that drives cognitive map learning is curiosity. The locale system is geared for "exploration or novelty-directed behaviour." [45] According to O'Keefe and Nadel, both novelty and anomaly trigger a "misplace" detection signal whose output "activates and directs the motor systems involved in exploration. This behaviour is directed toward the incongruous and new information can be incorporated into the map as a result of it." [46]

Third, the locale system is relatively protected against "interference" effects. If similar objects are encountered in a single environment, their neural representations are "isolated" from one another in virtue of their locations in the map. Furthermore, if the same object is encountered in two different spatio-temporal contexts, each appearance of the object will be stored in a separate neural representation. "Since each representation of a stimulus is encoded in terms of its spatial relations to other stimuli, an identical stimulus occurring in different parts of the same environment, or in totally different environments, will have distinct, and differentiable, representations in each case." [47] The locale system thus extracts and stores information about what is unique to particular episodes. Because each object is represented *within a particular context*, the particular features exhibited by the object in that context are represented and preserved. Subsequent exposure to the same object in a different context results in a *new* representation of the object. The features of

---

[45] Nadel, "Multiple Memory Systems: What and Why, an Update," p. 45.

[46] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 94.

[47] O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 95.

the first and second representations of the object are not blended together; the two representations remain distinct, neurally anchored in two different maps. "The hippocampal system emphasizes what is unique about a memory, and functions to separate memory traces on the basis of what distinguishes one from another." [48]

Fourth, the representational schema of the locale system is very flexible. Since place locations are stored in discrete, neurally realized representations, and since a map consists of a collection of such discrete representations, there is a great deal of *relational* spatial information encoded into each map. Using a cognitive map, an animal may choose shortcuts or detours from location A to location B, even though it does not have experiential knowledge of such pathways. "The maplike representations in the locale system are the basis for generating novel outputs, such as detours in mazes. This ability to generate novel output arises from the *flexibility* of the representation, which allows it to be used in novel ways." [49]

Animals who have bilateral hippocampal damage can still navigate their way through the environment, to some extent, thanks to the existence of the nonhippocampal taxon systems. A rat without a functioning hippocampal system would be deprived of its ability to form cognitive maps and would thus be "forced to rely on other inherently less flexible strategies to find its way: turn left at the junction, follow this odour, avoid that bright light." [50] The taxon systems operate on a different set of principles. Taxon behaviour is motivated not by curiosity, but by the goal of drive reduction (e.g. by the reduction of thirst or hunger). Nadel states that "taxon learning ... (is) assumed to be motivated by the traditional forces emphasized by Hull, and therefore to depend on the standard application of reinforcement." [51] As a result, taxon learning is gradual. Since

---

48 Nadel, "Multiple Memory Systems: What and Why, an Update," p. 45.

49 Nadel, "Multiple Memory Systems: What and Why, an Update," p. 45.

50 O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 175.

51 Nadel, "Multiple Memory Systems: What and Why, an Update," p. 45.

taxon learning consists of small and gradual changes to the system responsible for performance, it is more prone to interference effects than is hippocampal learning. There is no protective "off-line" representational storage in taxon systems. On the other hand, taxon learning is, in the long run, more stable. While the individual representations involved in hippocampal learning may be destroyed with relative ease, the multitude of small neural changes involved in taxon learning can be undone only gradually, through a lengthy process of "extinction".

For O'Keefe and Nadel, the signature feature of taxon learning is that it is context-free. This claim requires a bit of explanation. Hippocampal memory involves a memory for the context in which a memory was acquired, while nonhippocampal memory does not. This should not be confused with another context-related distinction between hippocampal and nonhippocampal learning. Once a hippocampal memory is acquired, it may be used flexibly, in novel contexts. As O'Keefe and Nadel stress, one of the critically important features of hippocampal memory is that a representation acquired in one spatio-temporal context can be combined with representations acquired in other contexts. Nonhippocampal skills, on the other hand, are relatively inflexible; they can be deployed only in contexts that are sufficiently similar to the acquisition context.

Before leaving O'Keefe and Nadel's theory of multiple memory systems, we should look briefly at their speculative remarks concerning how their theory might be modified to accommodate the data on *human* memory systems. In a relatively recent article in the journal, *Hippocampus*, Nadel describes the central difference between hippocampal memory in humans and animals. "I will toss in the towel and admit, as we did in 1978, that at least in the case of the human hippocampal system, there is more than merely spatial mapping going on."[52] He cites the relevant passage from the book, *The Hippocampus as a Cognitive Map.*

> The cognitive map in infra-humans should be viewed as a spatial map in which representations of objects experienced in the environment are ordered within a framework generating a unitary space. However, the central property of the locale system is its ability to order representations in a structured context. The development of objective spatial representations is not the only possible use for such a system ... mapping structures can represent verbal, as well as non-verbal systems. For both of these forms the locale system will be shown to be central to a particular form of memory: that concerned with the representations of experiences within a specific context.[53]

O'Keefe and Nadel, in broadening their functional description of the locale system, suggest that the human hippocampus is responsible for a non-spatial, episodic memory, the retention of an ordered narrative account of the individual episodes which constitute a lifetime. "The hippocampus is the neural substrate for one-trial episodic memories in which events are related to each other in a long map extending from the past into the future."[54]

On the issue of "verbal" memories, O'Keefe and Nadel offer a few admittedly speculative comments worth noting here. In humans, the right hippocampus is associated with spatial memory processing while the left hippocampus is associated with verbal memory processing. (Subjects with right unilateral hippocampal damage are impaired in spatial memory tasks but not verbal memory tasks. Subjects with unilateral left

---

52 Nadel, Lynn, "The Hippocampus and Space Revisited," *Hippocampus* 1.3 (1991), p. 227.

53 O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 381. Cited in Nadel, "The Hippocampus and Space Revisited," p. 227.

54 O'Keefe and Nadel, "A précis," p. 494.

hippocampal damage exhibit a reversed pattern of impairments.)[55] O'Keefe and Nadel hypothesize that the left hippocampus may be responsible for generating the linguistic equivalents of spatial cognitive maps. "We postulate that the left hippocampus receives information about linguistic entities and sets these, rather then items drawn from the physical world, into a mapping space."[56] The relationship between language and the hippocampus unclear. "To our dismay, there has been no systematic attempt to date to explore the possibility that the hippocampal formation has a central role to play in certain aspects of language. That there are verbal learning defects after damage to the hippocampus in the left hemisphere has been well-known for some time; the precise nature of these defects remains, however, unclear."[57] [58]

---

[55] Kolb, Bryan, and Ian Q. Whishaw, eds., Fundamentals of Human Neuropsychology, 2nd ed. (New York: W. H. Freeman and Co., 1985), p. 486.

[56] O'Keefe and Nadel, "A précis," p. 493.

[57] Nadel, "The Hippocampus and Space Revisited," p. 222.

[58] One final note on O'Keefe and Nadel is in order. Having promised to clarify the origins of the term "taxon memory", I will do briefly, though the explanation may muddy waters that should be kept clear. The hippocampal system in humans is often called the "declarative" memory system. Tulving has distinguished between two types of declarative memory: episodic memory and semantic memory. Episodic memory preserves information about particular events while semantic memory preserves context independent information about "facts". "I had a banana for breakfast this morning" is an example of an episodic memory. "Bananas are elongated, tubular fruits, typically yellow in colour" is an example of semantic memory.

In the case of humans, O'Keefe and Nadel argue that some of the taxon systems are devoted to semantic memory. They note that "taxon systems generalize over similarities" and may be responsible for producing "prototype" representations of objects. These taxon systems would thus be responsible for categorizing the world, providing a taxonomy of its contents. "The general absence of context information characterizes the memory-storage properties of the taxon systems. Concepts and categories, the look, the feel and sound of things, the goodness of badness of objects: all these are represented within the taxon systems. What is missing is the spatio-temporal context within which this knowledge was acquired." (O'Keefe and Nadel, *The Hippocampus as a Cognitive Map*, p. 100.)

In discussing human memory systems, O'Keefe and Nadel note that the distinction between locale and taxon memory bears a resemblance to Tulving's distinction between episodic and semantic memory. This claim muddies the waters, in this context, because the episodic - semantic distinction is typically used to differentiate between two types of declarative memory, *not* to differentiate between hippocampal memory and nonhippocampal memory.

For our purposes, it is best to focus on the account O'Keefe and Nadel provide of taxon systems in nonhuman animals. The following passage captures well the type of taxon system that is relevant in this context.

> Memory systems, concerned with the value of objects, with specific motor habits, with egocentric space, all are located outside the hippocampus, providing the basis for considerable learning in the animal without its hippocampus. [59]

A certain picture of hippocampal function emerges from the work of Hirsh, O'Keefe and Nadel. While we may not be able to extract a detailed account of the memory system responsible for skilled perceptual-motor behaviour, we can distinguish between the basic principles of operation that govern the hippocampal system and the basic principles that govern all nonhippocampal memory systems, including the system responsible for perceptual-motor skills. The hippocampal system operates on the principles of quick, one-trial acquisition rates and rapid extinction rates. It operates over distinct neurally realized representations. It appears to be implicated in relational processes, in processes in which two or more distinct neural representations are manipulated concurrently, in a way that respects the spatial or logical relations among the representations. The hippocampal system is not driven by reward reduction, but by general curiosity, a curiosity which is indifferent to the potential reward "value" of objects encountered in the environment.

The nonhippocampal systems, on the other hand, operate on the principles of slow acquisition and slow extinction. The mechanisms of nonhippocampal processing involve direct and gradual changes to the performance systems (e.g. the visual system, the motor system) but do not involve the creation or manipulation of discrete, transportable representations. Nonhippocampal systems are sensitive to the valence of objects encountered in the environment and are driven by traditional drive-reduction motivations. [60]

---

59 Nadel and O'Keefe, "The Hippocampus in Pieces and Patches," p. 384.

60 Objects have a particular valence, positive or negative, depending on whether they are desirable or noxious.

With this preliminary sketch of the differences between hippocampal memory and nonhippocampal memory in mind, we can turn to two additional accounts of hippocampal memory, accounts in which the Fodorian nature of hippocampal representations is made particularly clear.

Mishkin's Theory of Multiple Memory Systems

In the last ten years, Mishkin has developed Hirsh's peace proposal into a detailed theory of the differences between "representational" memory and "habit" memory. In a 1984 article, "Memories and Habits: Some Implications for the Analysis of Long Term Learning and Retention,"[61] Mishkin proposes a dual memory system model which he describes as being "most similar conceptually to the one advanced by Hirsh in 1974."[62] While Mishkin and Hirsh disagree on the precise neural substrates of the two systems, they both argue for the claim that the cognitivism - behaviourism debate may be settled by "mapping" these theories onto the two memory systems. As Mishkin writes in his 1984 article, "the dual-systems model of retention likewise suggests that we do not have to choose between behaviourist and cognitivist explanations of learning. Rather, the data from amnesic patients imply - and the neural model proposes - that both types of learning occur, with each neural circuit storing different products of experience."[63]

Mishkin describes the "two" memory systems as the "representational" memory system and the "habit" or "skill" system. While most researchers argue that the representational system is dependent on the hippocampus, Mishkin argues that only some forms of representational memory require the proper functioning of the hippocampus. Other forms of representational memory are dependent, on his view, on medial temporal

---

61 Mishkin, M., and H.L. Petri, "Memories and Habits: Some Implications for the Analysis of Long Term Learning and Retention," *Neuropsychology of Memory*, Ed. L. Squire and N. Butters (New York: Guilford Press, 1984), pp. 287-296.

62 Mishkin and Petri, "Memory and Habits," p. 288.

lobe structures other than the hippocampus. For that reason, I will now speak of medial temporal lobe (MTL) memory, not hippocampal memory, when referring to the neural substrate of the representational memory system.

The theoretical disagreements between cognitivists and behaviourists (with respect to the issues of learning and memory) can be outlined by examining their views on four components of the learning experience: sensory, motor, motivation, and satisfactional components. To be more precise, cognitivists and behaviourists, on Mishkin's account, disagree on (1) the nature of learned visual stimuli (the sensory component), (2) the nature of learned motor responses (the motor component), (3) the relative importance of needs versus motives in learning (the motivational component), and (4) the relative importance of reinforcements versus incentives in learning (the satisfactional component). Before turning to Mishkin's work on the memory systems, I will briefly review Mishkin's account of these four differences between cognitivism and behaviourism. This will provide us with a general idea of Mishkin's views on the differences between representational memory and habit memory.

The first area of disagreement concerns the nature of retained sensory information, particularly in the visual modality. The behaviourist argues that all of the constitutive *features* of a visual stimulus (e.g. size, shape, texture, hue, etc.) that are present during reinforcement will be, in some sense, "registered". Thus the animal learns about stimulus features, any one of which may play a causal role in eliciting the reinforced response in the future. The cognitivist, in contrast, argues that the animal retains information about meaningful *configurations* of stimulus features (e.g. salient objects). On this view, the animal does not perceive the world in terms of static bits of sense data, but in terms of meaningful objects. The visual trace of experience then is a configuration of stimulus features and it is this configuration that may be causally efficacious with respect to future

[63] Mishkin and Petri, "Memory and Habits," p. 288.

behaviour. Mishkin discusses Tolman in particular, and notes that for him, "stimuli were viewed as configurations to which meaning becomes attached, and which thereby provide information to the organism about its environment." [64] Thus cognitivists argue that animals learn "stimulus configurations " while behaviourist argue that animals learn "stimulus elements" or "stimulus features". Mishkin argues that a dual-systems model of memory can accommodate both of these views on the learned visual stimuli. As we will see shortly, Mishkin and his colleagues have shown that the two memory systems receive different types of visual information.

Second, behaviourists and cognitivists differ on the issue of learned motor activity. S-R theorists argue that animals learn specific motor responses. Mishkin reports that "the S-R behaviourist approach emphasized that learning is reducible to the formation of conditioned reflexes (Watson, 1930) or to the acquisition of simple movements (Guthrie, 1935) or habits (Hull, 1943). Given the necessary drive and reinforcement, response acquisition was regarded as essentially automatic, requiring awareness of neither the stimulus that evokes it nor the outcome to which it leads." [65] On this view, responses are shaped gradually via the incremental modification of S-R neural circuitry. Cognitivists, on the other hand, argue that animals may learn specific "acts". Mishkin writes, "By contrast, the cognitive approach, again typified by Tolman's views, regarded behaviour not simply as muscle contractions or glandular secretions, but as acts. The choice of the term 'act' was meant to imply that the behaviour is purposive or intentional, and is therefore directed toward some goal expected on the basis of past experience." [66] The acquisition of "acts" provides the basis for flexible, volitional behaviour. Thus behaviourists maintain that animals learn specific motor responses while cognitivists argue that animals learn intentional acts. Here again, Mishkin argues that the dual-memory systems model can

---

[64] Mishkin and Petri, "Memory and Habits," p. 290.

[65] Mishkin and Petri, "Memory and Habits," pp. 290 - 91.

[66] Mishkin and Petri, "Memory and Habits," p. 291.

accommodate both views. On Mishkin's account, the two memory systems are related to the motor systems in different ways.

Third, behaviourists and cognitivists differ on the relative importance of needs versus motives in the learning process. The behaviourist's explanatory framework is based on the idea that animals learn as a result of the imperative of drive reduction. "Hull (1943, 1951, 1952) a primary advocate of the S-R approach, initially emphasized the role of physiological needs or tissue deficits in the generation of drive (and drive stimuli) which in turn had to be reduced in order for learning to occur." [67] The cognitivist argues, however, that a great deal of learning occurs in the absence of immediate biological promptings. On this view, learning can result from "cognitive motivations" such as the desire to explore the environment or solve puzzles. Sated animals engage in exploratory behaviour; monkeys will master mechanical puzzles even when there is no reward for their labors. [68] Mishkin notes that when attention was drawn to the phenomenon of unrewarded motivation, the drive-reduction theory suffered a serious blow and was discounted in many quarters. He cautions, however, that both biological needs and cognitive motivations may play important roles in animal learning. "Perhaps burial of the drive reduction theory was premature. In the dual-systems model of retention, there is room for learning both through reduction of bodily needs and through the fulfillment of cognitive motives." [69]

A fourth difference between behaviourists and cognitivists is their views on the relative importance of reinforcements versus incentives in learning. Behaviourists focus on the efficacy of reinforcers in gradually shaping S-R associations. Cognitivists argue that learning can occur in the absence of reinforcers. "For cognitivist theorists such as Tolman ... reinforcement was not a necessary conditions for learning, as demonstrated by studies

---

[67] Mishkin and Petri, "Memory and Habits," p. 292.

[68] Mishkin and Petri, "Memory and Habits," p. 292.

[69] Mishkin and Petri, "Memory and Habits," p. 292.

on latent learning." [70] On the cognitivist view, behaviour is driven by "expectancies" generated when the animal consults its "cognitive map" of the environment. "According to the cognitivists, 'reinforcers' serve not to increase response probability, but to confirm or disconfirm an expectancy. The expectancy, in turn, provides the incentive for future acts. In short, behaviour is purposive or goal-directed in conforming with past learning." [71] Thus while the behaviourist focuses on the causal efficacy of reinforcers in shaping behaviour, cognitivists focus on the power of expectancies to generate purposive behaviour. Again, Mishkin argues that there is room for both views. "We suggest, again, that learning may involve both response strengthening and expectancy development, but that these two different storage processes are controlled by separate neural systems. The dual-systems model proposes that the response outcome can serve as the reinforcer that strengthens S-R associations within the habit system, but that the same anticipated outcome can also serve as an incentive within the (representational) memory system." [72]

Mishkin suggests, in very schematic terms, that the cognitive and behaviourist conceptions of learning can be mapped onto the two memory systems. The key to this mapping lies in the fact that for the cognitivists, all four components of the learning experience (sensory, motor, motivational, and satisfactional) are explicitly represented at the neural level. For the behaviourist, on the other hand, none of the elements is explicitly represented. With respect to the learning of visual stimuli, Mishkin notes that "in the habit system each of the components of a stimulus complex becomes connected by reinforcement to a specific response, as the behaviourists have argued, whereas in the (representational) memory system ... the neural representations of stimuli are stored as configurations to which meaning can subsequently become attached, as the cognitivists have proposed." [73]

---

70 Mishkin and Petri, "Memory and Habits," p. 293.

71 Mishkin and Petri, "Memory and Habits," p. 293.

72 Mishkin and Petri, "Memory and Habits," p. 292.

73 Mishkin and Petri, "Memory and Habits," p. 290.

With respect to the learning of motor abilities, the behaviorist's "responses" are the result of incremental changes to a performance system but the cognitivist's "acts" may be captured in full-fledged representations and thus come to play a role in complex, combinatorial memories. "When an act does enter into an associative memory, it must do so, according to the model, on the basis of an acquired connection with the stored neural representation of that act." [74] The behaviorist's biological "needs" create the right condition for learning, but the cognitivists "motives" are involved in complex associative learning - "they become integral components of the association, inasmuch as motives (like all other components in memory) are stored neural representations." [75] Finally, reinforcements drive the neural mechanisms that strengthen the bond between a stimulus and a response, but incentives themselves are another type of representations. "The incentive is presumed to be stored as a neural representation, which then becomes another component of associative memory." [76]

For the cognitivist, learning involves the acquisition of neural representations of stimulus configurations, acts, motives, and incentives in such a way that these representations can enter flexibly into various types of associative memories. They may become the objects of conscious manipulation. The ability to engage in flexible, volitional behaviour depends on this process of manipulating representations freely. For the behaviourist, an act of learning involves the associations of a specific motor response to a specific stimulus via the reinforcement of an S-R association for the sake of drive reduction. Learning is reducible, on this account, to the incremental modification of performance systems.

Mishkin's 1984 article on memories and habits provides only a tentative sketch of the similarities between cognitive memory and the representational memory system, on the

---

74 Mishkin and Petri, "Memory and Habits," p. 291.

75 Mishkin and Petri, "Memory and Habits," p. 292.

76 Mishkin and Petri, "Memory and Habits," p. 293.

one hand, and behaviourist memory and the habit memory system, on the other. We can now turn to other portions of Mishkin's work, including his most recent material, to see how well the olive branch hypothesis fares when the memory systems are studied in much greater detail.

I will begin this examination of Mishkin's views on memory by briefly introducing his views on vision. This foray into the literature on the vision, and its relationship to memory, is made necessary by the fact that cognitivists and behaviourists differ on the role of vision in learning. As we will see, current views on the nature of perception bear directly on the issue of the joint accommodation of behaviourism and cognitivism.

We thus begin with a brief detour through Mishkin's discussion of the difference between the visual input into the MTL memory system and the visual input into the habit system. He argues, in effect, that the representational memory system receives information about particular objects construed as wholes, while the habit memory system typically receives information about the constitutive visual attributes of objects (e.g. size, colour, shape).

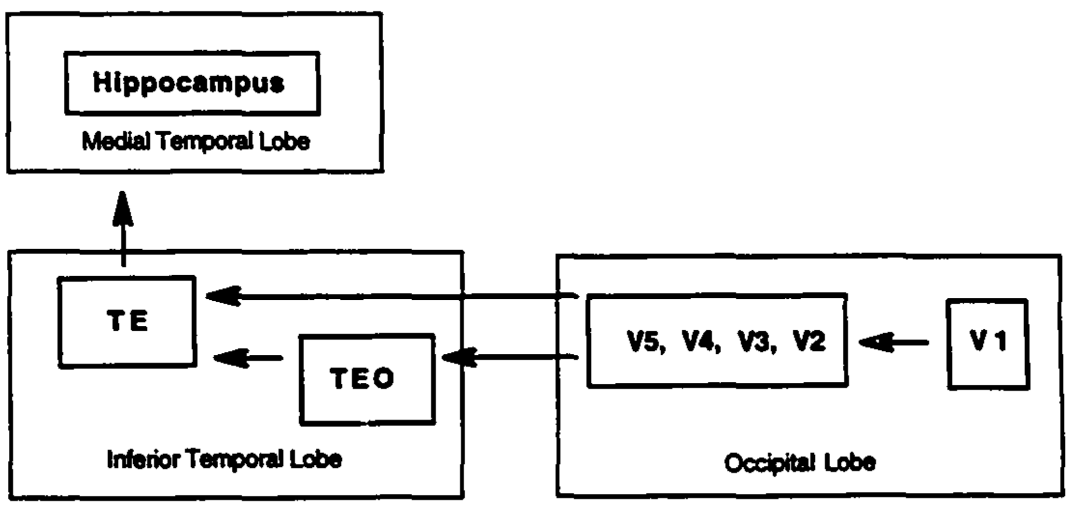

Figure 5.3: The Ventral Visual Pathway's Input into the Hippocampal System

Recall that there are two different cortical visual pathways, a dorsal occipito-parietal pathway devoted to spatial perception and the guidance of perceptual-motor behaviour, and a ventral occipito-temporal pathway devoted to object recognition. It is the ventral visual pathway that is relevant here and we will need to examine it in a bit more detail. The ventral visual pathway consists of a number of distinct processing areas: the primary visual cortex or "striate" cortex (V1), the prestriate areas (V2-V5), [77] and areas TEO and TE in the inferior temporal lobe. In the early stages of the pathway, neurons are highly specialized. They respond to stimuli in a very small portion of the visual field and then, only when the stimuli have certain specific properties. V3 cells, for example, are selective for orientation; they process information related to the form of the visual stimulus. V4 neurons are specialized to respond to both colour and form while V5 neurons are directionally sensitive - they serve to detect motion. In the later stages of the pathway, neurons respond to increasingly large areas of the visual field and to increasingly complex combinations of visual properties. Mishkin writes that "neurons in the pathway have 'windows' on the visual world that become progressively broader, in both their spatial extent and the complexity of the information they admit." [78]

Striate and prestriate areas process information about particular features of a visual stimulus, e.g. its orientation, texture, and hue. Area TEO is critical for processing information concerning the spatial characteristics of visual stimuli. Information from a multitude of prestriate and TEO neurons converges on single neurons in TE, enabling the TE neurons to collate a number of visual features present in a given visual stimulus. Mishkin argues that area TE, the endpoint of the ventral visual stream, is responsible for producing a full neural representation of the stimulus object. "The cells respond to progressively more of an object's physical properties - including its size, shape, colour and

---

[77] I limit my discussion to areas V1-V5 for the sake of convenience. There is now talk of at least twenty distinct visual areas and more will undoubtedly be identified.

[78] Mishkin, M., and T. Appenzeller, "The Anatomy of Memory," *Scientific American* 256.6 (1987), p. 82.

texture - until, in the final stations of the inferior temporal cortex, they synthesize a complete representation of the object." [79]

In the early 1980's Mishkin performed a series of neuropsychological studies in monkeys in which the capacity for the simple recognition of a previously seen object was tested after the selective ablations of different regions in the monkeys' ventral visual pathway. More specifically, Mishkin compared the performances of two groups of moneys: those who had received bilateral TEO lesions and those who had received bilateral TE lesions. The TEO monkeys were able to relearn the recognition task with relative ease post-surgically, and were able to perform nearly as well as normals on delays ranging from 10 - 120 seconds. "The (TEO) removal does not impair discrimination of an object's other, less complex features, which it is therefore reasonable to assume are processed in other subdivisions of the prestriate complex and then relayed forward to area TE, bypassing TEO." [80] TE monkeys, on the other hand, could not relearn the task even after 1500 trials. These results suggest that the representational memory system depends critically on visual input from TE, but not from TEO. Mishkin writes that "so specific and dramatic an impairment clearly indicates that the demands of the one-trial object recognition task approximates closely the functions of area TE." [81] Furthermore, Mishkin showed that when area TE is bilaterally disconnected from MTL, performance on a simple recognition test is severely impaired. This means that information from earlier processing stations in the ventral visual pathway can enter the MTL only via area TE. Since area TE is the endpoint of the ventral visual stream, the area in which various visual characteristics of a given object are brought together, Mishkin concludes that the representational memory system receives only highly processed visual information about objects, or "configurations" of visual stimuli.

---

[79] Mishkin and Appenzeller, "The Anatomy of Memory," p. 82.

[80] Mishkin, M., "A memory system in the monkey," *Phil. Trans. R. Soc. Lond. B* 298 (1982), p. 87

The habit system, on the other hand, receives projections from all areas of the ventral visual stream except the striate cortex. The habit system is centred in the basal ganglia, a collection of subcortical nuclei near the centre of the brain. Two of these sub-cortical nuclei, the caudate nucleus and the putamen, are collectively known as the striatum and serve as input modules of the basal ganglia, as gateways into the habit memory system.[82] St. Cyr and his colleagues have shown, using a retrograde tracing technique, that all areas of the ventral visual stream (except the striate cortex) send afferent fibers into proximal areas of the tail of the caudate nucleus. (In a retrograde tracing test, a tracing substance is injected into one location in order to identify the areas of the brain that project into that location.) We can assume, therefore, that both highly processed and barely processed visual information is sent into the caudate.

> Thus, following injections of the rostral tail of the caudate, the major labeled cortical field involved rostral inferior temporal area TE ... whereas for injections placed in more caudal portions of the tail, the labeled cortical region shifted to posterior portions of areas TE and into area TEO. As injections progressed yet more posteriorly into the genu of the caudate, the labeled region correspondingly shifted posteriorly into cortical areas V4, V3, and more sparsely, into V2.[83]

---

81 Mishkin, "A memory system in the monkey," p. 87.

82 Kandel, Eric, and James Schwartz, *Principles of Neural Science*, 2nd ed. (New York: Elsevier, 1985).

83 Saint-Cyr, J.A., L.G. Ungerleider, and R. Desimone, "Organization of Visual Cortical Inputs to the Striatum and Subsequent Outputs to the Pallido-Nigral Complex in the Monkey," *Journal of Comparative Neurology* 298 (1990), p. 133.

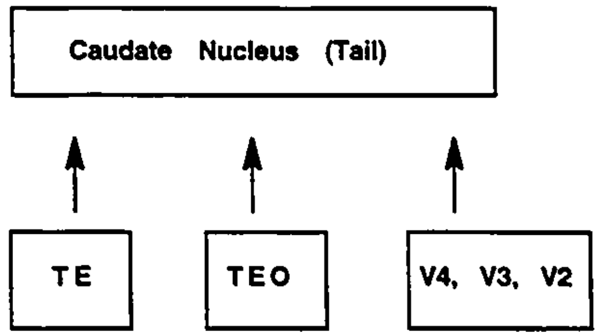

Figure: 5.4: The Ventral Visual Pathway's Input into the Basal Ganglia

These findings lend credence to Mishkin's claim that the dispute between cognitivists and behaviourists regarding the nature of visual stimuli can be resolved. Recall that cognitivists argue that learned visual stimuli consist of meaningful collections of "features", not individual features themselves. On this view, the animal does not retain information about isolated elements of sense data, but about objects, taken as wholes. This is consistent with the facts about the visual system's relationship to the representational memory system. Behaviourists, on the other hand, argue that animals retain information about isolated visual features of stimuli - their colour, size, or shape, for example. This view is consistent with the facts about the visual system's relationship with the habit memory system. Mishkin suggests that these data on the relationship between the two memory systems and the visual system indicate that both the cognitivists and the behaviourists were, in some sense, right.

> Perhaps a choice between the two positions is unnecessary. The dual-systems model suggests that stimuli can enter into learning in both ways - that is, both as elements and as configurations. We suggest that in the habit system each of the components of a stimulus complex becomes connected by reinforcement to a specific response, as the behaviourists have argued, whereas in the (representational) memory system (specifically within the sensory portion of that system located in the anterior temporo-insular region), the neural representations of

stimuli are stored as configurations to which significance can subsequently become attached, as the cognitivists have proposed. Thus, the model suggests that stimulus elements and configurations can enter into the stored product of experience in separate systems simultaneously. [84]

The Representational Memory System

The substrate of the representational memory system is a neural circuit consisting of three major processing centres: the medial temporal lobe (an area which includes the hippocampus), the medial thalamus, and the ventromedial frontal lobes. These three areas, together with their connecting fibers, are often referred to as the "limbic system". The medial temporal lobe (the MTL) contains two important sub-cortical nuclei - the hippocampus and the amygdala, a small, almond shaped nucleus located just anterior to the hippocampus. The amygdala is surrounded by the rhinal cortex and the entorhinal cortex, while the hippocampus is surrounded by the entorhinal cortex and, more posteriorly, the parahippocampal cortex (or parahippocampal gyrus, as it is commonly called).

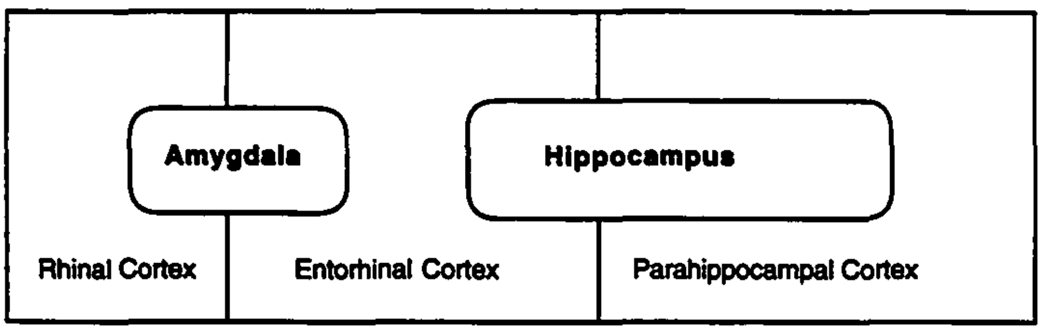

Figure 5.5: The Structures of the Medial Temporal Lobe

In the last ten years, there has been a great deal of debate concerning the functions of the various components of the MTL. For our purposes, it will suffice to think of the MTL as consisting of two distinct functional elements - the "front" (or "rostral") portion of the MTL and the "back" (or "caudal") portion of the MTL. The rostral MTL consists of

[84] Mishkin and Petri, "Memory and Habits," p. 290.

the amygdala and its surrounding cortices (the rhinal cortex and the anterior portions of the entorhinal cortex). The caudal MTL consists of the hippocampus -and its surrounding cortices (the posterior portions of the entorhinal cortex and the parahippocampal gyrus.) These two functional centres of the MTL serve as "headquarters" for two distinct neuronal "loops" within the limbic memory system. (That is to say, the one loop runs through the rostral MTL, through the medial thalamus, into the medial frontal lobes and back into the rostral MTL. The second loop follows the same course except that it begins and ends in the caudal MTL.)

	Rostral (Front) MTL	Caudal (Back) MTL
Sub-cortical Nucleus	Amygdala	Hippocampus
Cortical Tissue	Rhinal and Entorhinal	Entorhinal and Parahippocampal

Figure 5.6: The Rostral and Caudal Components of the MTL

The limbic memory system subserves a number of mnemonic functions, all of which can be grouped together under the general heading of "representational" memory. The general function of the limbic memory system is the acquisition and maintenance of neural representations of stimuli. These representations provide the resources necessary for the two major types of representational memory: recognition memory and associative recall memory. In recognition memory, a stimulus is perceived as familiar in virtue of the activation of its stored neural representation. In associative recall memory, the neural representation of a stimulus triggers the activation of the neural representation of some other, previously experienced stimulus. In what follows, I will provide a brief overview of the literature on recognition memory and associative recall memory.

Recognition Memory

The limbic memory system, or MTL system, is responsible for two types of recognition memory: object recognition and place recognition. In discussing the visual afferents to the MTL, we have already sketched the preliminary neural steps involved in object recognition. Information flows from the striate cortex, into the inferior temporal lobe until it reaches area TE, where many different types of visual information concerning a single stimulus or object are correlated. At some point in time during the perception of an object, neurons in the prestriate area, area TEO, and area TE are simultaneously activated. These active neurons collectively constitute what Mishkin calls a "neuronal ensemble". The neurons of the ensemble that reside in area TE then trigger the activity of the limbic memory system.

The MTL structures serve as the gateway to the limbic memory system. Information travels from the MTL to the medial thalamus and on to the ventromedial portions of the prefrontal lobe. The medial thalamus and the medial frontal lobes then activate another area of the medial frontal lobe - the basal forebrain. The basal forebrain is in turn reciprocally connected with the medial thalamus and the occipital-temporal visual areas. When activated, the forebrain releases a neurotransmitter, acetylcholine, at its terminal sites in the thalamus and occipital-temporal regions. This neurotransmitter then prompts a sequence of neurochemical events that serve to strengthen or weaken the synaptic connections among the neurons of the "neuronal ensemble" - the very neurons that activated the memory system in the first place. The end result of the process is that the neuronal ensemble is converted into a more stable structure - the cell assembly.

> The strength of many of the affected synapses is thereby increased, binding some small portion of the ensemble and limbic neurons into a more or less enduring,

lattice-like network. This network, which will be called a 'cell assembly' after Hebb, constitutes the stored representations of the stimulus. The cell assembly is presumed to be responsible for recognition. According to this proposal, visual recognition occurs whenever the neuronal ensemble activated by a retinal stimulus reactivates a previously formed cell assembly consisting of a small but critical subset of the neuronal ensemble. [85]

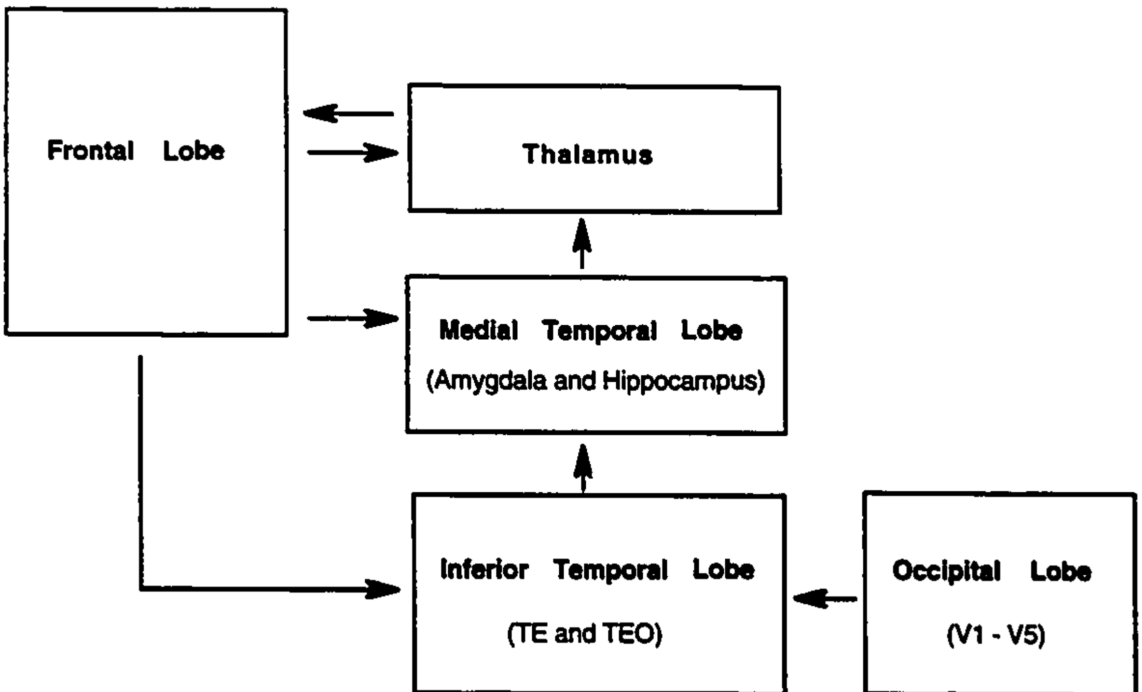

Figure 5.7: The Anatomical Substrate of Mishkin's Cell Assembly

The classic task used in studies of object recognition in animals is the Delayed Non-Matching to Sample task (DNMS). In this task, an animal is presented with a sample object placed over one of two foodwells. The monkey is trained to displace the object. Once the sample object is displaced, a curtain is lowered between the monkey and the foodwells and there is a delay ranging from a few seconds to several minutes. When the curtain is raised, both foodwells are now covered with an object. The sample object, which may or may not

[85] Mishkin, Mortimer, "Cerebral Memory Circuits," *Exploring Brain Functions: Models in Neuroscience*, Ed. T.A. Poggio and D.A. Glaser (New York: John Wiley and Sons, 1993), p. 117.

be in its original position, is now paired with a novel object. Only the foodwell beneath the novel object is baited with food. In order to learn how to succeed on the task, the monkey must be able to recognize the sample object as familiar and choose the *other* object. If the monkey displaces the "non-matching" object, he receives the food reward. The procedure is then repeated, using two new "trial-unique" objects. This process continues until the monkey has reliably learned to choose the novel object in each trial.

Mishkin and others have studied the performance of lesioned animals on the DNMS task in order to identify which MTL structures are implicated in recognition memory. The issue is controversial, but the various interlocutors in the debate are agreed on one general principle. An animal needs *either* an intact rostral MTL *or* an intact caudal MTL to succeed on the DNMS task. Severe impairments on DNMS are associated with conjoint bilateral damage to the rostral and caudal areas of the MTL. (The current debate centres on the issue of which elements within the rostral and caudal MTL are critical for DNMS. All are agreed that with respect to the rostral MTL, cortical areas, but not the amygdala, are implicated. With respect to the caudal MTL, some argue that both the hippocampus and the nearby cortical areas are implicated in recognition memory. Others (notably Mishkin) now argue that the hippocampus is not involved in recognition memory. He argues, that is, that conjoint damage to the cortical areas of the rostral and caudal areas of the MTL is the critical pathology in recognition memory impairments.) [86] Studies performed on rats and humans provide additional support for the view that the MTL is critical for object recognition. [87] According to Mishkin, destruction of the MTL precludes the possibility that

---

[86] Mishkin and his colleagues have recently written that "more recent experiments have shown, however, that the effect of that combined ablation on recognition memory may be attributable to the inclusion of the rhinal cortex, in the combined ablation. Indeed, selective rhinal cortex ablation, leaving intact the hippocampus and amygdala, also produces severe memory impairment in recognition memory." See Murray, E., Gaffan, D., & Mishkin, M. (1993). "Neural substrates of visual stimulus-stimulus association in rhesus monkeys," *Journal of Neuroscience*, 13 (10), 4549-4561.

[87] Aggleton, J.P., H.S. Blindt, and J.N.P. Rawlins, "Effects of Amygdaloid and Amygdaloid-Hippocampal Lesions on Object Recognition and Spatial Working Memory in Rats," *Behavioral Neuroscience* 103.5 (1989), pp. 962-74.

the neuronal ensemble (the active inferior temporal neurons) will be consolidated into a stable representation of the object. The neuronal ensemble will not, that is, be converted into a neuronal cell assembly.

The MTL is responsible for a second type of recognition memory: place recognition. Mishkin proposes that place recognition is subserved, for the most part, by the same neural processes that underlie object recognition. There are, however, two anatomical differences to keep in mind. (1) The visual information relevant to place recognition comes not from the ventral visual stream but from high-order spatial areas in the dorsal visual stream. (2) While object recognition requires only that *either* the rostral MTL *or* the caudal MTL be functional, place recognition requires the proper function of the caudal MTL, specifically, the hippocampus.

When the DNMS task is modified to test place recognition in rats, hippocampal rats are severely impaired. Dunnet has shown that hippocampal rats are severely impaired on a Delayed Non-Matching to Position task. [88] In this task, the rat is confronted with a device containing two levers, indistinguishable except for their positions. At first, only one lever is extended and the rat is trained to push it. After a delay, both levers are extended and the rat is rewarded only if he chooses the "novel" lever. Hippocampal rats showed delay dependent deficits on this Delayed Non-Matching to Position task.

According to Mishkin, then, there are two basic types of recognition memory - object recognition memory which depends on the availability of either the rostral MTL or the caudal MTL and place recognition memory which depends critically on the caudal MTL, particularly the hippocampus.

[88] Dunnett, S.B., "Comparative Effects of Cholinergic Drugs and Lesions of Nucleus Basalis or Fimbria-Fornix on Delayed Matching in Rats," *Psychopharmacology*, 87 (1985), pp. 357-363.

Associative Recall Memory

Once a stable cell assembly (a neural representation) is formed, it may be associated with other cell assemblies. In recall, the perception of a familiar stimulus triggers the memory of another, previously associated stimulus. This process is dependent on the availability of distinct representations of the two stimuli. According to Mishkin, such representations come in a variety of types. "The subsequent neural events with which the stored stimulus representations could become connected would likewise be stored representations, although these would be not only of other stimuli but also of places in the environment, or behavioural acts, or, finally, affective states." [89]

On Mishkin's view, the association of two cortical cell assemblies occurs via the activation of structures within the MTL. These MTL structures serve to link two independent representations together *indirectly*. This indirect association of cell assemblies allows for a certain degree of flexibility. If two cell assemblies were to be directly associated with one another, they could conceivably form one large cell assembly. "Compositionality", so to speak, would be lost. Because the association occurs via MTL structures, the two cell assemblies retain their freedom to associate independently with other assemblies.

> Linkage through limbic structures could help preserve the autonomy of individual cell assemblies and also provide the mechanism for associative flexibility. That is, each limbic structure could operate as a multiselection switch, triggering any of a potentially large number of associated cell assemblies, the ones selected depending on the particular instructions (i.e. other inputs) the limbic system is receiving currently. [90]

---

[89] Mishkin, M., B. Malamut, and J. Bachevalier, "Memories and Habits: Two Neural Systems," *The Neurobiology of Learning and Memory*, Eds. G. Lynch, J. L. McGaugh and N. M. Weinberger (New York: Guilford Press, 1984), p. 69.

[90] Mishkin, "Cerebral Memory Circuits," p. 119.

Tests on animals have shown that various sub-elements of the MTL system are specialized for particular types of associative memory. Mishkin has argued that (1) the amygdala is important in both cross-modal and object-reward associations, (2) the hippocampus is critical for object-place associations, and (3) the rhinal cortex is responsible for intra-modal associations.

The Non-Representational Memory System

Mishkin's interest in the non-representational memory system was sparked by unexpected results on an experiment he and his colleagues Malamut and Saunders reported in 1980. Mishkin was in the process of studying the performance of bilateral MTL monkeys on three types of memory tasks: Delayed Matching to Sample (DMS), object reward association, and "concurrent learning". It was expected that the monkeys would show severe impairments on all three tasks. To Mishkin's surprise, these MTL monkeys performed well on the task of concurrent learning. In order to appreciate why this result appears anomalous at first glance, the three tasks must be carefully compared.

The DMS, as its name suggests, is similar to the DNMS, except that the monkey is rewarded for choosing the familiar item, not the novel item. It thus tests the monkey's ability to remember which of two recently viewed items is familiar. The DMS paradigm uses "trial-unique" objects - the monkey encounters each pair of objects only once within a particular testing sequence. The task is thus highly focused. The only information relevant to the monkey's choice is the information concerning which of the two objects is familiar and which is not. If the monkey "stays" with the same baited stimulus, it "wins" - thus the task is described as a "win-stay" strategy task. (Compare this strategy with the "win-shift" strategy required in Delayed *Non* Matching to Sample.) The task thus draws upon the resources of the object recognition memory system, housed in the MTL - limbic circuitry. It is not surprising, therefore, that the monkeys without a functioning MTL memory circuit perform poorly on this task.

The object-reward association task tests the monkey's ability to remember which of two objects was baited a few seconds earlier. In this particular study, twenty pairs of trial unique objects were used. On any particular trial, the monkey was shown a pair of items, one at a time. Only one of these two items was baited. After a delay, the pair of objects was presented together, and the monkey receives a reward for choosing the previously baited object. The object-reward paradigm also promotes a win-stay strategy. Performance on this task requires that the object-reward pathway associated with the amygdala be intact. Thus monkeys with bilateral MTL lesions were, not surprisingly, severely impaired on this task.

The concurrent learning task appears to tap the same type of memory capacities as the DMS task and the object-reward task. In concurrent learning, the monkey is once again presented with a series of trial-unique object pairs and is "asked" to employ a win-stay strategy. Now the two items of the pair are presented simultaneously, one baited, one unbaited. The monkey is allowed to displace these objects and to see which of the two is baited. The pair of items is then taken away and a new pair is presented. Once again, only one of the items is baited. This procedure continues until all twenty pairs have been presented. At this point there has been no test of memory. The monkeys were merely exposed to the twenty pairs and allowed to discover which item in each pair was baited. The following day, however, the same twenty pairs were presented in the same serial order, with the same object in each pair is baited. The only difference between the presentation on the first and second days is that on the second day, the left-right positions of the objects within the pairs were varied pseudo-randomly. This procedure is repeated for a number of days. Normal monkeys learn which object within each pair is baited in about ten sessions. Much to the surprise of Mishkin and his colleagues, so did the MTL monkeys.

Note that the data from the DMS study indicates that these monkeys cannot recognize an object they have seen just seconds earlier. The data from the object-reward study shows that these animals cannot recall which of two items, seen just seconds earlier, was baited. Nevertheless, in concurrent learning paradigms, these monkeys are able to correctly choose the baited object in a pair after inter-trial intervals of 24 hours. The critical difference between concurrent learning, on the one hand, and object-reward and DMS on the other, is that only in the case of concurrent learning is the monkey exposed to the same object pairs repeatedly. The same object within a pair is baited every time. Learning can thus occur gradually over a number of trials. A memory system other than the MTL memory system must be responsible for the MTL monkeys' excellent performance on the concurrent learning task. According to Mishkin, this second memory system must be responsible for learning which occurs gradually over time and which involves the repeated association of stimuli and rewards. "It is a system for which the critical element is a stimulus-response repetition - exactly what is missing in delayed non-matching to sample (or delayed matching to sample)." [91] Mishkin argues that this second system is a "habit" memory system centred in the basal ganglia. On Mishkin's view, habit memory is a more basic type of memory than representational memory. "We call this kind of learning 'habit'. It is noncognitive: it is founded not on knowledge or even on memories (in the sense of independent mental entities) but on automatic connections between a stimulus and a response." [92] The representational system and the habit system thus "learn" in two very different ways. The representational system acquires information, information which it stores in discrete neural packages. The habit system, on the other hand, does not acquire "information" per se. Rather, it learns via the gradual, incremental modification of sensory systems and performance systems.

> The second learning process ... is viewed as involving instead the more gradual development of a connection between an unconditioned stimulus object and an approach response, as an automatic consequence of reinforcement by food. The

---

[91] . Mishkin and Appenzeller, "The Anatomy of Memory," p. 89. (My addition.)

[92] Mishkin and Appenzeller, "The Anatomy of Memory," p. 89.

product of this process is not cognitive information but a non-cognitive stimulus-response bond, that is, not a memory but a habit. Finally, what is stored in the habit formation system is not the neural representations of such items as objects, places, acts, emotions, and the learned connections between them but simply the changing probability that a given stimulus will evoke a specific response due to the reinforcement contingencies operating at that time. [93]

The representational system and habit system also differ with respect to their relationship to "consciousness". According to Mishkin, the end product of the MTL system is a neural representation that is available to consciousness, while the end product of the habit system is simply the neural modification of a sensory or performance system. These neural modifications are not themselves available to consciousness. "The product of habit learning is assumed to be a stimulus-response bond not accessible to conscious experience; it is only a tendency to respond in a particular way in a particular situation." [94] The habit system is activated by the appearance of a salient stimulus, but the animal need not consciously recognize the stimulus for it to be causally efficacious. "The characteristics of a stimulus are capable of triggering a response even though those characteristics need not be recognized - that is to say, no awareness is required." [95]

Mishkin argues that the basal ganglia would serve well as the "headquarters" for the habit system. The basal ganglia receive sensory input from all areas of the cortex and send projections into the frontal lobe, especially to those areas of the frontal lobes concerned with motor control and motor patterns. "Hence," argues Mishkin, "it is

---

[93] Mishkin, Malamut, and Bachevalier, "Memories and Habits: Two Neural Systems," p. 72.

[94] Petri, Herbert, and Mortimer Mishkin, "Behaviorism, Cognitivism, and the Neuropsychology of Memory," *American Scientist* 82 (1994), p. 36.

[95] Petri and Mishkin, "Behaviorism, Cognitivism, and the Neuropsychology of Memory," p. 36.

neuroanatomically suited to providing the relatively direct links between stimulus and action that are implicit in the notion of a habit." [96]

There are other reasons for thinking that the basal ganglia may be involved in habit memory, reasons having to do with the phylogenetic and ontogenetic development of the limbic system and the basal ganglia system. According to Mishkin, the behaviourist view of memory, or the S-R conception of memory, is particularly apt when speaking of less "highly developed" organisms. "There is one area, however, in which the behaviouristic positions will always remain unchallenged, and this is in its applicability across the entire phyletic scale. Even animals with the simplest nervous systems are capable of response adaptation; the acquisition of information or knowledge, by contrast, may require the evolution of a system analogous to the cortico-limbic-thalamic pathway of mammals." [97] The basal ganglia system is indeed more ancient, in evolutionary terms, than the limbic system. Furthermore, the basal ganglia system is functional, in any given animal, before the limbic memory system. While adult monkeys learn the DNMS task easily in fewer than 100 trials and perform well with various numbers of objects and delays of up to two minutes, infant monkeys do not perform well on the DNMS task until four to five months. It is not until they are approximately two years old that they perform as well as adults on this task. On tasks of concurrent learning, however, three month old monkeys perform just as well as adults, even with long lists and 24-hour inter-trial intervals. The young animal may rely exclusively on "habit" memory for a period of time until the limbic memory system is developed. We will look more closely at the structure and function of the basal ganglia at the end of this chapter.

The two memory systems provide the animal with different mnemonic capacities. Mishkin's analysis provides additional support for the views put forth by O'Keefe and Nadel. Mishkin argues that the representational memory system allows for fast, one-trial

---

[96] Mishkin and Appenzeller, "The Anatomy of Memory," p. 89.

learning via the storage of neural traces of stimuli, rewards, responses, etc. Furthermore, because it stores the information gleaned from experience in discrete neural representations, the MTL system is flexible. Representations can be combined and recombined indefinitely. The price to be paid for such speed and flexibility, however, is the loss of stability. Cell assemblies are relatively fragile. Interference during consolidation may weaken or destroy them. Even after consolidation, representational memories, if not rehearsed, may decay quite easily. Learning, as well as forgetting, occurs quite rapidly in the MTL system. The habit system, on the other hand, learns more slowly. Incremental neural changes to a performance system result, however, in a very stable form of learning. The numerous incremental changes involved in a habit are not readily undone. But the price for such stability, however, is a loss of flexibility. Thus the organism equipped with both a representational system and a habit system will have two distinct ways of encoding the lessons of experience - a fast, flexible, unstable system and a slow, inflexible, stable system, respectively.

The Multiple Memory Systems Theory of Cohen and Eichenbaum

Not all researchers describe the data on multiple memory systems in terms of the Tolman-Hull debate. In their book, *Memory, Amnesia, and the Hippocampal System*, Cohen and Eichenbaum offer a more "cognitivist" account of both hippocampal and nonhippocampal memory systems in that they describe both systems in representational terms. As we will see, however, they use the term "representational" far more loosely than many of their colleges. Furthermore, they draw a sharp distinction between the types of representations processed in the hippocampal and nonhippocampal systems. Only the representations in the hippocampal system are described in Fodorian terms. Their work is particularly valuable here because they take the time to distinguish carefully between

[97] Mishkin, Malamut, and Bachevalier, "Memories and Habits," p. 73.

processes that involve the manipulation of Fodorian neural representations and processes that may be described as "representational" in a far more general sense.

The organizing principle of the book is the distinction between hippocampal "declarative" memory and nonhippocampal "procedural" memory. The distinction was first formulated in these terms by Cohen in his Ph.D. dissertation, which was supervised by Squire. In 1980, Cohen and Squire published a brief account of the distinction in *Science*.

In 1980, research on multiple memory systems in humans was still in its infancy. Milner's work on H.M. had opened up a new field of research, but not much progress had been made. The general consensus, at that time, was that human hippocampal damage resulted in severe anterograde amnesia for facts and events but spared basic perceptual-motor skill acquisition. In their article in *Science*, Cohen and Squire argue that Milner's account of H.M.'s spared capacities was too limited, that H.M. was still capable not only of perceptual-motor skill learning, but of simple cognitive skill learning as well. "We now report that the class of preserved learning skills in amnesics extends beyond perceptual-motor tasks. Amnesic patients were able to acquire a mirror-reading skill that minimized perceptual-motor involvement and were able to retain it for more than 3 months." [98] The mirror-reading task involves the acquisition of the ability to read triplets of low-frequency words printed in mirror-image form with increasingly speed and accuracy. In some trials, the subject is presented with novel words, in other trials, previously seen word triplets are repeated. Both normals and amnesics were able to learn this task, though normals were more proficient with the repeated triplets than were the amnesics. "The facilitory effect (of the repeated words) was smaller for the amnesic patients than for the control subjects, which illustrates that, despite learning the mirror-reading skills normally, amnesics were

---

[98] Cohen, N. J., and L. L. R. Squire, "Preserved Learning and Retention of Pattern-analyzing Skills in Amnesia: Dissociation of Knowing how and Knowing that," *Science* 210 (1980), p. 208.

poor at remembering which words they had read." [99] After the mirror-reading task study was completed, both the controls and the amnesics were given a test of recognition memory in which they were asked to judge whether or not particular words had been used in the mirror-reading task. The amnesics performed very poorly. When questioned about the mirror-reading study, "none of the amnesic patients reported that words had been repeated during the task, even though by the end of session 4 the set of repeated words had been presented 20 times." [100]

There is a striking contrast between the amnesics' ability to acquire the mirror-reading skill and their inability to recognize the words used in the training sessions. Cohen and Squire argue that amnesics retain the ability to acquire both perceptual-motor skills *and* cognitive skills. They group these spared capacities together under the heading of "procedural" memory. Procedural memory, which is spared in amnesia, can then be compared with "declarative" memory, which is lost in amnesia.

Thirteen years later, the declarative-procedural distinction is used to provide the general framework of the theory Cohen and Eichenbaum present in *Memory, Amnesia, and the Hippocampal System.* In an introductory section titled, "A Note about Terminology: History and Baggage," Cohen and Eichenbaum note that "the choice of the terms ("declarative" and "procedural") turns out to be unfortunate, because it carries some unwelcome baggage from the meanings these terms have acquired from AI and from common language." [101] In order to avoid misunderstandings, therefore, it is important to keep in mind that "procedural memory" and "declarative memory" are technical terms for Eichenbaum and Cohen. The distinction is *not* equivalent to an early AI distinct between

---

[99] Cohen and Squire, "Preserved Learning and Retention of Pattern-analyzing Skills in Amnesia," p. 208.

[100] Cohen and Squire, "Preserved Learning and Retention of Pattern-analyzing Skills in Amnesia," p. 209.

[101] Cohen, Neal J., and Howard Eichenbaum, *Memory, Amnesia, and the Hippocampal System* (Cambridge, Mass.: MIT Press, 1993), pp. 89-90.

"knowledge represented in the form of routines or programs and ... knowledge represented as data structure." [102] Furthermore, a memory need not be "declarable" to be "declarative". Non-human animals have declarative memories and humans have declarative memories that cannot be verbalized. Cohen and Eichenbaum also note that "procedural memory should *not* be equated with memory for procedures." [103] With these caveats in mind, we can now turn to Cohen and Eichenbaum's account of the differences between hippocampal declarative memory and nonhippocampal procedural memory.

Throughout the dissertation, I have invoked the distinction between representational processes and non-representational processes, at both the psychological and neuroscientific levels. In order to accommodate the terminology used by Cohen and Eichenbaum, however, I will have to change course a bit. They use the term "representational" in a very expansive manner, as can be seen from the following comment they make in the introduction to their book.

> In neuroscience, research on learning and memory had been conducted largely within the context of plasticity. The goal of this work has been to characterize the way in which the brain is itself modified by - and thereby supports a lasting representation of - experience. [104]

If all non-transient neural modifications wrought by experience constitute "representations", there would be no sense in making a distinction between representational processing and non-representational processing. On this view of what counts as representational processing, all mnemonic processing is representational, by fiat. In the context of the argument I am developing, this use of the term is unfortunate. It doesn't much matter, however, in that while both declarative and procedural processing are representational, on the view of Cohen and Eichenbaum, the distinctions they draw

---

102 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 90.

103 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 90.

between these two types of processing are still quite germane. In the end, the problem concerning their use of the term "representational" is terminological and superficial.

Cohen and Eichenbaum argue that the two types of memory systems can be distinguished in terms of (1) "the type of representations they support" and (2) the (retrieval and encoding) processes that permit these representations to be built, stored, and accessed." [104] Their goal is to provide a detailed account of declarative and procedural memory, one which is founded upon but goes beyond the original account provided by Cohen and Squire.

The most important quality of declarative memory is that it is *relational.* A relational memory system is one that is capable of relating two or more distinct neural representations. The hippocampus is in an excellent position to mediate relational memory processes. It is located near the centre of the brain, both literally and figuratively. As we will now see, the account of relational processing provided by Cohen and Eichenbaum is quite similar to the account provided by Mishkin, though it goes beyond Mishkin's account in a few important respects. It therefore deserves a rehearsal here.

One way of understanding the relational property of hippocampal memory is to trace the inputs and outputs of the hippocampus. The hippocampus receives highly processed information from all the major cortical processing stations in the brain. Its output travels along reciprocal pathways back into these cortical processing centres. High-order visual, motor, proprioceptive, olfactory, auditory, and somatosensory centres have access to each other via indirect links through the hippocampus.

As a result of its location and pattern of cortical links, the hippocampus is well suited to the sort of processing involved in the "learning of relations among perceptually

---

[104] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 2.

[105] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 52.

distinct objects." [106] This claim is consistent with the fact that amnesics are impaired at learning new associations between words, names and phone numbers, stimulus objects, etc. The hippocampal memory system is also implicated, according to Cohen and Eichenbaum, in more abstract forms of relational mnemonic processing.

> The relational dimensions captured by or represented in declarative networks are at all levels of abstraction: They include strictly sensory relationships, such as relative size, colour, texture, shape, etc., more integrative perceptual relationships such as relative positions or objects in space and time; and higher-order relationships based on the temporal contiguity of objects and events. These latter involve both new causal coincidences, that is, the accidental conjunctions of objects or events, and causal contiguities, including the learned consequence of particular stimuli or events. [107]

The relational properties of hippocampal processing may be understood in anatomical terms. The key to appreciating hippocampal memory and its various uses is an understanding of the way in which the hippocampus links discrete neural representations. To take a relatively simple example, consider the way in which the hippocampus collates the sight and smell of a rose. As we have already seen, the final visual representation of the rose would be assembled in area TE of the ventral visual stream. In order for this visual representation to be consolidated into a long-term memory, there must be, for a certain critical period of time, active reciprocal connections between the representation in TE and the hippocampal memory system. After consolidation, the representation is "stored", so to speak, in area TE. A similar type of processing occurs in the olfactory system, which ultimately "stores" a representation of the rose's smell in the basal area of the frontal lobe. If the sight and the smell of the rose are experienced simultaneously, both neural representations will be active at the same time. The hippocampus can track such co-

---

[106] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 68.

activated neural circuits and link them together by "maintaining the coherence" of the cortical co-activations.[108]

> By virtue of the connections of these various cortical networks to the hippocampal system, the co-activations representing the outcomes of the difference analyses converges on hippocampal system networks. The hippocampal system is thereby in a position for mediating the representation of these conjunctions or co-activations.[109]

This sort of relational processing can be utilized to link the various aspects of a single subject, as well as the various components of a scene or event. If at some later time, one aspect of the object or event is re-experienced, the hippocampal system is able to retrieve the other aspects via the resurrecting of the original "co-activated" circuits. The smell of the rose reminds us of its visual form. The sight of the deserted street recalls to mind the lively scenes experienced there some other time.

Relational processing has two important and closely related attributes: (1) it is flexible and (2) it is promiscuous. Flexible processing, as we have seen in our discussion of Mishkin, involves the capacity to store and retrieve all manner of relational representations. Since individual representations remain distinct, they can enter flexibly into any number of associations; each representation is "promiscuously accessible to - and can be activated by - all manner of processes and processing modules; and can be manipulated and flexibly expressed in any number of novel situations, independent of the circumstances in which the information was initially acquired."[110]

---

107 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 62.

108 Squire, Larry, Neal Cohen, and Lynn Nadel, "The Medial Temporal Region and Memory Consolidation: A New Hypothesis," *Memory Consolidation*, Eds. H. Eingartner and E. Parker (Hillsdale, NJ: Erlbaum, 1984), Cited in Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 69.

109 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 69.

110 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 62.

One of the most important characteristics of relational hippocampal processing is the fact that it observes the principle of compositionality. Individual neural representations are not melded together: they remain distinct. The neural circuits subserving two distinct representations may be co-activated, but because the hippocampus is interpolated "between" the representations, each retains its own integrity.

> The important point is that a relational representation of scenes, events, and complex ideas in our proposed declarative memory system is *not* as blends, or configurals, and does *not* involve conjoining of the multiple individual stimuli, or constituent pieces of knowledge into unified knowledge structures. Rather, a relational representation preserves the status of the constituents of the larger structure while still permitting the larger structure to be appreciated. [111]

Compositionality is a critical feature of higher-order hippocampal processing. In the absence of the compositionality feature, relational processing would quickly bog down; the virtues of flexibility and promiscuity would be lost. Cohen and Eichenbaum note that compositionality engenders yet another critical feature of hippocampal processing: generativity. As is emphasized by O'Keefe and Nadel, the hippocampal system is responsible for generating novel output, for detecting new detour pathways that have never been simultaneously experienced. Cohen and Eichenbaum see the generativity of hippocampal processing as an important contribution to the generativity of even more abstract processes such as complex thought and language. Citing Fodor himself, they offer the following reading of the generative powers of hippocampal processing.

> This property of compositionality is critical to the generativity of language and, presumably, of thinking. This is because language and thought depend on the fact that some constituent pieces of knowledge can be used to construct any number of larger structures (complex ideas, sentences, and even books), while still

---

111 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 64.

maintaining their own identity, an identity that must remain *systematic* across the different larger structures they help to form.[112]

Hippocampal memory is thus relational, flexible, promiscuous, compositional, and generative. These properties can now be compared to the properties of the nonhippocampal system.

In hippocampal learning, various cortical processing centres receive sensory input, process it, and send the product of their labours into the hippocampal memory system for the purposes of both consolidation and relational processing with the outputs of other cortical systems. Nonhippocampal processing, on the other hand, is more localized. It occurs *within* individual cortical and sub-cortical processing stations; there is no representational output that could be cast free, so to speak, to travel to other areas of the brain. The output of nonhippocampal processing is performance.

The mechanisms of nonhippocampal processing are relatively basic. They consist, as we have seen, of the fine-tuning of neural circuits involved in the performance of ritualistic actions or habits. In addition, certain cognitive skills that are acquired via habitual practice, such as mirror reading, are also subserved by nonhippocampal memory system. Nonhippocampal processing does not involve the creation of distinct representations; it consists of the neural modification and fine-tuning of performance systems.

Eichenbaum and Cohen catalogue a number of nonhippocampal or "procedural" memory systems in the brain, systems involving both high-level cortical structures and low-level sub-cortical stations. We will look at one example of each.

The somatosensory cortex, in both humans and monkeys, is a horizontal strip of cortical cells in the parietal lobe. This area of the brain contains an inverted topographical

---

112 Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 64. (Emphasis added.)

map of the body's surface area. Touch the foot and neurons at the top of the map are activated; touch the head and cells at the bottom of the map are activated. Merzenich has shown that this somatosensory map is "plastic"; it undergoes modification as the result of the input patterns it receives due to experience. He notes that if the monkey is habitually exposed to a task in which the middle finger, but not the other fingers, must press a lever repeatedly, the portion of the map devoted to the middle finger will grow in size.[113]

Similar experience-dependent modifications to procedural systems have been documented in the auditory, visual, and motor cortices. "Thus cortical rearrangements may be a general consequence of experiences that support learning and memory."[114] Note that these systems are "representational" in a weak sense: configurations of neurons correspond to or represent certain features of the body of the world. These "representations", however, remain bound to the systems in which they are created.

More primitive, sub-cortical processing systems can also be mediated through experience. Eichenbaum and Cohen's discussion of the nictitating membrane response in rabbits is illustrative. Recall that nonhippocampal memory is associated with basic S-R learning, learning which often involves the cross-modal association of two different cues. This type of neural association is more basic than the types of associations supported by the hippocampus. Most importantly, it does not involve the *flexible* association of distinct representations. If a rabbit is exposed to a puff of air directed at its eyes, the nictitating membranes which protect the eyes close reflexively. If the rabbit is subject to a training paradigm in which the air puff is reliably preceded by an auditory cue, the rabbit will eventually exhibit a conditioned "blinking" response to the auditory cue. Note that in this example of procedural learning, cross-modal processing is required; the auditory cue must be associated with the motoric blinking response. In this case, however, it is the

---

[113] Merzenich, M.M., et al., "Somatosensory Cortical Map changes Following Digit Amputation in Adult Monkeys," *Journal of Comparative Neurology* 224 (1984), pp. 591-605.

cerebellum, not the hippocampus, that effects the association. While we may elect to speak of the association of two representations, one auditory, one motor, Cohen and Eichenbaum emphasize that these types of procedural representations differ profoundly from hippocampally generated representations.

> Thus, here, tuning of cerebellar networks can bring about representation of simple and highly specific *associations*. But by no means are these the same as the relational representations mediated by the hippocampal-dependent declarative system. ... (Hippocampal) representations are fundamentally flexible and promiscuous...., they can be accessed under a wide range of testing circumstances and once accessed can be used completely flexibly. Neither of these is in any way true of the procedural representation of simple associations. [115]

Procedural memory processes are system-dependent and thus inflexible. The so-called "representations" of these systems are non-transportable and are, as Cohen and Eichenbaum put it, "fundamentally individual". The expression of procedural learning is limited to contexts that are the same as or similar to the contexts in which the procedural skill was acquired. "(Procedural) memory can be expressed only *inflexibly*, only in a repetition of the original processing situation." [116] The inflexibility of procedural learning is, of course, relative. While procedural learning is gradual and is expressed in habitual actions and responses, a skill can be demonstrated in a variety of "similar" contexts. As Nadel points out, if you know how to play squash, you can play with approximately the same degree of finesse on any court. [117]

---

[114] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 76.

[115] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, pp. 81-82.

[116] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 73.

[117] Nadel, "Multiple Memory Systems: What and Why, an Update," p. 53.

The general principles of both hippocampal and nonhippocampal memory are now clear. The hippocampal system operates over distinct neural representations which are stored in various cortical areas throughout the brain. Hippocampal processing is strikingly similar to Fodorian processing. Both are flexible, compositional, promiscuous, and generative. Nonhippocampal systems, on the other hand, involve the gradual neural modifications of performance systems and are relatively inflexible. The do not involve the flexible manipulation of distinct representations. Before closing, we will take a brief look at the nonhippocampal system responsible for perceptual-motor skills - the basal ganglia system .

The Perceptual-Motor Skill System: The Basal Ganglia

The basal ganglia are a collection of sub-cortical nuclei that are implicated in the execution of complex motor sequences and in the acquisition of perceptual-motor skills. They receive input from all areas of the cerebral cortex and send projections, via the thalamus, back into the motor cortex and associated motor cortices (the premotor cortex and the supplementary motor cortex[118]), as well as into "prefrontal" areas of the frontal lobes.[119]

The basal ganglia consist of five sub-cortical nuclei: the caudate nucleus, the putamen, the globus pallidus, the subthalamic nucleus, and the substantia nigra. The globus pallidus and the substantia nigra each contain two functionally distinct parts. The globus pallidus is thus conventionally divided into an "external segment" and an "internal segment"; the substantia nigra into the "pars compacta" and the "pars reticulata".

---

118 The premotor cortex and supplementary motor cortex are "important for coordinating and planning of complex sequences of movement." Claude Ghez, "The Control of Movement," Kandel, Eric, James Schwartz, and Thomas Jessell, Eds. *Principles of Neural Science*, 3rd Ed. (New York: Elsevier, 1991), p. 539.

These nuclei are classified into three groups: "input nuclei" (which receive input from the cerebral cortex), "output nuclei" (which send projections back into the motor cortices and the prefrontal cortex), and the "intrinsic nuclei" (which receive projections from the input nuclei and send projections either back into the input nuclei or on to the output nuclei). [120] The input nuclei are the caudate nucleus and the putamen, collectively known as the "striatum". The intrinsic nuclei are the external segment of the globus pallidus, the subthalamic nuclei, and the substantia nigra pars compacta. The output nuclei are the internal segment of the globus pallidus and the substantia nigra pars reticulata.

INPUT NUCLEI	INTRINSIC NUCLEI	OUTPUT NUCLEI
Caudate Nucleus	Globus Pallidus - External Segment (GP - E)	Globus Pallidus - Internal Segment (GP - I)
Putamen	Substantia Nigra - Pars Compacta (SN - PC)	Substantia Nigra - Pars Reticulata (SN - PR)
(Striatum)	Subthalamic Nucleus	

Figure 5.8: The Nuclei of the Bsal Ganglia

There are two basal ganglia pathways involved in the control of complex movements: a "direct" pathway and an "indirect pathway". The direct pathways streams from various regions of the cerebral cortex into the striatum (the caudate nucleus and the putamen) and then proceeds through the output nuclei (GP - I and SN - PR), through the thalamus, and back into the motor areas of the frontal lobe. The indirect pathway is similar, except that there is a small "loop" in the pathway involving the external segment of

119 The term "prefrontal cortex" is misleading. The prefrontal cortex is *in* the frontal lobe, anterior to the motor cortices. It is thought to be involved in a number of cognitive functions, including the planning of goal-directed behaviour.

120 John Martin, *Neuroanatomy Text and Atlas* (New York: Elsevier, 1989), pp. 268-69.

the globus pallidus and the subthalamic nucleus. In both pathways, the activity of the striatum is influenced by input from the substantia nigra pars compacta.

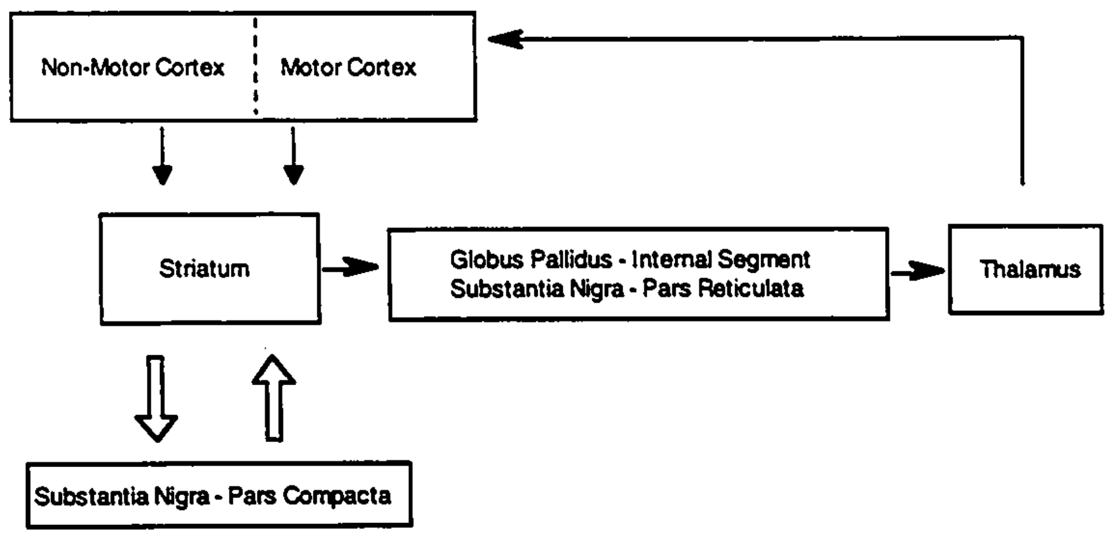

Figure 5.9: The Direct Motor Pathway of the Basal Ganglia

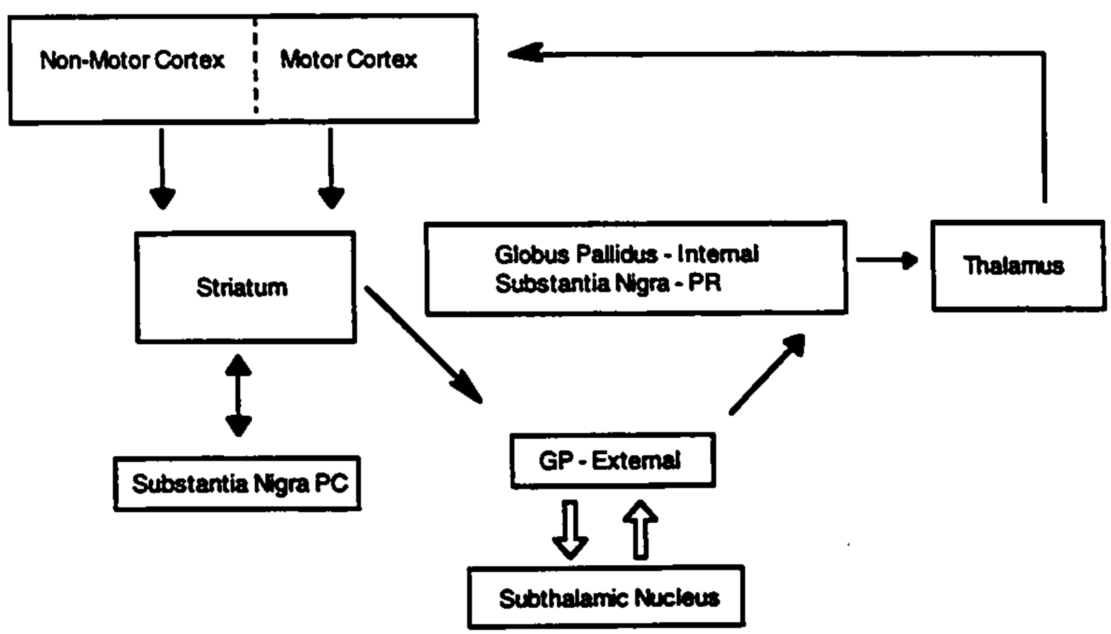

Figure 5.10: The Indirect Motor Pathway of the Basal Ganglia

The ability to control complex movements is dependent upon the proper functioning of both the direct and indirect pathways. The two pathways work together, so to speak, to modulate activity in the motor cortices. While activation of the direct pathway excites the neurons of the motor cortices, the activation of the indirect pathway inhibits them. When both pathways are functioning properly, complex movements are well-organized and fluid. Both the direct and indirect pathways depend on input from the substantia nigra pars compacta (SN - PC). The SN - PC contains cells which deliver an important neurotransmitter, dopamine, into the striatum. If dopamine levels in the striatum fall, voluntary movement becomes increasingly difficult.

Damage to the basal ganglia is associated with a number of motor deficits. In Parkinson's disease, for example, there is widespread cell death in the SN - PC. This leads to a serious decrease in the dopamine levels in the striatum. Parkinson's patients suffer from akinesia (difficulty in making voluntary movements), bradykinesia (slowness in movements), "cogwheel" rigidity, and resting tremors. Huntington's disease, which is associated with damage to the striatum and the frontal cortex, is marked by chorea (uncontrollable jerky movements) and progressive dementia. Damage to the subthalamic nucleus is associated with "hemiballism", a condition in which the patient makes sudden "ballistic" movements of the limbs on the side of the body contralateral to the lesion.

As we have seen, Mishkin argues that the basal ganglia are well-situated to serve as the headquarters of a memory system responsible for the acquisition of perceptual-motor skills. We are now in a better position to appreciate his proposal concerning the neural mechanisms involved in the acquisition of a "habit" or "skill". Recall that for Mishkin, skill acquisition involves the neural association of a stimulus and a response. Let's consider an example in which a motor response becomes associated with a particular *visual* stimulus. Visual information from the occipitotemporal visual pathway is fed into the striatum of the basal ganglia. If the organisms makes a motoric response to the stimuli, information from the motor cortices (and from the parietal lobe) also flows into the striatum. On Mishkin's view, if the behaviour is "successful" or "rewarded", the neurotransmitter dopamine will play the role of a neural "reinforcer" and will strengthen the connections between the co-activated visual and motor pathways. If the stimuli and response are associated a sufficient number of times, the perceptual-motor performance systems are gradually fine-tuned. The motoric response becomes increasingly refined and automatic.

> A stimulus-elicited neuronal ensemble in the occipitotemporal pathway would activate an array of synapses in the caudal neostriatum, at the same time that the neostriatal-to-premotor circuit was weakly assisting the ... motor system in evoking a previously learned response. On each occasion that a particular S-R conjunction was followed by reward-elicited activation of the dopaminergic system (i.e. reinforcement), the currently active array of visuo-neostriatal synapses would be strengthened incrementally by cellular mechanisms analogous to those described earlier for the formation of cell assemblies. Conversely, each time a particular S-R conjunction failed to be followed by a dopaminergic activation, due to the absence of reward (i.e. extinction), the currently active array of synapses would be greatly weakened. ... In time, as a consequence of the repeated reinforcement or extinction of particular S-R conjunctions, the learned motor response would be consistently evoked in the presence of certain stimuli and consistently suppressed in the presence of others. [121]

John Gabrieli argues that Mishkin's hypothesis is supported from data on the spared capacities of patients suffering from Parkinson's Disease and Huntington's Disease. [122] He notes, for example, that both Parkinson's patients and Huntington's

---

[121] Mishkin, "Cerebral memory circuits," p. 121.

[122] Gabrielli, John. "Contribution of the Basal Ganglia to Skill Learning and Working Memory in Humans," Houk, James, Joel David, and David Beiser, Eds., *Models of Information Processing in the Basal Ganglia* (Cambridge, Mass.: MIT Press, 1995), p. 278.

patients are unimpaired on tests of recognition memory but severely impaired on tests of perceptual-motor skill acquisition. Some of the impairment on perceptual-motor skill learning tasks is due to a primary deficit in motor coordination. Gabrielli argues, however, that these deficits in skill acquisition cannot be accounted for solely in terms of motor deficits. These patients are impaired in the *learning and retention* of skills. He notes that Parkinson's patients are impaired on the rotary pursuit task and that Huntington's patients are impaired on both the rotary pursuit task and on the mirror reading task. These are tasks on which H.M., despite his profound amnesia, demonstrates normal learning.

> Patients with basal ganglia diseases have shown impairment on skill-learning tasks despite having better recall for their skill-learning experience and better recognition of the testing materials than amnesic patients who learn these tasks normally. Thus, the skill learning deficit in HD (Huntington's Disease) (and) PD (Parkinson's Disease) ... patients cannot be accounted for by a generalized memory problem. Rather, these findings point toward a specific contribution of the basal ganglia to skill-learning in humans. [123]

Until quite recently, the basal ganglia were thought to be exclusively devoted to the control of complex movements. This view of the basal ganglia is currently being revised. The connections between the basal ganglia and the non-motoric prefrontal cortex are being more closely scrutinized. Gabrielli argues that the connections between the striatum and the frontal lobe suggest that the basal ganglia may be involved in "working memory", a phenomena that is typically associated with prefrontal cortex. He defines "working memory" as "a multicomponent psychological system that supports the temporary storage, manipulation, and transformation of information needed to perform cognitive tasks." [124] If a subject is required, for example, to remember the temporal order of a sequence of events, she must have the capacity to keep several items of information in "working memory" until

[123] Gabrielli, " Contribution of the Basal Ganglia to Skill Learning," pp. 278-79.

the task is complete. James Houk, among others, argues that the basal ganglia are critically important for the task of placing information gleaned from experience into working memory. [125] Gabrielli summarizes many current hypotheses regarding basal ganglia function in the following passage.

> One may speculate that an essential contribution of the basal ganglia to human learning and memory is to support the speeded execution of component processes of a multistep cognitive or motor action. When that support is lost due to a basal ganglia disease, components are executed too slowly to accomplish either the smooth sequence of movements that characterizes perceptual-motor skill or the rapid sequencing of thoughts that characterizes flexible working memory capacities." [126]

The function of the basal ganglia is thus not limited to the control of complex movements. The relationship between the basal ganglia and both perceptual-motor skill learning and working memory is currently under investigation. The next few years will undoubtedly provide new insights. I suspect that this research will prove to be philosophically important. For now, it is enough to note that perceptual-motor skill learning is indeed associated with a nonhippocampal memory system, one which operates via gradual adjustments to perceptual and motor performance systems. The basal ganglia appear to provide the neural mechanisms required for the orchestration of these gradual modifications of performance systems.

---

124 Gabrielli, " Contribution of the Basal Ganglia to Skill Learning," p. 283.

125 Houk, James, "Information Processing in Modular Circuits Linking Basal Ganglia and Cerebral Cortex," James Houk, Joel David, and David Beiser, Eds., *Models of Information Processing in the Bsal Ganglia* (Cambridge, Mass.: MIT Press, 1995), pp. 3-9.

126 Gabrielli, " Contribution of the Basal Ganglia to Skill Learning," p. 190.

Conclusion

In *Memory, Amnesia, and the Hippocampal System*, Cohen and Eichenbaum laud the growing trend toward cooperation among cognitive scientists and neuroscientists. One benefit of such cooperation, on their view, is that a knowledge of neural processing principles will prevent cognitive scientists from proposing biologically implausible cognitive mechanisms. It is important "to ascertain whether or not the brain systems exist that actually possess the machinery and anatomical connections needed to perform the computations proposed." [127]

In his Representational Theory of the Mind, Fodor offers an account of cognitive processing which is, as it turns out, consistent with the principles of a singularly important neural system - the hippocampal system. Whether or not the details of Fodor's theory survive close scrutiny is not at issue here. It is clear, however, that the neural mechanisms needed to support a flexible and generative representational system do indeed exist. Furthermore, neuroscientists agree that this system may be responsible for just the sort of cognitive processing described in Fodor's RTM.

Fodor's arguments for applying the RTM to perceptual motor skills are, however, quite weak. Furthermore, we have good neuroscientific reasons for supposing that perceptual-motor skill acquisition and implementation have a distinct etiology, one which does not involve the manipulation of Fodorian representations. It is worth noting that the RTM's scope may be further limited by future research on the cognitive capacities of the basal ganglia systems. Not all habits are perceptual-motor. If a cognitive skill is acquired over time, through repetitive practice, and is deployed only in particular contexts that resemble the context in which the skill was acquired, it is reasonable to hypothesize that such a skill is supported by neural systems other than the hippocampal system. A Fodorian analysis of the skill would therefore be inappropriate. After all, Fodor himself cautions that

[127] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 6.

the RTM should be applied to a cognitive system only if the behaviour generated by the system is genuinely caused by the manipulation of distinct representations. The moral of our story, then, is that philosophers of mind and cognitive scientists will find that neuroscientific research can profitably be used in the development and refinement of our philosophical and psychological theories.

In this chapter, I have shown that the processing mechanisms of nonhippocampal memory systems are not plausibly characterized in terms of the manipulation of neurally realized Fodorian representations. Given that perceptual-motor skills are dependent on nonhippocampal memory circuits, we can conclude that the manipulation of Fodorian representations is not an element of the etiology of perceptual-motor skills.

Conclusion

The central claim of this dissertation is that perceptual-motor skills are not caused by representational mental processes. One might think that I am belaboring an obvious and commonsensical intuition. We should remember, however, that the issue at hand is the scope of cognitivism. Whether or not perceptual-motor skills can be characterized in representational terms is a critical source of conflict between neo-Heideggerians and cognitivists. What I have shown is that neuroscience vindicates the neo-Heideggerians. If common sense is vindicated at the same time, so much the better. The point is that the scope of cognitivism is not as extensive as its proponents would have us believe. By way of conclusion, I will offer a chapter-by-chapter summary of the thesis, a review of its central argument, and a few speculative comments on how my work here might be developed into a more expansive research project.

The distinction between high-order rational thought and low-order perceptual-motor skills has played an important role in a number of debates in psychology, philosophy, and neuroscience. In the first part of the thesis, Chapters One and Two, I outlined how this distinction figures in the debate between cognitivists and phenomenologists. Descartes' representational theory of the mind elicited the critical responses of phenomenologists, such as Merleau-Ponty and Heidegger, who argue that Cartesian representationalism is problematic because it overlooks the primordial skills of the engaged agent. These skills involve the capacity to discern the salient or "value-laden" properties of objects and events and the ability to deploy perceptual-motor skills in maneuvering through the environment and in manipulating objects encountered within it. The same debate is currently being replayed in conversations between cognitivists like Fodor and neo-Heideggerian phenomenologists such as Dreyfus and Taylor. At issue is the scope of representational theories of the mind. Cognitivists tend toward theoretical imperialism and argue that all types of intelligent behaviour can be explained within the framework of representationalism. Contemporary phenomenologists, on the other hand, argue that certain types of engagement skills cannot be explained in representational terms.

Dreyfus offers two types of arguments against Fodor. First, he claims, following Heidegger, that a cogent phenomenological demonstration of the nature of engagement skills will reveal the implausibility of the cognitivist claim that all intelligent behaviour is representational. Implausibility is, however, in the eyes of the beholder. Cognitivists privilege "scientific" or "causal" accounts of behaviour and tend to remain unmoved by phenomenological persuasion. Since Dreyfus' phenomenological account of skilled perceptual-motor comportment is not couched in causal terms, it is not always "appreciated" by its intended audience - cognitive scientists. This is why a neuroscientific account of the non-representational nature of perceptual-motor skills is of great strategic value.

Dreyfus' second argument is based on the fact that the theory of skill-pragmatism successfully predicts the pattern of failures manifest in AI research. As it turns out, it is precisely those skills which neo-Heideggerians label "non-representational" that have proven to be difficult or impossible to simulate computationally. Dreyfus' analysis of the literature on AI is an important indication that the scope of cognitivism may not be as far-reaching as its proponents suppose.

Part One of the thesis summarizes the philosophical differences between Fodor and Dreyfus and emphasizes the role that perceptual-motor skills play in the theories of each. In Part Two, I argue that Fodor's analysis of perceptual-motor skills is inconsistent with contemporary neuroscientific accounts of multiple memory systems. Chapters Three through Five are thus devoted to various aspects of the literature on multiple memory systems.

In Chapter Three I investigate Bergson's distinction between what he calls "representational memory" and "habit memory". On Bergson's view, representational memory stores the events of a lifetime in a series of distinct representations. It operates "indifferently" and is not sensitive to the valence of particular objects and events. These representations are immaterial and reside in consciousness. Habit memory is a corporeal phenomenon. It is always practical, always directed toward action. Experience is "stored" in an increasingly diverse set of adaptive motor-mechanisms which unfold "automatically" in appropriate situations. Furthermore, in conjunction with practical perception, habit memory is sensitive to the adaptive significance of objects in the environment. Bergson's account of memory types is consistent with the principles of skill-pragmatism, in that it portrays perceptual-motor skills in non-representational terms. Furthermore, his account of memory types turns out to be consistent with contemporary theories of multiple memory systems. As a result, Bergson's work may be read as an important link between neo-Heideggerian phenomenology and contemporary neuroscience.

In Chapter Four, I review two important historical events in the literature on multiple memory systems: the Tolman - Hull debate and Milner's early work on hippocampal amnesics. Tolman and Hull offer competing accounts of animal learning. Hull focuses on low-order perceptual-motor habits and argues that they, as well as all other types of intelligent behaviour, can be explained within the framework of S-R theory. Tolman focuses on more complex forms of "cognitive map" learning and argues that they cannot be explained unless we permit ourselves to posit internal representational states. The Tolman - Hull debate is important because it provides the theoretical framework for a number of contemporary theories of multiple memory systems. Scientists now argue that, in a sense, both Tolman and Hull were right; there are enough neural circuits to go around, enough to keep both cognitivists and behaviourists happy.

The contemporary theories of multiple memory systems which are said to vindicate both Tolman and Hull are all dependent on the ground-breaking work of Milner. It was

Milner who demonstrated that the hippocampus is a critical structure in mnemonic processing. Furthermore, she demonstrated that hippocampal amnesics retain the capacity to acquire and deploy new perceptual-motor skills. Her work thus provides support for a crucial premise in my argument: perceptual-motor skills are not dependent on the hippocampal memory system. In distinguishing between the lost and spared capacities of human amnesics, Milner inaugurated the new field of multiple memory systems research.

In Chapter Five, I survey several contemporary theories of multiple memory systems and argue that, in combination with Milner's work, they provide the necessary ingredients for a forceful argument against Fodor's claim that perceptual-motor skills can be explained within the framework of his RTM. Here's the argument.

P1. The hippocampal memory system operates over neurally realized Fodorian representations.

P2. Nonhippocampal memory systems are not plausibly characterized in terms of the processing of neurally realized Fodorian representations.

P3. Both humans and non-human animals retain the capacity to acquire, refine, and deploy perceptual-motor skills in the absence of a functional hippocampal memory system.

P4. Learned behaviours are dependent on either the hippocampal system or the nonhippocampal systems or on both.

P5. Perceptual-motor skills are learned behaviours.

C. Perceptual-motor skills are dependent on neural memory systems that are not plausibly characterized in terms of the processing of neurally realized Fodorian representations.

The first premise is supported by the work of several contemporary neuroscientists, but it is Cohen and Eichenbaum who make the point most clearly. They argue that the hippocampal system operates over relational representations which have Fodorian properties. Each representation is discrete in that it retains its own integrity even when combined with other representations. Through the mechanisms of the hippocampus, individual representations throughout the brain may enter flexibly and promiscuously into novel combinations with other representations. The hippocampus thus serves as a grand central station: it couples and uncouples representations from all cortical processing centres in the brain. Because each representation retains its own integrity during these coupling processes, second-order representations may be said to have internal structure and transportable parts - the two properties associated with "Fodorian" representations. As a result, we may conclude that the hippocampal system traffics in neurally realized Fodorian representations.

The second premise is supported by the theories of Hirsh, O'Keefe and Nadel, Mishkin, and Cohen and Eichenbaum. In all cases, these neuroscientists argue that the mechanisms of nonhippocampal memory systems involve the gradual adjustment or fine-tuning of performance systems, including both perceptual and motor systems. The Fodorian-style representations implicated in hippocampal processing are not implicated in the mechanisms of the nonhippocampal systems.

The third premise is supported by Milner's work on amnesia, in which she demonstrates that perceptual-motor skills can be acquired and retained in the absence of a functional hippocampal system. The fourth premise is the claim that there are essentially two distinct types of mnemonic processing: hippocampal processing and nonhippocampal processing. If a behaviour is learned, it must be dependent on one or both of these processing types. The fifth premise is straightforward: perceptual-motor skills are paradigmatic learned behaviours. Since P4 indicates that learned behaviours are dependent on either the hippocampal system and/or the nonhippocampal systems, and since P3 eliminates the possibility that perceptual-motor skills are dependent on the hippocampal system, we may conclude that perceptual-motor skills are dependent on nonhippocampal memory systems. Given P2, we can conclude further that perceptual-motor skills are dependent on neural memory systems that are not plausibly characterized in terms of the processing of Fodorian representations.

This argument allows us to consider neo-Heideggerian phenomenology from a new perspective. As Heidegger and his philosophical descendants emphasize, many of our everyday coping skills are dependent on a familiarity with the world, a familiarity which cannot be characterized in representational or propositional terms. As Dreyfus notes, "our familiarity does not consist in a vast body of rules and facts, but rather consists of dispositions to respond to situations in appropriate ways." [1] Taylor supports Dreyfus' view, noting that "Dreyfus tirelessly points out how implausible it is to understand certain of our intelligent performances in terms of a formal calculus, including our most common everyday ones, such as making our way around rooms, streets, and gardens, picking up and manipulating the objects we use, and so on." [2] The sense of familiarity that underpins such everyday coping skills is the product of a certain type of experience. More precisely, it is a product of the neural changes that experience brings about in nonhippocampal systems. We can and should, of course, continue to investigate the nature of engaged coping skills in non-neuroscientific terms. There are many aspects of engaged coping that will not submit to a neuroscientific analysis. Our claim that these skills are not representational can be supported, however, by those neuroscientific theories which indicate that nonhippocampal memory is *non-representational.*

Perceptual-motor skills figure naturally in the accounts of both phenomenologists and neuroscientists. There are two other aspects of the neo-Heideggerian project that deserve mention here, despite the fact that they are difficult to translate into neuroscientific terms. My comments here will thus be admittedly speculative. The first is the phenomenon of "significance relationships" between engaged agents and the world; the second is the issue of consciousness.

---

1 Dreyfus, *A Commentary*, p. 117.

Neo-Heideggerian arguments against cognitivism typically feature a claim about the inability of cognitivists to account for "significance relationships", relationships in which the agent is sensitive to the *salience* or *value* of objects or events. For the most part, neuroscience is not the appropriate tool for gaining an understanding of significance relationships. *Au fond*, however, our most basic significance relationships are generated by the necessity of appreciating the salience of objects with biological survival value. There is a general tendency among multiple memory systems theorists to depict the hippocampal system as being relatively "disinterested" in the valence of environmental objects. Tolman, O'Keefe and Nadel, for example, emphasize that cognitive mapping behaviour is motivated by general curiosity, not by the need to service pressing biological needs. Nonhippocampal systems are, in general, depicted as being more sensitive to the immediate reward value of stimuli. They are characterized as being more "egocentric" and less "objective" than hippocampal systems. We might hazard a guess, then, that the capacity to appreciate the very basic significance features of objects in the environment is the province of nonhippocampal systems.

With respect to the issue of consciousness, neo-Heideggerians often refer to the "pre-reflective" nature of coping skills. In engaged coping, agents do not pause to represent to themselves explicit goals or courses of action. Their behaviour is not deliberative. According to Dreyfus, much of our everyday coping behaviour goes through without conscious deliberation or awareness. One of the distinguishing features of nonhippocampal systems, as Mishkin, Eichenbaum, and Cohen point out, is that it does not produces a transportable neuronal "output" and, *a fortiori*, it does not produce an output that can be made available to consciousness. The hippocampal system, on the other hand, is said to generate output in the form of transportable representations. Only the hippocampal systems could "send" the fruits of its labour into the neural systems responsible for conscious awareness. There may be an important connection to be made

---

2 Taylor, "Overcoming Epistemology," p. 470.

between the differential access hippocampal and nonhippocampal memory systems have to the neural systems subserving conscious awareness, and the distinction between engaged coping and deliberative reasoning. These comments concerning "significance" and "consciousness" are clearly speculative; they are meant to suggest that the distinction between hippocampal and nonhippocampal memory will most likely prove to be philosophically important in ways not yet known.

My argument against Fodor is inspired by research in philosophy, psychology, and neuroscience. I will close by offering one final speculative suggestion, a suggestion as to how my analysis of multiple memory systems might be extended to incorporate material from the ongoing debate between classical AI theorists and connectionists.

In his book, *Microcognition*, Andy Clark argues that the debate between classical AI theorists and connectionists suffers from its own brand of theoretical imperialism. [3] Classical AI theorists argue that the mind is a virtual machine that utilizes serial, symbol-manipulation mechanisms. According to Clark, some classicists (e.g. Simon and Langley) are "explicit about their belief that the symbol processing architecture they investigate has all the resources to model and explain all aspects of human thought." [4] Connectionists argue that the mind is a virtual machine that operates on the principles of parallel distributed processing (PDP). Many connectionists have been tempted by the view that PDP models provide all the resources necessary for the explanation of intelligent behaviour. Clark refers to the view that one's preferred type of cognitive architecture provides all the

[3] Clark, Andy, *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing* (Cambridge, Mass.: MIT Press), 1990.

[4] Clark, *Microcognition*, p. 16.

necessary resources for such explanatory projects as the "uniformity principle". He urges us to "resist the uniformity assumption in all its guises" [5] and to think of the mind as consisting of "a multitude of possible virtual computational architectures adapted to various task demands." [6]

The symbol system approach of conventional AI theorists is associated with representational theories of the mind. Clark defines a classical system as "one that posits syntactically-structured symbolic *representations* and that defines its computational *operations* to apply to such representations in virtue of their structure." [7] His invocation of Fodor's account of classical theories is particularly useful. Fodor writes that "classical theories ... posit mental representations (data structures) with a certain form. Such representations are *syntactically structured*, i.e., they are systematically built by combining atomic constituents into molecular assemblies, which (in complex cases) make up whole data structures in turn. In short, they posit *symbol systems* with a combinatorial syntax and semantics." [8] On Clark's view, this type of architecture is well-suited for certain types of explanatory projects, namely the explanation of what he calls "recent achievements". As he makes clear, "recent" is here to be read in both phylogenetic and developmental terms. Clark is referring to the cognitive achievements manifest in such higher-order tasks as conscious planning, problem solving, and logical inference, tasks which involve "complex sequential operations that may require a system to follow explicit rules." [9] For these sorts of tasks, Clark claims that conventional symbol-processing architectures are the most appropriate form of modeling. "Where the conscious-reasoning aspects of these two tasks is concerned, the standard architecture of the classical cognitivist models offers an excellent

---

5 Clark, *Microcognition*, p. 128.

6 Clark, *Microcognition*, pp. 128-129.

7 Clark, *Microcognition*, p. 20.

8 Fodor, J., and Z. Pylyshyn, "Connectionism and Cognitive Architecture: A Critical Analysis," *Cognition* 28 (1988), pp. 3-71. Cited in Clark, *Microcognition*, p. 19.

design-oriented aid to their solution. ... The architecture is perfectly suited to the sequential application of explicit rules to an ordered series of symbol strings." [10]

PDP architectures are based on a very different type of mechanism. There are no discrete representations in a PDP net. To whatever extent it is appropriate to speak of representations at all, they are embodied in distributed modifications of "synaptic" connections within the net. According to Clark, this type of architecture is best suited for modeling basic perceptual-motor capacities. "PDP approaches deploy a means of encoding and processing information that is particularly well-suited to evolutionarily basic tasks like low-level vision and sensorimotor control." [11]

The similarities between conventional architectures and hippocampal memory are clear. They employ similar mechanisms: discrete representations are processed in such a way that their integrity is preserved even when they are combined or associated with other representations. Both are associated with the linear processing that underlies evolutionarily advanced tasks such as reasoning. The representations of both conventional architecture and MTL memory are relatively unstable - local damage can destroy a particular representation entirely. There are also important similarities between PDP architectures and striatal memory. Both rely on a process in which the synaptic connections among nodes (or neurons) are gradually modified. Both are relatively stable - the integrity of the net often survives local damage. Both are said to subserve basic evolutionary skills, primarily in the perceptual-motor domain. Clark specifically mentions the suitability of PDP nets for modeling the spared capacities of human amnesics. He notes that "some amnesics ... *can* learn things by very dense repetition of the appropriate experiences." [12] He notes that in the case of PDP learning, repeated exposure to the same situation allows

---

9 Clark, *Microcognition*, p. 127

10 Clark, *Microcognition*, p. 127.

11 Clark, *Microcognition*, p. 104.

12 Clark, *Microcognition*, p. 101.

the net to adjust itself gradually to the task at hand; the net is gradually fine-tuned to respond appropriately to the features of the training situation that remain constant over time. Clark thus claims that there is a certain "fit" between PDP memory and non-representational memory. Some neuroscientists have drawn a similar conclusion. In a passage titled "Modeling Procedural Memory," Eichenbaum and Cohen note that "the tuning and biasing of various brain processors in accordance with real-world regularities of the inputs to which they are exposed finds a very comfortable home in current connectionist (or neural network) models. [13]

Perhaps one day we will be in a position to correlate more precisely the theories of perceptual-motor skills found in phenomenology, cognitive science, and neuroscience. There are clear indications that Heideggerian coping, the perceptual-motor skills of the nonhippocampal memory systems, and the skills associated with PDP architectures have much in common - enough in common to warrant a cooperative interdisciplinary research project. In his recent book, *Consciousness Reconsidered*, Owen Flanagan provides a useful methodological suggestion. [14] He argues that, in our study of consciousness, we should adhere to the principles of what he calls the "natural method" - first, treat the findings of phenomenologists, cognitive scientists, and neuroscientists with equal respect and second, attempt to bring these three types of accounts into reflective equilibrium. This thesis represents an initial attempt to apply Flanagan's method to the topic of perceptual-motor skills.

The significance of the distinction between hippocampal and nonhippocampal memory is yet to be fully appreciated. In the disputes between phenomenology and philosophical cognitivism, behaviourism and psychological cognitivism, and classical AI

---

[13] Cohen and Eichenbaum, *Memory, Amnesia, and the Hippocampal System*, p. 82.

[14] Flanagan, Owen, *Consciousness Reconsidered* (Cambridge, Mass.: MIT Press, 1992), p. 11.

and connectionism, the distinction between hippocampal memory and nonhippocampal memory serves an important peacekeeping role. In ways yet unknown, the distinction may prove invaluable in protecting us against the temptation of what Clark calls the "uniformity principle", the temptation of theoretical imperialism.

BIBLIOGRAPHY

Aggleton, J.P., H.S. Blindt, and J.N.P. Rawlins. "Effects of Amygdaloid and Amygdaloid-Hippocampal Lesions on Object Recognition and Spatial Working Memory in Rats." *Behavioral Neuroscience* 103.5 (1989): 962-74.

Ammons, R. B. "Psychology of the Scientist: IV. Passages from the 'Idea Books' of Clark L. Hull." *Perceptual and Motor Skills*, 15 (1962): 807 - 82.

Amsel, A., and M. E. Rashotte. *Mechanisms of Adaptive Behaviour: Clark L. Hull's Theoretical Papers, with Commentary*. New York: Columbia University Press, 1964.

Bergson, Henri. *Matière et mémoire*. Paris: Presses Universitaires de France, 1911.

Bergson, Henri. *Matter and Memory*. Trans. N.M. Paul and W.S. Palmer. New York: Zone Books, 1991.

Clark, Andy. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, Mass.: MIT Press, 1990.

Cohen, Neal and Larry Squire. "Preserved Learning and Retention of Pattern-analyzing Skills in Amnesia: Dissociation of Knowing How and Knowing That." *Science* 210 (1980): 207-210.

Cohen, Neal, and Howard Eichenbaum. *Memory, Amnesia, and the Hippocampal System*. Cambridge, Mass.: MIT Press, 1993.

Corkin, Suzanne. "Acquisition of Motor Skill after Bilateral Medial Temporal Lobe Excision." *Neuropsychologia* 6 (1968): 255-265.

Corkin, Suzanne. "Tactually-guided Maze Learning in Man: Effects of Unilateral Cortical Excision and Bilateral Hippocampal Lesions." *Neuropsychologia* 3 (1965): 339-351.

Dennett, Daniel. "A Cure for the Common Code?" *Brainstorms: Philosophical Essays on Mind and Psychology* Cambridge, Mass.: MIT Press, 1981: 90-108.

Dennett, Daniel. *The Intentional Stance*. Cambridge, Mass.: MIT Press, 1987.

Dewey, John. *Human Nature and Conduct. An Introduction to Social Psychology*. London: George Allen and Unwin, 1922.

Dreyfus, Hubert. *Being-in-the-World: A Commentary on Heidegger's Being and Time. Division I*. Cambridge, Mass.: MIT Press, 1991.

Dreyfus, Hubert. *What Computers Still Can't Do*. Cambridge, Mass.: MIT Press, 1993.

Dreyfus, Hubert. "The Hermeneutic Approach to Intentionality." Montreal: Society for Philosophy and Psychology, 1992.

Dunnett, S.B. "Comparative Effects of Cholinergic Drugs and Lesions of Nucleus Basalis or Fimbria-Fornix on Delayed Matching in Rats." *Psychopharmacology* 87 (1985): 357-363.

Eichenbaum, Howard. "The Hippocampal System and Declarative Memory in Humans and Animals: Experimental Analysis and Historical Origins." *Memory Systems 1994*. Ed. Daniel Schacter and Endel Tulving. Cambridge, Mass.: MIT Press, 1994: 147-201.

Fodor, Jerry. *The Language of Thought.* Cambridge, Mass.: Harvard University Press, 1975.

Fodor, Jerry. "Special Sciences, or The Disunity of Science as a Working Hypothesis." *Readings in Philosophy of Psychology.* Vol. I. Ed. Ned Block. Cambridge, Mass.: Harvard University Press, 1980: 120 - 33.

Fodor, Jerry. "The Appeal to Tacit Knowledge in Psychological Explanation." *Representations: Philosophical Essays on the Foundations of Cognitive Science.* Cambridge, Mass.: MIT Press, 1981: 63-78.

Fodor, Jerry. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge, Mass.: MIT Press, 1987.

Fodor, Jerry. "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade Mecum." *Mind* 94 (1985): 76-100.

Fodor, Jerry. *Representations: Philosophical Essays on the Foundations of Cognitive Science.* Cambridge, Mass.: MIT Press, 1981.

Gabrielli, J.D. "Contribution of the Basal Ganglia to Skill Learning and Working Memory in Humans." Houk, James, Joel David, and David Beiser, Eds. *Models of Information Processing in the Basal Ganglia.* Cambridge, Mass.: MIT Press, 1995: 277-294.

Gabrieli, J.D., N. J. Cohen, and S. Corkin. "The Impaired Learning of Semantic Knowledge Following Bilateral Medial Temporal-lobe Resection." *Brain and Cognition* 7 (1988): 157 - 177.

Claude Ghez. "The Control of Movement." Kandel, Eric, James Schwartz, and Thomas Jessell, Eds. *Principles of Neural Science.* 3rd. Edition. New York: Elsevier, 1991: 533-547.

Goodale, Melvyn A., and A. David Milner. "Separate Visual Pathways for Perception and Action." *Trends in Neurosciences* 15.1 (1992): 20-25.

Heidegger, Martin. *Being and Time.* Trans. John Macquarrie and Edward Robinson. New York: Harper and Row, 1962.

Heidegger, Martin. *The Basic Problems in Phenomenology.* Trans. Albert Hofstadter. Revised ed. Bloomington: Indiana University Press, 1982.

Heidegger, Martin. *The Metaphysical Foundations of Logic.* Trans. Michael Heim. Bloomington: Indiana University Press, 1984.

Hirsh, Richard. "The Hippocampus and Contextual Retrieval of Information from Memory: A Theory" *Behavioral Biology* 12 (1974): 421-444.

Houk, James. "Information Processing in Modular Circuits Linking Basal Ganglia and Cerebral Cortex." James Houk, Joel David, and David Beiser, Eds. *Models of Information Processing in the Basal Ganglia..* Cambridge, Mass.: MIT Press, 1995: 3-9.

Houk, James, Joel David, and David Beiser, Eds. *Models of Information Processing in the Basal Ganglia.* Cambridge, Mass.,: MIT Press, 1995.

Hsiao, S., and R. L. Isaacson. "Learning of Food and Water Position by Hippocampus Damaged Rats." *Physiol. Behav.* 6 (1971): 81 - 83.

Hull, Clark. "Autobiography." *A History of Psychology in Autobiography.* Eds. E. G. Boring, H. S. Langfeld, H. Werner, & R. M. Yerkes . Worcester, Mass.: Clark University Press, 1952.

Hull, Clark. *Principles of Behaviour.* New Haven: Yale University Press, 1943.

Isaacson, R. L., and D. P. Kimble. "Lesions of the Limbic System: Their Effects upon Hypotheses and Frustration." *Behavioural Biology* 7 (1972): 767 - 795.

Kandel, Eric, and James Schwartz. *Principles of Neural Science.* 2nd ed. New York: Elsevier, 1985.

Kendler, H. "The Influence of Simultaneous Hunger and Thirst Drives upon the Learning of Two Opposed Spatial Responses of the White Rat." *Journal of Experimental Psychology* 36 (1946): 212 - 220.

Kendler, H. *Historical Foundations of Modern Psychology.* Philadelphia: Temple University Press, 1987.

Kimble, D. P. "The Effects of Bilateral Hippocampal Lesions in Rats." *Journal of Comparative Physiological Psychology* 56 (1963): 273 - 283.

Kolb, Bryan, and Ian Q. Whishaw, Eds. *Fundamentals of Human Neuropsychology.* Second ed. New York: W. H. Freeman and Co., 1985.

Lakoff George. *Women, Fire, and Dangerous Things.* Chicago: University of Chicago Press, 1987.

Marsden, C.D. "The Mysterious Motor Function of the Basal Ganglia: The Robert Wartenberg Lecture." *Neurology* 32 (1982): 514-539.

Martin , John. *Neuroanatomy Text and Atlas.* New York: Elsevier, 1989.

Merleau-Ponty, Maurice. *Phenomenology of Perception.* Trans. Colin Smith. London: Routledge & Kegan Paul, 1962.

Merleau-Ponty, Maurice. *Phénoménologie de la perception.* Paris: Gallimard, 1945.

Merzenich, M.M., et al. "Somatosensory Cortical Map changes Following Digit Amputation in Adult Monkeys." *Journal of Comparative Neurology* 224 (1984): 591-605.

Milner, Brenda. "Further Analysis of the Hippocampal Amnesic Syndrome: 14-Year Follow-Up Study of H.M." *Neuropsychologia* 6 (1968): 215-234.

Milner, Brenda. "Intellectual Function of the Temporal Lobes." *Psychological Bulletin* 51.1 (1954): 42-62.

Milner, Brenda. "Les troubles de la mémoire accompagnant des lésions hippocampique bilatérales," P. Passouant (Ed.) *Physiologie de l'hippocampe.* Paris: Centre Nationale de la Recherche Scientifique, 1962.

Milner, Brenda. "Psychological Defects Produced by Temporal Lobe Excision." *Res. Pub. Assoc. Nerv. Ment. Dis.* 36 (1958): 244-257.

Mishkin, M., and T. Appenzeller. "The Anatomy of Memory." *Scientific American* 256.6 (1987): 80-89.

Mishkin, M., B. Malamut, and J. Bachevalier. "Memories and Habits: Two Neural Systems." *The Neurobiology of Learning and Memory.* Ed. G. Lynch, J. L. McGaugh and N. M. Weinberger. New York: Guilford Press, 1984.

Mishkin, Mortimer, and H.L. Petri. "Memories and Habits: Some Implications for the Analysis of Long Term Learning and Retention." *Neuropsycholoy of Memory.* Eds. L. Squire and N. Butters. New York: Guilford Press, 1984.

Mishkin, Mortimer. "Cerebral Memory Circuits." *Exploring Brain Functions: Models in Neuroscience.* Ed. T.A. Poggio and D.A. Glaser. New York: John Wiley and

Sons, 1993.

Mishkin, Mortimer. "A Memory System in the Monkey." *Phil. Trans. R. Soc. Lond. B* 298 (1982): 85-95.

Murray, E., D. Gaffan and M. Mishkin. "Neural Substrates of Visual Stimulus-Stimulus Association in Rhesus Monkeys." *Journal of Neuroscience* 1[illegible] (10) (1993): 4549-4561.

Nadel, Lynn and John O'Keefe. "The Hippocampus in Pieces and Patches." *Essays on the Nervous system: A Fetschrift for Professor J. Z. Young*. Eds. R. Bellairs and E.G. Gray. Oxford: Clarendon Press, 1974: 365-390.

Nadel, Lynn. "Multiple Memory Systems: What and Why, an Update" Memory Systems 1994. Eds. Daniel Schacter and Endel Tulving. Cambridge, Mass.: MIT Press, 1994: 39-63.

Nadel, Lynn. "The Hippocampus and Space Revisited." *Hippocampus* 1.3 (1991):221-229.

O'Keefe, J., and J. Dostrovsky. "The Hippocampus as a Spatial Map: Preliminary Evidence from Unit Activity in the Freely Moving Rat." *Brain Research* 34 (1971): 171-175.

O'Keefe, John, and Lynn Nadel. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press, 1978.

O'Keefe, John, and Lynn Nadel. "Précis of O'Keefe and Nadel's *The Hippocampus as a Cognitive Map*." *The Behavioural and Brain Sciences* 2 (1979): 487-533.

Penfield, Wilder, and Brenda Milner. "Memory Deficit Produced by Bilateral Lesions in the Hippocampal Zone." *A.M.A. Archives of Neurology and Psychiatry* 79 (1958): 475-497.

Penfield, Wilder, and Brenda Milner. "Memory Defect in Bilateral Hippocampal Lesions." *Psychiatric Research Reports* 11 (1959): 43-52.

Petri, Herbert, and Mortimer Mishkin. "Behaviorism, Cognitivism, and the Neuropsychology of Memory." *American Scientist* 82 (1994): 30-37.

Polster, Michael R., Lynn Nadel, and Daniel L. Schacter. "Cognitive Neuroscience Analyses of Memory: A Historical Perspective." *Journal of Cognitive Neuroscience* 3.2 (1991): 95-116.

Saint-Cyr, Jean A., Leslie G. Ungerleider, and Robert Desimone. "Organization of Visual Cortical Inputs to the Striatum and Subsequent Outputs to the Pallido-Nigral Complex in the Monkey." *Journal of Comparative Neurology* 298 (1990): pp. 129-156.

Schacter, Daniel. "Implicit Memory: A New Frontier for Cognitive Neuroscience." *The Cognitivie Neurosciences*. Ed. M. Gazzaniga. Cambridge, Mass.: MIT Press. 1995: 815-824.

Schacter, Daniel L., and Endel Tulving. Eds. "What are the Memory Systems of 1994?" *Memory Systems 1994*. Cambridge, Mass.: MIT Press, 1994.

Scoville, William, and Brenda Milner. "Loss of Recent Memory after Bilateral Hippocampal Lesions." *J. Neurol. Neurosurg. Psychiat.* 20 (1957): 11-21.

Spence, K.W., G. Bergman, and R. Lippit. "A Study of Simple Learning Under Irrelevant Motivational-Reward Conditions." *Journal of Experimental Psychology* 40.5 (1950): 539 - 551.

Squire, Larry R. *Memory and Brain*. Oxford: Oxford University Press, 1987.

Squire, Larry, Neal Cohen, and Lynn Nadel. "The Medial Temporal Region and Memory Consolidation: A New Hypothesis." *Memory Consolidation.* Ed. H. Eingartner and E. Parker, Hillsdale, NJ: Erlbaum, 1984.

Taylor, Charles. *Human Agency and Language: Philosophical Papers.* Vol. I. Cambridge: Cambridge University Press, 1985.

Taylor, Charles. "Cognitive Psychology." *Human Agency and Language: Philosophical Papers.* Vol. I. Cambridge: Cambridge University Press, 1985: 187-212.

Taylor, Charles. "Overcoming Epistemology." *After Philosophy: End or Transformation.* Ed. Kenneth Baynes, James Bohman, and Thomas McCarthy. Cambridge, Mass.: MIT University Press, 1987: 464 - 488.

Taylor, Charles. "The Concept of a Person." *Human Agency and Language: Philosophical Papers.* Vol. I. Cambridge: Cambridge University Press, 1985: 97-114.

Taylor, Charles. "To Follow a Rule..." Cahiers d'épistémologie. Cahier no. 9019. Montréal: Université du Québec à Montréal, 1990.

Thomas, Garth. "Cognition, Memory, and the Hippocampus." *The Behavioural and Brain Sciences* 2 (1979): 515 - 16.

Tolman, Edward. *Purposive Behaviour in Animals and Men.* New York: Century, 1932.

Tolman, Edward. " Cognitive Maps in Rats and Men." *Behaviour and Psychological Man: Essays in Motivation and Learning.* Berkeley: University of California Press, 1966: 242-43. [Originally published in *Psychological Review* 55.3 (1948): 189-208.]

Tolman, Edward. "A Behaviourist's Definition of Consciousness." *Behaviour and Psychological Man: Essays in Motivation and Learning.* Berkeley: University of California Press, 1966. [Originally published in *Psychological Review* 34 (1927): 433 - 439.]

Tolman, Edward. "A Behaviouristic Theory of Ideas." *Behaviour and Psychological Man: Essays in Motivation and Learning.* Berkeley: University of California Press, 1966. [Originally published in *Psychological Review* 33 (1926): 352-369.]

Tolman, Edward. "Behaviourism and Purpose." *Behaviour and Psychological Man: Essays in Motivation and Learning.* Berkeley: University of California Press, 1966. [Originally published in *Journal of Philosophy* 22 (1925): 36-41.]

Tolman, Edward. "Gestalt and sign-gestalt." *Behaviour and Psychological Man: Essays in Motivation and Learning.* Berkeley: University of California Press, 1966. [Originally published in *Psychological Review* 40 (1933): 391 - 411.]

Tolman, Edward. "Habit Formation and Higher Mental Processes in Animals, Part 2." *Psychological Review* 25 (1928): 24-53.

Tolman, Edward. "There is More than One Kind of Learning." *Psychological Review* 56 (1949):144-155.

Tulving, E., and S.A. Madigan. "Memory and Verbal Learning." *Annual Review of Psychology* 21 (1970): 437-84.

Warrington, Elizabeth and Larry Weiskrantz. "A New Method of Testing Long-term Retention with Special Reference to Amnesic Patients." *Nature* 217 (1968): 972-974.

Warrington, Elizabeth and Larry Weiskrantz. "The Amnesic Syndrome: Consolidation or Retrieval." *Nature* 228 (1970): 628-630.

Weiner, W.J., and A.E. Lang. "An Introduction to Movement Disorders and the Basal Ganglia," *Movement Disorders: A Comprehensive Survey*. Ed. W.J. Weiner and A.E. Lang. Mount Kisko, NY: Futura Publishing Co. 1989: 1-22.