To Julie, Maya and Theo

The Philosophy of Cognitive Science

M. J. CAIN

polity

Copyright © M. J. Cain 2016

The right of M. J. Cain to be identified as Author of this Work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

First published in 2016 by Polity Press

Polity Press
65 Bridge Street
Cambridge CB2 1UR, UK

Polity Press
350 Main Street
Malden, MA 02148, USA

All rights reserved. Except for the quotation of short passages for the purpose of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

ISBN-13: 978-0-7456-4656-5
ISBN-13: 978-0-7456-4657-2 (pb)

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Cain, M. J.
The philosophy of cognitive science / Mark Cain.
pages cm
Includes bibliographical references and index.
ISBN 978-0-7456-4656-5 (hardcover : alk. paper) -- ISBN 0-7456-4656-5 (hardcover : alk. paper) -- ISBN 978-0-7456-4657-2 (pbk. : alk. paper) -- ISBN 0-7456-4657-3 (pbk. : alk. paper) 1. Philosophy and cognitive science. 2. Philosophy of mind. I. Title.
BD418.3.C34 2015
128'.2--dc23
2015008643

Typeset in 9.5 on 12 pt Utopia by
Servis Filmsetting Ltd, Stockport, Cheshire
Printed and bound in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY

The publisher has used its best endeavours to ensure that the URLs for external websites referred to in this book are correct and active at the time of going to press. However, the publisher has no responsibility for the websites and can make no guarantee that a site will remain live or that the content is or will remain appropriate.

Every effort has been made to trace all copyright holders, but if any have been inadvertently overlooked the publisher will be pleased to include any necessary credits in any subsequent reprint or edition.

For further information on Polity, visit our website: politybooks.com

Contents

Acknowledgements

Writing this book has taken much longer than I anticipated when I first sketched its outline several more years ago than I care to dwell upon. I therefore need to thank my editors Emma Hutchinson and Pascal Porcheron at Polity for their continued patience, support and encouragement. Thanks also to my philosophy colleagues at Oxford Brookes University, Steve Boulter, Constantine Sandis, Dan O'Brien and Bev Clack. Rumour has it that they find many of my philosophical views slightly bonkers, but that hasn't stopped them providing a warm and stimulating environment to work in. I've also had the pleasure of supervising the PhD research of Dana Vilistere, Annie Webster, Marcus Westberg, Andrew McIntyre, Colin Beckley, Linda Fisher-Hoyrem, Esther Shallan and Istvan Zardai, all of whom have provided very helpful feedback on my developing ideas. Massive thanks to my daughter Maya Reid-Cain for helping me wrestle with the various diagrams used in this book and for keeping her disdain at my IT ineptitude in check. Most of all thanks to my wife Julie and children Maya and Theo for all their love and support. This book is dedicated to them.

1 Cognitive Science and the Philosophy of Cognitive Science

1 Introduction

This is a book about the philosophy of cognitive science. That topic immediately raises two questions: what is cognitive science and what is the philosophy of cognitive science? In one respect the answers to these questions are obvious: cognitive science is the scientific study of cognition, and the philosophy of cognitive science is that branch of philosophy that addresses philosophical questions generated by the scientific study of cognition. But these answers are hardly illuminating as they raise a number of subsidiary questions. What is cognition? What are the key assumptions and methods adopted by those who attempt to study cognition scientifically? How does cognitive science relate both to other sciences and to our commonsense understanding of ourselves as minded agents? When and how did cognitive science emerge as a distinct discipline? How does the philosophy of cognitive science relate to other branches of philosophy such as the philosophy of mind and the philosophy of science? In this chapter I will address these general questions and so provide the foundations for the more specific discussions of the later chapters.

2 Cognition

What exactly is cognition? In order to answer this question it is helpful to begin from a commonsense perspective. According to commonsense – or at least the commonsense perspective of most twenty-first-century Westerners – human beings can be distinguished from inanimate physical objects in having a mind. What, according to commonsense, is involved in our having minds? Here is a by no means exhaustive list: we think; we perceive the external world by means of our senses; we experience pain and other sensations when our bodies are appropriately stimulated; we experience moods such as depression and light-heartedness; we experience emotions such as anger, joy and jealousy; we are conscious both that we think and feel and how we specifically think and feel; we act on the basis of our decisions and intentions, which in turn often reflect how we perceive the world to be, what we think and what we want; when we

act we are conscious of what we do without having to rely on our external sense organs; we recollect our past thoughts, actions and experiences; we imagine particular scenarios; and so on.

In so describing the activities that are central to human mentality from the commonsense perspective I did not use the term 'cognition'. But if one were to ask which of the above activities were best described as involving cognition, most people would answer 'thinking'. So one might say that, to a first approximation, cognition is thinking. But what exactly is thinking? Thinking is a mental process or activity that results in having a thought. Such thought processes range from the intellectually demanding and abstruse (as when one thinks when doing philosophy) to the banal and everyday (as when I address the question of whether I have enough time before my next lecture to buy a coffee). Some concern matters of great importance, others not. Some issue in thoughts that have an immediate impact on action and others don't. Some issue in thoughts that become deeply entrenched and influence many subsequent thought processes, while others issue in thoughts that are fleeting and inconsequential. Some concern counterfactual or hypothetical matters (would I have been late for my lecture had I queued for a coffee?) whereas others are straightforwardly factual.

If thinking is a mental process that issues in a thought, then what exactly is a thought? In everyday talk the term 'thought' is most commonly used to refer to the mental state of considering a particular hypothesis or answer to a question or that of committing oneself to a particular hypothesis or answer. Such mental states are, or are closely related to, beliefs. For when one considers a particular answer to a given question one is paving the way to holding a particular belief, and when one commits oneself to a particular answer one has thereby acquired a belief. Beliefs are examples of what philosophers call propositional attitudes, as believing that p involves adopting the belief attitude to the proposition p. But there are propositional attitudes that are not beliefs, for, just as one can believe that p, one can desire that p, intend (to make it the case) that p, hope that p, fear that p, expect that p, and so on. Consequently, although it slightly strains everyday usage, one might say that thoughts are propositional attitudes, so that thinking is the mental process that results in the acquisition of a propositional attitude, be it a belief, a desire, an intention, or whatever.

In virtue of being relations to propositions, propositional attitudes have meaning or semantic properties. They are therefore akin to declarative sentences of a natural language. Just as such sentences are about particular things (or types of thing) or states of affairs and represent them as being a particular way, so do propositional attitudes. For example, my belief that aardvarks eat termites is about aardvarks and represents them as being termite eaters. What a propositional attitude is about and how it represents its

object are elements of its meaning, but philosophers tend to use the term 'content' when talking about the meaning of propositional attitudes. So, for example, the sentence 'aardvarks eat termites' has a particular meaning, and the belief that aardvarks eat termites has a particular (corresponding) content. There is a further commonality between declarative sentences and propositional attitudes. Sentences are made up of simpler components, namely words, and the meaning of a sentence is a product of the meaning of the words that belong to it and the way they are put together (the syntactic structure of the sentence).[1] It is natural to think that something similar is true of propositional attitudes. The simpler components of propositional attitudes are concepts, and the content of a propositional attitude is a product of its constituent concepts and the way they are put together. Hence, one cannot form the belief that aardvarks eat termites without employing each of its constituent concepts.

I have suggested that, to a first approximation, cognition is thinking. Allied with the idea that thinking is a mental process that results in the formation of propositional attitudes, the implication might seem to be that much of what goes on in the mind falls outside of the domain of cognition. This is because philosophers standardly distinguish between propositional attitudes and other mental states such as sensations, emotions and perceptual experiences. However, even if we want to uphold such a distinction (say, on the grounds that such mental states differ from propositional attitudes in that they have an essential phenomenal, qualitative or 'what it is like' aspect), it would perhaps be a little rash, for several reasons, to conclude that only propositional attitudes and the processes that generate them count as cognitive. First, consider perception. Perception is a process where the external world stimulates an individual's sensory organs, resulting in a perceptual experience. In the case of vision, light reflected off external objects is focused onto the retina, a light-sensitive surface at the back of the eye that sets off a mental process which results in a visual experience. Even if, as many philosophers have thought, perceptual experiences are very different from propositional attitudes in not involving the deployment of concepts and in having an intrinsic qualitative character, they are like them in a key respect. For perceptual experiences are representational in the respect that they are typically about objects located in the external world and represent them as being a particular way. For example, my current visual experience was not only caused by the computer directly in front of me but is also about that computer and represents the outer world immediately before me as containing a grey, rectangular object. In virtue of their being representational, it seems reasonable to think of perceptual states as belonging to the domain of cognition if propositional attitudes are paradigmatic cognitive states.

A second point relates to the nature of the mental process of thinking.

Typically, when one acquires a propositional attitude by means of thinking, a process takes place in one's mind that is extended in time. This process has stages each of which involves moving from one or more propositional attitudes to another, where the earlier propositional attitude(s) in each stage justify the later. Here is an example. Feeling a little de-energized, I wonder if I have enough time to buy a coffee before my next lecture. I start from the beliefs that it will take me ten minutes to walk from the café to the lecture hall and that, given the current length of the queue, it will take me five minutes to purchase a coffee. Given that my lecture starts at 1.00 p.m., I infer that I must leave the café by 12.45 p.m. at the latest to be on time. I look at my watch and come to believe that it is 12.50 p.m. and draw the implication that I cannot buy a coffee and arrive at my lecture on time. Given my strong desire not to be late, I decide to give coffee a miss.

When we think, we are often aware of the stages we go through to reach the propositional attitude that is the end product of the thinking process. Or, if the thought process is so quick and routine that we don't have such awareness, we can retrace our steps by deliberately seeking to justify our conclusion and so gain awareness of our thought process retrospectively. But, with respect to perception, things don't seem to be like that. I open my eyes, orient them to the world and have a perceptual experience without having any awareness of executing an extended process the earlier elements of which justify the later elements. So there would appear to be a substantial contrast between thinking and perception once again. But on second thoughts the contrast might not be so great. Just because we are not aware or conscious of any stages of inference in the process of perception, it does not follow that there are no such stages once we countenance the possibility that much might go on in the mind that is unconscious. Indeed, as we shall see in later chapters, in the 1950s and 1960s psychologists began to hypothesize that perception does involve unconscious inference.[2] If one takes this idea seriously, then one has grounds for including perception in the domain of cognition. Moreover, similar points can be made concerning other mental states and processes. Suppose I am in a café having a cappuccino. I reach out to grasp the cup, raise it to my lips and take a sip. This is a routine everyday event, and when it takes place I am not aware of my having done much in the way of thinking. But if we try to make sense of how we carry out such actions it is clear that an awful lot of complex internal processing must be involved. Once again, in the 1950s and 1960s the idea developed that our actions are driven by thought-like inferential processing that draws upon and coordinates constantly changing information about the state of the external world and one's own body coming from a range of distinct sources. So we might have grounds for regarding action and motor control as belonging to the domain of cognition.

A third point is that, even if there is a significant difference between thinking and other aspects of the mind such as perception and action, the latter must interface with the former. Consider perception first. Perceptual experiences might not be propositional attitudes or the upshot of thought processes, but they do play a key role in determining our propositional attitudes. For we routinely form beliefs and update our stock of beliefs on the basis of our perceptual experiences. Were this not the case, then perception would be of little use to us, as its value resides in its ability to provide us with knowledge about the world that we can utilize in deciding how to act so as to satisfy our needs and desires. For example, suppose you are in the supermarket looking for a tomato and while in the fruit and veg section looking at a tomato you have a visual experience as of a spherical, red object. If this experience is going to help you to satisfy your desire for a tomato, you must take it at face value and come to believe that there is a spherical, red thing before you, and you must further infer the belief that that spherical red thing is a tomato. All this clearly requires the output of perception to be taken up by thought processes. A parallel point can be made about motor control and action. Even if the movements that we make when we act are not directly driven by thought processes, those movements must be routinely and systematically motivated by our propositional attitudes. Were this not the case then our propositional attitudes would be robbed of their central function of enabling us to act so as to satisfy our desires. For example, there wouldn't be much benefit in desiring an elderflower pressé and (correctly) believing that there was a bottle of elderflower pressé in the fridge if that belief and desire pair were incapable of making an impact upon how one acted. In short, then, it is a fundamental property of thinking that it interfaces with both perception and action. This suggests that one should resist drawing a fundamental divide between perception and action, on the one hand, and thinking, on the other, and view only the latter as belonging to the domain of cognition.

I began this section by asking what cognition is and put forward the suggestion that cognition is thinking understood in commonsense terms as a process involving inferring propositional attitudes from other propositional attitudes. However, I argued that, although such thought processes are clear-cut cases of cognition, there are other processes that take place in the mind, namely those involved in perception and action and motor control, whose similarities and relationship to thinking are such as to imply that they count as cognitive processes. Thus, if cognitive science is the scientific study of cognition, it should be concerned as much with perception and action as it is with central cases of thinking.

3 Science and cognition

In the previous section I gave a preliminary account of cognition. That account implies that cognitive science is the scientific study of such phenomena as thinking, perception and action. Thus, the concern of the cognitive scientist is to explain how we humans are able to think, perceive and act, to uncover precisely what goes on within our minds when we exercise such cognitive capacities. If successful, cognitive science will reveal the core properties of humans that enable us to cognize and thus differentiate us from all the inhabitants of the universe that are not capable of cognition. Put this way, it is clear that cognitive science rests on the assumptions both that cognition is the kind of phenomenon that is amenable to scientific investigation and that we humans are cut out to execute such a scientific investigation. However, such assumptions are not universally held within the philosophical community. In this section I will describe some of the most prominent reasons for scepticism about the viability of cognitive science.

A first reason for scepticism relates to a metaphysical view about the mind most associated with Descartes (1985) and is generally known as substance dualism. According to Descartes, the human individual is a two-component system consisting of a body and a mind. The body is an inhabitant of the physical world and so, reflecting the nature of the physical world in general, is a mechanical system whose behaviour is governed by laws of nature that can be stated in mathematical terms. Moreover, the body is essentially spatially extended; that is, it has a spatial location and takes up physical space. In virtue of all this, the human body is precisely the kind of system whose nature and workings can be investigated from the scientific perspective. The mind, on the other hand, is essentially a thinking thing. It does not inhabit the physical world and, in virtue of having free will, it is not a mechanical system governed by laws of nature. Nevertheless, the mind is harnessed to the physical body and engages in a systematic two-way process of causal interaction with it. For example, the body can affect the mind, as when the physical stimulation of the sense organs causes a perceptual experience, and the mind can affect the body, as when a decision to execute a particular action causes the body to move. For Descartes, in virtue of being a fundamentally different kind of thing than the body, the mind, and therefore cognition, is not amenable to scientific study. If Descartes is correct on this point, cognitive science is not a viable enterprise.

How should we respond to this Cartesian line of thought? Descartes produced a number of arguments for dualism, but the general view of the philosophical community is that none of these are successful. Moreover, it is widely held that dualism cannot make sense of the existence of causal

interaction between the mind and the body: if mental phenomena reside outside of the physical domain, how could they cause or be caused by physical phenomena taking place within the body? Indeed, as we shall see, most mainstream cognitive science is underwritten by a commitment to the anti-dualist view that the mind is in some substantial respect the brain, so the study of cognition is the study of the workings of a physical system.

A second source of scepticism regarding the viability of cognitive science accepts that cognition is the kind of phenomenon that could in principle be investigated scientifically but questions whether we humans are up to the task of conducting such an investigation. One way of characterizing this line of thought is in terms of Noam Chomsky's (2000) distinction between problems and mysteries. A problem is a question that is difficult to answer but which we have some hope of answering. A mystery is a question which we have no hope of answering because it is beyond our cognitive powers to do so. Which questions are problems and which mysteries can be species relative. For example, questions we humans find easy to answer might be utterly mysterious to a rat. We humans are bound to have cognitive limitations which make some questions that we can frame mysterious to us. One popular suggestion as to what might be a mystery for us is phenomenal consciousness. Colin McGinn (1989b) argues that phenomenal consciousness, the 'what it is like' aspect of our experiences (Nagel, 1974), is a physical phenomenon but that we are incapable of explaining how physical phenomena can generate consciousness.[3] Now I haven't characterized phenomenal consciousness as belonging to the cognitive domain, but it wouldn't be outlandish to argue that questions about human cognition are mysteries for us. For, in attempting to answer such questions, we are attempting to use our cognitive powers to understand our cognitive powers.

I don't think we should expect to be able to produce a knock-down response to this kind of scepticism. Perhaps the most sensible response would be to argue that we should evaluate the power of the objection in light of the success of our ongoing attempts to explain cognition. Thus, the power of the scepticism would be undermined to the extent that cognitive scientists produced theories and explanations that were productive and successful when judged by the criteria operative in other scientific domains.

A third source of scepticism concerning the viability of cognitive science relates to my characterization of the discipline in terms of our commonsense understanding of ourselves. For, one might argue, the phenomena that seem significant from a commonsense perspective might not be tractable from a scientific perspective, in that they are too complex and messy in being the products of many distinct but interacting factors, each of which belong to different scientific domains. Oddly enough, this idea can be

found in the work of Chomsky, who, to my mind, has made one of the most valuable contributions to cognitive science.

Suppose that I utter a particular sentence. To understand me you will have to cognize what individual words I produced and how they were put together to make the sentence in question (that is, the grammatical or syntactic structure of the sentence). This is no mean feat. For one thing, there is no neat correspondence between words and sounds. This is because distinct utterances of one and the same word can be very different at the sonic level; think of how different the word 'dog' sounds when uttered by a small child, a young woman and an elderly man. And, with respect to syntactic structure, that is not an immediately perceivable property of our utterances, as is indicated by the phenomenon of structural ambiguity. For example, the sentence 'he chased the dog with the stick' could have a structure such that it means that the dog had the stick or one such that it means that the man had the stick. But suppose that you overcome such problems and work out what individual words I uttered and how I structured them in the sentence. If you know the meaning of the individual words of the sentence, you will be able to further cognize the literal meaning of my utterance, assuming that the literal meaning of a sentence is exhaustively determined by the meaning of its component words and its syntactic structure.

However, we normally understand more than the literal meaning of the sentences we hear, and successful communication depends upon this. For we appreciate the communicative intentions of the speaker (Grice, 1975). For example, suppose that you come into my office and I say 'There's a nice fresh breeze coming through the door.' I might be using that sentence with the literal meaning that it has to say any one of several things. I might be aiming to state a fact that does not require you to do anything. Or I might be aiming to point out that you have left the door open and be requesting that you close the door. If communication between us is to be successful, you must appreciate what I am aiming to do in producing the sentence – what my communicative intentions are. But if you are to do this you need to draw upon an appreciation of a potentially wide and disparate range of information. In this case such information might relate to my facial expression, my tone of voice, the temperature conditions in the room, your history of interactions with me, the likelihood of my desiring privacy for our meeting, the conventions governing meetings in my institution, and so on. Now much of this potentially relevant information lies outside of the body of linguistic knowledge that one might think we need to be capable of producing syntactically well-formed sentences and appreciating the literal meaning of the sentences we encounter. Chomsky's point is that, because understanding communicative intentions involves appreciating such a potentially wide and disparate range of information from outside the linguistic domain, it is going to be impossible for us to explain how we understand communicative intentions.

A related point is that sciences carve the world into domains of enquiry within which the scientist has some hope of making progress. But in the real world such domains often interact with one another, so that phenomena that are salient from a commonsense perspective are often massive interaction effects from the scientific perspective. Consequently, explaining such phenomena would require us to engage in 'the science of everything' – something that is not possible for us.

Chomsky's own engagement in cognitive science suggests a response to this sceptical line of thought that involves conceding the point that lies at its core. Our commonsense conception of ourselves as cognizers motivates our engagement in cognitive science and provides the discipline with a starting point of questions to address. But it is an open question as to which of those questions remain on the agenda and what questions are to replace those that are abandoned. How these issues are to be settled is something that cannot be determined in advance of our actual engagement in cognitive science. This situation with respect to cognitive science is typical of science in general, so it motivates scepticism concerning the viability of cognitive science only if it motivates scepticism about the viability of science in general.

4 Science

The upshot of the discussion of the previous section is that it would be somewhat hasty to conclude that cognitive science is not viable. But this does raise a further question: what characteristics should we expect cognitive science to have in virtue of being a science? In this section I will address this question.

Science is one of the great triumphs of Western civilization and has provided us with a systematic body of knowledge of the workings of the natural world. I am very wary of attempting to provide a general account of the nature of science. Rather, I will describe some prominent features of science that will prove to be very relevant when understanding the core commitments of cognitive science and when addressing philosophical questions about cognitive science.

First, most scientists assume that happenings in that portion of the natural world with which they are concerned are not entirely random and irregular; rather, they are governed by laws. Thus, one of the main goals of science is to discover the laws that govern the workings of the natural world. Here are some examples of such laws of nature: freely falling bodies accelerate at a uniform rate (Galileo); planets have an elliptical orbit (Kepler); the volume occupied by a gas is inversely proportional to the pressure on

it (Boyle); the strength of the gravitational attraction between two bodies depends on the product of their masses and is inversely proportional to the square of the distance between them (Newton). Note that some laws (particularly those operating at the quantum level) are probabilistic rather than deterministic. That is, they are of the form that, if x is the case, then, with probability P (where P is less than 1), y will be the case. And some laws are *ceteris paribus* (all else equal) laws. Such laws have the form that, if x occurs, then y will occur, all else equal. Hence, the claim that such a law holds is not necessarily undermined by the observation that an x has occurred without being followed by a y. For, all else might not have been equal in such a case. Jerry Fodor (1987) has provided a well-known example of a *ceteris paribus* law from geology: a meandering river erodes its outer bank, *ceteris paribus*. A case where a meandering river might fail to erode its outer bank would be one where the bank had been concreted over; in such a case all else wouldn't be equal. Given the general role of laws in science, we should seriously entertain the possibility that the cognition is law governed and that one of the main tasks of cognitive science is to uncover such laws.

Second, another of the major goals of science is to explain features of the natural world. This involves not merely describing the natural world but saying why it is the way it is and how things work in the natural world. Laws are relevant here as they have an important role in explaining natural events. For example, Galileo's law plays a key role in explaining why a cannonball and a marble when dropped from a particular height at a given time will hit the ground at the same point in time, in violation of our commonsense expectations that the cannonball, in virtue of being heavier, would hit the ground first. Indeed, explaining why a particular law holds typically involves appealing to more basic laws. For example, Newton appealed to his laws of motion and the inverse square law in order to explain Kepler's law of planetary motion. My initial characterization of cognitive science highlighted the role of explanation, as I said that cognitive science is concerned with explaining our cognitive abilities.

Third, science is an empirical discipline in that observation and experiment play a central role in the scientific project. This contrasts science with any purely *a priori* or armchair activity. One prominent account of the precise way in which science is empirical was developed by Karl Popper (1959). Popper was concerned with the so-called demarcation problem, with specifying precisely what distinguished science from non-science. He argued that it was a mark of a scientific theory that it was falsifiable – that is, a scientific theory has implications as to what observable phenomena will take place in particular circumstances such that there are, in principle, observations we could make that would definitively show the theory to be false. Thus, the scientist proceeds by a two-stage process of conjecture and refutation. This involves producing a conjecture or hypothesis in order to

explain some target phenomenon. The observational implications of the conjecture are then worked out and an experiment conducted in which the relevant observations are made. If the observations are inconsistent with the observational implications of the conjecture, then the conjecture is rejected as false and the scientist goes back to the drawing board. If the observations do not falsify the conjecture, then the scientist conducts further experiments in order to test and falsify the conjecture.

Fourth, in the course of explanation, scientists often postulate theoretical entities. These are entities that are not observed but are invoked to explain phenomena that are observable. For example, suppose a biologist wanted to explain why organisms generally bear many similarities to their parents but have some differences from them. The explanation proffered by the biologist will appeal to genes. The key point is that the genetic explanation[4] is not produced on the basis of observing genes and their activities. Rather, the genes are postulated in order to make sense of the otherwise mysterious observed phenomena. Other prominent examples of theoretical entities are atoms, quarks and photons. Given the role of theoretical entities in science, we should expect cognitive scientists to postulate theoretical entities in theorizing about human cognition.

Fifth, although science is empirical in nature, the role of observation in science is not as straightforward as implied by Popper's work as described above. W. V. O. Quine's 'The Two Dogmas of Empiricism' – first published in 1951 – is probably the most important philosophical article written in English since the Second World War. In this paper Quine seeks to undermine two central theses of the empiricism of his day. The second of these theses is that the meaning of a sentence is its method of verification – that is, the empirical means of determining its truth-value. Thus, any meaningful sentence can be translated without loss into a sentence about experience or observations (this doctrine is known as reductionism). In rejecting verificationism, Quine champions a position that has become known as the Duhem–Quine thesis.[5] It has this name because Quine's basic idea was anticipated by the French physicist and philosopher of science Pierre Duhem (1954), in his book *The Aim and Structure of Physical Theory*.

Suppose that one wants to determine by empirical means whether a particular (declarative) sentence is true or false. One will then have to make relevant observations using one's senses. It is natural to think that, with respect to any true sentence, it is in principle possible to establish that it is true by making relevant observations; and that, with respect to any false sentence, it is in principle possible to establish that it is false by making relevant observations. Applied specifically to science, the idea would be that, with respect to any sentence expressing a scientific theory, it is in principle possible to establish whether it is true or false by observational means. Quine's basic point is that this is not the case because it

is not possible to verify a sentence or theory in isolation. This is because individual sentences or theories do not by themselves have the kind of observational implications that enable them to be verified. Rather, it is only collections of sentences or theories that have such implications. To see this, consider a topical example. Attention Deficit Hyperactivity Disorder (ADHD) is a condition attributed to many children in the Western world today. The condition manifests itself in a child's inability to concentrate on schoolroom tasks and a tendency to be disruptive. What causes ADHD and why is it so prevalent today? One theory is that it is the result of a deficiency in Omega 3. How are we to verify this theory or the sentence that expresses it? Here is an answer. Give children with the condition an Omega 3 supplement and observe their subsequent behaviour. If the children are observed to undergo an improvement in their powers of attention and concentration, then that tells in favour of the theory, suggesting that the theory may well be true or is a serious candidate for truth. If, however, the children are observed to undergo no such improvement, then that tells against the theory, suggesting that it is (probably) false. The problem is that attempting to verify the theory/sentence in this way involves making a whole load of assumptions that are independent of the theory/sentence. For example, the assumption that the supplement contains Omega 3 in a form that can be readily absorbed by the body; the assumption that if the condition is due to a deficiency in Omega 3 then it can be remedied by taking a supplement; assumptions about how concentration/attention powers manifest themselves in observable behaviour; and so on. Making such assumptions will involve committing oneself to the truth of substantial scientific theories. If one doesn't make these assumptions, then one will regard the above way of verifying the theory as illegitimate. If one makes alternative assumptions, then one will be committed to adopting a different means of verifying the theory. And if one makes no supplementary assumptions at all, then one will have no idea how to verify the theory, as the theory in itself will not tell one how to go about verifying it.

Now suppose that one does make the above described assumptions and that no improvement in the children is observed. Does this tell against the theory? According to Quine, the answer to this question is negative. For, in order to deal with the recalcitrant observational data, one can hold onto the theory but reject some of the associated assumptions. Thus Quine says, 'any statement can be held true come what may, if we make enough adjustments elsewhere in the system' (1951: 43). All this is captured in the Duhem–Quine thesis, according to which:

> any theory can be held onto in the face of recalcitrant observational data by making suitable adjustments elsewhere in one's system of commitments.

The Duhem–Quine thesis is widely held by philosophers of science. With respect to cognitive science, an implication of the thesis is that it may well be difficult to adjudicate between different theories about our cognitive lives, as their advocates disagree on the significance of the empirical data they gather.

A sixth feature of science concerns its relationship to commonsense. If a human individual is going to prosper in a challenging and potentially dangerous world, they will need to have a battery of concepts in order to categorize the phenomena with which they interact. And they will need to employ those concepts to form knowledge or beliefs concerning how such phenomena behave. Armed with such knowledge, they will then be able to predict and explain worldly events and so enhance their survival prospects. When philosophers and cognitive scientists talk about commonsense, they are talking about the relatively untutored and unsophisticated conceptual schemes and associated belief/knowledge systems that ordinary people routinely employ in their everyday lives. Commonsense is therefore distinguished from mature and sophisticated scientific theories. However, that is not to say that commonsense bears none of the characteristics of science. Questions about commonsense are prominent within cognitive science, and a widely held view represents it as bearing some of the characteristics of mature science. First, it has several components each of which relate to a distinct subject matter, so that commonsense is underwritten by the assumption that there are different kinds of phenomena in the world that behave in their own distinctive way. In other words, commonsense, like science, has its component disciplines.[6] Moreover, some of them correspond to prominent scientific disciplines. For example, there is commonsense physics that is concerned with inanimate physical objects, commonsense psychology that is concerned with minded agents, particularly humans, and commonsense biology that is concerned with biological entities such as animals. Second, commonsense concepts can be quite abstract in the respect that they group together phenomena that differ widely in terms of their perceivable properties, such as their shape, size and colour, and in the respect that they refer to the unobservable. For example, commonsense physics employs a quite general concept of a physical object that is utilized in such beliefs that unsupported physical objects fall to the ground. And commonsense biology employs the concept of a hidden essence, in that we think of creatures as having characteristics we cannot directly perceive that determine the kind of thing they are and which are causally responsible for the non-essential surface characteristics that we can perceive (Keil, 1989).

One quite natural view of the relationship between science and commonsense is that the former is born of the latter. Science is a cultural phenomenon which in an organized and rigorous form has a relatively

short history of approximately 500 years and so is very much pre-dated by commonsense. Commonsense constituted the starting point for science, in that it provided it with its core questions and conceptual scheme. For example, if commonsense physics assumes that physical objects causally interact with one another and are governed by generalizations that advert to their physical properties, then those assumptions generate questions for the physicist – questions as to the underlying nature of these physical properties (for example, what is heat?), the underpinnings of the generalizations (for example, why do unsupported bodies fall?) and the identity of any further physical properties and generalizations not currently recognized by commonsense. However, even if science is born of commonsense, it doesn't follow that in its mature form a science shares many of the characteristics of its parent. Science is a self-conscious research endeavour that is driven by a relentless search for progress and the truth in a way that commonsense is not. As a result, science often exposes the limitations of commonsense: the parochial and inadequate nature of its conceptual scheme, the falsity of its assumptions, the limits of its explanatory powers and ambitions (Churchland, 1979). To see this, just think of how relativity theory and quantum mechanics have left commonsense physics behind. In short, then, individual sciences have a basis in commonsense but, in their mature form, have often moved a considerable distance from their parent.

In this section I have described a number of key characteristics of science in general. In virtue of its status as a science, we should expect cognitive science to share these characteristics. Thus we should expect cognitive science to (i) seek to uncover laws governing the workings of the cognitive mind; (ii) utilize such laws to explain cognitive phenomena; (iii) appeal to unobservable entities in its laws and explanations; (iv) utilize observation and experimentation; (v) confirm theories in a Duhemian–Quinean manner; and (vi) be born of yet to have moved beyond commonsense psychology.

5 The birth of cognitive science

So far I have been talking quite generally about cognitive science, but it is now time to be a little more specific about its origins and core commitments. Attempts to study cognition from a scientific perspective are hardly new. For example, David Hume ([1738] 1978) described the project executed in his work *A Treatise of Human Nature* as being an attempt to do for the mind what Newton had done for the external physical world. Indeed, some of the concepts and ideas employed by Hume have their analogues in contemporary cognitive science. Nevertheless, it is generally assumed that cognitive science came into existence as a discipline with a distinct identity only in the late 1950s and early 1960s with an intellectual turn known as the cognitive revolution.

The cognitive revolution was a revolt against a movement in psychology that dominated that discipline, at least in the English-speaking world, throughout the first half of the twentieth century. The movement in question was behaviourism, which had J. B. Watson (1913) as one of its pioneers and B. F. Skinner (1953) as its most prominent advocate in its maturity. Psychology as the study of mind and behaviour had become established as a discipline studied in universities in the late nineteenth century.[7] During its infancy, psychology witnessed several different approaches, chief among them being introspectionism and psychoanalysis. Introspection is the process of observing one's own mental states and so involves the mind turning its gaze in on itself. Hence, introspectionism is an approach in psychology that relies upon an individual's testimony as to what they have introspected. For a psychoanalyst such as Freud, the limitation of introspection is that much of what goes on within the mind is unconscious. Hence, psychology should involve postulating hidden phenomena from a less subjective third-person perspective.

For the behaviourist, neither introspectionism nor psychoanalysis constitutes an acceptable means of conducting psychology. The problem with them is their failure to be scientific, as they focus their concern on phenomena that are either subjective or hidden, with the upshot that their theories are unconstrained by observable evidence. In order to be scientific, the behaviourists thought, psychology must focus on observable phenomena and so turn its attention away from inner, mental phenomena.

The core concepts employed by the behaviourists were those of stimulus, response and reinforcement, and they can be understood by considering the work of B. F. Skinner. For many years Skinner worked with animals such as rats and pigeons where his central concern was to control and predict their behaviour. In a classic Skinnerian scenario, a rat would be placed in a box containing a bar that can be pressed. In this context, a stimulus is some feature of the environment that can impinge upon the rat. When the rat is first placed in the box it is not disposed to behave in any particular way in response to the stimulus constituted by the bar. Suppose that, behaving randomly, it presses the bar, and further suppose that a food pellet is dispersed as a result of this behaviour. The dispersal of the food pellet will act as a further stimulus that will serve to reinforce the behaviour of bar-pressing in response to being confronted by the bar in the box; that is, it will increase the probability or frequency of the rat making such a response to the stimulus in future interactions. In this way the rat's behaviour can be conditioned or trained. If the dispersal of the food pellet were made conditional on the presence of some other stimulus such as the flashing of a light, then the rat can be conditioned to press the bar only when the light flashes. And if the dispersal of the food pellet is further made conditional on the bar being pressed with some specific force, then the rat

can be conditioned to press the bar with that specific force when the light flashes. In short, the rat's behaviour at any given point will be a product of the current stimulus and its history of reinforcement and so can be controlled by reinforcement and predicted on the basis of knowledge of the stimulus and history of reinforcement.

Skinner's success in training animals led him to argue that human behaviour is not fundamentally distinct from that of rats and pigeons, in the sense that it too is a product of the current stimulus and history of reinforcement. Thus, in his book *Verbal Behaviour*, Skinner (1957) argued that language learning is a matter of reinforcement; so, for example, learning the English word 'horse' is a matter of being reinforced to behave by uttering 'horse' in response to horses.

In 1959 Noam Chomsky published a blistering review of *Verbal Behaviour* which is often characterized as a seminal event in the history of cognitive science, in that it revealed the bankruptcy of the behaviourist approach and the need for an alternative that focused on our inner mental workings. Chomsky's review is lengthy, subtle and intricate and gestures towards several of the ideas for which he was to become famous. I can hardly do justice to it in this context and, accordingly, will restrict myself to highlighting just a few of his points. First, Chomsky argues that Skinner's success in studying animals doesn't carry over to the study of humans. If we understand terms such as 'stimulus', 'response' and 'reinforcement' literally, then those terms don't apply to humans, and if we understand them metaphorically they just turn out to be mentalist terms in disguise. Second, when it comes to animals, sometimes their behaviour is learned but sometimes it is a product of a genetically controlled process of maturation. So, in advance of studying the relative contributions of the kind of factors highlighted by Skinner, on the one hand, and the organism's contribution to learning, on the other, it is plain dogmatic to adopt the behaviourist approach and ignore the internal workings of the organism under study. Third, with respect to language learning, it is clear that Skinner offers little insight. Children typically learn language quickly and reliably without being provided with much in the way of organized reinforcement. Hence, if we are to understand language acquisition, we must uncover the internal mechanisms that facilitate learning.

The general moral of Chomsky's review is that behaviourism must be replaced by a serious study of the contribution made by the internal workings of the human organism to behaviour and learning. This call did not fall on deaf ears and proved to be a huge factor in motivating the emergence of cognitive science. But it could only motivate the emergence of cognitive science because of the availability of a number of ideas that could be brought together to provide cognitive science with a body of core assumptions and methodological principles. Those ideas came from different

disciplines: some were of considerable vintage and some had a presence in more than one discipline; it was their coalescence that gave birth to cognitive science as a recognizable coherent discipline. I will now describe the ideas in question and how they cohere so as to ground cognitive science during its infancy.

A first idea was widely endorsed by philosophers during the modern period of the seventeenth and eighteenth centuries. This is the idea that the mind is populated by representations.[8] A representation is a symbol residing in the mind, and mental representations are involved whenever a person is in a particular mental state (for example, has a belief or a perceptual experience) or executes a mental process (for example, thinks or recollects a past event). Just as external representations such as spoken or written words, pictures or maps have meaning or content, so too do (mental) representations, and the content of a mental state is inherited from the content of the representation involved in having it. One prominent question about mental representations is why they have the content that they have. One answer[9] is that mental representations are imagistic in nature and their meaning is grounded in resemblance, so that, for example, the representation involved in having thoughts about horses means *horse* in virtue of resembling a horse.

The idea that representations are involved in perception and cognition was endorsed in the nineteenth century by many psychologists who were influenced by philosophers of the modern period. For example, Wundt was influenced by Kant and William James was heavily influenced by Hume.

However, the idea that representations are involved in cognition became unfashionable in the first half of the twentieth century in both philosophy and psychology. Of course the behaviourists would have no truck with representations, given their inner, mental nature. Moreover, they had a specific problem with them, namely, that the postulation of representations inadvertently involves a postulation of homunculi, little intelligent agents residing in the mind. This kind of objection was independently developed by philosophers such as Wittgenstein (1953) and Gilbert Ryle (1949),[10] from where it takes the following. External representations do not mean what they mean in and of themselves but rather mean what they mean in virtue of how we use and understand them. As contemporary philosophers put it, their meaning is derived rather than original (Searle, 1992). But such use and understanding is a mental phenomenon that exhibits intelligence. The meaning of a mental representation would similarly be a product of its intelligent use and understanding, so that to postulate such a representation would be inadvertently to postulate an intelligent agent in the mind (a homunculus) to provide it with its meaning. But that would raise the question of the basis of the intelligence of the homunculus. If we postulate further representations in its head, then an infinite regress looms, and if

we leave its intelligence unaccounted for, then our attempt to understand or explain cognition counts as circular. It is no good responding to this objection by pointing out that, if the representations are images, then their meaning is grounded in resemblance, a relationship that is objective and does not require intelligence. This is because, as Wittgenstein argued, resemblance is not an objective relation at all and resides very much in the eye of the beholder. For example, a picture of an old man walking up a hill resembles the phenomenon of an old man sliding down a hill just as much as that of an old man climbing a hill. What settles the ambiguity is how the picture is used and understood. In the face of such worries, the postulation of representations fell out of favour and was to be resurrected only with the emergence of a second idea.[11]

The second idea involved in the birth of cognitive science comes from logic and mathematics and is associated particularly with Alan Turing. Turing invented a very simple, abstract computing machine that has become known as a Turing machine. He didn't actually build any models of this machine; rather, he theorized about such machines and attempted to prove logico-mathematical theses about them. What he established in this way is that any computable function can be computed by an appropriate Turing machine. The term 'function' is to be understood in the mathematical sense as a mapping of numbers onto numbers. For example, addition is a function which maps the numbers 2, 2 onto 4, the numbers 6, 3 onto 9, and so on (whereas subtraction maps those same orderings of numbers onto 0 and 3 respectively). A computable function is one that can be mechanically computed in a finite number of steps. What this means is that there is a procedure that, given as input any items in the domain, can mechanically work out in a finite number of steps what value in the range the function maps those items onto.[12] A universal Turing machine is a Turing machine that can be programmed to mimic any other Turing machine. What Turing further established is that any computable function can be computed by a universal Turing machine.

The idea that Turing's work suggests is that the mind is a Turing machine or ensemble of such machines, so that cognition is a form of computation that involves the mechanical manipulation of symbols. To see how Turing's work suggests this idea we need to know a little bit more about Turing machines. A Turing machine consists of an infinitely long tape divided into squares. Such squares can be filled with a symbol such as '0' or '1' or remain blank. The machine also consists of a read–write head which scans squares on the tape one at a time. When it scans a square, the read–write head is able to detect whether it is blank or filled with a '0' or a '1'. How it responds to what it detects will depend upon its internal state, for the read–write head is capable of being in a number of distinct internal states. Whatever internal state it is in, the response will have a common

complex form that has a number of elements: it will involve (i) either leaving the symbol unchanged or deleting it and replacing it with some other symbol; (ii) moving one square to the left, one square to the right or halting; (iii) remaining in the same state or changing into some other state. A Turing machine's machine table is a specification of how it responds to any possible symbol for each of the internal states it is capable of being in. This table can be thought of as a series of rules or instructions that constitute the machine's program and which is hard-wired into the machine. A machine table is such that, when a Turing machine is fed a section of tape, it will go through a procedure of scanning the squares of that section, changing some of the symbols that it so detects and changing from one internal state to another as it does so, until finally it comes to a halt. What the machine will have done is taken input in the form of a section of tape with symbols printed on it and produced as output a section of tape with symbols printed upon it. These strings of symbols can be regarded as representations of numbers in the binary notation. Suppose a particular machine has a table such that, whenever it is fed a section of tape with n distinct numbers printed on it in binary notation, it responds by producing a section of tape with the number that is their sum printed on it in binary notation. Then the machine will be an adding machine that computes the addition function. In short, then, what Turing machines do is take symbols as input and produce symbols as output and, in so doing, compute mathematical functions.

What Turing's work suggested to philosophers and psychologists interested in cognition was a way of resurrecting the old idea that cognition involved representations without running into the homunculus problem. Solving mathematical problems is precisely the kind of thing we do when we cognize. As Turing machines solve mathematical problems by means of manipulating symbols without relying on intelligence or insight, they are hardly homunculi. So, if one thinks of the mind as a computer akin to a Turing machine, then one can have representations without homunculi.

A third idea that contributed to the birth of cognitive science relates to the metaphysics of the mind. Philosophers have long been concerned with the metaphysical question of the nature of the mind and its relationship to the physical. Earlier I described Descartes' substance dualist view and characterized it as problematic for anyone who wants to study the mind and cognition from a scientific perspective. During the nineteenth and twentieth centuries dualism waned in popularity among philosophers and materialism or physicalism became the orthodox view. But to say that the mind is physical or material isn't in itself to present a fully fledged theory. However, appropriate theories that could underpin cognitive science gradually emerged.

Even Descartes thought that there was a particular intimate relationship between the mind and the brain, in that the interface between the two

systems lay at the pineal gland in the brain. Developments in neurophysiology in the nineteenth century suggested an even closer relationship, especially with the discovery – by the likes of Broca and Wernicke – that distinct components of the brain were directly associated with particular cognitive capacities. In short, the more we came to understand about the workings of the brain, the clearer it became that there was a direct and intimate relationship between the mind and the brain. It is a short step from this to the conclusion that in some substantial respect the mind is the brain, for, one might ask, how could there be such an intimate connection were the mind not the brain? It is no good answering this question by postulating a causal relationship that implies a distinction between the two, as that raises the problem of how two fundamentally different kinds of thing (the physical brain and the immaterial mind) could causally interact.

However, the theory that the mind is the brain can take forms which are problematic for cognitive science. To see this, consider a popular mid-twentieth-century view known variously as type-type physicalism, the mind–brain identity theory and central state materialism.[13] The basic idea is that types of mental state are identical to types of brain state, or, alternatively, mental properties are identical to neural properties. Here the term 'identical to' means 'one and the same as'. These mental–neural identity relationships were conceived as being akin to such familiar identity relationships as those holding between water and H_2O and heat and mean kinetic energy. Thus, with respect to any particular type of mental state M (e.g., pain, the belief that coffee contains caffeine), answers to such questions as 'What is M?', 'What is it to be in M?' and 'What do all creatures that are in M have in common in virtue of which they are in M?' will appeal to a particular type of neural state or property. The type identity theorists held that these identity relationships needed to be discovered by scientists, but as an example they often claimed that pain is C-fibre firing.

The type-type identity theory implies that, in essence, mental states are neural states. Thus, if I describe one of my internal states as a desire for a cup of coffee, I haven't described that state in terms of what, from the scientific perspective, are its core properties. Rather, I have described it from a commonsense, pre-scientific perspective. That I do this is a product of my own ignorance, so that, when science corrects that ignorance and finds a means of understanding the brain from the neural perspective, talk of beliefs, desires, and the mental and cognitive in general should be abandoned and replaced by purely neural talk. This implies that we should be aiming for a neuroscience unsullied by mental and cognitive talk rather than a cognitive science that inherits our commonsense perspective of ourselves as cognizers.

In short, then, cognitive science needed a respectable metaphysical view about the mind that, though anti-dualist, contrasted with the type-type

identity theory. Such a theory developed in the early 1960s out of a recognition of a key failing of the type-type identity theory. This failing was identified by Hilary Putnam (1967), who accused the type-type identity theory of being chauvinistic in that it denied mentality of systems (be they earthly creatures, extraterrestrials or inorganic machines) physically unlike humans. For the theory implies that only systems with brains similar to ours are capable of sharing any aspect of our mental life. Putnam pointed out that this implication is highly implausible, as we are confident that many animals share aspects of our mental lives (for example, feel pain and have thoughts) even though they have central nervous systems that are very different from ours.

In the light of this failing, Putnam sought to develop a non-chauvinistic theory, and what he came up with was functionalism. According to functionalism, mental types are functional types. Here the term 'functional' is to be understood in causal rather than teleological terms. What the particular instances (the tokens) of any distinct mental type have in common in virtue of which they belong to that type is the functional or causal role that they play. It would perhaps help to consider an example of a mental state, namely pain, that isn't a cognitive state in order to illustrate the functionalist approach. Pain plays a distinctive causal role in our internal economies, a role that is specified by generalizations such as these: pain is caused by bodily damage or certain kinds of nerve stimulation; pain causes worry; pain causes moaning, groaning and crying; and so on. According to the functionalist, occupying this causal role is not a contingent feature of pain; rather, it is part of its essence. Thus to be in pain just is to instantiate an internal state that occupies the relevant causal role (which stands in the appropriate causal relations to inputs, outputs and other mental states). Generalized, the view is that each distinct type of mental state, including cognitive states, plays a particular distinctive causal role that is central to its identity and that to instantiate a state of any given mental type is to instantiate an internal state that plays the appropriate role in one's internal economy. Most functionalists are physicalists in that they regard minded systems as being complex physical systems and hold that the token states that occupy these roles are internal physical states of such systems. However, mental states are multiply realizable at the physical level. That is, for any given type of mental state M, the physical state that occupies the M role in one system may vary considerably from that which occupies the M role in another system. For example, the pain role may be occupied by C-fibre firing in humans, O-fibre firing in octopuses, and yet another state in a Martian with a silicone-based chemistry. Thus, the chauvinism problem that dogged the type identity theory is avoided.

An important feature of the causal roles of mental states is that they involve interacting with other mental states in distinctive ways; for example,

pain causes worry, a belief that one's body is in danger of being damaged, a desire for the pain to stop, and so on. Moreover, how a mental state manifests itself in behaviour will depend upon what other mental states one is in due to the fact that mental states cause behaviour in concert with other mental states (again, this is central to their causal role). This implies that one can't characterize a mental state without referring to other mental states.

In actual fact, Putnam appealed to Turing machines in developing his own version of functionalism. Particular Turing machines are defined by their machine table, which specifies the relations between possible symbolic inputs, symbolic outputs and internal states. Turing machines are multiply realizable at the physical level in that it is possible to build a given type of Turing machine out of different materials (all such physically divergent machines will satisfy the same machine table). With respect to the states of the machine, they are implicitly characterized in terms of their relations to inputs, outputs and one another and not in terms of the physical form they take in any particular concrete machine. Putnam argued that mental states are Turing machine states. In developing functionalism along these lines, Putnam made clear its affinity with the idea that cognition involves computation and so revealed its value as the metaphysical theory of the mind needed by cognitive science.

A fourth idea that played a key role in the birth of cognitive science relates to consciousness and self-knowledge. Descartes is associated, in addition to dualism, with the view that we have a thoroughgoing consciousness or knowledge of the contents of our own minds. On this view, if I am thinking a particular thought, I cannot but know that I am thinking that thought. From the commonsense perspective this view is very appealing, as we do seem to have a direct and immediate means of access to the contents of our own minds that we do not have to the contents of the minds of our fellows, which makes us authoritative about what we think.

This idea of the extent of our self-knowledge was, in effect, attacked by Freud in his postulation of the unconscious. For Freud, mechanisms of repression ensure that many of us have beliefs and desires that, because of their potentially repellent or unsettling nature, are pushed into parts of the mind to which we have no direct access. Nevertheless, such unconscious mental states can manifest themselves in behaviour in such a way that a skilled third party – such as a trained analyst – who has knowledge of the individual's history can uncover them. In short, Freud's work suggests that there is much in an individual's mind that is unconscious and that it is possible to know from a third-person perspective that an individual has a particular mental state even when they do not have such knowledge from the first-person perspective. Both Freud's work and the psychoanalytic movement that it spawned are in many ways highly controversial.

However, they have made the idea that we have unconscious mental states both familiar and popular. In fact, I would go so far as to say that Freud's work has impacted on commonsense in such a way that it is part of our contemporary commonsense vision of ourselves that some of our mental states are unconscious.

Although cognitive science and Freudian psychoanalysis are in many ways poles apart, the rise of the latter in the late nineteenth and early twentieth century popularized the concept of the unconsciousness and, in so doing, paved the way for the idea that many of the processes and states involved in cognition are unconscious. This idea is a core idea of cognitive science.

In this section I have described four ideas, namely, that the mind is inhabited by representations, that cognition involves computation, that mental states are functional states, and that much of what goes on in the mind when we cognize is unconscious. We have seen that some of these ideas very much pre-date cognitive science and come from very different sources. These ideas made cognitive science possible, and their coming together in the late 1950s and early 1960s both gave birth to cognitive science as a distinct and self-conscious discipline and provided it with its core theoretical assumptions.

6 Interdisciplinarity

Cognitive science is often described as interdisciplinary in nature, where the main contributing disciplines are usually identified as psychology (particularly cognitive and development psychology), Artificial Intelligence, neuroscience, philosophy and linguistics. Cognitive science is interdisciplinary for two reasons. First, the core ideas that came together at its birth originate from different disciplines, so that early cognitive scientists were thereby engaging with and combining ideas that have different disciplinary homes. Second, the core ideas provide a picture of cognition which implies that several traditionally distinct disciplines have major contributions to make. Cognitive psychology has a role to play as it is the branch of psychology concerned with cognition, along with those parts of developmental psychology that are concerned with how our cognitive capacities develop from birth onwards. Neuroscience has a role to play in revealing how cognitive processes are ultimately implemented in the brain and in placing constraints on higher-level theories of cognitive processing that are not directly concerned with their neural implementation. Artificial Intelligence is the project of programming computers to behave in a way that would count as exhibiting intelligence if done by a human individual. The strengths and limitations of particular AI programs can suggest hypotheses as to how we cognize and provide evidence concerning the plausibility of such hypotheses.

The role of linguistics in cognitive science is a little less direct and obvious and has to do with the work of Noam Chomsky and the approach in linguistics that he initiated.[14] We have already seen the significance of Chomsky's attack on Skinner and his demand that we concern ourselves with the contribution of the organism in studying linguistic behaviour. But he also went on to develop a view of language that implies that linguistics, in virtue of being the study of language, belongs to cognitive science. For Chomsky, language is not a social entity such as a body of social conventions or practices that exist externally to the minds of individual speakers of the language. Rather, an individual's language is an internal state of her mind. Thus, in studying language, the linguist is studying the mind. Moreover, Chomsky demands that linguists produce theories of language which explain language acquisition, so that linguistics is also concerned with cognitive development, in particular, the process by means of which a child moves from the state where she appears to have no grasp of language to that where she is a competent mature speaker.

Finally we come to the place of philosophy in cognitive science. Cognitive science is concerned with many of the issues that have been central problems throughout the history of philosophy, issues as to the place of the mind in the physical world, the nature of thinking, the relationship between mind and language, the role of learning in cognitive development, and so on. But this is not a straightforward case where philosophical speculation flourished in the absence of scientific insight only to lose its relevance at the hands of scientific advance. For, as we have seen, many of the core concepts and insights lying at the basis of cognitive science came from philosophy. Indeed, as will be made clear throughout this book, philosophy has continued to play such a role and so has an especially important role in adjudicating competing claims as to the importance of distinct bodies of data coming from different disciplines. Hence, philosophy is an important partner in the cognitive scientific enterprise.

7 Conclusion

In this chapter I have provided a general account of the nature of cognitive science and how it emerged in the late 1950s and early 1960s. I have indicated that in its early incarnations it was based on a commitment to the idea that cognition is a form of computation. In the next chapter I will examine in some detail this commitment to computationalism and how an alternative version of the computationalist vision emerged in the 1980s in the form of connectionism. I will then consider some recent challenges to this shared conception that internal representations play a central role in cognition.

2 Representation and Computation

1 Introduction

In the previous chapter we saw that, at its inception, cognitive science was committed to the idea that cognition involves the manipulation of representations by means of computation. These representations and representation-manipulating computational processes are implemented or realized in the brain. The upshot of this is that explaining a particular cognitive capacity, such as visual perception, object recognition, high-level reasoning, action planning, language development, understanding the mental states of another person, and so on, involves identifying the representations and computational processes involved whenever that capacity is exercised. But what general form do these representations and computational processes have? Within cognitive science there are two broad answers to this question, reflecting a divide between two competing approaches. The first answer is associated with an approach widely known as classical computationalism. As its name suggests, classical computationalism dominated cognitive science during the early decades of its existence.[1] The second answer is associated with an approach known as connectionism, which, though having its origins in work in the 1950s,[2] came to the fore only in the 1980s with the publication of Rumelhart and McClelland's connectionist 'bible' (Rumelhart et al., 1986; McClelland et al., 1986). Both connectionism and classical computationalism are very much alive today. However, recent years have seen the emergence of a family of approaches offering a radical alternative to both classical computationalism and connectionism, an alternative that questions the role of mental representations in cognition. In this chapter I will examine each of these three perspectives on cognition, beginning with classical computationalism.

2 Classical computationalism

A classical computer is a machine that takes structured language-like symbols as input and produces structured language-like symbols as output. This input–output profile is mediated by the application of rules which