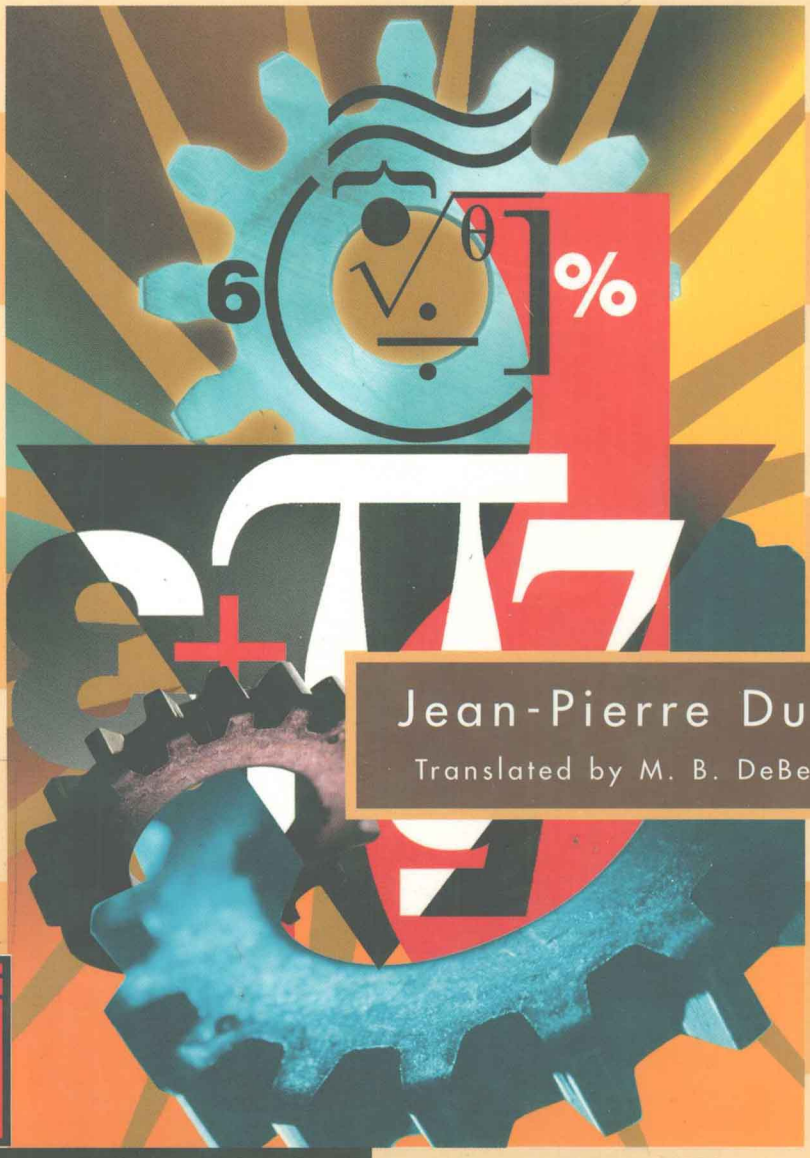
THE MECHANIZATION OF THE MIND

ON THE ORIGINS OF COGNITIVE SCIENCE



Jean-Pierre Dupuy

Translated by M. B. DeBevoise

NEW FRENCH THOUGHT

Jean-Pierre Dupuy

The Mechanization of the Mind

ON THE ORIGINS OF COGNITIVE SCIENCE

Translated by M. B. DeBevoise

NEW FRENCH THOUGHT

PRINCETON UNIVERSITY PRESS · PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press, 3 Market Place,
Woodstock, Oxfordshire OX20 1SY

Translated from Jean-Pierre Dupuy, *Aux origines des sciences cognitives*.

Library of Congress Cataloging-in-Publication Data

Dupuy, Jean-Pierre. 1941–
[Aux origines des sciences cognitives. English]
The mechanization of the mind : on the origins of cognitive
science / Jean-Pierre Dupuy ; translated by M.B. DeBevoise.
p. cm. — (New French thought)
Includes bibliographical references and index.
ISBN 0-691-02574-6 (alk. paper)
1. Cognitive science—History. I. Title. II. Series.
BF311.D84513 2000
153—dc21 00-032675

Publication of this book has been aided by the French Ministry of Culture

This book has been composed in Adobe Bauer Bodoni

The paper used in this publication meets the minimum requirements of
ANSI/NISO Z39.48-1992 (R 1997) (*Permanence of Paper*)

www.pup.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

The Mechanization of the Mind

NEW FRENCH THOUGHT

SERIES EDITORS

Thomas Pavel and Mark Lilla

TITLES IN THE SERIES

Jacques Bouveresse, *Wittgenstein Reads Freud: The Myth of the Unconscious*

Jean-Pierre Dupuy, *The Mechanization of the Mind: On the Origins of Cognitive Science*

Marcel Gauchet, *The Disenchantment of the World: A Political History of Religion*

Marcel Gauchet and Gladys Swain, *Madness and Democracy: The Modern Psychiatric Universe*

Blandine Kriegel, *The State and the Rule of Law*

Mark Lilla, ed., *New French Thought: Political Philosophy*

Gilles Lipovetsky, *The Empire of Fashion: Dressing Modern Democracy*

Pierre Manent, *The City of Man*

Pierre Manent, *An Intellectual History of Liberalism*

Alain Renaut, *The Era of the Individual: A Contribution to a History of Subjectivity*

Pierre Rosanvallon, *The New Social Question: Rethinking the Welfare State*

Jean-Marie Schaeffer, *Art of the Modern Age: Philosophy of Art from Kant to Heidegger*

In Memory of Jean Ullmo

For Heinz von Foerster

Preface

TO THE ENGLISH-LANGUAGE EDITION

My interest in cybernetics dates from 1976, the year I was fortunate enough to meet the founder of its second phase of activity, Heinz von Foerster, who, as the secretary of the last five Macy Conferences, was responsible for preparing the transcripts of these meetings. Elsewhere I have recounted the circumstances of our encounter, which was decisive in orienting my future research.[1]

At the time, von Foerster was nearing the end of a career almost wholly devoted to what he called "second-order cybernetics," whose home was the Biological Computer Laboratory he had founded two decades earlier at the University of Illinois. Thanks to von Foerster I came into contact with two scientists whose work on self-organization in living organisms had drawn its initial impetus from his own thinking: the French biophysicist Henri Atlan and the Chilean neurophysiologist Francisco Varela.[2] The wave of orthodox cognitive science had not yet reached the shores of France, and so it was the theory of self-organizing systems that first brought together French researchers interested in problems of cognition, not only from the point of view of the life sciences but also from that of the social sciences. Two conferences that I organized during the summer and fall of 1981, one at Cerisy-la-Salle in Normandy, the other at Stanford University, demonstrated the fruitfulness of this approach.[3]

In the meantime the École Polytechnique in Paris had decided, at the urging of the eminent physicist and philosopher of science Jean Ullmo, to set up a center for philosophical research having a strong epistemological component. I was called upon to organize and direct this center, which was to be known as the Centre de Recherche en Épistemologie Appliquée (CRÉA), with a staff chosen from among those who had contributed to the success of the two conferences just mentioned, in particular Atlan and Varela. It is entirely natural, then, that I should dedicate this book to the memory of my teacher, Jean Ullmo, and no less natural that I dedicate it also to Heinz von Foerster, whose contagious enthusiasm had a great influence upon the theoretical orientation that I was led to give CRÉA at the time of its founding.

Theories of self-organization were therefore part of the center's research program from the beginning. At the same time it was clear that,

despite their very great interest, they represented only one branch of cognitive science, and a quite marginal one at that. The authorities responsible for supervising scientific research in France had become aware, albeit somewhat belatedly, of the importance of a program of research that cut across long-established disciplinary boundaries, stimulating dialogue and debate among neurophysiologists, linguists, philosophers, economists, anthropologists, and psychologists. CRÉA quickly became one of the leading research centers for cognitive science in France, placing great emphasis on philosophical issues. Its originality by comparison with similar centers elsewhere in the world, particularly in the United States, derives from the fact that it provides a forum for quite varied approaches and paradigms. Both analytic philosophy of mind and language and naturalized phenomenology play an important role in the work of the center; neither one shrinks from looking to mathematical models for inspiration—in this respect following the example of the first cybernetics. As a result of circumstance and institutional politics, research in cognitive science in France came to be focused primarily on neurobiology and artificial intelligence, with the human and social sciences taking a back seat. Here again CRÉA occupies a special place by virtue of the importance it attaches to social cognition, and therefore to social and political philosophy, as well as to the social sciences proper, chiefly economics and anthropology. At the Macy Conferences, as we shall have occasion to see, these last two disciplines conceived their future development in terms of the mechanistic view of the world championed by cybernetics.

It was in this context that I felt the need to reflect upon the origins of cognitive science, all the more since my own work—but also to some extent the force of circumstance—had made me one of its architects. Although the history of science and ideas is not my field, I could not imagine adopting Alfred North Whitehead's opinion that every science, in order to avoid stagnation, must forget its founders. To the contrary, it seems to me that the ignorance displayed by most scientists with regard to the history of their discipline, far from being a source of dynamism, acts as a brake on their creativity. To assign the history of science a role separate from that of research itself therefore seems to me mistaken. Science, like philosophy, needs to look back over its past from time to time, to inquire into its origins and to take a fresh look at models, ideas, and paths of investigation that had previously been explored but then for one reason or another were abandoned, great though their promise was. Many examples could be cited that confirm the usefulness of consulting history and, conversely, the wasted opportunities to which a neglect of history often leads. Thus we have witnessed in recent years, in the form of the theory of deterministic chaos, the rediscovery of Poincaré's dazzling intuitions and early results concerning nonlinear dynamics; the return to macroscopic physics, and

the study of fluid dynamics and disordered systems, when previously only the infinitely small and the infinitely large had seemed worthy of the attention of physicists; the revival of interest in embryology, ethology, and ecology, casting off the leaden cloak that molecular biology had placed over the study of living things; the renewed appreciation of Keynes's profound insights into the role of individual and collective expectations in market regulation, buried for almost fifty years by the tide of vulgar Keynesianism; and, last but not least, since it is one of the main themes of this book, the rediscovery by cognitive science of the cybernetic model devised by McCulloch and Pitts, known now by the name of "neoconnectionism" or "neural networks," after several decades of domination by the cognitivist model.

The reasons for my own interest in the history of cognitive science are essentially philosophical. The fact that for a number of years I have directed a major center of research in cognitive science does not in any way imply that I am a convinced materialist. Although I am a rationalist, I nonetheless believe in the autonomy of the soul. I wholly subscribe to Thomas Nagel's view that "a solution to the mind-body problem is nowhere in sight."[4] Nonetheless, as I try to show in the first chapter of this book, the apparent inevitability of a materialist and mechanist solution was not due to any unanswerable philosophical argument but rather to the fact that it was rooted in a conception of the world and of knowledge that had a long history, going back to Hobbes and Vico. Accordingly, a purely philosophical critique of the mechanist materialism of cognitive science does not suffice to undermine its foundations. To do this, it is necessary to know where it came from—hence the present inquiry.

In 1983, the Centre National de la Recherche Scientifique (CNRS) asked me to head up a study group, under the auspices of its Science-Technology-Society program, to investigate the history of theories of self-organization. I put together at once a team whose members included Isabelle Stengers, a chemist as well as a philosopher and historian of science, and the author of a recent book with Ilya Prigogine, whose work on far-from-equilibrium systems had won him the Nobel Prize in chemistry six years earlier.[5] Together we were able to exploit the very rich work done at the Biological Computer Laboratory in Illinois, the mecca of the second cybernetics, Stengers retracing the history of the notion of self-organization in the physical and chemical sciences while I reserved for myself the task of analyzing the transactions of the Macy Conferences, to which I was able to obtain access owing to the generosity of Heinz von Foerster.

On learning that I had embarked on this project, von Foerster gave me the liveliest encouragement, for he was dismayed that the Macy Conferences had so far failed to arouse curiosity among scholars. This situation

was rapidly changing, however. An American historian of science, Steve Joshua Heims, had recently decided to reexamine the history of this period as well. I therefore got in touch with him, and we met for the first time the following year, 1984, in Boston. Heims and I remained in close contact during the entire time I worked on the original edition of this book. In 1985 I invited him to give a series of talks at CRÉA in Paris, and in our final CNRS report, transmitted in November that year in the form of two special issues of *Cahiers du CRÉA*, I included an essay by Heims that summarized the argument of his forthcoming book.[6] Without his careful historical research, I could not have successfully completed my own work.

The years went by. In France they were marked above all by the vigorous development of cognitive science and the emergence of CRÉA as an important center. In 1991 Heims published his long-awaited work on the Macy Conferences, *The Cybernetics Group*.[7] Three years later the results of my own research appeared under the title *Aux origines des sciences cognitives*.[8]

The situation of French thought in America, no less than that of American thought in France, is very much a part of the story told in this book. Since 1981 I have divided my time between research at the École Polytechnique and teaching at Stanford University. So that the American reader may better appreciate the perspective I bring to the history of cognitive science, let me quote from a talk that I delivered a few years ago at Stanford on the possibility—and the desirability—of going beyond the schism between the "two cultures."[9] Entitled "Beyond the Dualism between the Cultured Ignorami and the Hidebound Savants," it began by inquiring into what I call, with only mild exaggeration, the "schizophrenia" of American academic life:

> I have been asked to describe what it is like to be divided between two worlds. The division in my case is geographical, since I commute between California and France; but it is also cultural, since even at Stanford I find myself straddling the humanities, dominated by French poststructuralism, and philosophy and social science, dominated by American neopositivism. More fundamentally, I find myself divided—indeed torn—between a number of conflicting allegiances: between my background in logic, mathematics, and physics and my identity as a philosopher committed to the human sciences; between my need to think in terms of formal models and my deeply held conviction that literature is a superior form of knowledge to science; between the two ways of doing philosophy today: "Continental" philosophy—profound, rich, meaningful, but too often willfully obscure, elitist, and, at times, dishonest—and "analytic" philosophy—rigorous, egalitarian, democratic, but too often shallow and tedious—the one pointing toward literature, the other toward science; and, finally, between the narrow professionalism of

American academics, who devote themselves to knowing everything about "fields" so restricted that they often border on nothingness (the hidebound savants of my title), and the distinguished dilettantism of many French intellectuals, who tend to know almost nothing about everything (the cultured ignorami, or "foggy froggies").

Though I am torn, I refuse to be forced to choose between the Scylla of French intellectualism and the Charybdis of American academicism. From the unusual and rather uncomfortable vantage point I occupy on an American university campus, I observe the following oddity: on the one hand, students of literature are initiated into the mysteries of French-style "deconstruction," taught to celebrate the death of the human subject and to repeat ad nauseam that man is not his own master and that such awareness as he may have of his own affairs is severely limited by a sort of tyranny of the unconscious; while at the same time their fellow students in the economic, political, and cognitive sciences learn to systematically reduce social institutions to voluntary agreements between fully conscious and free individuals. It is fortunate for the stability of the system that these students practically never talk to each other—no more often, in fact, than do their professors.

Opposing the rationalist individualism of the American humanities and social sciences, including cognitive science, to the deconstruction of metaphysical humanism that animates the human sciences in France runs the risk of combining the worst aspects of French and American thought. Even if it is institutionally embedded in the heart of the American academy, such a distinction is not tenable philosophically. One of my aims in the present book is to establish just this, by showing that cognitive science represents both the highest expression of Western humanism and the source of its ultimate condemnation.

In addition to Steve Heims, I wish to express my gratitude to my colleagues at CRÉA, whose thinking about cognitive science, in both their published research and the many informal conversations I have been fortunate to have with them over the years, has greatly contributed to my work: Daniel Andler, Paul Dumouchel, Pascal Engel, Françoise Fogelman, Pierre Jacob, Maurice Milgram, Jean Petitot, Joëlle Proust, François Recanati, Dan Sperber, and Gérard Weisbuch; to those who have shown confidence in me by supporting my work and showing great patience when it has gone more slowly than I would have liked: Jean-Michel Besnier, François Gèze, and Dominique Wolton; to my research partners during the first phase of this project, from whom I borrowed a number of insights and discoveries: Isabelle Stengers, Pierre Lévy, and Pierre Livet; and, finally, to Henri Atlan, Francisco Varela, and Heinz von Foerster, who first introduced me to the ideas of cybernetics.

The ethical questions that are posed—and will continue to be posed—by new technologies of the mind, to which I briefly refer in the introduction, have been the object of much debate within a study group that Monique Canto-Sperber and I have recently formed at the École Polytechnique to examine issues in moral philosophy, the Groupe de Recherche et d'Intervention sur la Science et l'Éthique (GRISÉ). I wish to thank Monique from the bottom of my heart for all the encouragement and advice she has given me throughout the course of my work.

The English-language edition of this book is an extensively revised version of the original French edition. It was made possible by the interest shown in my work by three people: the psychotherapist and communication theorist Paul Watzlawick, a disciple of Gregory Bateson and, by virtue of this, an heir to the cybernetic tradition;[10] Michael Arbib, of the University of Southern California, an eminent brain researcher and one of the very few cognitive scientists who, as a faithful disciple of Warren McCulloch, has had the courage to declare—and to brilliantly illustrate in his own work—that the cybernetic paradigm remains a promising alternative, which artificial intelligence and cognitivism have not succeeded in rendering obsolete;[11] and Thomas Pavel, of the University of Chicago, whose efforts to give a fuller and more accurate picture of current French thought than the small portion of it that is known to the literature departments of American universities presently permits are remarkable. To each of them I wish to express my thanks.

Last, but not least, I would like to express my profound gratitude to my translator, Malcolm DeBevoise. "Translator" is not really the right word. Faced with a difficult and dense French text, he repeatedly insisted on clarifications while criticizing my arguments point by point. The result is a book that is very different from the original, clearer, more complete, more candid. In France it is customary for authors not to reveal their motives any more than is strictly necessary, for fear of multiplying openings for critics; in America, I have learned, candor is a mark of the respect an author owes his readers. In this, and in other things, Malcolm has been a demanding tutor.

Jean-Pierre Dupuy
Paris, October 1999

Contents

The Mechanization of the Mind

The Self-Mechanized Mind

From 1946 to 1953 ten conferences—the first nine held at the Beekman Hotel at 575 Park Avenue in New York, the last at the Nassau Inn in Princeton, New Jersey—brought together at regular intervals some of the greatest minds of the twentieth century. Sponsored by the Josiah Macy, Jr. Foundation, these meetings have since come to be known as the Macy Conferences. The mathematicians, logicians, engineers, physiologists, neurophysiologists, psychologists, anthropologists, and economists who took part set themselves the task of constructing a general science of how the human mind works. What brought them together, what they discussed, and what came of a collaboration unique in the history of ideas—these things form the subject of the present volume.

Every group of this kind adopts a code name as a way of affirming its identity. In the case of the Macy Conferences it was "cybernetics." Today this name has fallen out of fashion, to say the least. Since 1954 the project undertaken by the Cybernetics Group[1] has been carried on under a series of different names, ultimately coming to be known as "cognitive science." Why cognitive science today is ashamed of its cybernetic heritage is one of the chief questions I wish to address.

The Cybernetic Credo

The Cybernetics Group drew exceptional energy and passion from two convictions that were shared by most of its members and that were so novel at the time that the simple fact of defending them made one part of an elitist avant-garde, worshipped by some and demonized by others. These two convictions were based on logical and scientific discoveries that had been made in the immediately preceding decades, the consequences of which the members of the Cybernetics Group intended to exploit to the fullest. In very general terms, which will need subsequently to be made more precise, they held that:

1. Thinking is a form of computation. The computation involved is not the mental operation of a human being who manipulates symbols in applying