

A (Cybernetic) Musing: Ashby and the Black Box

Ranulph Glanville¹

Introduction

Over the years that I have written this column, I have often referred to the *Black Box*. I believe the Black Box is such a powerful device, that it is time to explore it seriously, in its own right; for it allows us that most magical of tricks, a way of acting confidently with/from the unknown/unknowable. From this position and with this device, we can build an understanding of a world regardless of whether we can honestly regard it as “real” or existing knowably independent of our knowledge of it. We can also build descriptions of the world that, ultimately, are based not in presumed knowledge but in ignorance.

I shall do this by means of an argument based, in the first instance, on the work of Ross Ashby, as far as I know one of the few scholars to write extensively and seriously about what the Black Box is. Before that, I will provide some context. In a later paper I will explore the consequences and use of the Black Box, and some of the extensions to and critical sequitors of the arguments presented here. This paper is intended to explain what the (classic) Black Box is. The interested reader will understand from this argument how radically misused the notion has often been, even before this is argued and explained in the follow up paper.

At the end of this paper, I shall show how, in developing his understanding of the Black Box, Ashby joins that band of early cyberneticians whose cybernetics was already, in effect what became called second-order cybernetics.

Which Black Box?

The term *Black Box* is used confusingly. Before I begin, I need to clarify what I intend to write about.

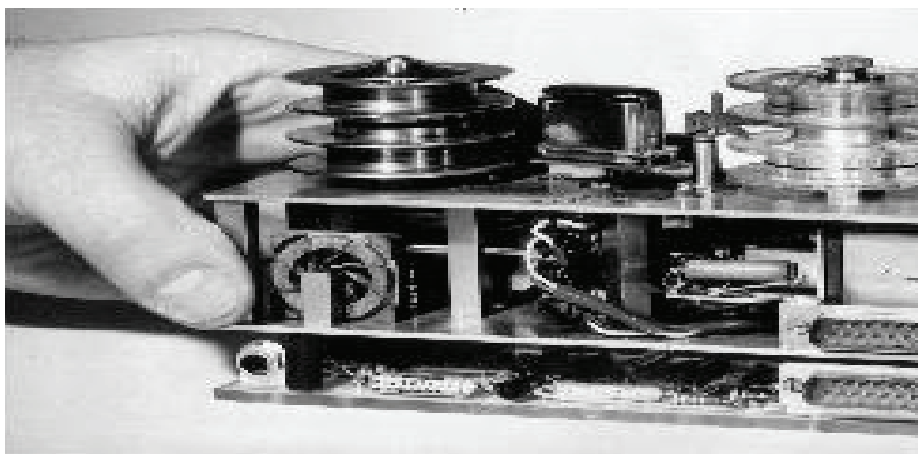
The use of the term I will explore is associated with the great Scottish mathematician and theoretical physicist, James Clerk Maxwell.² Apparently, Maxwell first used the term, in this sense, in 1871 in his *Theory of Heat*.³

1. CyberEthics Research, Southsea, UK; Email: ranulph@glanville.co.uk

2. Interestingly, a google search on the terms Black Box Maxwell scarcely mentions this (earliest) use of the Black Box, in the first 200 returns.

3. I owe this information to Dr Albert Mueller of the Institute for Contemporary History at the University of Vienna. Although Maxwell is regularly credited with the invention of this use of the term *Black Box* finding a reference is difficult indeed, and so I am specially indebted to Dr Mueller for this information.

Amongst the uses we will not consider but which I mention in passing, is that of the Black Box as an alternate/synonym for a radiant Black Body in thermodynamics, and the recent development of the Black Box flight recorder. This device – properly known in the first instance as a *Flight Memory Unit* – was created in response to the Comet jet crashes of the early 1950’s by Dr Dave Warren of the (Australian) Aeronautical Research Laboratories in Melbourne, in 1957 and has since become universally adopted not only in aeroplanes, but in many other vehicles. Although this is the best known use of the term Black Box, it is, in fact, usually bright orange!



Dr Warren’s original “Flight Memory Unit” prototype (source *The Black Box Flight Recorder* (2000, downloaded from: www.kidcyber.com.au, April 22, 2007)

The American psychologist B. F. Skinner also used the Black Box concept in a manner that has led to major misunderstandings—as I shall explain in the next section.

A short historical prelude: The Black Box and behaviour

While it is difficult to access Maxwell’s origination of the concept, we are not left without guides. Both Wiener and Ashby mention the Black Box and its significance. In fact, Ashby does more than mention it: he devotes his paper “General Systems Theory as a New Discipline” (Ashby, 1958/1991; a paper without a single reference in it)⁴ almost entirely to a discussion of the Black Box; and there are 17 acknowledged entries dated between 1951 and 1956 in the notebooks that make up his journal.⁵

-
4. Ashby studies the Black Box extensively in his *Introduction to Cybernetics* (Ashby 1956). However, those studies are more technical and less philosophical. I have not found the account in *Introduction* so compelling as the account in the paper I quote from so extensively in this column.
 5. Ashby’s journal was given by his family to the British Museum. A web site is currently being finalised. I am fortunate enough to have been given access to this site, but cannot reference it, since it is not yet in the public domain. This is one reason I do not quote from or reference the journal.

In fact, it was a remark at the beginning of the Black Box section of the paper that first really excited my interest. Here is Ashby introducing the Black Box notion:

there is ...no finer approach than that given by the so-called Problem of the Black Box. It arises in electrical engineering, but its range is really far greater—perhaps as great as the range of science itself. (Ashby in Klir, 1991, p. 252)

If anything, I have come to believe that Ashby under-, rather than overstates the range of application of the Black Box, for here he talks of it as an existing machine that cannot be opened, whereas I believe we are all familiar with a Black Box approach as part of the conceptual development we go through. I believe it is interesting and valuable to look at the way we learn about the world as a Black Box type of exploration.⁶ If I want to argue this, I need do no more than refer to the amazing way I can understand (at least, present an understanding of) how you think, and even how I do, without recourse at all to any form of mechanical examination of what goes on in the brain. As we shall see, Ashby also understood this.⁷

This may seem a bizarre, if familiar position: and perhaps particularly odd coming from a second-order cybernetician. It's the position of B. F. Skinner and his school of behaviourist psychology; and, indeed, Skinner did write of the Black Box (e.g., Skinner, 1984). But is this really so strange? Cybernetics is, after all, a subject deeply tied in to the concept of behaviour, and of explaining (accounting for) behaviour through mechanism.

What makes the difference between a cybernetic position and Skinner's is not the behavioural interest, not the use of the Black Box, but how we interpret the nature of what the Black Box approach can tell us. Skinner may have chosen to present the behaviour associated with a Black Box as mind-less and mechanical (automatic), but for the cybernetician who takes into account the observer there is a big difference. The behaviour is not of the Box alone, but is shared between Box and Investigator (observer). As Ashby writes:

for what *we* are considering can be viewed as a compound system, composed of Box and Investigator. He acts on the Box when he stimulates it, and the Box acts on him when it gives him a dial-reading as observation. Thus each acts on the other. (Ashby in Klir, 1991, pp. 252–253, emphasis in original)

Furthermore, the Black Box, as we shall see, is an unknown. The behaviour that results from the “interaction” (Ashby's word) between Box and Investigator is

6. However, see later Ashby quote from *Introduction to Cybernetics*.

7. Luhmann also came to describe the experience of communication between 2 entities through the image of the Black Box. Soeren Brier, our editor, provides this quote: “Two Black Boxes, by whatever accident, come to have dealings with one another. Each determines its own behaviour by complex self-referential operations within its own boundaries. Each assumes the same about the other. Therefore, however many efforts they exert and however much time they spend, the black boxes remain opaque to one another” (Luhmann, 1995, p. 109.). Luhmann's views were developed from Ashby's (see next Ashby quote, pp. 252–253), and my own work which I will introduce in a later paper.

described (generally) by the Investigator: We do not know what happens in the Box. Furthermore, we cannot know what is “going on in the head” of anyone—but we are getting ahead of ourselves.

Finally, for Skinner it seems the explanation of the behaviour should be taken as a precise 1:1 mapping of the mechanism – the map was the territory. In cybernetics, we try not to confuse these two, but to recognise that an explanation is precisely an explanation and not what is explained.

Introducing the Black Box

Ashby:

We imagine that the Investigator has before him a Black Box that, for any reason, cannot be opened ... In its original, specifically electrical, form, the problem was to deduce the contents in terms of known elementary components. Our problem however is somewhat wider. The questions we are interested in ... are such matters as:

What *general* rules of strategy should guide the exploration, when the Box is not limited to the electrical but may be of any nature whatever?

When the raw data have been obtained from the outputs, what operations should *in general* be applied to the data if the deductions made are to be logically permissible?

What can in principle be deduced from the Box's behaviour and what is fundamentally not deducible? (Ashby in Klir, 1991, p. 252)

And then:

Whether the Box is behaving in a machine-like way does not require study of its internal details ... to find something of the connections does not demand the opening of the Box. (Ashby in Klir, 1991, p. 253)

Later:

Thus, Black Box theory leads us naturally into the theory—most important for those who study the brain—of the mechanism that, for whatever reason, is not wholly accessible. (Ashby in Klir, 1991, p. 256)

And:

You will have noticed that a good deal of what I have had to say has not been concerned directly with the Black Box but rather with what the Investigator can or cannot achieve when faced with one. *We*, the systems theorists, have in fact been studying, not a Black Box, but a larger system composed of two parts, the Black Box and the Investigator, each acting on the other. ... Thus, if the Investigator is a scientist studying the Box, *we* are metascientists, for we are studying both; we are working at an essentially different level. (Ashby in Klir, 1991, p. 257)

Ashby's Black Box

What does Ashby tell us about the Black Box?

First, he makes a claim for the generality of the Black Box. He tells us that the Black Box need not remain in the realm of Electrical Engineering or deal with the behaviour of discrete components we are already familiar with. And he tells us that it may be applied in a wide range of situations. For reasons I will explore in the follow up paper, I specially like this notion, and would extend it towards the universal. If this generality is not clear in the quotes above, consider this further quote (from Ashby's *Introduction to Cybernetics*):

Back Box theory is, however, even wider in application than these professional [engineering, RG] studies. The child who tries to open a door has to manipulate the handle (the input) so as to produce the desired movement at the latch (the output); and he has to learn how to control the one by the other without being able to see the internal mechanism that links them. In our daily lives we are confronted at every turn with systems whose internal mechanisms are not fully open to inspection, and which must be treated by the methods appropriate to the Black Box. (Ashby, 1956, p. 86)

Second, he tells us the Black Box can remain unopened: there is no need to open the Box to determine (model) the behaviour that the observer observes and, to some extent at least, instigates. Ashby is quite specific about this. It is interesting, in this light, to note he never mentions, in the paper, journal or "Introduction," whitening the Black Box. Whitening is the term used when explaining what "really" is in the Black Box: the term White Box implies that the Box is somehow fully known, and, in that sense, is opened, stripped back and revealed. Leaving aside epistemological questions of whether we can ever "really know," that is, whether there is not always another question that may be asked, another recursion that may be applied, it is often held that the point of examining a Black Box is to know what goes on in it. I maintain that to attempt to open the Box is both improper and, in principle, impossible. Ashby's approach is a proper precursor to this position: He insists we do not need to open the Black Box, and (many would say fortunately) that it is not always possible to think we could—as for instance is the case with the head of another human. Two more quotes are appropriate, even if coming in a little from the periphery:

That homo has a brain, no more entitles him to assume he knows how he thinks than possession of a liver entitles him to assume that he knows how he metabolises. (Ashby, 2004)

And:

A man no more knows how he thinks, just because he has a brain in his skull, than he knows how he makes blood, because he has marrow in his bones. (Ashby, 2004)

Third, the Black Box is a device for changing input (value) into output (value). This means that there is an input and an output; and that there is, also, a before and after (the input becomes the output). Associated with this is the notion of behaviour, which is recognised by the observer whose job it is to try to relate (find an apparently causal connection between) them. Thus, central to the notion of the Black Box are the notions of behaviour and of regularity/pattern. The importance of behaviour makes the

Black Box a seemingly ideal device for Skinnerian psychologists (I will argue more fully than above, in a follow up, that their interpretation is completely inappropriate). The importance of pattern, of finding the general, brings us into other Ashby realms, such as variety: but it also places Ashby in the same camp as Piaget.

Fourth, not everything is deductible. Ashby argues that there are situations in which Black Boxes are not determinable through observation. It is, I will argue, but a small step to say that none are. Ashby intimates this when he talks of the history and unopenability of the Black Box. If we take these points seriously, then we will also want to move beyond Ashby to the point where we deny that using the Black Box to develop explanations ever leads to certainty—though it may lead to a high degree of likelihood. That is essentially what Popper (1963) insisted, a mere decade after Ashby's paper was published.

Ashby and second-order cybernetics

What should we learn about Ashby's thinking in general, from his view of the Black Box?

Consider perhaps the most startling statement in the paper, the quote given above from p. 257 of the paper. This tells us that already in 1958,⁸ Ashby had understood what later came to be called second-order cybernetics. In his description of the Black Box and its Investigator, he describes the recursion of observation that is at the heart of both von Foerster's and Maturana's work, and which (as it occurred in Pask's extension of Loefgren [1968] in his classes in 1971–2) was also central to Pask's arguments towards second order cybernetics. Ashby is, here, explicitly talking not only of the observer, but the observer of the observer. Ashby is observing observing, (von Foerster, 2003)

There is, I think, an irony that Ashby, perhaps the most classically mechanistic of all early cyberneticians, and the one who created the most thorough mathematical description of the subject and its working, is also already in 1958 an embryo second-order cybernetician. I have made a similar argument concerning Pask (Glanville, 2005): His *An Approach to Cybernetics* (Pask, 1961) is full of statements, coming out of his experience designing truly interactive machines, that are precursors to both his own later development and the general development of second order cybernetics. In my view, Pask was always at least on the edge of second-order cybernetics. This view can be extended to Bateson and Mead (the official mother of second-order cybernetics), whose understanding was, I maintain, always of a second-order cybernetics. This is why, with the exception of the briefing paper von Foerster commissioned (Mead, 1968) they didn't mention it. What other sort of cybernetics was there?

If I am correct in my assertions, it is clear that, from the earliest days, at least several of those involved in cybernetics were taking what when fully articulated

8. In fact, this point is made in Ashby (1956). So 1958 (when this paper was first published) is not the dawning of this understanding in Ashby's work.

became called a second-order cybernetic position as their position.⁹ That is to say that, for many (specially those whose roots were in the arts, humanities and social sciences), second-order cybernetics was implicit in first order cybernetics, but for some it was explicit. Or, perhaps, there was little difference. Maybe we should no longer maintain the distinction.

Is there confirmation of the claim about Ashby's cybernetics? Yes, in the first quote I gave. Here, Ashby insists that whatever is found out about any Black Box is not from the Black Box alone, but from the observer's (Investigator's, using his term) interaction with the Black Box. Thus, in regard to the Black Box, Ashby's position is exactly that of second order cybernetics: that we cannot ignore the observer. As Maturana famously says:

Everything said is said by an observer.

To which von Foerster retorts:

Everything said is said to an observer.

Summary: Introducing the Black Box

In Ashby's usage, the Black Box is, then, a unopened device, observed by an observer (or Investigated by an Investigator) that exists between behaviours observed at an input and (assumed-to-be consequent) behaviour observed at an output, that is taken to cause the change of input into output. By changing the behaviours we input and observing the corresponding output behaviours, we test the understanding we have of the connection between input and output, taken to be the mechanism assumed to be inside the closed Box. The Box does not need to be opened—indeed some would suggest that it cannot be—that it is essentially a conceptual device, a *gedenken* experiment, though the argument to support this position has to await the follow up paper. The power of the Black Box is that we do not need to know what is in it. Nor do we need to understand the nature of the observed change: through using the Black Box we are enabled to work with the unknown and, some would argue, the unknowable. It is a truly extraordinary invention.

References

- Ashby, W. R. (1956). *Introduction to cybernetics*. London: Chapman and Hall
- Ashby, W. R. (1991). General systems theory as a new discipline. reprinted in G. J. Klir (Ed.), *Facets of Systems Science* (pp 249–257). New York: Plenum Press. (Originally published in 1958 in *General Systems*, 3, 1-6)
- Ashby, W. R. (2004). *Aphorisms* (collected by John Ashby). Retrieved from <http://www.cybsoc.org/ross.htm> April 22, 2007.
- Foerster, H. von. (2003). *Understanding understanding*. New York: Springer.
- Glanville, R (2007). "An approach to cybernetics." In R. Glanville & K. H. Mueller (Eds.), *Gordon Pask: Philosopher Mechanic*. Vienna: edition echoraum. (Originally published as *Lerner ist Interaktion: "Gordon Pask's An Approach to Cybernetics."* In D. Baecker (Ed.), *Schlüsselwerke der Systemtheorie* (pp.75-94) published by Verlag für Sozialwissenschaften of Wiesbaden in 2005)

9. I once asked Gordon Pask, the only person I knew well who knew Wiener, if Wiener would have considered it impertinent to modify his cybernetics. Pask's response was that Wiener knew there was another, crucial step to take, but that he did not know how to take it himself.

- Loefgren, L. (1968) An axiomatic explanation of complete self-reproduction. *Bul Math Biophys*, 38 (3).
- Luhmann, N. (1995). *Social systems* (J. Bednarz & D. Baecker, Trans.). Stanford, CA: Stanford University Press.
- Mead, M (1968) The cybernetics of cybernetics, in H. von Foerster, J. D.White, L. J. Peterson, & J. K. Russell. (Eds.), *Purposive systems*. New York: Spartan Books.
- Pask, G. (1961.) *An approach to cybernetics*. London: Methuen.
- Popper, K. (1963). *Conjectures and refutations*. London: Routledge and Kegan Paul
- Skinner, B. F. (1984). Selection by consequences. *Behavioral and Brain Sciences*, 7, 477-510.