



Available online at www.sciencedirect.com

**ScienceDirect** 

Cognitive Systems

Cognitive Systems Research 43 (2017) 119-124

www.elsevier.com/locate/cogsys

The Passage of a Transdisciplinary Movement through an All-Too-Brief Moment in Time. Review of *The Cybernetics Moment: or Why We Call Our Age the Information Age*, Ronald R. Kline. Johns Hopkins (2015). 336 pp.

Book review

Action editor: Peter Erdi

Peter Cariani

Center for Hearing Research, Boston University, United States

Received 30 December 2016; accepted 31 December 2016 Available online 27 February 2017

## 1. Review

The Cybernetics Moment focuses on the history of cybernetics and information theory in the US as mid-20th century research programs during the period of roughly 1940–1975. Kline's scholarly yet highly readable book pursues three major intertwined threads: cybernetics, information theory and the information age. The text mainly presents the interactions between major players in these fields in a roughly chronological order. For the most part, the book's trails revolve around Norbert Wiener, Claude Shannon, and the Macy conferences on cybernetics (1946–1953). To lesser degrees the discussions include Warren McCulloch, Walter Pitts, Warren Weaver, Gregory Bateson, Margaret Mead and British cyberneticians and information theorists W. Ross Ashby, Grey Walter, Donald MacKay, Gordon Pask, Stafford Beer, and Colin Cherry.

Author Ronald R. Kline is a senior historian of technology and Professor of Ethics of Engineering at Cornell University who has written two other books on technology and society. *The Cybernetics Moment* is an extensivelyresearched, well-crafted history of these fields that will be essential reading for scholars doing serious intellectual archeology of the post-war technology research in the US. Scholars will appreciate the seventy-five pages of the book filled with endnotes and copious citations, but will

http://dx.doi.org/10.1016/j.cogsys.2016.12.004 1389-0417/© 2017 Elsevier B.V. All rights reserved. bemoan the thin index and lack of a separate bibliography. Here is where the text-searchable, electronic Kindle edition could come in handy. The wealth of details that Kline has amassed will create many small surprises and revelations in the minds of those already familiar with cybernetics and information theory. But because relatively little effort is devoted to introducing the essential ideas of these movements, readers less familiar with these fields may find the going a bit tougher.

The book complements a relatively small but growing historical literature on cybernetics research that includes Steve Heims' John von Neuman and Norbert Wiener: From Mathematics to Technologies of Life and Death (MIT Press, 1982) and Constructing a Social Science for Postwar America: The Cybernetics Group, 1946–1953 (MIT Press, 1991), Paul Edwards' The Closed World: Computers and the Politics of Discourse in Cold War America (MIT Press, 1996), Flo Conway and Jim Siegelman's biography Dark Hero of the Information Age: In Search of Norbert Wiener, the Father of Cybernetics (Basic Books, 2005), Andrew Pickering's The Cybernetic Brain: Sketches of Another Future (University of Chicago, 2011), Eden Medina's Cybernetic Revolutionaries: Technology and Politics in Allende's Chile (MIT Press, 2011), and Tara Abraham's recent Rebel Genius: Warren S. McCulloch's Transdisciplinary Life in Science (MIT Press, 2016). Cybernetics also figures in histories of cognitive science such as Jean-Pierre Dupuy's On the Origins of Cognitive Science: The Mechanization of the Mind (MIT Press, 2009) and Margaret Boden's

E-mail address: cariani@bu.edu

impressive two-volume *Mind as Machine: A History of Cognitive Science* (Oxford University Press, 2006).

By comparison, aside from James Gleick's *The Information: A History, a Theory, a Flood* (Random House, 2011) there appear to be few books that cover the history of information theory apart from its mathematics. No full biographies of Claude Shannon seem to have been written thus far, with the closest attempt to date being Erico Guizzo's master's thesis *The Essential Message: Claude Shannon and the Making of Information Theory* (MIT, 2003). In Guizzo's thesis, I was surprised to learn about Shannon's successful use of mathematics to beat the house in Las Vegas casinos. Kline's coverage of Claude Shannon and his relationship to co-founder of the field of information theory Norbert Wiener makes a significant contribution to the early history of that field.

What exactly is (or was) cybernetics? This has been a perennial ongoing topic of debate within the American Society for Cybernetics throughout its 50-year history. Because the Kline book, being an academic history, does not attempt to suggest such a definition, what follows is my best provisional description.

Although the coining of the term "cybernetics" is widely attributed to Norbert Wiener because he used it as the title of his popular 1948 book, the word has a much older history reaching back to Plato, Ampere ("Cybernetique = the art of growing"), and others. "Cybernetics" comes from the Greek word for governance, *kybernetike*, and the related word, *kybernetes*, steersman or captain.

The movement that was to become cybernetics arose in the 1930s and 1940s as a social network bound together by common concepts about functional organization. Its ideas emerged from engineering, biology, psychology, neuroscience and the social sciences. The core concepts involved natural and artificial systems organized to attain internal stability (homeostasis), to adjust internal structure and behavior in light of experience (adaptive, self-organizing systems), and to pursue autonomous goal-directed (purposeful, purposive) behavior.

In my opinion, the essential basis of cybernetics involves circular causation in the form of feedback control (feedback to state) and adaptive self-construction (feedback to structure). The archetype of a cybernetic system is a purposive system with internal goals whose behavior is steered by feedback from processes that evaluate the effects of its actions vis-a-vis the goals.

The full scope of cybernetics has encompassed both simple feedback control mechanisms and more complex brainlike and society-like self-organizing systems in which the structure of the system itself is subjected to feedback control. In such purposive systems the ends adapt the means to fit current goals and situations. Cybernetics in this view is the study, design, fabrication, and use of purposive observer-actor/percept-action systems. Cybernetics as such is a transdisciplinary science of effective action that is essential for understanding functional organizations of different types of systems: biological organisms, minds and brains, autonomous robots, interactions between individuals and groups, human organizations, and societies. Cybernetics encompasses so-called second-order systems in which networks of self-modifying purposive systems mutually interact, cooperate and compete, and co-evolve to construct new internal mental structures. Second-order cybernetics is required for designing systems that can construct their own meanings and pursue their own purposes in an open-ended way. We shall need second order cybernetics or something like it if we want to build machines that can come to understand our meanings and purposes, such that we can interact with them on the basis of mutually evolved understandings, in a manner similar to the way which we humans deal with each other.

Cybernetics, from its inception, was concerned with effective action within and on the physical world, and it took its inspiration from wartime analog feedback-controlled fire control mechanisms, from homeostasis, adaptation and evolution in biological organisms, and from the capabilities of human and animal minds for self-direction and open-ended learning. It looked forward to possible worlds of sensory, motor, and cognitive prosthetics, brain-like goal-directed self-organizing devices, autonomous robots, mixed artificial-biological cybernetic organisms (cyborgs), and coevolving human-machine and human-human cooperations.

Information theory, on the other hand, was more narrowly conceived as a mathematical theory of efficient and reliable signal transmission, making it much easier to define and delimit its scope. As Kline argues, in contrast to cybernetics, its main proponents, Claude Shannon and others, also did considerable "boundary work" to keep its aims relatively pure.

Information theory was concerned with the quantification of variety in a signal that could take different possible alternative distinguishable states, also bringing into account the respective probabilities of those states. Information could be viewed epistemologically as a reduction in uncertainty by an observer (Wiener, Ashby) or ontologically as a measure of the information capacity (variety) inherent in a particular set of signals (Shannon). Communication theory involved efficient encoding of signals and quantification of the reliability of faithful transmission of messages through noisy channels. The signals could be discrete strings of symbols or time-series of continuous values. Information and communication theory eventually led to effective, optimal techniques for data compression and encryption.

Kline's book presents the histories of cybernetics and information theory in rough chronological order. The first chapter "War and Information Theory" opens with the publication of Wiener's 1948 landmark, widely-read book *Cybernetics, or Control and Communication in the Animal and the Machine*, whose title popularized the term "cybernetics" and whose contents included expositions of both measures of information and cybernetic feedback control. In the same year Shannon published his seminal paper on the quantification of information, entitled "A Mathematical Theory of Communication" in the Bell System Technical Journal (27(3): 379–423, 1948). The two theories are formally similar, despite differences in interpretation, such that Wiener and Shannon together are credited with founding this mathematical theory of information. However, there were substantial differences in their visions.

Shannon conceived of his theory of communication as dealing with only the recognition of the forms of sent messages (syntactics), eschewing all considerations related to their meanings (semantics) or purposes (pragmatics) of the messages. The constrained scope of the theory enabled the properties of the sender and receiver to be well-defined such that the theory could be formalized.

Wiener, MacKay, Miller and others sought a broader theory that could include the meanings and purposes of the message for the receiver (i.e., the effect of the message on the receiver's subsequent behavior and/or belief state and its effect vis-a-vis the goals of the sender or receiver). On the positive side, this points to a universalistic theory signals and symbols having syntactic, semantic, and pragmatic relations, opening the theory to application on all purposive, meaning-making systems. On the negative side, the more expansive scope entails many more challenges.

The second chapter ("Circular Causality") covers the Macy conferences on cybernetics and its inclusion of information-theoretic topics for discussion after 1949. Ongoing frictions between Wiener and McCulloch were evident, even before their complete parting of ways, which occurred for personal reasons shortly after McCulloch moved to MIT in 1951 (see Conway and Siegelman's Wiener biography for details).

The next two chapters ("the Cybernetics Craze" and "The Information Bandwagon") deal with the blossoming of the two movements from relative obscurity to mass popularization. The newly acquired popularity created stresses as various people and causes sought to identify themselves with these rising stars. Kline discusses the problems such movements can have in maintaining focus and rigor by establishing clear boundaries. Information theory as a field of study maintained a much more limited scope that was tied to specific, shared operational definitions and mathematical formalisms, whereas cybernetics with its more universalistic and philosophical orientation and its more diverse community resisted such restrictions.

The next two chapters ("Humans as Machines" and "Machines as Human") deal with cybernetics and information theory vis-a-vis the social sciences (Karl Deutsch, Herbert Simon, Talcott Parsons, Gregory Bateson), the emerging cognitive sciences (George Miller, Roman Jacobson), biological cybernetics (Heinz von Foerster), bionic prostheses (Norbert Wiener), and cyborgs (Manfred Clynes, Nathan Kline). These sections contain a wealth of details about these areas of research, circa 1950–1965. There are some notable gaps that are created from the Wiener- and MIT-centric focus of the narrative. Discussions of Ashby's systems theory, his homeostat device and his two books, *Design for a Brain* (Chapman & Hall, 1952) and *Introduction to Cybernetics* (Chapman & Hall, 1956) are given relatively short shrift compared to Wiener's *Cybernetics*. In his systems theory, Ashby outlined a general theory of empirical models, which was elaborated and refined by the late George Klir (*Architecture of Systems Problem Solving*, Plenum, 1985).

Self-organizing systems was perhaps the most visionary subfield of cybernetics research. Two conferences on *Self-Organizing Systems* were held in 1959 and 1961 that were funded by the Office for Naval Research (Marshall Yovits and S. Cameron, editors of proceeding volumes that were published in 1960 and 1962). In my opinion, along with the proceedings of the 1958 *Mechanization of Thought Processes* meeting, the two proceedings volumes of these meetings contain some of the most far-reaching papers of the era. These include papers on self-organizing electrochemical devices (Gordon Pask), neural networks (Frank Rosenblatt), adaptive neural timing nets (Donald MacKay), oscillatory neural networks (Peter Greene), as well as adaptive programming systems (John Holland) and neuromimes (Leon Harmon, H.D. Crane).

Finally, the many intersections of cybernetics and information theory with neuroscience are also largely missing. These have yet to be properly surveyed by historians. Many of the cybernetics people were psychologists, psychiatrists, and neuroscientists working directly in clinical, laboratory, and theoretical domains. The first Hixon Symposium on "Cerebral Mechanisms and Behavior" in 1948 are as cybernetic in spirit as any of the Macy proceedings. The proceedings volume (Cerebral Mechanisms of Behavior, Jeffress, L. ed., Wiley, 1951) contained McCulloch's paper "Why the Mind is in the Head" on oscillatory scanning mechanisms and networks of inverse-feeback loops. It also contained John von Neuman's paper "The General and Logical Theory of Automata" which covers wide ground, from analogical vs. digital coding principles to self-constructing selfcomplexifying self-reproducing machines.

The collected papers of Warren McCulloch in *Embodiments of Mind* (MIT Press, 1965) is perhaps the best place to begin for excavating linkages between cybernetics and brain theory. Efforts are underway to reprint the long out-of-print four-volume complete collection of his works (*Collected Works of Warren S. McCulloch*, Rook McCulloch ed., Intersystems Press, 1989). I once queried Jerry Lettvin about the contributions of cybernetics to neuroscience, and he pointed to the work of Werner Reichardt on elucidating opto-motor feedback control loops in the fly (see Reichardt's chapter in the volume *Sensory Communication*, edited by Walter Rosenblith, MIT Press, 1961), a highly successful research program that continues to this day. Other works in this tradition would be Grey Walter's *The Living Brain* (W.W. Norton, 1953, 1963), Michael

## Arbib's *The Metaphorical Brain* (Wiley, 1972) and Gerd Sommerhoff's *Logic of the Living Brain* (Wiley, 1974).

Early on information theory was applied to problems of neural coding and the quantification of the performance of human perceptual systems in signal detection and recognition tasks. In 1952 in their paper entitled "The Limiting Informational Capacity of a Neuronal Link" (Bull. Math. Biophys. 14, 127-135), MacKay and McCulloch used information theory to quantify and compare the respective informational capacities of neural coding schemes based on spike rates vs. temporal codes using reasonable neural parameters. A common error, present in Kline's book as well as many other historical treatments, is to assume that the pulsatile, all-or-nothing character of action potentials (spikes) implies that the brain operates using digital signals, in a manner distantly similar to a digital computer. The formal neural networks that had been proposed by McCulloch and Pitts in 1943, which implemented Boolean functions using pulsatile signals and summing-thresholding elements, certainly contributed to this view.

However, the MacKay and McCulloch paper belies the assumption that leading neuroscientists of the period held this view of brains-as-digital-signaling-systems. Whether a neural code is analogical, with continuously variable states, or digital, with discrete states, depends on how the spike train signals are used by the system (whether a signal is functionally analog or digital depends upon how it is interpreted by its receiver). The paper considers temporal codes based on the time durations between spikes, interspike intervals, that can vary continuously, and therefore function as analog signals. Such interspike interval codes are found in auditory, somatosensory, visual and other sensory systems. Neural codes based on average firing rates become progressively more continuous in character as the number of spikes in their temporal integration window increases. Again, depending on the nature of the receiving network, slightly different spike counts are likely to function more as continuous gradations, than as discrete, categorical alternatives.

From the wartime development of radar came signal detection theory, which involves optimal procedures for detecting a signal embedded in noise. William Siebert used signal detection theory to test whether optimal use of information present in prospective firing rate and temporal codes in the auditory nerve would be sufficient to account for the acuity of pitch perception, making it possible to rule out of hand candidate codes that were informationally inadequate to explain perception. Others used signal detection theory to quantify the quality of perceptual discriminatory capacities in psychophysical experiments (see Green, D.M., and Swets, J. A. *Signal Detection Theory and Psychophysics*, Wiley, 1966).

In Kline's view, the "cybernetics moment" lasted roughly from World War II to the mid-1970s, with the movement heading into decline by the mid-1960s ("Cybernetics in Decline"). Kline discusses at length the ins and outs of the use of the label "cybernetics," a debate which continues to this day within the American Society for Cybernetics. Interestingly, John McCarthy, an early proponent of the digital computational paradigm invented the term "artificial intelligence" to avoid associations with cybernetics and its analog feedback mechanisms. Cybernetics also inspired mass popularity through works of science fiction, which undermined its respect within the hard sciences. The lack of a clearly defined focus meant that anyone could project onto the movement whatever they liked and call it cybernetics, thereby gaining cachet for themselves but degrading the field's credibility. Kline describes the disdainful reactions of Wiener to L. Ron Hubbard's dianetics, which assumed that minds literally are computers, and Alfred Korzbski's general semantics, a psychiatric self-improvement/therapy program.

Ironically, as cybernetics was losing ground in the US, it boomed in the Soviet Union as a universal science (see Slava Gerovitch's *From Newspeak to Cyberspeak: A History of Soviet Cybernetics*, MIT Press, 2002). Kline references a purported "cybernetics gap" that created both opportunities and difficulties for continued US funding. This perceived gap seems to have provided the impetus for the Central Intelligence Agency (CIA) to fund a professional organization for cybernetics. In 1964 the American Society for Cybernetics (ASC) was founded, with funding from the CIA and Warren McCulloch as its first president. Norbert Wiener, who would have certainly opposed such an alliance, had died that spring.

Kline notes that McCulloch was fiercely anticommunist, but some of the details of his musical tastes in Tara Abraham's recent biography suggests that he had sympathies, not unlike George Orwell, to the anti-Stalinist anti-Fascist Republican loyalists in the Spanish Civil War. Jerry Lettvin, a close friend of Warren McCulloch and Walter Pitts and one of my mentors at MIT, also had obvious sympathies for the Spanish anarchists in that struggle (as freshmen he had us reading Peter Kropotkin's *Mutual Aid*, on the evolution of cooperation).

Kline discusses various other travails of the cybernetics movement, but there is not much in the book that details its gradual loss of military funding from the mid-1960s onwards. One structural reason why cybernetics declined in the mid-1960s may be related to its competition with and displacement by the rising research program of symbolic artificial intelligence (AI). Whereas cybernetics was conceived in terms of analog feedback mechanisms and autonomous robots, inspired by biological organisms and brains, pragmatist in philosophy, and oriented towards effective action in the physical world, symbolic AI was conceived in terms of symbolic computations and digital computers, inspired by logic and formal systems, platonist in philosophy, and oriented towards symbolic representations in virtual worlds.

By the end of the 1960s, US funding for research in biological cybernetics, bionics, self-organizing systems, neural

networks, and trainable machines had all dried up, in large part due to the Vietnam War era Mansfield Amendment, which prohibited military funding for long-range research. Active attempts to defund competing paradigms by influential proponents of symbolic AI, such as Marvin Minsky and John McCarthy, may have also played a crucial role, Heinz von Foerster, head of the major laboratory for cybernetics outside MIT, the Biological Computer Laboratory (BCL), reportedly believed this. Funding for symbolic AI work continued until it met its own first funding winter in the middle of the 1970s. It took two more decades for some of these allied research programs, such as neural networks, trainable machines, and parallel computers to be revived, and then only after the widely perceived failure of AI research on expert systems had occurred. The subfields of self-organizing systems and bionics were never revived.

Following the cessation of funding, the community that remained identified with cybernetics worked primarily in fields related to management cybernetics, conversation theory and human-machine learning systems, radical constructivism, social models, and family therapy. As Kline puts it, cybernetics re-invented itself under the banner of "second-order cybernetics," a.k.a. the "cybernetics of cybernetics" and the "New Cybernetics." Inspired by von Foerster, Mead, and Bateson, the focus of this community shifted to theoretical frameworks that take the role of the self-constructing observer-actor explicitly into account.

Oddly missing is an early document of this shift, Herbert Brun and Stephen Sloan's massive, graphically dense *The Cybernetics of Cybernetics* (Second, reprinted edition, 1995, Future Systems). Published in 1972 with funding from Stewart Brand, it is a visionary work similar in style to his own *1968 Whole Earth Catalog* and to Theodor Nelson's 1974 *Computer Lib/Dream Machines: You Can and Music Understand Computers Now* (1974, self-published; revised edition, Tempus Books of Microsoft Press, Redmond, WA, 1987). Brand's vision of "machine cybernetics" is contrasted with the "organic cybernetics" of Gregory Bateson in the last chapter of Kline's book ("Two Cybernetic Frontiers").

Key works of this era include Humberto Maturana and Francesco Varela's Autopoiesis and Cognition: the Realization of the Living (Cohen, R. S. and Wartofsky, M. W. eds. D. Reidel, 1973, 1980); collections of von Foerster's papers in Understanding Understanding: Essays on Cybernetics and Cognition (Springer, 2001); Gordon Pask's The Cybernetics of Human Learning and Performance: A Guide to Theory and Research (Hutchinson Educational, 1975); Conversation, Cognition, and Learning (1975), and Conversation Theory (1976); and Ernst von Glasersfeld's Radical Constructivism: A Way of Knowing and Learning (Routledge, 1996).

In the meantime, the classical techniques of feedback control, sensory substitution, and design of autonomous robotic systems continued to be refined in the context of traditional engineering disciplines, but these were no longer being developed under the rubric of cybernetics. Many ideas from cybernetics have likewise since been incorporated into disciplines of biology, psychology, neuroscience, and the social sciences, albeit shorn of their past transdisciplinary origins.

The third major thread of the book, on the etymology of the term "Information Age" is taken up in a chapter entitled "Inventing an Information Age." Its inclusion seemed to me to be a forced marriage of themes. The subtitle of this book *Why We Call <u>Our</u> Age the Information Age* echoes the title of the Wiener biography, *Norbert Wiener: Dark Hero of the Information Age*. But if one knows anything at all about cybernetics and information theory, this linkage brings immediate puzzlement, and the mystery of the reason for this juxtaposition is not addressed or alleviated until the eighth chapter, 200 pages into the book. I cringed the first time I saw the subtitle, because it feeds the popular, but deeply misguided, conflation of cybernetics with computers and mass information technologies.

Indeed, what *does* "The Information Age" have to do with either information theory or cybernetics per se? Cybernetics is about purposive, feedback controlled systems whereas information theory concerns the quantification, efficient coding, and reliable transmission of information. It seems on its face that neither cybernetics nor information theory is why we call our era "The Information Age." Why we call our times by this name has more to do with what the information means, how we use it, and how freely available it is to us for those uses.

The state we find ourselves in, with its massive availability of information, is much more closely related to Vannevar Bush's vision of digitized, automated, and free access to text, images, movies, sounds, and data that he outlined in his famous essay ("As We May Think", *The Atlantic, July 1945)* than it is to information theory. That theory per se, while being an enabler, is not the cause of our present information age. It's the meaning and purpose of the information, not its quantity, reliability, or compressibility.

Today we find ourselves awash in information precisely because of its many uses and also because of its free (or cheap) availability. Efficient data compression that has been made possible by information theory has exponentially expanded accessibility by lowering its cost, in turn making cell phones as well as digital music and video economically viable. But I would argue that Moore's Law, the falling price of its storage, and expanding availability and variety through massive networking have been at least as important as efficient coding in making information "too cheap to meter" and too useful to ignore.

This book is certainly a treasure trove of historical details that have been gathered into a coherent, sensible, and I think highly accurate narrative, but the work is meant for academic historians and those already familiar with cybernetics and its history. It will certainly be indispensable for those scholars doing serious intellectual archeology of post-war development of technology.

From my perspective, as a non-historian trying to learn and leverage deep principles from cybernetics for my own work, a major failing of this book and many other academic histories of technology is that these books rarely explicate the underlying concepts. It is probably asking way too much of the already Herculean feats of assembling, presenting and organizing the mass of historical data, but these histories rarely attempt to do double duty as intellectual histories. There is no one place in the book where even a provisional definition of cybernetics is laid out or pointers into the cybernetics literature are given. A consequence is that nothing is done to mitigate accumulated confusions about these fields, such as the common conflation of cybernetics with computers, AI, and all digital information technologies. As Paul Pangaro remarked in 1988, "Cybernetics is simultaneously the most important science of the age and the least recognized and understood. It is neither robotics nor freezing dead people. It is not limited to computer applications and it has as much to say about human interactions as it does about machine intelligence." (reprinted on the American Society for Cybernetics website, http://www.asc-cybernetics.org/foundations/definitions.htm).

Without a unifying conceptual basis, what one is left with is an uneasy feeling that cybernetics, for contemporary historians, is whatever Norbert Wiener chose to include in his 1948 book, and that *The Cybernetics Moment* is not about a movement or a cohering set of principles, but more about the historical period when Wiener's book had real currency. In this diminished view, the "Cybernetics" of *The Cybernetics Moment* is a reference to the book rather than to the movement.

What I would like to see from academic historians are volumes that one could give to a student or colleague who is curious about what cybernetics is (or was, or could

be), such that they will gain at least some of the profound insights of that intellectual ferment. As a corrective, in my opinion, the best book for conveying the conceptual breadth and flavor of the universalist, classical feedback control worldview would be Pierre de Latil's Thinking by Machine: A Study of Cybernetics (Houghton Mifflin, 1956). The best sourcebook for seminal papers in cybernetics and systems theory would be Walter Buckley's Modern Systems Research for the Behavioral Scientist (Aldine, 1968), and the best introductory textbook would be Gerald Weinberg's An Introduction to General Systems Thinking (Dorset House, 2001). Ashby's two books, Design for a Brain and An Introduction to Cybernetics, and Pask's An Approach to Cybernetics (Hutchinson, 1961) provide a much more coherent picture of cybernetics than Wiener's Cybernetics. My choices for comparable entry-points for information theory, would be John Pierce's An Introduction to Information Theory: Symbols, Signals, and Noise (Dover, 1980) and Colin Cherry's On Human Communication (MIT Press, 1966).

The purpose of remembering the past is to anticipate the future, as the saying goes. Histories of technology should be indispensable for those who seek to create new ones. I feel that the core ideas of cybernetics, which show us how to think about, design, and build purposive systems are essential if we are to proceed from the age of meaningless information to the age of meaning and purpose. We will certainly need cybernetic principles to understand minds and brains. I still hold the hope and belief that the Cybernetics Moment lies not in our past but in our future. And that the future will include an age of purposive systems that act autonomously to communicate and cooperate with us in ways similar to how we now do amongst ourselves. We are not there yet, not by a long shot, but eventually such an age will surely come about, once we clearly see our path forward to it.