

The Dis-automating Automaton: Intelligence Between Automaticity and Disruption

Alan Díaz Alva

Book review of David W. Bates's *An Artificial History of Natural Intelligence: Thinking With Machines From Descartes to The Digital Age*

Alan Turing's pathbreaking 1950 essay "Computing Machinery and Intelligence" is commonly read as a foundational text for the field of research known as artificial intelligence. In this text, so the standard interpretation goes, Turing conceptualises human thought as a mechanical process which can be simulated by a machine, and devises the famous "imitation game" (later called the "Turing test") which nonchalantly sidesteps thorny philosophical questions about what constitutes 'thought' and 'intelligence' by reducing the issue to whether a human can be tricked, in a text-based interaction, into believing that their machinic interlocutor is in fact another human. What was originally a curious thought experiment eventually became the official benchmark for the Loebner Prize, a competition—or a crude publicity stunt for some—in which various chatbots competed for the rather gimmicky title of being judged the most 'human-like.' Although the Prize is defunct since 2020 (and would arguably be redundant in our post-ChatGPT era), the notion of the Turing test remains a shorthand for two overlapping ideas: first, the mechanisable character of the human mind; and second, the rather circular definition of artificial intelligence as the successful imitation of human behaviour that is judged to be intelligent.

This understanding of artificial intelligence gleaned from Turing's paper has been influential as it has been polemic, setting the template for early cognitive science and AI research, shaping cultural imaginaries of artificial intelligence, and reinforcing the idea that the latter is tantamount to the mimicry of certain observable behaviours. To this extent, both boosters and critics tend to adhere to this interpretation of Turing. In his own reading of "Computing Machinery and Intelligence," however, David W. Bates draws entirely different conclusions. In the twenty-fifth (and, in my opinion, one of the most interesting) chapter of his most recent book enticingly titled *An Artificial History of*

Natural Intelligence: Thinking with Machines from Descartes to the Digital Age, Bates argues the following:

Rather than simply claiming to understand human thinking as just another mechanical process, so that it could be successfully simulated with another machine, [...] Turing was clearly interested in quite a different goal—namely, to conceptualize and construct *intelligent* machines. Or to put it another way, Turing was looking closely at human intelligence for clues of how to transform uninteresting machines into intelligent ones. What made this project so challenging for Turing is that he clearly recognized that the most interesting aspects of human intelligence clearly resisted any easy reductionist account, an account that would of course seem to be necessary for the purely mechanical simulation of the highest forms of human thinking.¹

In other words, while he recognised that certain functions of the human mind operate mechanistically, and can thus be reproduced artificially, Turing was more interested in those aspects of the human mind that *resist* mechanisation—an interest already found in his earlier work on computability, where he agreed with Gödel and the mathematical intuitionists in holding that the mind has a special capacity to ‘see’ proofs that cannot themselves be derived within the formal systems that generate them. Near the end of “Computing Machinery and Artificial Intelligence,” Turing offers a less frequently cited definition of intelligence, one which reflects these concerns: “Intelligent behavior presumably consists in a departure from the completely disciplined behavior involved in computation, but a rather slight one.”² This definition echoes Wolfgang Köhler’s influential interwar study of the intelligence of higher apes, where he argued that intelligent behaviour is expressed in the capacity to take ‘productive’ detours [*Umwege*].³ In his search for the possibility of a machine capable of productive deviations and divergences from established routines, Turing envisioned a machine “capable of *transcending its own mechanisms*,” a machine whose intelligence would stem from its “*failing to be a proper machine*.”⁴ In this way, he was already anticipating the paradigm of artificial intelligence that succeeded the first wave of so-called symbolic AI, namely, that of machine learning and today’s deep neural networks.

Fascinating as the re-readings of famous figures—ranging from Charles Darwin and André Leroi-Gourhan to Charles Babbage and Turing himself—and the various forays into more obscure ones—such as Duke George Campell or Lazarus Gieger—might be, it is

1 David W. Bates, *An Artificial History of Natural Intelligence: Thinking with Machines from Descartes to the Digital Age* (Chicago: The University of Chicago Press, 2024), 261.

2 Alan M. Turing, “Computing Machinery and Intelligence,” *Mind* 59 (1950): 459

3 Bates, *An Artificial History of Natural Intelligence*, 169–172.

4 Bates, *An Artificial History of Natural Intelligence*, 279.

often difficult to see what holds all twenty-nine chapters of *An Artificial History of Natural Intelligence* together, and to discern the theoretical intervention that the book, beyond its already immense value as a treasure trove in the history of ideas, is trying to make. Thus, I begin with Bates' reading of Turing not because it plays a particularly prominent role in the book itself. Instead, it illustrates what I see as the thread running through most (if not all) of the chapters, even if this thread can easily slip through our fingers as we follow Bates in his inspection of the various nooks, crannies, and hidden compartments of what constitutes a meticulously researched and painstakingly erudite (if often somewhat chaotic and even a bit frustrating) inquiry into the intellectual history of the relationship between *automaticity* and *disruption*.

Why call this an 'artificial history' of 'natural intelligence'? Bates situates himself squarely within the long tradition of thinking about the "originary technicity"⁵ of the human, a theme that recurs in German philosophical anthropology, contemporary strands of the philosophy of technics, the paleoanthropology of Leroi-Gourhan, and in the discipline of cognitive archaeology. The main idea is that, if there is something distinctive or special about human intelligence, it lies in its artificiality: rather than being immaterial or reducible to the neurophysiology of the brain, intelligence is mediated through, and has co-evolved with, artefacts, technology, and material culture more generally. From this perspective, as Bates notes in the introduction, the very notion of 'natural intelligence' is a misnomer (and the title of the book itself a catachresis): "There is no natural intelligence. All intelligence is artificial."⁶ Hence, while the title might give the impression that it is somehow a book about AI technologies, there is in fact little discussion of these apart from a handful of chapters. Instead, Bates displaces the notion of 'artificial intelligence' from its status as a label for contemporary technological developments to function as a heuristic device for his historical inquiry. He is interested in "the conceptualization of machine intelligence, not as defined in AI research as a form of imitation or simulation, but instead as the philosophical and anthropological theorization of the essential technicity of intelligence itself."⁷ In this way, the goal of the book is to trace a conceptual trajectory through "the question of intelligence as it emerges as the very mark of the artificial."⁸ It is worth noting that this trajectory unfolds in such a way that, somewhat unfashionably in today's zeitgeist, the recurring question of whether we can attribute intelligence to non-human beings—a question whose parameters are so slippery so as to often seem unproductive—is largely sidestepped. Bates is interested on what constitutes *human* intelligence, and, more importantly, what this question implies for how we understand what the human is in the first place.

5 See Arthur Bradley, *Originary Technicity: The Theory of Technology from Marx to Derrida* (Hampshire: Palgrave Macmillan, 2011).

6 Bates, *An Artificial History of Natural Intelligence*, 13.

7 Bates, *An Artificial History of Natural Intelligence*, 293.

8 Bates, *An Artificial History of Natural Intelligence*, 13.

If one had to name a philosophical influence underlying Bates' project this is undoubtedly Bernard Stiegler—a rather straightforward inference to make, considering that the book is dedicated to his memory. Stiegler's influence can be felt throughout, not only in the aforementioned concern with the artifactuality of all things human, but also in the conceptual nexus mentioned above, namely, that of automaticity and disruption (or dis-automation). Much of Stiegler's later thought is intensely preoccupied with the question of how to craft the tools, both theoretical and practical, necessary to foster pockets of dis-automating contingency in a society increasingly determined by algorithmic systems of prediction and anticipation; in other words, with the prospect disrupting the algorithmic automatisations of decision-making that is creeping on in every aspect of our everyday lives.⁹ Bates inherits this concern from Stiegler: a preoccupation with the possibility of crafting, or finding, "autonomy within automaticity."¹⁰

One of book's most interesting aspects is that, instead of circumscribing automaticity to the recent phenomenon of technological systems in the age of automation, Bates pieces together a longer story in which the human is viewed as the being forced to pave a way between the various types of automatisms that constitute its very nature. The human, as Bates notes in the introduction, is the peculiar being that finds itself at the intersection of *biological* and *technical* automatisms:

I hope at least to make the case that there is a critique of the automatic era that is possible from within the domain of technology and from within the domain of the *automatic* in particular. [...] What grounds critique in this space is the special character of human beings—poised between the *automaticity* of the organic and physical world and the *automaticity* of its own technical being. The human is not outside natural or artificial life. But thinking, I argue [...] is not possible except in that *gap* between the two.¹¹

Readers familiar with Stiegler will immediately recognise the strong pharmacological tone of this passage. The originary technicity of the human, or its fundamental prostheticity, is at once a vector of automatism and of potential disruption. Technogenesis through exosomatisation and externalisation is what enabled the *homo sapiens* to partially transcend the mechanisms of its biological nature by the technical construction of "cultural niches,"¹²

9 See Bernard Stiegler, *The Automatic Society Volume 1: The Future of Work*, trans. Daniel Ross (Cambridge: Polity, 2016).

10 Bates, *An Artificial History of Natural Intelligence*, 309.

11 Bates, *An Artificial History of Natural Intelligence*, 12–13.

12 Gary Tomlinson, *Culture and the Course of Human Evolution* (Chicago: The University of Chicago Press, 2018), 51; Oliver Schlaudt, *Das Technozän: Eine Einführung in die evolutionäre Technikphilosophie* (Frankfurt am Main: Klostermann Rote Reihe, 2022), 41.

while at the same time this becomes the source of new social and cultural automatisms. Some of these might be beneficial or perhaps even necessary—since “culture preserves itself through the *automatization* of learned behaviours”¹³—whereas others bind us to destructive sociotechnical systems that seem to escape our control.¹⁴

Bates’ book weaves a fragmentary narrative composed by episodes of variable length, from vignettes to article-length analyses, where intelligence oscillates between automatism and its interruption. If the Turing chapter narrates a case in which this struggle unfolds within the specific context of computing machinery—or more precisely, of their theoretical conception—the configurations of this nexus will vary widely across the book. The scope and variety of the material analysed is formidable, ranging from German Idealism to post-Darwinian evolutionary theory, psychology, and neuroscience. The structure of the book itself, however, is mostly chronological. Part I focuses on the canon of modern philosophy, discussing how figures such as Descartes, Spinoza, Hume, and Kant discussed various bodily and cognitive automatisms, as well as the ways in which the human mind is able to escape mechanistic determinisms. Part II deals with the nineteenth century emergence of the brain as an object of physiological study, and how this research interfaced with contemporary developments in evolutionary theory, and mechanical computing. What we find here is the positioning of the brain, with the discovery of the plasticity and undifferentiated nature of the human cerebral cortex, “as something capable, at one and the same time, of both automaticity and the very *disruption* of that automaticity,”¹⁵ along with various early reflections (which would foreshadow Turing’s own), by figures such as Babbage and Ada Lovelace, on the (im)possibility of computing machines to simulate this amphiboly. Part III (the most wide-ranging) covers a lot of ground in the twentieth century, from Gestalt psychology to philosophical anthropology, phenomenology, and further developments in brain research. Lastly, the fourth part overlaps chronologically with the previous one, while mostly focusing on postwar cybernetics, artificial intelligence, computer science, but also including one chapter devoted to Leroi-Gourhan.

It goes without saying that this brief outline cannot even come close to doing justice

13 Bates, *An Artificial History of Natural Intelligence*, 290.

14 Bates quotes Alfred N. Whitehead’s rather famous claim that “It is a profoundly erroneous truism [...] that we should cultivate the habit of thinking of what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them.” Alfred North Whitehead, *An Introduction to Mathematics* (London: Oxford University Press, 1948), 41–42. On the other hand, drawing from Marxist philosopher Søren Mau, we can argue that it is precisely the need for exosomatic tools and social relations—which constitute the ‘double mediation’ of social (re)production—that allows for another automatism to engrain itself at the core of the human metabolism, namely, that of the mute compulsion exerted by the logic of value. Søren Mau, *Mute Compulsion: A Marxist Theory of the Economic Power of Capital* (London: Verso, 2023), 89–122.

15 Bates, *An Artificial History of Natural Intelligence*, 123.

to either the sheer transdisciplinary breadth of authors and theories covered or to the multitude of resonances that emerge when they are placed side by side. It is left to the reader to discover these, either by reading the book from cover to cover or, perhaps more conveniently, by dipping in and out the various chapters as they see fit. So, instead of further expanding on its contents, I would like to conclude by commenting on another central aspect of Bates' work that, I believe, merits reflection.

As any historian will agree, while there are intersubjective and disciplinary rules and methods that guarantee a basic degree of objectivity, constructing a historical narrative always implies selection and interpretation. Things get a bit looser when it comes to intellectual history, where the domain of empirical factuality shrinks to a minimalist, source-bound objectivity, and the realm of hermeneutics gains considerable latitude. *An Artificial History of Natural Intelligence* is no exception, with the added difficulty that, as Bates himself acknowledges in the introduction, the project of assembling an intellectual history of the artefactuality of human intelligence is both vast and, until now, largely unexplored. Consequently, the constellation the book offers is inevitably fragmentary, selective, and contingent. While much of the text maintains the rather distanced tone of historical survey, the rationale behind the selection of its various components is clearly laid out in the more programmatic, bookending chapters.

In the introduction, Bates echoes the Stieglerian diagnosis of the automatic age in which our decision-making capacities have been dismantled, both practically through algorithmic technologies and their interplay with neuroscientific theories that reduce the human to neural automatism, but also theoretically by recent strands of contemporary theory that reduce the human to a node in networks of power and control. Bates therefore argues that we find ourselves in a "crisis of *decision*,"¹⁶ one which is technologically induced but also inadvertently reinforced by the erosion of the idea of human agency in the wake of theoretical trends such as poststructuralism, posthumanism, and actor network theory. Within this context, the theoretical wager propelling Bates' historical inquiry is the that the current conjuncture demands nothing short of "a rethinking of the human," and his book contributes to this purpose by seeking to "recuperate what we can call *autonomy* from within the historical and philosophical and scientific establishment of the automatic age."¹⁷

These statements throw into relief the kind of narrative—one often told against the grain—that *An Artificial History of Natural Intelligence* constructs from its materials: a story of human exceptionalism. This is not the familiar, long-criticised form of chauvinist exceptionalism grounded in essentialist, transcendental, or divine claims. Rather, in this picture, the human's exceptionality, insofar as it is "part of nature, yet forever outside of

16 Bates, *An Artificial History of Natural Intelligence*, 4.

17 Bates, *An Artificial History of Natural Intelligence*, 5–6.

the natural,” stems from its intrinsic artefactuality: technology is itself “the marker of the human exception.”¹⁸ The image of the human culled from the complex constellation of thinkers and theories assembled here, emerges as that of a being singularly capable of consciously (or at least intentionally) disrupting the automatisms that otherwise govern nature, including its own. Crucially, this capacity does not rely on some inscrutable internal source of volition, but often operates in ways that involve “the essential technical dimension of human thought itself.”¹⁹

One might agree with the attempt to delineate what distinguishes the human, as the technogenetic disautomating automaton, from other non-human animals. Against the anthropocentrism-panic that permeates many strands of contemporary theory, I strongly believe that drawing such distinctions is not tantamount, and does not necessarily lead, to rendering them hierarchical. Moreover, abstaining from making such distinctions can often have pernicious political consequences when it comes to understanding our place in the unfolding ecological catastrophe. That said, while reading Bates’ book, I couldn’t help wondering about alternative ways of narrating a story that might illuminate our algorithmically automatised present. There is a sense in which, I think, the ‘critical humanist’²⁰ position that focuses on the exceptional capacity of the human for disrupting or de-automating risks, if not backsliding into a version of what Leif Weatherby has recently called “remainder humanism”—a defensive theoretical attitude which situates the ‘human’ in the shrinking area of those things that computers cannot (yet) do—then perhaps remaining fixated on the “deconstructed-yet-active binary between the human and the machine”; a position that might be “diagnostically precise but lack attention to the ‘restricted economies’ of formal and predictive systems.”²¹ While Bates recognises the “alien” and “inhuman”²² logics of sociotechnical systems, throughout the book, the tendency is rather to look at them through the lens of prostheticity. To what extent are we living in a non-correlational and post-prosthetic phase of the history of technology and media?²³ Shouldn’t we consider the purported capacity of the human mind for autonomy within the mechanisms of computation alongside the becoming-autonomous of computation itself? Even if one is inclined to be somewhat sceptical of recent attempts to

18 Bates, *An Artificial History of Natural Intelligence*, 8, 11.

19 Bates, *An Artificial History of Natural Intelligence*, 9.

20 Bates, *An Artificial History of Natural Intelligence*, 6.

21 Leif Weatherby, *Language Machines: Cultural AI and the End of Remainder Humanism* (Minneapolis: University of Minnesota Press, 2025), 24.

22 Bates, *An Artificial History of Natural Intelligence*, 342.

23 See Mark B. N. Hansen, *Feed-Forward: On the Future of Twenty-First Century Media* (Chicago: The University of Chicago Press, 2015); R. Joshua Scannel, “Terra Ignota: Noncorrelation and Computational Agency,” in *My Computer Was a Computer—Catalyst: M. Beatrice Fazi*, ed. David Cecchetto (Noxious Sector Press, 2022).

marry pancomputationalism with what might be described as ‘panintelligentism’,²⁴ are we doing ourselves a disservice by reverting to a (revised) human exceptionalism? However, no single book can do everything, and, important as I believe these questions are, the fact that Bates does not address them is perhaps more a trade-off than a shortcoming.

Bibliography

- Aguera y Arcas, Blaise. *What Is Intelligence?: Lessons from AI About Evolution, Computing, and Minds*. Cambridge, MA: The MIT Press, 2025.
- Bates, David W. *An Artificial History of Natural Intelligence: Thinking With Machines From Descartes to the Digital Age*. Chicago: The University of Chicago Press, 2024.
- Bradley, Arthur. *Originary Technicity: The Theory of Technology from Marx to Derrida*. Hampshire: Palgrave Macmillan, 2011.
- Hansen, Mark B. N. *Feed-Forward: On the Future of Twenty-First Century Media*. Chicago: The University of Chicago Press, 2015.
- Mau, Søren. *Mute Compulsion: A Marxist Theory of the Economic Power of Capital*. London: Verso, 2023.
- Scannel, R. Joshua. “Terra Ignota: Noncorrelation and Computational Agency.” In *My Computer Was a Computer—Catalyst: M. Beatrice Fazi*, edited by David Cecchetto, 13–32. Noxious Sector Press, 2022.
- Schlaudt, Oliver. *Das Technozän: Eine Einführung in die evolutionäre Technikphilosophie*. Frankfurt am Main: Klostermann Rote Reihe, 2022.
- Stiegler, Bernard. *The Automatic Society Volume 1: The Future of Work*. Translated by Daniel Ross. Cambridge: Polity, 2016.
- Tomlinson, Gary. *Culture and the Course of Human Evolution*. Chicago: The University of Chicago Press, 2018.
- Turing, Alan M. “Computing Machinery and Intelligence.” *Mind* 59 (1950): 433–60.
- Whitehead, Alfred North. *An Introduction to Mathematics*. London: Oxford University Press, 1948.
- Weatherby, Leif. *Language Machines: Cultural AI and the End of Remainder Humanism*. Minneapolis: University of Minnesota Press, 2025.

24 See Blaise Aguera y Arcas, *What Is Intelligence?: Lessons from AI About Evolution, Computing, and Minds* (Cambridge, MA: The MIT Press, 2025). For Arcas, every living thing is a computer, and the only difference between, say, the intelligence of a bacteria and ours is merely one of scale.