2008

Cybernetics's Reflexive Turns

Klaus Krippendorff
University of Pennsylvania, kkrippendorff@asc.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/asc_papers

Part of the Communication Technology and New Media Commons, Engineering Commons, Epistemology Commons, Neuroscience and Neurobiology Commons, Other Social and Behavioral Sciences Commons, Philosophy of Language Commons, and the Sociology Commons

Recommended Citation

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/asc_papers/136
For more information, please contact repository@pobox.upenn.edu.
Cybernetics’s Reflexive Turns

Abstract
In the history of cybernetics there have been several attempts by cyberneticians to put themselves into the circularities of their theories and designs, invoking a shift from the cybernetics of mechanisms to a cybernetics of cybernetics. The latter is the title of a book chapter by Margaret Mead (1968) and of Heinz von Foerster’s (1974) edited compilation of articles on cybernetics. Foerster introduced the concept of second-order cybernetics which may have overshadowed or sidelined other reflexivities. I am attempting to recover four reflexive turns, describe their origin, implications, and suggest ways in which they continue what Karl Müller (2007) calls an unfinished revolution. These turns are not discussed here in their historical succession but as conceptual expansions of cybernetics.

Keywords
cybernetics, reflexivity, observation, participation

Disciplines
Arts and Humanities | Communication | Communication Technology and New Media | Engineering | Epistemology | Neuroscience and Neurobiology | Other Social and Behavioral Sciences | Philosophy of Language | Sociology

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/asc_papers/136
Cybernetics’s Reflexive Turns

Klaus Krippendorff

In the history of cybernetics there have been several attempts by cyberneticians to put themselves into the circularities of their theories and designs, invoking a shift from the cybernetics of mechanisms to a cybernetics of cybernetics. The latter is the title of a book chapter by Margaret Mead (1968) and of Heinz von Foerster’s (1974) compilation of articles on cybernetics. The latter introduced the concept of second-order cybernetics which may have overshadowed or sidelined other reflexivities. I am attempting to recover three additional reflexive turns, describe their origin, implications, and suggest ways in which they continue what Karl Müller (2007) calls an unfinished revolution. These turns are not discussed here in their historical succession but as conceptual expansions of the horizon of cybernetic inquiries and designs.

(1) Cognitive autonomy and the observer

There is widespread consensus that nervous systems are operationally closed, that humans are cognitively autonomous beings, responding in their own terms to what should be regarded as perturbations from their environment. This is the starting point of Ernst von Glasersfeld’s (1995, 2008) radical constructivism, Humberto Maturana and Francisco Varela’s (1980) biology of cognition (also Maturana, 2000; Maturana & Poerksen, 2004), Varela’s (1979) principles of biological autonomy, Gordon Pask’s (1961) cybernetics of the observer, and underlies von Foerster’s (1974, 1979) second-order cybernetics. The earliest mechanical simulation of cognitive autonomy is Ross Ashby’s (1952) homeostat. All of these cyberneticians draw conclusions from the acknowledgement that whatever is outside our nervous system is accessible only through that nervous system, and cannot be observed directly and separated from how that nervous system operates.

1. Gregory Bateson Professor for Cybernetics, Language, and Culture, The Annenberg School for Communication, University of Pennsylvania. Email: kkrippendorff@asc.upenn.edu
There are significant differences among these theorists. Ashby designed his homeostat in an effort to understand how a brain could adapt to its environment, known only through how it disturbs brain activity. Von Glasersfeld, educational psychologist with a Piagetian orientation, roots his radical constructivism in (subjective) experiences and describes how individuals learn by constructing realities, one on top of or in place of another, whose criterion is not truth but the viability of the observer. Maturana and Varela speak of the nervous system in mechanistic terms as strictly structurally determined, presupposing the autopoiesis (continuous self-production) of the organism. For Maturana (Maturana & Poerksen, 2004, p. 35), observers act in the awareness of distinguishing what they are observing. Von Foerster, the most pronounced critic of the enlightenment project’s objectivism, suggests that believing one could observe the world without the realization of being an observer is but an illusion. His reflexivity consists of observers entering their domain of observation and describing their own observing, not what causes them. Like von Glasersfeld, von Foerster considers the reality we see as a construction or invention. Being not determined by external causes, the reality that is seen could be otherwise, thus manifesting human agency. Despite these differences, acknowledging cognitive autonomy leads to the common conception of the rather abstract notion of an observer who attends to his or her own observing, understands his or her understanding, and describes that process in his or her own terms.

I should mention that the idea of second-order cybernetics developed only slowly. In his Cybernetics of Cybernetics or the Control of Control and the Communication of Communication, von Foerster (1974) merely distinguishes first and second-order cybernetics by defining the former as the cybernetics of observed systems, and the latter as the cybernetics of observing systems, offering no further articulations. Much of his reflexivity emerged later, in provocative aphorisms and interviews (Foerster, 1979, 1992; Foerster & Poerksen, 2002, pp. 109-111). In his 2003 collection of essays on cybernetics and cognition, he suggested that the essentials of a theory of the observer had been worked out (Foerster, 2003, p. 285).

I agree with Karl Müller's (2007) characterization of second-order cybernetics as an unfinished revolution in the conception of science and of the world more generally. I am suggesting, though, that further development of cybernetics’s reflexivity is constrained by three commitments inherent in second-order cybernetics.

1. The commitment to cognitivism – taking the autonomy of individual cognition as the exclusive source of explanation, in effect ignoring the social nature of human beings, reducing it to mutual accommodations of individual conceptions (Glasersfeld, 2008), or merely affirming that human beings live in language
2. The commitment to the primacy of observation – identifying human beings as observers who seek to understand their experiences of observing (the world), rather than, say, as engineers of technology, governors of their society, encouraging meaningful practices of living, including for others
3. The commitment to rely on a *representational language* (Rorty, 1989) – consisting of *descriptions*, not necessarily of reality, but acceptable within a community of (scientific) observers.

It should be noted that cognition, the process presumed to underlie the attention paid by observers to their observations, is a highly individualistic construction. This is not to say that second-order cybernetics has no place for language, which is a collective artifact. Indeed, von Foerster playfully extended Maturana’s proposition that “everything said is said by an observer” by adding “everything said is said to an observer” (Foerster, 1979). This echoes Maturana and Varela’s (1988) conception of scientific observers as providing explanations and hypotheses that are testable by the phenomena the environment continues to bring to the observer’s attention and acceptable by a community of scientific observers. However, the conception of language underlying these assertions is descriptive, not of an external world but of the individual understandings shared within a scientific community.

Moreover, descriptions are not meant to act on what they describe. For example, Maturana and Varela’s conception of autopoiesis as an account of the process of living of living systems has no necessary relation to how living beings operate. Living beings existed before the concept of autopoiesis was invented in the 1970s and are not assumed to have changed since the inception of this explanation. Evidently, descriptions do not change the phenomena they describe only an observing theorist’s understanding of them.

I maintain that these three commitments keep second-order cybernetics with one foot in the enlightenment project for a positive science, which is committed to providing truth-verifiable descriptions, in second-order cybernetics, however, from the perspective of the rather abstract concept of a self-reflecting observer.

(2) Participation – use, design, and conversation – instead of observation

The second reflexivity arises by shifting attention from the limiting concept of an observer to that of a participant. Consider three situations.

*First.* Just like von Glasersfeld, Maturana, Varela, Pask, and von Foerster, Ashby too was preoccupied with cognition, with the mysterious complexity of the human brain. However, already in his *Design for a Brain*, Ashby (1952) positioned himself not as an observer but as an experimenter with mechanisms he treated as black boxes (Ashby, 1956), that is, presuming no knowledge of their design. In this capacity, he acknowledged that acts of experimentation rendered experimenters as participants in a larger system composed of them and their objects of interests. Of this larger system, he concluded, experimenters have access only to the observable consequences of their own actions. The remainder of the system is left unobserved, hypothetical or hidden from view. Thus, experimenters do not merely reflect on their own observations, they must reflect on their participation in systems they actively vary for observable
effects—always acknowledging that an understanding of these systems can never be complete, that one’s horizon is always expandable through further actions.

I feel compelled to note that participants rely on what observers regret as undesirable and unintentional observer effects or biases. Heisenberg’s uncertainty principle, for example, suggests that whenever observations affect the observed, one can no longer determine with accuracy the properties of the object of one’s attention before it was observed. He showed that in quantum physics, one can determine either the location or the momentum of particles, but not both simultaneously. Heisenberg’s principle states a limit of the enlightenment ideal of accurate observation. Ashby, however, did not subscribe to this ideal. He acknowledged without regret to be the source of disturbing the object of his attention and constructed models that accounted for both the variations he introduced and the changes he could subsequently observe. By modeling the observable consequences of his actions, Ashby entered the system he constructed as an active participant, not merely as a detached observer, describing or explaining his observations.

Human-computer interfaces provide a good example. What users come to know about their personal computer is limited to the options they were willing and have the time and resources to explore. For competent users, who have learned to navigate a computer along desired paths, that limit may be unproblematic. It may be noticed however, when experiencing their interface to break down, observing different users, or learning something about its architecture.

Second. Designing and building artifacts amounts to a participation that is quite unlike using or experimenting with them. While users’ competence results from developing conceptions that are adequate for interacting with what they face, designers’ competence is demonstrated by actions that realize intended artifacts in their environment, artifacts whose known structure and functionality conforms to how their designers conceived of experiencing them. Here, human agency shifts from cognitive accommodation to an existing environment to adapting that environment to one’s conceptions. Introducing known structures and functionalities into their environment, designers have a far more extensive grasp of what their artifacts can do than users could acquire by examining the options they afford. Despite this extensive grasp, designers too face the limitations of their horizon. Of a working artifact not everything is determined by design. Artifacts may break down unexpectedly, exhibit structurally unanticipated behaviors, may be used in unintended ways, or develop a life of their own, especially when others are involved. Thus, what can be known by design is always only part of the system involving designers, users of designs, and the environment they experience.

With an interest in brains, which are not designed, Ashby built mechanical artifacts to simulate certain brain functions, convinced he could learn something about the structure of the brain from the knowledge of how his models were designed. But the structure of a model is always only one of several hypotheses of the makeup of the modeled.
In his epistemological explorations of black boxes, Ashby encountered informational or computational limits of experimentation or use by building artifacts that exceeded these limits. Subsequently, von Foerster (1984) called the latter “non-trivial machines.” The distinction between trivial and non-trivial machines points to fundamental differences between observation, use, and design (Krippendorff, 2008b). Clearly, observers are at the mercy of how the world manifests itself: unfolding sequentially. Only the most trivial machines turn out to be intelligible by observation. Non-trivial machines, contemporary computers, for example, are designed and manufactured for use and the act of designing leaves their designers with insights into possibilities of use far exceeding any users’ ability to explore all alternative paths. Ashby and von Foerster are clear in relating the distinction between trivial and non-trivial machines to human abilities – observation, experimentation or use, and design. Their accounts did not, however, extend to the language that made this comparison possible, which is my next point.

*Third.* The use of language, languaging, is essentially embodied in their speakers and participatory. Most cyberneticians are quick to affirm the importance of language, for example in drawing distinctions and understanding. When von Foerster (1979) playfully extended Maturana’s proposition “everything said is said by an observer” by adding “everything said is said to an observer,” he embedded the description of observations in a communicational paradigm. Von Glasersfeld, following Piaget, demonstrated how the grammatical structures of linguistic expressions reflect cognitive constructions much like Benjamin Lee Whorf (1956) did. Maturana (2000) defined language as the consensual coordination of consensual coordination of behavior, introducing coordination as the defining function of language. However, detailed conceptions of structures in languaging, what a language does and how it is altered in conversations is rarely found in the discourse of cybernetics.

To me, a cybernetic conception of language needs to be built on conversation, echoing Pask (1975) but avoiding his computer metaphors and embracing instead Maturana’s consensual coordinations, Austin’s (1962) performatives, C. Wright Mills’s (1940) and Marvin Scott’s (1968) accountability, Searle’s (1968) speech acts, and, most importantly, Ludwig Wittgenstein’s (1958) language games (see also Krippendorff, 2008c). I take conversation as

1. Constituting itself in the practices of its participants, what they say and do, especially affirming their participation in the conversation
2. Involving naturally embodied and mundane verbal, gestural, and mediated interactions or turns, not requiring any theory of conversation to engage in one
3. Self-organizing. Participants mutually regulate each others conversational practices without outside intervention, instruction, or management
4. Preserving dialogical equality among participants (e.g., Holquist, 1990, pp. 40-66), which entails contributing in response to the contributions by others and maintaining mutual accountability, the possibility of holding each other accountable and a commitments to offer accounts when requested (Krippendorff,
2008c) or deemed constructive, especially including the willingness of (re)articulating one’s understanding of other participants’ understanding (second-order understanding, see Krippendorff, 1993). This would exclude the role of (detached scientific) observers.

5. Producing a diversity of jointly constructed artifacts: coordinated understandings, plans of action, designs, and novel use of language

Cybernetically informed family systems theory, for example, has embraced this reflexive turn by considering therapists as dialogically equal participants in conversations with their clients (Anderson, 1997). I have proposed a conception of social organizations as reconstitutable networks of conversation through which members of organizations move more or less freely (Krippendorff, 2008c).

As Gregory Bateson (1972) taught us, we always participate in the circuitry of the world. Acknowledging one’s participation in a larger system, whether as explorers, designers, or constituents of social formations, is a reflexive turn that reveals reality, its parts, and the self as interactively or dialogically constructed, and admits that individual knowledge is necessarily incomplete, expandable with efforts. By contrast, observers are more likely to take the position of masters of their observations and building monological models of their world. One could say that participation is based on know-how, observation on know-what (Krippendorff, 2008b).

(3) **Realizing human agency in the relativity of discourses**

Ashby argued that cybernetics had its own foundation, distinct from established disciplines like physics, psychology, and general systems theory, to name but three, and began to develop that foundation in his *An Introduction to Cybernetics* (Ashby, 1956). While his project could only be accomplished in language—he wrote the text—he did not see his writing as an essential part of this foundation. I consider Ashby’s *Introduction* as constituting an inter-disciplinary discourse for cybernetics.

But what is discourse? Dictionaries tend to define discourse an extensive body of writing. I am relying on a more inclusive definition.

To me, *discourse* is systematically constrained conversation (Krippendorff, 2009, pp. 217-234). It exhibits five defining properties:

1. Like conversation, a discourse *produces artifacts*, be they textual matter (designs, theories, and explanations), particular practices (rituals, legal decisions), a system of tangible artifacts (technologies), or constructions of reality (ranging from facts to coherent universes)
2. It requires a discourse community whose members are committed to maintain, rearticulate, reconstruct, elaborate, or invent these artifacts
3. Members of its discourse community institute their recurrent practices, for example, in the form of requiring the use of particular methods, providing regular publications, manning educational programs, certifying of membership, and setting standards for communication within the community
4. It maintains a boundary at which decisions are made regarding who or what belongs and who or what does not. This binary distinction may be made on various grounds, a paradigm, set of axioms, epistemology, or the kind of artifacts it produces.

5. It justifies itself – its artifacts and practices – to other discourses, so as to sustain it in the continuous flow of discourse practitioners, available resources, and reality constructions supplied by or to other discourses.

Physics, for example, is a discourse that constructs a universe that is coherent and can be observed and experimented with by the methods of physics, which presumes, among other things, that this universe is causally determined, has no agency, and does not understand how it is being investigated. Within its boundary, the discourse of physics is self-organizing in that it is physicists who educate new physicists and determine whether the theories proposed by physicists are consistent, valid, and further the standing of the discourse. Psychology is another discourse. It creates models of individual human beings’ behavior, now largely of human cognition. Biology is a discourse. It creates theories purporting to explain the operation of living systems as distinct from systems of interest to other disciplines. Cybernetics is a discourse as well – cyberneticians decide who is one, settle on their own conceptual repertoire in conversations, and introduce circular systems into their environment – although in the past, it failed to develop strong institutions and its inter-disciplinarity and widespread fascination with its artifacts has left its boundary open for other discourses to forage what suits them.

Owing to the ability of humans to travel across discourses, discourses are not entirely independent from one another. Some discourses mine other discourses for their ideas, as computer science, artificial intelligence, and biology did from first-order cybernetics and Luhmann’s (1995) sociology has done from second-order cybernetics, autopoiesis in particular. Some colonize other discourses as cognitive science is attempting to do regarding linguistics and psychology; or the chemical/pharmaceutical discourse does regarding psychiatry – encouraging psychiatrists to administer drugs in place of engaging in talk. There are also discourses that assume foundational status, for example, by insisting that everything real is physical, physics considers itself foundational to all other “less perfect” or “underdeveloped” natural sciences. Similarly, by claiming that all social phenomena are ultimately economic transactions, economics assumes a foundational role in relation to other social sciences. In view of these interactions, discourses that do not claim to be foundational to other discourses often struggle to maintain their autonomy.

One manifestation of human agency is the ability to reflect on one discourse from the perspective of another. Foundational discourses tend to prevent that reflection, equating the artifacts they construct with the universe that exists. Physicists, for example, construct so-called laws of nature that are presumed to underlie nature as observed. This conviction discourages physicists to go outside their discourse and experience the agency they always have in moving from one discourse to another.
unable to acknowledge their artifacts as human constructions. I am invoking Bateson’s (1979) insight that information arises from multiple descriptions, Werner Ulrich’s (1996, 2000) boundary critique in critical systems thinking, Wittgenstein’s (1958) account of meaning as the awareness of alternative ways of “seeing as,” and suggest that the belief in a singular truth amounts to being entrapped in one privileged discourse and hence blind to it.

Biology is a discourse, of course. It encourages conversations that follow its (bio)logic for explaining the operation of living beings. Maturana and Varela (1980) take autopoiesis – a biological artifact invented to explain the process of living of living beings – as a prerequisite for all other features of living beings: reproduction, cognition, adaptation to their environment, and speaking a language. The belief that autopoiesis is primary, prior to, or a prerequisite for the human use of language distinguishes biological foundationalism (reiterated in Maturana, 2000, 2008) from non-foundationalist acknowledgements that humans have the option of discursively constructing themselves, each other, their bodies, their worlds, including their vocabularies and conversations. Not that Maturana and Varela are wrong. Their biology is impeccable. However, taking the reflexivity of Heidegger (Dreyfus, 1992) to heart, one can also argue that it is our attention to our being-with-others that makes us, our body and relationships to each other, real for us. Biology is limited to what its discourse highlights. Our body does more, is far richer than our explanations of it. Furthermore, it presupposes linguistic competencies to enter the discourse of biology and rearticulate its artifacts, including autopoiesis, comprehension of which is subordinate to linguistic competence and entirely optional to the living of living beings, as above suggested.

To preserve this reflexive turn, I am suggesting that cybernetic discourse not be grounded in foundationalist conceptions such as of physics, biology, cognitive science or semiotics, that it should rely instead on a conversational conception of human beings, replace observers by participants and the representational conception of language by constructive, interactive, or performative ones, which are capable of acknowledging human agency and the ability of redesigning reality.

(4) Inclusion or social contextualization of cybernetics

Von Foerster cites Margaret Mead’s (1968) talk on the cybernetics of cybernetics as the starting point of second-order cybernetics. In personal communication, von Foerster claimed to have clarified Mead’s “somewhat confused ideas.” However, rereading her talk (at which I was present) reveals that she may have had something quite different in mind.

As an anthropologist, Mead was of course keenly aware of the role of language in directing the attention of its speakers. She characterized “cybernetics as a way of looking at things and as a language for expressing what one sees” (1968, p. 2). This quote could be interpreted in support of the cognitivism in von Foerster’s second-order cybernetics. But then she expressed concerns about what the “language of
cybernetics” did to society, how society becomes organized as a result of implementing cybernetic systems, calling on cyberneticians to assume responsibilities for how these systems are changing society in unprecedented ways.

Mead’s talk was delivered during the cold war. She saw how the Soviet Union embraced cybernetics quite unlike the U.S. did, fearing “the possibility that the Soviet system may become totally cyberneticized, in the technical sense, as a way of controlling everything within its borders and possibly outside, with thousands of giant computers linked together in a system of prodigious and unheard-of efficiency” (1968, p. 3). In contrast and regarding the West, she spoke of the “possibility of using cybernetics as a form of communication in a world of increasing specialization” in the hope of being able to develop sophisticated ways of handling the complexities that cybernetically designed systems are introducing in society (1968, pp. 4-5). She saw that the automation of social institutions made society less governable—a phenomenon that we now know in retrospect eventually undermined the Soviet Union and networked individuals and institutions in unprecedented ways, opening cyberspaces for us to explore but also causing global instabilities for which we have no ready-made answers. Mead’s proposal to account for the organizational/technological consequences of implementing cybernetic ideas was one of expanding the discourse of cybernetics to embrace the context of its social consequences. It amounts to suggesting that cybernetics take up what sociologists push to their margins as the unintended consequences of actions, and economists consider the externalities of their theories.

Warren McCulloch (1945) had discussed a similar reflexivity before Mead did, suggesting that physicists account for themselves in terms of the discourse of physics, which would, he thought, transform physicists into neurophysiologists. Mead, however, was an anthropologist, not a logician, and advocated a move from the early first-order cybernetics to where it becomes embodied. Keenly aware of how the application of cybernetic discourse was complexifying society, she turned McCulloch’s logical inclusion problem into an inter-disciplinary challenge.

The reflexive turn that Mead advocated renders cybernetics as a socio-technological discourse that acknowledges the discourses of other disciplines, addresses the social embeddedness of their artifacts, and calls on cyberneticians to not only attend to how they contribute to the ongoing transformations of society but also be accountable for their effects. Mead recognized that all sciences are part of society. One would hope that all sciences therefore address the consequences of their products. However, the convenience of partitioning the world into separately manageable fields, the widespread illusion of objectivity, and the blindness of established sciences to their own discourse, leaves cybernetics as a unique inter-discipline, able to address the very socio-technological complexities it gives birth to. One might add that at the time Mead made her proposal, the cybernetic revolution, commonly but mistakenly called the information revolution, had barely begun and the complexities it created were difficult to foresee.
The emerging cybernetic epistemology

These four reflexive turns jointly encourage a new epistemology:

- an epistemology of participation in social systems under continuous reconstruction (rearticulation and redesign) by their human constituents, able and willing to hold each other accountable for what they contribute, say and do, and how they move through discourses and networks of conversations among them.

This cybernetic epistemology (theory of knowing) does not claim to represent anything outside of it (an ontology or description of reality, absent its knowers). It must merely be livable by knowledgeable actors in and on their environment. It acknowledges various roles of language – networks of conversation and multiple discourses – in how the human constituents of social systems coordinate themselves in the creation of artifacts that constitutively enter and alter the very systems of which they are a part.

Specifically, it recognizes that promoting theories, proposing courses of actions, designing and building artifacts commensurate with one’s understanding of a system of interest, and communicating with other constituents of that system changes the system as understood.

By locating abstractions and meta-perspectives in the language used in conversations and discourses, which are self-organizing, it is an epistemology that is embodied in human practices.

As such, it abandons the unattainable and unethical God’s eye view of the universe, common to enlightenment scientists and the adherents of universalism, for the admission that one can experience only one’s own contributions to a continuously evolving social reality.

Conversely, it suggests that observing social systems from their outside precludes access to what constitutes them: participation. Knowing a social system is demonstrable only by constructively participating in it.

It invokes a communication ethic that grants all human constituents of social systems a measure of dialogical equality (participation), adequate agency (choices) and accountability (see Krippendorff, 2008c), cognitive autonomy (holding individually constructed and potentially unlike conceptions), and an ability to communicate with one another (coordinating each others’ understandings).

It affirms that social organizations are repeatedly reconstitutable by their members and, once enacted, under continuous reconstruction by their constituents.

It favors the construction of realities that preserve the circularities of participation in networks of conversation.

While participation is common and unavoidable in everyday social life, the difference between an unreflective being-with-others and cybernetically guided participation lies not in claiming privileged access to reality and truths but in
accepting accountability to those who may have to live in the realities one’s discourse is projecting, supporting, or bringing into being.

The reflexive turns that various cyberneticians have been taking for some time—not always fully articulated or formalized—suggests a convergence on a definition:

Cybernetics – an inter-disciplinary discourse that brings forth radically reflexive realities.

The attention to reflexivity, which includes a host of well-known circular causalities, self-references, recursions, and autonomies, moves cybernetics from a science of “control and communication” to an effort by the cybernetic community to come to grips with the challenges that its discourse is creating in the world.

References


Forsythe, K. (2008). *City*. 60 cm x 60 cm, acrylic on canvas with encaustic.