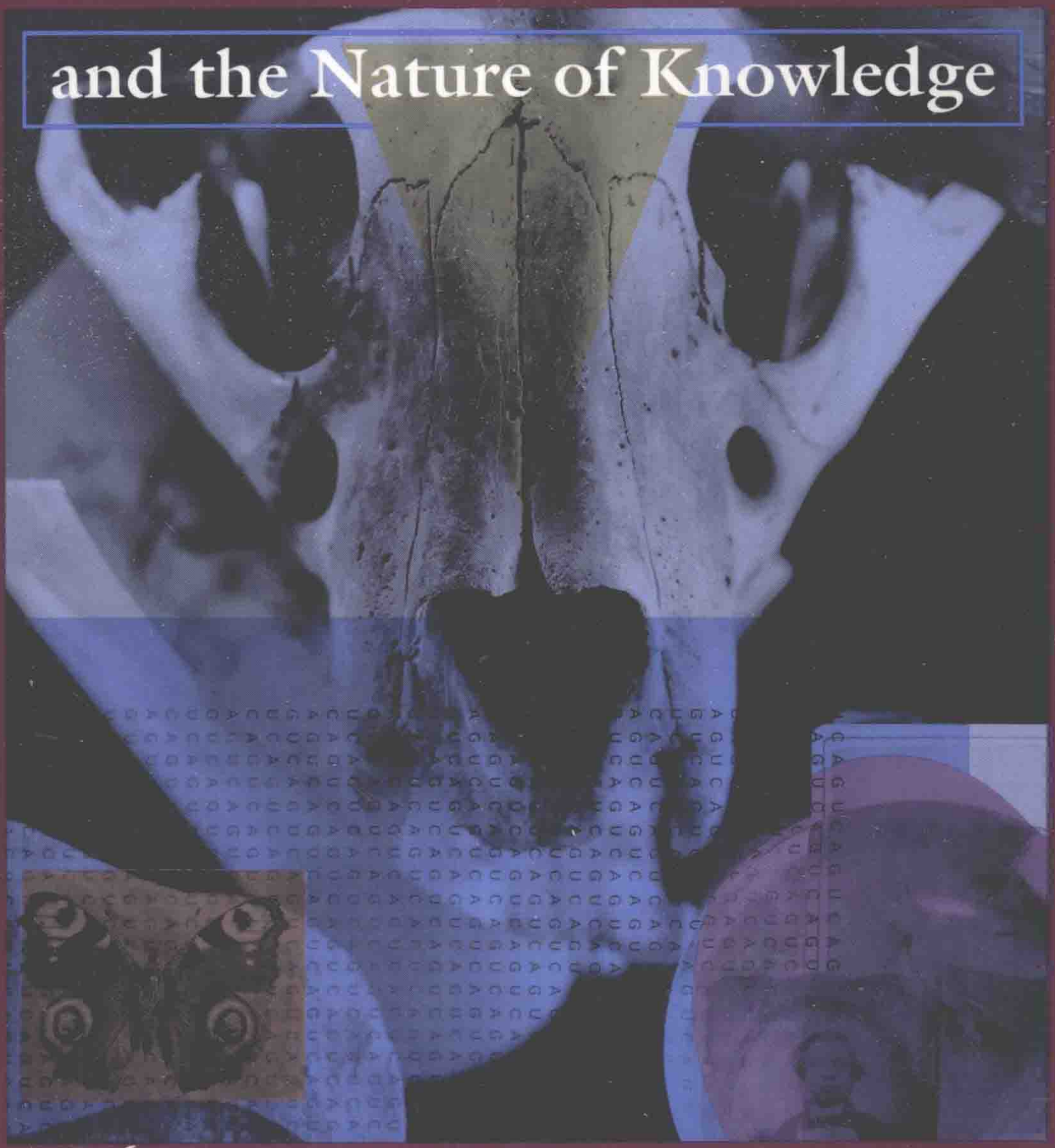
DARWIN MACHINES

and the Nature of Knowledge



HENRY PLOTKIN

HENRY PLOTKIN

DARWIN MACHINES AND THE NATURE OF KNOWLEDGE

HARVARD UNIVERSITY PRESS
CAMBRIDGE, MASSACHUSETTS

Printed in the United States of America
Second printing, 1997

First published in Great Britain by Penguin Books Ltd

First Harvard University Press paperback edition, 1997

Library of Congress Cataloging-in-Publication Data

Plotkin, H. C. (Henry C.)
Darwin machines and the nature of knowledge / Henry Plotkin.
p. cm.
Includes bibliographical references and index.
ISBN 0-674-19280-X (cloth)
ISBN 0-674-19281-8 (pbk.)
1. Knowledge, Theory of. 2. Evolution. 3. Evolution (Biology)
I. Title.
BD161.P6 1994 93-39328
121—dc20 CIP

Darwin Machines and the Nature of Knowledge

For Victoria, Jessica and Jocelin

Preface

To know something is to incorporate the thing known into ourselves. Not literally, of course, but the knower is changed by knowledge, and that change represents, even if very indirectly, the thing known. This is ancient folklore but also a commonplace assumption about knowledge that we all make without really thinking about it. It surfaces in various religious and psychodynamic precepts too: know what you fear, for in so doing you take the feared thing into yourself, it becomes a part of your substance, and in that way you control and conquer it. Are these merely primitive folk-tales conjured up by high priests and thaumaturges? Or plain fanciful nonsense? Not at all. There is more than a germ of truth in these claims, which, though they lack the precision of proper analysis, now find support in the science of knowledge. It is just such a science of knowledge that is presented in outline in this book; and, as we shall see, what comes out of it is a smartened-up version of the notion that knowledge is indeed a kind of incorporation of the world, but one of a special sort.

At the heart of the chapters that follow, then, is this proposition that what we commonly call knowledge, be it the name of a friend, the face of an acquaintance or the position of the kitchen in one's house, is a form of incorporating the thing known into the knower. However, it is also something that bears a very close relationship to a much more widespread property of living things, namely the organizational and structural harmony that exists between life and the world in which it has its being. First exposure to this idea leads many to judge it rather weird, and perhaps difficult to understand. It certainly is not an intuitively obvious

connection to make. But part of the pleasure of science is the making of unexpected and far-from-obvious links that, when unravelled in greater detail, lead to better understanding of the world. Unfortunately, the fun and fascination of science are often obscured to the uninitiated – that is, to the non-scientist or the not-yet-quite-scientist – by jargon, mathematics and abstract visual depictions.

However, it is not only the arcana of science that the outsider finds impenetrable, but also the insistence of so many scientists in presenting their material to one another in a tone of determined, unrelieved tediousness. Yet the boring appearance of modern science is completely at odds with just how exciting and gripping it really is, particularly in the making of these strange yet profound connections and hence in the unexpected and fascinating ways in which it carves up the world – ways that often seem entirely at variance with ordinary experience.

However odd or difficult the idea that knowledge, human knowledge, what you and I know of the world, is closely connected to a very fundamental feature of all living things might seem at first, I will show that it none the less is important because it tells us significant things about why we ever came to know anything, and how we do so. And what it tells us makes contact with what psychologists and neuroscientists are beginning to understand about our capacity for knowing. In other words, the theory does one of the things that theories ought to do, which is that it makes contact with what scientists of different sorts are thinking and demonstrating. It is also, like so much of science, a very exciting way of thinking, and not actually difficult at all if presented in an approachable way. The object of this book is to do just that – to present this idea, and some that flow from it, in a manner that makes it available to everyone with an interest in how and why we, both as species and as individuals, ever came to know anything about our world and ourselves.

Having written this book for many different kinds of people, some of whom have little knowledge of science, I have tried to keep technical jargon to a minimum and provide a definition

when first an unfamiliar word is used. Certain words and phrases, however, appear repeatedly throughout the book (such as 'evolution', 'phenotype', 'heuristic'), and for this reason I have included a glossary that should make unnecessary repeated thumbing back through pages to find those first definitions and explanations, and also serves as a further aid to understanding. The glossary lists the key technical words and phrases that are essential to a science of knowledge in general, as well as to the specific approach of this book. Where my usage is idiosyncratic I give both my meaning and more traditional usage.

The suggested reading at the end of all but the final chapter is meant to serve as a guide to those not familiar with these issues but who find themselves fascinated and want to know more. Some of my readers, though, will be evolutionary biologists, psychologists and neuroscientists of various kinds. These cognoscenti may want to delve deeper into some matters that are not within their areas of expertise, and for them I have provided rather more extended bibliographies for each of Chapters 2 to 6. In order to avoid the visual clutter that chapter-by-chapter reading lists impose, these are presented together at the back of the book.

I am grateful to Celia Heyes, David Hull, Kevin Laland, John Morton and Stephen Walker, all of whom read the original manuscript. Their reactions ranged from encouraging noises to detailed comments and criticisms. The book might have been better had I acted on all of the latter, but ignoring good suggestions is an author's privilege – or vanity. It is certainly the case that the book has benefited greatly from the changes made in the light of their suggestions. I am also grateful to Judith Flanders for her skilful editing.

London
1993

Introduction

The notion of harmony between the organization and structure of individual animals and plants and the world in which they live is almost as old as recorded thought. This is not surprising. The apparent fit, the matching, of living things to the features and conditions of their world – be it the streamlined body shape of animals that move rapidly through water, the camouflaging coloration by which so many animals merge into their surroundings, or the cutting and slashing teeth of carnivores – is a readily observed characteristic of life forms, that immediately impresses itself upon us humans with our special talent for detecting correlated patterns in the world about us. The matching is a result of living creatures somehow incorporating into themselves those aspects of the world that are matched. This is the source of the sense of harmony between the organization of living things and the world about them.

Sometimes the organization is extraordinarily complex, as in the classic example of the human eye; in other cases it may be simple but highly effective, as in the tuning of the hearing receptors of moths to just that frequency of ultrasound used by the bats that prey on them. But simple or complex, match they do. This harmony is the very stuff of so many popular natural-history films and books. It astonishes and delights us, and has been doing so for thousands of years. We call these seemingly clever and often beautiful forms of organization *adaptations*.

Adaptations are the most enduring and powerful focus for attention and study in all of biology. They have been at the centre of metaphysical arguments about harmony and order in the universe, and they have been persuasively used in arguments for the existence of a divine Creator. Prominent in the writings and

thoughts of nineteenth-century naturalists, they were a powerful and central spur to the formulation of the theory of evolution that we now call Darwinism or neo-Darwinism. Adaptations are macro-features of the organization of living things and, we now know, are formed by a very long process of interaction between the environment and successions of organisms that make up lineages of organisms extending over thousands, hundreds of thousands and millions of years. They are crucial determinants of whether organisms survive and reproduce or not. It is the study of adaptations that defines biology as different from other natural sciences. And being central to biology they are central also to the human sciences that are not reducible to chemistry; because however special we might think ourselves – and we are indeed special in our unparalleled abilities in just those areas studied by the social sciences – we are, when all is said and done, just a particular form of animal. We human beings are a finely woven cloth of adaptations, as are all other animals.

How, then, can we make that connection between adaptation and knowledge? We do so through a two-track argument. The first is that the human capacity to gain and impart knowledge is itself an adaptation, or a set of adaptations. To the scientifically literate this may not seem to be a startling claim. But it does have specific and interesting implications. We simply will not understand human rationality and intelligence, or human communication and culture, until we understand how these seemingly unnatural attributes are deeply rooted in human biology. They are, I will argue, the special adaptations that make us special. What is unarguable is that they are the products of human evolution, whether adaptations or not. There really are no substantive alternative ways of understanding our extraordinary capacity for knowledge. Creationism of some kind is, of course, another way to look at these things. Perhaps we are able to know what we know because our Creator gave us this ability. But such a view would turn the clock back 140 years. No scientist now accepts creationism as an explanation of anything. Alternatively, there is the curious and indefensible stance of some scientists who study

humanity that either we do not need to understand our human qualities in terms of our biology (they somehow just are, and the why and wherefore do not matter!), or if biological recognition is to be given then it must not be in terms of the widely accepted evolutionary principles of selection and descent. None of these are credible, persuasive or interesting positions to take. We are clever animals, and that cleverness needs to be understood, indeed can only be understood, using the same analytical tools and principles that we use to understand our size, shape, gait or metabolism.

The second track of the argument is the one that many find strange and difficult, and one which has already been partially given in the Preface. It is that adaptations are themselves knowledge, themselves forms of 'incorporation' of the world into the structure and organization of living things. Because this seems to misappropriate a word, 'knowledge', with a widely accepted meaning – knowledge usually just being something that only humans have somewhere in their heads – it makes the argument easier if the statement reads 'adaptations are biological knowledge, and knowledge as we commonly understand the word is a special case of biological knowledge'. The line of reasoning that is presented in greater detail throughout the following pages goes like this: the relationship of fit between parts of the organization of an organism, its limb structure for instance, and some feature or features of the world in which it lives, such as the terrain or medium through which it must move, is one in which that organization is in-formed by the environment.

This is the only way to understand the effectiveness of adaptations. If adaptations were formed by mere chance, then the extent to which they work would reflect those same improbable odds. But adaptations, by definition, almost always work. This is because of the in-formed nature of adaptations resulting from the in-forming relationship between that adaptation and its environment. This in-forming relationship between parts of organisms and their world is knowledge, or biological knowledge if one prefers. Now, when a person comes to have knowledge of a particular thing, for example

the layout of the keys on the keyboard of a computer (the 'qwerty' keyboard), then the brain state that represents the keyboard is a particular form of organization that also bears a relationship of fit to a feature of the world, the qwerty keyboard, just as the coloration of an insect bears a relationship of fit to the colours of its surroundings. Human knowledge conforms to the relational quality of fit that adaptations have.

So if adaptations are knowledge, and if what we commonly call knowledge (or better, our ability to gain knowledge) is an adaptation, then what in ordinary everyday life we call knowledge is actually a special form of this much wider phenomenon, what I am here calling biological knowledge. Human knowledge is just one kind of a much wider biological knowledge. And science itself is a very special kind of human knowledge. A science of knowledge, then, is a particular kind of knowledge about a special case, human knowledge, that is part of a wider form of knowledge, biological knowledge. If these claims elicit the image of nesting and recursion, of wheels within wheels, that, as we shall see, is entirely appropriate. What follows, then, is the unpacking of this seemingly contorted argument. I will show that our human capacity for knowledge is deeply and ineradicably rooted in human biology; and human biology, of course, is itself rooted in a conceptually and historically wider biology. When we come to know something, we have performed an act that is as biological as when we digest something.

The structure of this book is simple. The first chapter deals with common-sense and philosophical approaches to knowledge and shows that we have now reached the point where a third source of understanding knowledge is possible, and that is through science. I am no philosopher, and this book is not a philosophical tract. But for well over 2,000 years the study of knowledge has been the exclusive property of the philosophers, and they have made important discoveries, the most important and uncomfortable being that human knowledge is never infallible. Now that a science of knowledge is becoming possible it is of real interest to see what light science throws on the age-old philosophical problem of the fallible nature of human knowledge. I do not believe we

can write a science of knowledge without at least passing reference to traditional philosophy.

Science, though, is certainly what this book is about, and in my view such a science must be based on evolutionary theory, which is the central theorem of all biology. Chapter 2 presents a non-technical, contemporary view of evolutionary theory. To have merely referred the reader to the many excellent accounts that are readily available would have left the book incomplete and unable to stand on its own as a single introductory source to the science of knowledge. Chapter 3 is a different affair altogether. It departs from standard evolutionary theory and presents the notion of universal Darwinism and 'Darwin machines', that happy phrase of the American neurobiologist William Calvin. In recent years Richard Dawkins has written some splendid things about universal Darwinism. But his message has been the more restricted one that wherever in the universe life has evolved, it has done so by the processes of Darwinian evolution. The version of universal Darwinism adopted here is that in addition to biological evolution as it is normally understood, Darwinian evolution is also operating to produce the transformations in time that we see in certain other spheres, such as immune system function and even the way science itself operates. This is not a new position. It began with Darwin himself and with his friend T.H. Huxley. It has been developed by a number of eminent scholars in the last 140 years. The point of devoting Chapter 3 to universal Darwinism is to prepare the reader for the idea that, in at least some respects, our brains are Darwin machines too, and the way in which we gain knowledge is another form of universal Darwinism.

Chapters 2 and 3 provide the essential conceptual tools for understanding adaptations as knowledge and knowledge as adaptive. Chapters 4 and 5 use these tools to provide that understanding. Chapter 4 considers instinct as unlearned and unthinking behaviour that, none the less, is knowledge in the same way that the camouflaging coloration of an insect constitutes knowledge of its surroundings. Chapter 5 then deals with the reason why a small number of animals have evolved the capacity for altering their behaviour in the light of past experience – that is, why learning and intelligence ever evolved at all.

I am sometimes asked by colleagues and students how the derivation of learning and intelligence, as expounded in Chapters 4 and 5, makes a difference to our understanding of learning and intelligence as we see it now in ourselves and other intelligent animals. One of the answers that I give is that if the arguments of Chapter 5 are correct, and of course I believe that they are, then all animals that can learn and think are born knowing what it is that they must learn and think about. Chapter 6 applies this very important lesson to what we know about knowing in its most obvious and commonly understood form, namely human knowledge. It considers, among other things, the way in which we come to master language, recognize significant people in our lives, reason, react emotionally and share knowledge through culture. It is an explication of what psychologists call domain-specific cognitive function or modularity.

Finally, Chapter 7 returns to philosophy. As already indicated, not only is a science of knowledge now possible, but science has something to tell the classical philosophers of knowledge. And so we return to where we seem to have started. But it is not the same place, because what comes between the first and last chapters is a knowledge of knowledge that traditional philosophy has never had available to it.

A science of knowledge is the first and essential part of a more general project to write a proper science of human beings. We are not able to do that yet, but some time in the future we surely will be. Central to such a science will be a proper understanding of our extraordinary capacity for gaining and communicating knowledge; knowledge that must be understood first as a part of our nature, and only after that as an issue in nurture. Something like the theory presented in the following pages will be a part of that science.

Contents

Preface
ix
Introduction
xiii

1 *The Problem with Knowledge*
1
2 *Evolutionary Theory*
22
3 *Universal Darwinism*
59
4 *Behaviour without Thought*
102
5 *The Evolution of Intelligence*
134
6 *Aspects of Human Knowledge*
179
7 *The Philosophical Problems in Perspective*
228

Glossary
246
Bibliography
257
Index
264

I

The Problem with Knowledge

Knowledge is what gives our lives order. We know who our neighbours are; and which of them can be relied upon for help. We know who we love. We know something of the spatial arrangement of our world so that we can get to and from work, school, shops and our friends' houses. And we know how to manipulate objects: shoelaces get tied, words get written and meals get cooked. We also use knowledge of language to communicate with one another. Without knowledge we could not live, for our world would seem too disordered and we would lose the stable psychological framework that is indispensable to our survival.

This notion, that knowledge is not merely what we know but an indispensable part of our lives, is essential to an understanding of that knowledge. However, if we are really to understand the nature of knowledge, then we are going to have to delve much deeper into the nature of all living things, go beyond an understanding of knowledge of what is or can be known in ordinary, common-sense human terms, and realize that knowing *is* living and surviving. It is central to our lives, and not just our lives but to all life. When we have done that, when we realize how human knowledge is related to other fundamental biological phenomena, and when we have an appreciation of what knowledge is in any living creature, then we will have something approaching a correct understanding of human knowledge, which is an extension of the order of all of life.

These very grand claims for what is usually thought of as a rather ordinary if almost exclusively human characteristic come collectively under the rather indigestible label of evolutionary epistemology. Epistemology is a branch of philosophy concerned

with the validity of knowledge – it deals with such questions as how we can know anything, and how we can be certain that what we know is true. Evolutionary epistemology is, in simple terms, the biological study of knowledge. More specifically, it is the study and understanding of knowledge through the use of evolutionary theory. The phrase 'evolutionary epistemology' was first coined in 1974 by Donald T. Campbell, an American psychologist. It is a rather wider and more inclusive phrase than the related 'genetic epistemology' of Jean Piaget, the Swiss philosopher and psychologist. As we will see in a later chapter, genetic epistemology is actually a variant form of the more inclusive evolutionary epistemology. For that reason, when we have to use a wide label, it will be the latter.

'Evolutionary epistemology' is a regrettable phrase because it is both pompous and portentous – it threatens people with an intellectual mugging by a philosopher. However, it does have currency within cognitive science and philosophy. More importantly, the phrase is an explicit reminder that a science of knowledge must be grounded in evolutionary theory. It informs us that knowledge is a problem in evolutionary biology. So it does important work, and we will keep it and use it, if sparingly.

Before developing these ideas about the biological nature of knowledge, which all subsequent chapters will be concerned with, I want first to review some common-sense approaches to knowledge, and to contrast these with philosophical and scientific views. The reason for doing this lies in the fact of our ordinary understanding of knowledge being so different from both scientific and philosophical approaches to the problem, and in the philosopher's approach to knowledge in turn being so different from that of the scientist. If one is to understand what a scientist thinks knowledge is, and most non-scientists think that scientific claims about knowledge are exceedingly strange, then one must realize that scientists themselves appreciate the oddity of their position. This is best done by this scientist, the author, making it clear that he does have an appreciation of what the man or woman in the street thinks knowledge is, and also what philosophers think it is – or indeed, in the philosopher's case, whether knowledge is