

The ape that made up its mind

Review of *An Artificial History of Natural Intelligence: Thinking with Machines from Descartes to the Digital Age* by David W. Bates. 408pp. University of Chicago Press.

Keith Frankish

There is something exceptional about human intelligence. Somehow we *homo sapiens* developed the ability to detach ourselves from the demands of the present and find mental space to reflect on the past, anticipate the future and adapt to new conditions. As a result we have transformed ourselves and our environment in ways that no other creature has done. Yet, like all other animals, we are essentially biological machines, equipped by evolutionary processes with specific cognitive capacities and shaped by individual histories of environmental stimulation. How, then, did we manage to free ourselves from the moment and act creatively? This is the topic of *An Artificial History of Natural Intelligence* by David W. Bates.

The author frames his investigation as a “rethinking of the human” in the face of its dismantling by, on the one side, the social sciences, which treat the human as socially constructed, shaped by economic, political and social systems; and, on the other, the cognitive sciences, which picture the human brain as a complex information processor whose operations are largely automatic. These trends have created a crisis, Bates argues. Our “automatic age” threatens to undermine the possibility of *decision* (by which, I take it, he means distinctively human intervention) and we need to recuperate a notion of human autonomy.

Bates accepts that the mind sciences cannot readmit any spiritual or transcendental element. Instead he locates the specialness of human intelligence in its *artificiality*. We didn't evolve to become special; we made ourselves special, and we made ourselves special through our technology (using the term in a broad sense to include culturally transmitted artefacts and skills of all kinds, including linguistic ones). Our species is distinguished by its symbiosis with technology—its technicity. Hence the inversion in the book's title: the history of natural human intelligence is a history of artifice. There may be no ghost in the machine, but there is, we might say, an *artificer* there. This perspective, Bates suggests, may enable us “to take back technology, to make the decision for humanity and against the acceleration of automation and automatic governmentality”.

He doesn't argue directly for this view, but develops it through studies of key thinkers and topics ranging from the seventeenth century to the present day. The book shows how the entangled themes of automaticity, autonomy and technology resurfaced

and mutated in different theoretical contexts as thinkers sought to find a place for human self-determination within the modern scientific world-view, continually conceptualizing human intelligence as both like and radically different from the technologies of their time.

The first part discusses early-modern philosophers from René Descartes to Immanuel Kant. Topics include Descartes's insights into our ability to interrupt and regulate our own automatic processes; the notion of a "spiritual automaton" (developed in different ways by Baruch Spinoza and G. W. Leibniz), which mirrors the bodily automaton, but follows its own logic; and eighteenth-century accounts of the ontogenetic development of rational judgement through internal processes of cognitive self-organization. (There is a detailed "epigenetic" reading of Kant.) Bates also shows how writers throughout the period stressed the importance of artificially training the mind and using external aids to augment it.

In the second part of the book, the author reviews nineteenth-century work in mechanical computing, neuroscience and evolutionary anthropology, discussing Charles Babbage, Ada Lovelace, Charles Sherrington, Charles Darwin, James Mark Baldwin and many others. He explains how attempts to mechanize reasoning highlighted the limits of computational thought, how neuroscientists linked creative thinking to the plasticity and undifferentiated nature of the human cerebral cortex, and how anthropologists located language, tool use and cultural processes at the core of human exceptionalism.

The book's third part focuses on early-twentieth-century work on biological organization and the role of technology in human life. Topics covered include the Gestalt psychologists' work on the nature of insight, Jakob von Uexküll's Kant-inspired theories of biological self-organization, neurophysiological studies of the brain's capacity to respond to shock and injury with adaptive systemic reorganization, and Alfred Lotka's view of artefacts as exosomatic organs that externalize and enhance our biological capacities and develop their own evolutionary dynamic. Bates also explores prewar German philosophical work by Ernst Cassirer, Max Scheler, Martin Heidegger and Helmuth Plessner, which located technicity at the heart of human phenomenology.

The final part of the book covers postwar work in cybernetics, computer science and artificial intelligence, as well as more recent work on neural networks and predictive processing. Familiar themes recur with a new urgency. Bates shows that the pioneers of AI recognized the limits of the systems they were building and saw that intelligent machines would need to possess a plasticity that allowed for adaptive reorganization in response to error, breakdown and interference. (There is a particularly insightful chapter on Alan Turing's views on this topic.) Postwar thinkers also reflected on the accelerating externalization of human mental abilities in exosomatic form. Some

dreamt of using technology to liberate human creativity (a dream encapsulated in Vannevar Bush's imaginary Memex machine), but the dangers of incorporating ourselves into vast networks of industrialized technology were also obvious. Many of the book's themes come together in an interesting chapter on André Leroi-Gourhan's views on the origins of tool use and the development of humans as a technical species.

Bates leaves readers to join the dots, but the picture that emerges is this: humans were able to escape the present by creating artefacts that acted as exosomatic organs, externalizing and extending their mental and physical capacities. These artefacts could be shared, replicated and refined, initiating processes of technological evolution that operated on a far shorter timescale than genetic evolution and became the dominant factor in human development. The biological basis for human specialness was not an increased neural specialization, but precisely the opposite—an increased plasticity and lack of differentiation in the human cerebral cortex, which made it open to developmental shaping by culturally transmitted resources.

Being human, then, requires not an escape from automaticity, but the cultivation of an automaticity that is both human-created and human-creating. The special character of the human, Bates says, lies in its unique position “poised between the *automaticity* of the organic and physical world and the *automaticity* of its own technical being”. The implication is that we must strive to preserve that balance, freeing ourselves artificially from our organic being, but maintaining control of the artificial processes that free us.

This is an impressive and important book. Readers will learn much from the author's sensitive and insightful readings of the scientific and philosophical literature, and philosophers of mind will benefit from his refreshing prioritization of explanatory concerns over ontological ones. Bates acknowledges that his survey is selective and personal, and there are, of course, omissions. There is, for example (to exhibit my own biases), no discussion of Daniel Dennett's work, which bears directly on Bates's concerns, or of dual-process theories in cognitive and social psychology, which focus on the distinction between automatic and controlled reasoning (a distinction I read as one between natural and artificial reasoning processes). But these omissions are more than compensated for by extended discussions of writers less well known to anglophone philosophers. (I am grateful to Bates for introducing me to Lotka and Leroi-Gourhan.)

This isn't a book for the casual reader looking for easy insights. It requires attention and effort. But that does not mean that it is only for academic readers. The topic is too important. Technological evolution is accelerating rapidly in response to economic and political forces, and to its own internal logic. (Consider what may happen to language—the fundamental human artefact—if we use large language models to externalize our writing skills. As future models are trained on the outputs of earlier ones, the obscure

logic of machine learning may transform our language into something alien and incomprehensible.)

As David W. Bates shows, we cannot remove ourselves from the world of technology; it is our world. But we can try to guide its evolution to protect the things we value. This is something we all need to think about, and *An Artificial History of Natural Intelligence* provides essential tools for the task.