'Purposeful behaviour, Expectations, and the Mirage of Social Justice: The Influence of Cybernetics on the Thought of F. A. Hayek.'

Paul Lewis
Department of Political Economy
King's College London
paul.lewis@kcl.ac.uk

Draft 3rd June 2016: Prepared for the 2016 History of Economics Society meeting, Durham NC.

Abstract

This paper examines the influence of cybernetics on the development of Friedrich Hayek's postwar work in theoretical psychology, economics, and social theory.

The first part of the paper centres on an examination of various drafts of Hayek's 1952 book *The Sensory Order*, which reveals that, in developing his account of the working of the mind, Hayek drew on the seminal work of British cybernetician Kenneth Craik (one of the founders of the discipline of cybernetics). It is argued that in particular that Hayek drew on Craik's work in developing his account of how the pattern of impulses proceeding through the neurons in a person's brain constitutes a *model* of the external world that enables the person to form expectations about future events in his/her environment, on the basis of which they are able to form plans that are likely to come to fruition. It is also argued that a reading of the work of some of the other founders of cybernetics, such as Norbert Wiener and Warren McCulloch, helped Hayek to extend his account of the working of the mind so as to encompass purposeful human action, most notably through Hayek's appropriation from cybernetics of the notion of 'negative feedback'.

The second part of the paper investigates the extent to which ideas drawn from cybernetics influenced Hayek's economics and social philosophy. On the basis of an inventory of those places in his writings on economics and social theory where Hayek mentions cybernetics, it is argued first of all that the key elements of Hayek's finished account of the generation of the market order or catallaxy were well known to him, and had been marshalled by him into a coherent account of how such order is possible in decentralised market economies, independent of his acquaintance with cybernetics. That is to say, it appears that cybernetics did not shape the development of Hayek's economics in a significant way. However, the evidence also suggests that Hayek did make made significant use of ideas drawn from cybernetics in developing his broader social theory, in particular his critique of social justice. More specifically, it is argued that the notion of negative feedback played an important role in Hayek's efforts to justify his claim that the market process invariably involves the systematic disappointment of people's expectations and that, as a result, the outcome of the market is one to which the notion of social justice is inapplicable. In this way, cybernetics and, in particular, the notion of negative feedback, did play an important role in Hayek's broader social theory, by providing him with a conceptual bridge through which he could (attempt to) connect his account of the market order with his critique of social justice.

Finally, the consistency of the ways in which Hayek uses cybernetics both with the subjectivist account of human action that informs his economics, and also with his account of the notion of 'spontaneous order', is considered.

Electronic copy available at: http://ssrn.com/abstract=2790169

'Purposeful behaviour, Expectations, and the Mirage of Social Justice: The Influence of Cybernetics on the Thought of F. A. Hayek.'

Paul Lewis
Department of Political Economy
King's College London
paul.lewis@kcl.ac.uk

Draft 3rd June 2016: Prepared for the 2016 History of Economics Society meeting, Durham NC.

1. INTRODUCTION[1]

The goal of this paper is to ascertain the extent to which Friedrich Hayek's postwar work on the mind and the market was shaped ideas drawn from the then emerging discipline of cybernetics. It is well known that, after World War Two, Hayek began increasingly to conceptualise the mind and the market as complex, adaptive systems (that is to say, as systems that consisted of a structured set of parts of elements, whether they be neurons or people, that are related to one another, and so interact with each other, in particular ways; that display emergent properties, such as the capacity to create the sensory world we perceive and to act in a purposeful fashion, in the case of people, or to coordinate the activities of a multiplicity of individuals, in the case of the market economy); and that are adaptive in the sense of adjust to the broader environment in which they are found via an evolutionary process involving variation, selection, and differential rates of reproduction) (Caldwell 2004a; Gaus 2006; Lewis and Lewin 2015). At around the same time as Hayek began to develop these aspects of his work, namely the mid-1940s, a group of scholars drawn from a variety of disciplines including mathematics, engineering, physiology and psychology embarked on work that led to the creation of a new discipline of cybernetics aimed at understanding the behaviour of physical, chemical and biological systems. The *dramatis personae* in the story of the development of cybernetics included a fascinating cast of characters. Perhaps the best known was American mathematician Norbert Wiener, who in addition to coining the term 'cybernetics' wrote the first systematise treatise outlining its key claims (Wiener 1948a). Other prominent figures included Mexican physiologist Arturo Rosenbleuth, American neuropsychologist Warren McCulloch, British experimental psychologist Kenneth Craik, and British psychiatrist Ross Ashby. These scholars developed a set of concepts, perhaps most notably those of 'feedback', 'information', and 'control', along with the notion of mental 'models', which they used not only to understand physical systems but also, as we shall see, to analyse the human nervous system and purposeful human action.

Several distinguished commentators have suggested that concepts drawn from cybernetics helped to shape Hayek's thinking from the late 1940s onwards, as he was developing his account of the mind, and the market, as complex systems. For example, in a recent survey of the connections between Austrian economics and complexity theory, Barkley Rosser has argued that, "[C]ybernetics … was very much on the mind of Hayek as he developed his views on a complexity approach … He was influenced by the work of the founder of cybernetics, Norbert Wiener" (2015: 597). In a similar vein, historian of economics Philip Mirowski (2007: 366) has argued that, after the end of World War Two, "Hayek began to endorse various aspects of the 'cybernetics' project, which sought to reduce thought to mechanism." Commenting on Hayek's work in the 1950s on complex systems, the editor of Hayek's *Collected Works*, Bruce Caldwell, states that, "It was during this period that Hayek began citing work in fields like cybernetics and systems theory, which themselves investigated similarities among systems studied by a

[1] An earlier version of this paper was presented at George Mason University. I'm grateful to the participants in that seminar, in particular Peter Boettke and Deirdre McCloskey, for their comments. I'm especially grateful to the graduate students who were present, who offered me detailed, thoughtful, and very helpful written feedback. I'm also indebted to Ted Burczak, Bruce Caldwell and Jerry Gaus for valuable comments and advice. None of those mentioned is responsible for any shortcomings in what follows.

Electronic copy available at: http://ssrn.com/abstract=2790169

variety of disciplines" (Caldwell 2004: 362; also see p. 303). Similar views have recently been expressed by Ulrich Witt (2013: 126), who contends that, "Hayek had read widely on new developments in systems theory, cybernetics, and the biophysics of self-organization and the emergence of order. These readings seem to have led him to the conclusion that the spontaneous order arising from the interactions in free markets is a case of 'emergence' in the sense of the theory of self-organizing systems by R. Ashby, H. von Foerester and N. Wiener." Finally, commenting on Hayek's theory of expectations formation, Butos and Koppl write that, "The generation of knowledge and the actions induced by expectations involve a cybernetic relationship between mind and environment" (1993: 312).

However, insightful though these remarks are, their authors do not elaborate in detail on the nature and significance of cybernetics' influence on Hayek. It is one thing to comment that Hayek had read, cited, or begun to endorse work carried out under the banner of cybernetics. It is another to explore, systematically and in detail, precisely what Hayek took from those works and what difference they made to his thought. Therein lies the contribution of this paper, which aims to build on the insights of the scholars mentioned above by exploring precisely how Hayek used ideas drawn from cybernetics, both in his theoretical psychology and in his economics. It should then be possible to ascertain what difference, if any, those ideas made to Hayek's thinking and thereby to assess their importance for the development of his work on the mind and market as complex systems.[2]

There are several reasons why the account provided below will be of interest to historians of Hayek's economics. First, as just above, several commentators have suggested that cybernetics helped to shape Hayek's account of the economy as a complex system, without however explaining their claims in detail. The analysis presented in this essay uses archival evidence as well as a detailed examination of Hayek's published works to ascertain how ideas drawn from cybernetics shaped Hayek's postwar work on the economy. Second, the analysis presented here also shows how Hayek attempted to use cybernetics to connect his economics to his broader political philosophy, something that should be of interest to readers of a journal of the history of political economy.

Third, whilst some of the analysis presented below ostensibly concerns Hayek's theoretical psychology, it has important ramifications for Hayek's economics and, therefore, for those interested in the history of his writings in that field. One of the key ideas that Hayek takes from cybernetics is that the pattern of nerve impulses passing through people's brains at a given moment in time constitutes a 'model' of their circumstances that enables them to 'try out' in their minds the various courses of action that are open to them without actually having to act. In this way, as will be explained below, the capacity of the brain to 'model' its environment enables people to form expectations of the likely consequences of those actions, on the basis of which they can decide what to do, thereby facilitating conduct that is more likely to be well-adapted to the context in which it takes place and so more likely to come to a successful conclusion. This concept of the 'model', whose origins are discussed below, has been invoked in debates over the (dis)similarities between the varieties of subjectivism to which Hayek and Keynes subscribed, with it being argued—not least in this journal—that the way the Hayekian theory of expectations-formation is grounded in his account of the brain as providing a 'model' of the world renders it superior to Keynes's view of expectations as depending on animal spirits (Butos and Koppl 1993, 1997). In exploring the origins of the concept of the 'model', therefore, this paper investigates one of the building

[2] Part way through writing this essay, I was sent an interesting paper that also examines the relationship between Hayek and cybernetics (Olivia 2016). The two papers of course address the same broad topic, and inevitably cover some of the same ground (most notably concerning how Hayek used the notion of negative feedback to develop the account of motor control found in *The Sensory Order*). However, there remain significant differences between them, including: the systematic and sustained use in the current paper of comparisons between the various drafts of *The Sensory Order* in order to identify precisely when, and from which cyberneticians, Hayek drew various ideas; the account, found below, of the importance of the work of Kenneth Craik in furnishing Hayek with the idea, central to his account of purposeful action, that people form forward-looking 'models' of their circumstances; the useful analysis provided by Olivia of Hayek's ultimately unsuccessful attempt to develop a theory of communication in his *Within Systems* manuscript (Hayek n.d.); and the different approaches adopted in the two papers to the analysis of the influence of cybernetics on Hayek's work on economics and social theory, with Olivia focusing more on explaining why Hayek did not develop a fully-fledged cybernetic model of the economy, while the current paper examines how Hayek used the notion of negative feedback to connect his analysis of the market process with his broader social philosophy.

blocks of the Hayekian theory of expectations-formation, a topic of considerable significance that continues to draw attention from contemporary economists interested in Hayek's account of the merits of the market process (Carabelli and De Vecchi 2001; Burczak 2001; Koppl and Butos 2001; Festré and Garrouste 2016).

Fourth, and relatedly, an understanding of the influence of cybernetics on Hayek's thinking can illuminate one of the issues at the heart of a debate between two prominent historians of economic thought concerning the extent to which Hayek's embrace of evolutionary ideas from around 1960 onwards represents a natural continuation of, or a rupture with, his earlier postwar view of the economy as a complex system (Caldwell 2004a; Mirowski 2007). One of the reasons why the capacity to model the world is important, Hayek argues, is that it confers an evolutionary advantage on those being that possess it. This argument—which, it will be shown below, Hayek takes from one of the seminal texts in cybernetics, namely Kenneth Craik's *The Nature of Explanation* (1943)—is significant for historians of economics for two reasons: first, because it suggests that Hayek embraced evolutionary themes earlier than one of the protagonists in that debate claims, thereby providing support for those who emphasise continuity rather than rupture in Hayek's postwar thinking; and, second and relatedly, because it indicates that, at least in the case of Craik's influence on Hayek, the contrast between case evolutionary and complexity arguments can be overdrawn.

Fifth, and finally, the evidence marshalled below can also be used to consider the merits of the claims advanced by various commentators on Hayek's work who have called into question either, first, the coherence of Hayek's use of cybernetics with other aspects of his economics, or, second, the consistency with which he uses ideas drawn from cybernetics. An example of the first kind of criticism is exemplified by Bruce Caldwell, who has expressed concerns that, in drawing on cybernetics, Hayek was importing into his theoretical psychology a mechanistic approach that was incapable of doing justice to the importance of the subjective dimension of purposeful human action (Caldwell 2004a: 300). A second criticism of this kind has been advanced by Geoff Hodgson, who has argued that Hayek's use of ideas drawn from cybernetics betrays an incoherence in his efforts to theorise the notion, central to his account of the market process, of 'spontaneous order'. According to Hodgson, Hayek "linked the concept of spontaneous order together with 'autopoieses, cybernetics, homeostasis, spontaneous order, self-organisation, synergetics, systems theory' (Hayek, 1988, p. 9) as supposedly allied and similar ideas … However, this list of topics is not itself conceptually or theoretically homogeneous. It further betrays a serious shortcoming of Hayek's work: a lack of clarity about the crucial concept of spontaneous order." The second kind of criticism has been advanced by Ulrich Witt (2013), who has observed that Hayek's use of ideas drawn from cybernetics is inconsistent in the sense that of being limited to the notion of negative feedback, ignoring without good reasons the concept of positive feedback. As we shall see, the analysis presented below can shed some light how on how well-grounded are these claims, thereby helping to demonstrate the significance of what follows for those interested in Hayek's economics.

2. CYBERNETICS: A SUMMARY

The term *cybernetics* was coined in 1947 by the American mathematician Norbert Wiener to describe the study of the communication and manipulation of information in the service of the control and guidance of physical, chemical or chemical systems (Wiener 1948a: 13). The actual inception of cybernetics as a discipline is usually dated to a few years earlier, to the simultaneous but coincidental publication in 1943 of three classic works: Rosenbleuth, Bigelow and Wiener (1943); McCulloch and Pitts (1943); and Craik (1943). A seminal, book-length treatment of the subject, in which key concepts were set out and a basic vocabulary fixed, was published in 1948, in the form of Norbert Wiener's *Cybernetics* (Wiener 1948a).[3]

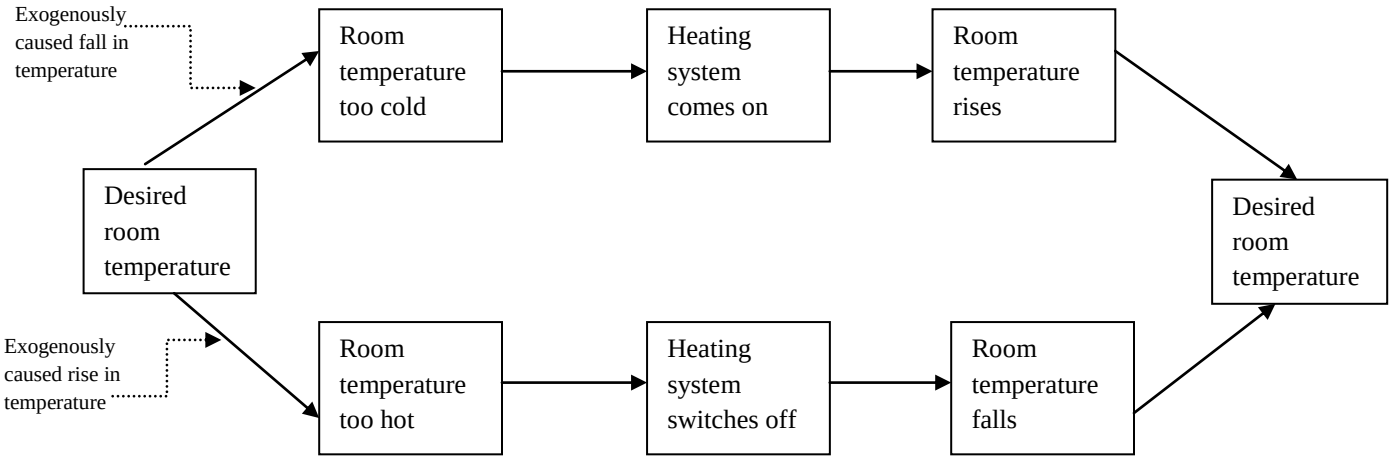
[3] For histories of cybernetics, see Heims (1991), Boden (2006: 168-236) and Dupuy (2009).

This section of the paper outlines the key features of cybernetics, drawing exclusively on the works on cybernetics cited by Hayek, in order to prepare the ground for the subsequent analysis of the influence of cybernetics on Hayek's work. In later sections of the paper, where that influence is discussed in detail, the relevant sections of key works in cybernetics upon which Hayek drew will be summarised.

2.1 Cybernetic systems

A *cybernetic system* is a machine that produces an action in response to an input of information and which includes the results of its own action in the new information by which it modifies its subsequent behaviour. The action of the machine is directed towards achieving a particular target output or goal and is therefore, in a sense to be elaborated below, 'purposeful.' For example, the goal or purpose of a central heating system is to maintain the inside of a house at a particular temperature (Wiener 1948a: 115). If the internal temperature differs from the preset target, then that deviation is detected by sensors within the thermostat, which turns the heating on or off according to whether the actual temperature is too low or too high. The former case is illustrated the top branch of Figure 1, the latter case by the bottom branch.

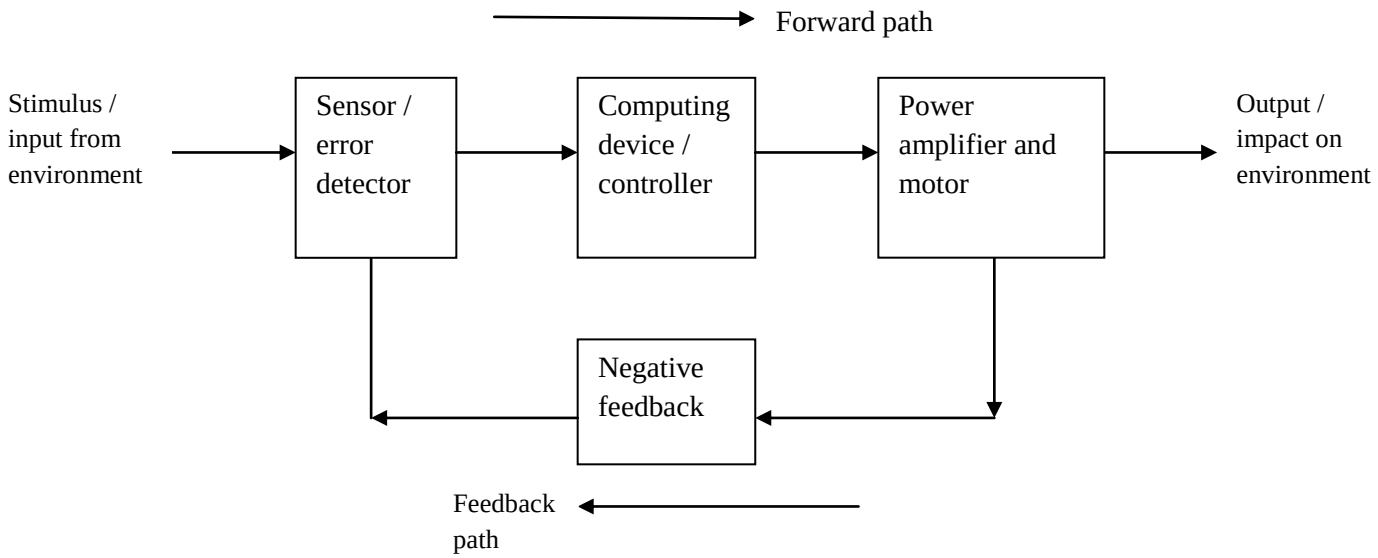
Figure 1: A simple illustration of negative feedback



In this way, provided that the system is well designed, it will maintain the temperature within the house at the desired level.

The key components of a cybernetic system, all of which are implicit in the account just given, are as follows (see Figure 2). First, there must be a sensory device for detecting the disturbance or deviation from the target to be countered. Second, there needs to be a computing device for determining the right kind of response. Third, there needs to be an amplifier, which ensures that the deviation elicits from the system a response equal to or greater than itself and an effector or motor mechanism for making that response.

Figure 2: An example of a simple negative feedback system (modified from Frank 1948: 192).



This diagram also illustrate another key idea, namely that of feedback.[4] A *feedback loop* of the kind illustrated in Figure 2 is a circular arrangement of causally connected elements in which an initial cause (the 'input') propagates around the elements in the loop, with each element affecting the next, until the final element (the 'output') causally influences (or 'feeds back' the effect) to the first element in the cycle. The existence of such feedback loops means that a system is capable of modifying its own conduct as a result of its own previous behaviour.

Cybernetics conceptualises feedback as involving the transmission of information about the outcome of any process or activity. In particular, *negative feedback*, or information about error, is the difference between the actual output of the system (e.g., the actual temperature in a house) and its target output (the desired temperature, at which the thermostat is set).[5] Once the gap between the actual and the target states of the system has been detected by the receptors within the cybernetic system, the machine can take corrective behaviour (involving, in our example, the heating system turning on or off as appropriate). The feedback is negative in the sense that the action taken by the system is the opposite of the stimulus that evoked the response. If the outcome of the corrective behaviour is that the actual temperature falls below the desired level, then that outcome (or system 'output') will also be detected (thereby itself becoming an 'input' to the process) and will in turn precipitate an appropriate response (with the heating coming in order to raise the temperature).[6] In this way, errors in performance are fed back as part of the input of the system in such a way that they oppose the signal responsible for the error, thereby reducing its magnitude (Rosenbleuth *et al*. 1943: 19; McCulloch 1948: 261; Wiener 1948a: 13; Ashby 1949: 717).[7]

---

[4] For more on the relevant notion of 'control', see McCulloch (1948: 260).

[5] The sense in which notion of 'information' is used in cybernetics is discussed in Wiener (1948a: 15-19, 1948b: 202-06) and Ashby (1949: 717-18, 721).

[6] There is a second kind of feedback that is distinct from negative feedback, namely positive feedback: "The feedback ... is positive [if] the fraction of the output which re-enters the object has the same sign as the original signal input. Positive feedback adds to the input signals, it does not correct them (Rosenbleuth *et al*, 1943: 19; also see Wiener 1948: 113-36 and Ashby 1949: 717, 721). Positive feedback tends to amplify, rather than dampen, the value of the output of the system by adding the value of the output to the system input quantity. As a result, the system will accentuate, rather than offset, external disturbances.

[7] Two varieties of cybernetic system may be distinguished, according to whether the value of the target variable remains constant or changes over time. In *homeostatic* systems, the value of the target variable is constant, as in the case of those metabolic systems in the human body that the value of certain key variables—body temperature, for example, and blood sugar—within the range required to sustain life (Cannon 1932).

2.2 Cybernetics and purposeful behaviour

The scholars who developed cybernetics argued that systems of the kind described above can be said to exhibit *purposeful*, goal-directed behaviour. As the authors of one of the three founding works of cybernetics put it, "the term 'purposeful' is meant to denote that the act or behaviour [of a system] may be interpreted as directed to the attainment of a goal—i.e., to a final condition in which the behaving object reaches a definite correlation in time or in space with respect to another object or event." All such purposeful behaviour "may be considered to require negative feedback. If a goal is to be attained, some signals from the goal are necessary at some time to direct the behaviour" (Rosenbleuth *et al*., 1943: 18, 19). On this view, the traditional opposition between mechanism and teleology—that is, between the mechanical behaviour of inanimate machines and the purposeful, goal-driven behaviour of living beings—is false, so that scientists can speak, without confusion or contradiction, of 'teleological mechanisms'. So long as the behaviour of the mechanisms in question is controlled by "the error of the reaction—i.e., by the difference between the state of the behaving object at any time and the final state interpreted as the purpose", then the term 'teleology' can legitimately be used to describe it: "Teleological behaviour thus becomes synonymous with behaviour controlled by negative feedback" (Rosenbleuth *et al*., 1943: 23-24).

The cyberneticians argued that purposeful human action can be understood in terms of what certain machines do, with concepts and terminology drawn from engineering being used to analyse and explain human conduct. On this view, man-made, self-correcting systems such as computing machines and guided missiles can act as material models for purposeful human action. A *material model* is "the representation of a complex system by a system which is assumed simpler and which is also assumed to have some properties similar to those selected for study in the original complex system" (Rosenbluth and Wiener 1943: 317-18; also see Wiener 1948b: 208). In using material models, the scientists who founded cybernetics were using actual, physical mechanisms of some complexity, whose workings they understood well, as the basis for understanding purposeful human action. In doing so, they were able to "assist the scientist by replacing a phenomenon in an unfamiliar field by one in a field in which he is more at home" (Rosenbleuth and Wiener 1945: 217; also see Frank 1948: 191). This approach, its advocates argued, "reveals that a uniform behaviouristic analysis is applicable to both machines and living organisms, regardless of the complexity of the behaviour ... The broad classes of behaviour are the same in machines and in living organisms" (Rosenbleuth et al., 1943: 22; also see Ashby 1949: 716).

3. THE IMPACT OF CYBERNETICS ON HAYEK'S THEORETICAL PSYCHOLOGY: FROM *WHAT IS MIND?* TO *THE SENSORY ORDER: MODELS, EXPECTATIONS AND PURPOSIVE BEHAVIOUR*

3.1 Hayek's Theoretical Psychology: *The Sensory Order*

The impact on Hayek's thinking of ideas drawn from cybernetics can first be seen in his work on theoretical psychology, as set out in his 1952 book *The Sensory Order* (Hayek 1952). The task Hayek sets himself in that book is to explain why the phenomenal (subjective, mental) picture of the world provided by our senses differs from the physical order revealed to us by the natural sciences. Objects that resemble each other in sensory terms may display very different physical relations to each other, while objects that appear altogether different to us may display very similar physical properties. There are, therefore, two different orders: a physical order, which is revealed to us by the natural sciences; and a phenomenal, or mental, or sensory, order which we experience as individuals. The goal of theoretical psychology, as Hayek conceives it, is to explain how the neurons that comprise the human central nervous system form a

---

Systems in which the value of the target variable changes over time, as is true for example of a radar-guided anti-aircraft gun tracking the location of an aircraft, are known as *servo-mechanisms* (McCulloch 1948: 262; Rosenbleuth *et al*., 1943: 19; Wiener 1948a: 11-13, 133).

classificatory structure that is capable of discriminating between different physical stimuli so as to give rise to the sensory qualities we actually experience (Hayek 1952: 2-8, 13-19, 37-40).

For Hayek, the human central nervous system consists of a hierarchical network of interconnected nerve fibres. Neurons can generate outgoing electrical impulses if they are stimulated sufficiently by incoming impulses, and it is through the transmission of such impulses that neurons interact with each another (Hayek 1952: 42, 55-64). Neurons in which impulses often occur together tend to become connected, so that over time the nervous system acquires a structure, in which the position of any one neuron is defined by its connections to other nerve fibres. In Hayek's view, it is this structure that accounts for the key features of our mental experiences (Hayek 1952: 12). To see why, note that the (primary) nerve impulse generated by a particular external stimulus will in turn stimulate neurons connected to those along which that primary impulse is transmitted. The external stimulus thus leads to the generation within the central nervous system of an induced pattern of (secondary) nerve impulses, described by Hayek as the *following* of the initial impulse, characteristic not only of the external stimulus but also of other stimuli that have accompanied it in the past. And it is by classifying external events according to the followings they trigger that the central nervous system differentiates between them and thereby creates distinct sensory data. External events are classified as the same, and are experienced as having the same sensory qualities, if they trigger an identical set of neural events or following (Hayek 1952: 48-54, 62-78). For Hayek, then, external events stimulate the growth of an organised structure of nerve fibres that reproduces, albeit only imperfectly, the patterns of events occurring in the external environment, in the sense that the structure of connections between the neurons corresponds topologically to the structure of the relations between some of those external stimuli. Indeed, on Hayek's account, the neural order—that is, the set of connections between nerve fibres in the brain, and the impulses proceeding in them—that is produced in this way just is the sensory order of phenomenal experience (Hayek 1952: 40).

3.2 The Writing of *The Sensory Order*: A Brief Chronology

The origins of this account of the mind lie in a student paper Hayek wrote in 1920. The paper in question was entitled, *Beiträge zur Theorie der Entwicklung des Bewusstseins*, or 'Contributions to a Theory of How Consciousness Develops' (Hayek 1920, [1920] forthcoming). In it, Hayek describes, in terms closely resembling those used in the final version of *The Sensory Order*, how what he terms "elementary consciousness phenomena", such as sensory qualities and emotions, are generated (Lewis 2016a: 120-23).[8] However, Hayek laid aside the *Beiträge* in September 1920 returned to it only in the early-to-mid 1940s, driven by a desire to justify his own preferred subjectivist approach to economics in the face of claims that references to subjective states of mind were unscientific and so had no place in 'scientific' economics (Caldwell 2004a: 137, 2004b: 246-48).

In the course of transforming his 1920 paper into the published version of *The Sensory Order*, Hayek wrote two draft manuscripts. The first, which Hayek produced in 1945, is entitled, *What is Mind? A Working Hypothesis on the Origin of Sensory Qualities and Abstract Concepts* (Hayek 1945). It will be referred to hereafter as *WIM'45*. A second, longer draft was completed on December 25 1947 and bears a similar title to the first, but a different subtitle, namely *What is Mind? An Essay in Theoretical Psychology* (Hayek 1947). It will be denoted in what follows as *WIM'47*. A comparison of these two drafts, with each other, with the 1920 student paper, and with the final version of *The Sensory Order*, makes it possible both to trace out the development of Hayek's ideas and also to shed light on some of the influences that shaped his thinking.

Both in his 1920 student paper and also in *WIM'45* Hayek's focus is on explaining the generation of the sensory qualities and emotions people experience. In neither manuscript does he attempt to explain

[8] The fact that Hayek had set out the key elements of his account of the brain as a classificatory system, composed by a structured arrangement of neurons, by 1920—that is to say, before he began to develop his own research on economics—casts doubt on Mirowski's claims that Hayek's "prior image of the market provided the metaphoric template for his theory of the structure of the brain" and that "the formal template for his understanding of neurons was his own previous *Pure Theory of Capital*" (Mirowski 2002: 239, 240).

purposeful human action. The most notable difference between the *Beitrage* and *WIM'45* lies in the way that in the 1945 draft Hayek uses the metaphor of a 'map' to describe how the structure of neural connections in the brain reproduces, if only roughly, the patterns of events in the external world (Hayek 1945: paragraphs 151-62; also see Hayek 1952: paragraphs 5.17-5.32).

The notion of the 'map' provides the basis for what is, for our purposes, a much more important development in Hayek's theory, which sees him use ideas drawn from cybernetics to extend his analytical framework beyond the study of sensory qualities and emotions to encompass the explanation of purposeful human action. The development in question centres on Hayek's introduction of the notion of the 'model' into his theoretical scheme, which he does in paragraphs 193-228 of *WIM'47*. None of those paragraphs was present in *WIM'45*; most were added to the manuscript between 1945 and 1947.[9] Hayek begins by explaining that the semi-permanent structure of neurons that constitutes a person's 'map' of the external world is static because it "represents the kind of world in which the organism has been moving in the past, but not the particular environment in which it is moving at the moment" (Hayek 1947: paragraph 195; cf. Hayek 1952: 5.40-5.49).

The picture of the world in which the person currently finds him- or herself is provided by the pattern of active nerve impulses currently proceeding through the person's brain. Those impulses constitute a *model* of the situation in which the person currently finds him- or herself in two ways. First, they give rise to a classification of incoming stimuli, thereby generating "a concrete picture of the particular environment in which the individual is behaving at the moment" (Hayek 1947: paragraph 199). Second, that representation is *forward-looking* in that it draws attention to events that have in the past tended to accompany the ones currently being experienced and which therefore might accompany them now, so that the person begins to expect them to occur on this occasion too (Hayek 1947: paragraphs 204-11; cf. Hayek 1952: 5.50-62). To see why, note that the nerve impulses generated by an external stimulus in turn stimulates neurons connected to those along which that primary impulse is first transmitted. By generating these secondary impulses, or "chains of associative processes" as Hayek also terms them, the individual's model will continually tend to "run ahead of the actual facts of the environment into associatively evoked representations of the various possible consequences of the existing situation", generating expectations of certain possible future stimuli that typically accompanied the ones currently being experienced (1947: paragraphs 204, 208; cf. Hayek 1952: paragraphs 5.50, 5.58-59). The model will thus function as an "apparatus of orientation" by representing both "the environment of the moment" and also "possible future developments in that environment" (Hayek 1947: paragraphs 204, 207; cf. Hayek 1952: 5.51, 5.59).

In a sub-section of *WIM'47* entitled 'Purposive Behaviour', Hayek explains how, using the notion of the 'model', his account of the mind can be deployed to explain goal-oriented human action (Hayek 1947: paragraphs 217-28; cf. Hayek 1952: paragraphs 5.68-77). "Purposive behaviour," Hayek (1947: paragraph 218) writes, is that which is "'directed' towards a future 'goal', i.e. towards an event which is not present in the existing situation but which may be brought about by a series of more or less complex actions of the organism." He then argues that behaviour of this kind can be produced by a structured arrangement of neurons of precisely the kind he has been describing:

> If we conceive of the higher nervous centres as producing a model of the external world which represents things that might come about as the result of the existing situation as well as that situation itself, there exists no difficulty of principle [in] accounting for such behaviour. If the model or representation can perform or predict the effects of different courses of action and pre-select amongst those results that which in the existing state is 'desirable', it will also be capable of directing the organism to that course of action which has thus been 'mapped out' for it

[9] These paragraphs closely resemble paragraphs 5.40-5.91 of *The Sensory Order*.

> Whatever the 'desired result', in order that it should become effective as a cause it must be represented in the field of the perceptual representation of the environment, i.e. it must be one of the innumerable potentially present 'associations' which give the representations of the environment its significance or meaning. Among the many possible results of the given situations this desired result will be singled out by the fact that it will also be part of the field of the 'urge' or 'drive' [to satisfy the organism's needs] and that its representation will therefore be enforced by the convergence upon it of two different streams of nervous energy which will raise the strength and distinctiveness of its representation above that of all other possible results and thus make it a cause of further processes. (Hayek 1947: 219-20; cf. Hayek 1952: paragraph 5.68-69).

In other words, if the events modelled by the brain include the actions a person might take, then the representations thereby generated can include the potential consequences of those actions. In this way, the model enables a person to 'try out' in his or her mind various possible courses of action, exploring their consequences until the one that is desired, given the person's current situation, is identified, at which point action designed to realise that outcome will be taken.

Significantly, in outlining how notion of the model can be used to analyse purposive human conduct, Hayek refers on no fewer than three occasions to a 1943 book entitled, *The Nature of Explanation*, written by K.J.W. Craik (Hayek 1947: 219, 223, 226). Indeed, Craik's book is the only work referenced by Hayek in the eight paragraphs in *WIM'47* where he sets out the explanation of purposeful action summarised above (Hayek 1947: paragraphs 212-27; also see Hayek 1947: p. 151 and Hayek 1952: 5.68-5.75). As we shall see, it is from *The Nature of Explanation* that Hayek takes the idea that the brain forms forward-looking models of the world. Given that Craik's book is commonly held to be one of the three founding works of cybernetics, it is in affording Hayek the concept of the model that the first significant impact of cybernetics on his theoretical psychology is to be found.

3.3.1 Kenneth Craik on the Nature of Thought: The Significance of the Model

Kenneth Craik (1914-45) was a British experimental psychologist who worked in Cambridge in the late 1930s and early 1940s, prior to his tragic death in a road traffic accident in 1943. While less well known than his American counterparts such as Norbert Wiener and Warren McCulloch, Craik is celebrated as one of the founders of cybernetics (Heims 1991: 43; Boden 2006: 201-18).[10] The 1943 book cited by Hayek in *WIM'47* and *The Sensory Order* was the only work of significant length Craik published in his lifetime. In it, and especially in chapter 5 where he outlines his "Hypothesis on the Nature of Thought" (pp. 50-61), Craik argues that thought involves a person constructing a model of external reality in his or her head and then manipulating that model so as to derive predictions about alternative possible futures. More specifically, there are three main stages to the process of thought:

> (1) 'Translation' of external processes into words, numbers or other symbols,
> (2) Arrival at other symbols by means of a process of 'reasoning', deduction, inference, etc., and
> (3) 'Retranslation' of these symbols into external processes (as in building a bridge to a design) or at least recognition of the correspondence between these symbols and external events (as in realising that a prediction is fulfilled). (Craik 1943: 51.)

In the case of human beings, these three processes "become the translation of external events into some kind of neural patterns by stimulation of the sense-organs, the interaction and stimulation of other neural patterns as in 'association', and the excitation by these of effectors or motor organs" (Craik 1943: 53). These neural patterns represent external events in the sense that both the relations between the neural

[10] For more on Craik and his work, and for a more detailed account of his influence on Hayek than considerations of space allow here, see Lewis (2016b).

events, and also the causal mechanisms and processes that generate them, are analogous to the relations between the events, and the underlying causal mechanisms and processes, in the external world. The neural events, etc., therefore provide a "model" of those external events, etc., because in essential ways the neural events, etc., work in the same way as the external events, etc. they represent (Craik 1943: 59):

> By a model we thus mean any physical or chemical [including neurophysiological] system which has a similar relation-structure to that of the process it imitates. By 'relation-structure' ... I ... mean ... the fact that it is a physical working model which works in the same way as the process it parallels, in the aspects under consideration at any moment. (Craik 1943: 51.)

Craik describes this "power to parallel or model external events" as "the fundamental feature of neural machinery" and, indeed, as the "essential feature" of human thought (1943: 52, 57).

"One of the most fundamental properties of thought," Craik (1943: 50) argues, is its power of predicting events." This predictive power stems from the use of models: "Only this internal model of reality—this working model—enables us to predict events which have not yet occurred in the physical world, a process which saves time, expense and even life" (1943: 82). By making it possible to try out, and to consider the consequences of, alternative courses of action without actually having to carry them out, the use of internal models enables people to adapt their actions so that they fit their circumstances and so have a better chance of coming to fruition:

> If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilise the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it. (Craik 1943: 61; also see pp. 7, 52, 59.)

Once a goal has been selected through this modeling process, the person's actions are governed by the discrepancy between the actual state of the world and that specified in the model (that is, by negative feedback). Craik speculates, moreover, that the capacity to model the external world in this way is likely to confer an evolutionary advantage on those creatures that possess it, thereby helping to explain how such capacities arise in the first place (1943: 60-61).

What makes Craik so important for the development of cybernetics is that the attributes he regarded as the hallmarks of human thought—the translation of external events into an internal model that can be used to explain and predict events and thereby to guide action—were also, as he saw it, displayed by certain kinds of machine:

> Surely, however, this process of prediction is not unique to minds … A calculating machine, an anti-aircraft 'predictor', and Kelvin's tidal predictor all show the same ability. In all these cases the physical process which it is desired to predict is *imitated* by some mechanical device or model which is cheaper, or quicker, or more convenient in operation. (Craik 1943: 51.)

Craik goes on to describe how, in case of Kelvin's tidal predictor—a machine devised by Lord Kelvin to predict the patterns of tides in harbours—the working of the mechanical model exemplifies the three-stage process that, in Craik's view, constitutes reasoning: "representation [of the external process] by symbols, calculation, and retranslation into events" (1943: 57; also see 51-52, 56, 59-61).[11]

[11] Craik was writing before the development of digital computers, so of necessity had to base his ideas on analog devices.

What this shows is that Craik was one of the first scientists to argue that many of the principles of brain function are common to all information processing systems, including certain kinds of machine:

> My hypothesis, then, is that thought models, or parallels, reality—that its essential feature is … symbolism, and that this symbolism is largely of the same kind as that which is familiar to us in mechanical devices which aid thought and calculation. (Craik 1943: 57.)

On this view, thought is constituted by the manipulation of symbols and therefore is something that can be undertaken not only by people but also by machines.

3.3.2 The Influence of Cybernetics, Part I: Craik's Influence on Hayek's Theoretical Psychology

In this section we consider the influence of Craik's work on Hayek's theoretical psychology. In a footnote appended to paragraph 219 of *WIM'47*, at the outset of his explanation of how the notion of a model can be used to help explain purposive human action, Hayek writes as follows: "In connection with the following discussion of the model making possible purposive action compare particularly K.J.W. Craik, The Nature of Explanation, Cambridge 1943, esp. pp. 51-63" (Hayek 1947: paragraph 219, p. 151; cf. Hayek 1952: paragraph 5.68, p. 124 n. 1, and Hayek n.d. paragraph 23). The pages in Craik's work to which Hayek refers are, of course, those on which Craik sets out his central "Hypothesis on the Nature of Thought", as described above. In the paragraphs that follow, Hayek makes use of key ideas found in Craik's book: the hypothesis that the operation of the nervous systems produces "a model of the external world which represents things that might come about as the result of the existing situation as well as that situation itself … [and that as a result] can pre-form or predict the effects of different courses of action" (Hayek 1947: paragraph 219; cf. Hayek 1952: 5.68); the claim that this model "anticipates or tries out possible developments" (Hayek 1947: paragraph 222; cf. Hayek 1952: 5.71); and the view that "such guidance by a model, which reproduces and tentatively tries out the possibilities of the surroundings, can produce action which is purposive to any desired degree" (Hayek 1947: paragraph 226; cf. Hayek 1952: 5.75). In addition, later in *WIM'47*, Hayek refers to Craik's book when describing thought as "consisting in the building of models of the material processes which they 'reproduce'" (Hayek 1947: paragraph 354a and p. 151; cf. Hayek 1952: 8.49-50). Such references, to ideas that were not present in *WIM'45*, and which appear in Hayek's work only after he had read Craik's book, suggest that it is to Craik that Hayek owes the concept of the 'model.'[12]

Additional evidence is provided by a later, unpublished essay, namely 'Within Systems and about Systems', where Hayek again invokes Craik's notion of the model. More specifically, Hayek writes that the process of classification through which sensory qualities are created and purposeful action generated "will be assumed to 'parallel' or 'model' sequences of events in the environment" and that the classificatory apparatus "may … be regarded as embodying a theory of the external world which enables it to predict (= produce representat[ions] of) the events which the former will cause in the environment" (Hayek n.d.: paragraph 12). In a footnote, Hayek writes of the term 'parallel' that it is, "The expression used by K.J.W. Craik, 1943" (Hayek n.d.: paragraph 12). What we see here is Hayek using Craik's work to argue his own account of the brain as an instrument of classification implies—"may … be regarded as embodying"—the view, taken from Craik, that the essence of human thought lies in the construction and manipulation of internal models that represent, and facilitate predictions about, the external world.

Traces of Craik's influence can also be found in Hayek's discussion of the evolutionary advantages enjoyed by beings possessing the capacity to 'model' the world. In chapter 5 of *The Sensory Order*, Hayek sets out to "construct[.] a model of the model-object relationship itself" (1952: paragraph

[12] Further support for this claim is found in the way Hayek also uses Craik's work to deal with two additional points that, he believes, must be addressed in showing how his theory can be used to explain purposive behaviour. Both concern the neural processes through which people select one particular outcome to pursue and decide a concrete course of action by which to achieve that goal. Hayek draws on Craik's work in order to give an account of the neural mechanisms through which these issues are resolved in practice (1947: paragraphs 222-23, 205-07; cf. Hayek 1952: paragraphs 5.71-72 and Hayek n.d. paragraph 31). For details, see Lewis (2016b).

5.77). That is, he aims to explain how the facility to model the world arises. In the next paragraph, he states the explanation will be an evolutionary one: "Our task is … to show in what sense it is possible that within parts of the macrocosm [that is, the physical order] a microcosm [that is, the neural order] may be formed which reproduces certain aspects of the macrocosm and through this will enable the substructure of which it forms a part to behave in a manner which will assist its continued existence" (1952: paragraph 5.78). The continued existence of such structures, Hayek goes on to argue, will be promoted if "they can respond appropriately to certain events, and even in some measure anticipate their occurrence," a capacity that will be facilitated if the organism in question has "[s]uch an internal structure which … we have called a model" (paragraph 5.83, 5.85). "The model-building by such an apparatus of classification simplifies and extends the scope of successful adaptation", Hayek concludes, by facilitating "the prediction of events which are important for the persistence of the structure" (1952: paragraph 5.90). It is perhaps not too far-fetched to interpret this argument as reflecting Hayek's reading of Craik's claim that "it is much more illuminating to regard the growth of symbolising [i.e. modelling] power from this aspect of the survival value", emphasising how "natural selection causes the survival of certain organisms—those, for instance, in whom the passage of the 'monitoring' nerve impulse results in such activity of the whole organism as will tend to preserve it" (Craik 1943: 60).[13]

It was noted earlier that Craik believed that the hallmark of human thought, namely the translation of external events into an internal model that can be used to explain and predict events and thereby to guide action, are also displayed by certain kinds of machine. Hayek had already offered a few brief remarks on the question of whether his account implies that the reaction of the human central nervous system classifies stimuli should be regarded as 'mechanical' in *WIM'45*, arguing that "it would seem to follow from our thesis that the difference between purely 'mechanical' and mental processes is not one of kind but merely one of degree" (Hayek 1945: paragraph 165). Hayek elaborates on this issue in *WIM'47* (Hayek 1947: paragraphs 212-216; cf. Hayek 1952: paragraphs 5.63-5.67), summarising his views as follows:

> That such guidance by a model, which reproduces and tentatively tries out the possibilities of the surroundings, can produce action which is purposive to any desired degree is shown by the fact that machines could be produced on this principle (and some, like the predictor for anti-aircraft guns, or the automatic pilots for aircraft, have actually been constructed) which would show all the characteristics of purposive behaviour. Such machines would of course be comparatively primitive and restricted in their range of operations compared with the central nervous system … But with regard to purposiveness this difference would be merely one of degree and not of kind. (Hayek 1947: paragraph 226; cf. Hayek 1952: 5.75 and Hayek n.d. paragraph 50).

Hayek (1947: p. 151) adds a footnote to this passage in which he references p. 51 *The Nature of Explanation*, where Craik writes about the capacity to develop predictive models of the external world as follows:

---

[13] The fact that Hayek draws from Craik in this way has some implications for a recent debate over the development of Hayek's economics after World War Two. In reviewing Bruce Caldwell's *Hayek's Challenge* (Caldwell 2004a), Philip Mirowski maintained that Caldwell underestimates the discontinuity involved in Hayek's transition from the early 1950s, when he "began to endorse various aspects of the 'cybernetics' project, which sought to reduce thought to mechanism", and his later "appeal to 'evolution' to explain how an ineffable complex order, … The Market, could have come about" (Mirowski 2007: 366). The fact that, at least in part because of his reading of Craik's work, Hayek was discussing the evolutionary benefits deriving from man's ability to model his environment in the mid-to-late 1940s, suggests that there was more continuity between that stage of Hayek's thought and his later commitment to evolutionary thinking than Mirowski acknowledges. The point is only reinforced by the way that in one of Hayek's later essays, namely "Notes on the Evolution of Systems of Rules of Conduct", he outlines the evolutionary advantages conferred on human beings by the fact that the brain contains a model of its environment in terms that are almost identical to those used by Craik (Hayek [1967] 2014: 285). For more on this, see Lewis (2016b).

> Surely, however, this process of prediction is not unique to minds, though no doubt it is hard to imitate the flexibility and versatility of mental prediction. A calculating machine, an anti-aircraft 'predictor' … show the same ability. (Craik 1943: 51.)

It appears, therefore, that reading Craik's work helped Hayek to develop his comparison between the operation of the nervous system and mechanical processes, in two main ways: first, by providing Hayek with an account of purposive action, centring on the formation and use of models, that could be applied to explain the behaviour both of people and also of machines, thereby underlining the similarities between them; and, second and relatedly, by furnishing Hayek with a fund of examples of machines—such as the anti-aircraft predictor—whose behaviour was guided by 'models' of its environment and which therefore appeared to display something like purposeful action.[14]

3.4.1 The Influence of Cybernetics, Part II: Norbert Wiener, Warren McCulloch, and Ross Ashby

The second way in which cybernetics influenced Hayek's theoretical psychology also centres on how his theory of the nervous system as a classificatory instrument can be used to provide an account of purposive human action. This, Hayek states, requires an analysis of the interplay between sensory and motor processes in the human nervous system (Hayek 1952: paragraphs 4.1-4.10). Hayek undertakes the requisite analysis in chapter 4 of *The Sensory Order*, one section of which in particular bears the imprint of ideas drawn from cybernetics, namely Section 6, on "Patterns of Motor Responses" (Hayek 1952: paragraphs 4.48-4.55). There is no discussion of the issues in question in *WIM'47*; the relevant paragraphs were added by Hayek between 1948 and the final publication of the book. In this section of the book, as we shall see, Hayek draws on the work of Nobert Wiener, Warren McCulloch, and Ross Ashby, in order to elaborate on how his theoretical scheme can be used to understand the interplay between sensory and motor nerve firings that governs purposeful behaviour.[15]

3.4.2 Cybernetics and Voluntary Human Action: The Importance of Negative Feedback, the 'Hierarchical' Nature of Motor Control, and 'Self-Organisation'

Wiener and Ashby argued that voluntary human action involves circular neural processes, characterised by negative feedback. In this case, the output of the nervous system is the movement of a person's muscles, while the inputs are the nerve impulses caused by the visual and proprioceptive stimuli set in train by those muscular actions. For example, when a person extends her arm and hand to pick up a pencil, she is continuously informed—by visual or proprioceptive stimuli impinging on her sense organs—about how close her hand is to the object. On the basis of those messages concerning the distance by which she has failed to pick up the pencil, her motion proceeds "in such a way that … the amount by which the pencil is not yet picked up is decreased at each stage" (Wiener 1948a: 15, 1948b: 212). In this way, negative feedback enables the person to adjust her movements so as to grasp the pencil smoothly and effectively. Hence, Wiener concludes, "Feedback is vital to all voluntary activity" (1948b: 212; also see Ashby 1947a: 139, Ashby 1949: 721-22 and McCulloch 1948: 261-62).

Ashby also examined how the principle of negative feedback can be used to understand the neurological processes through which a person exercises motor control over the movement of his/her hand in order to pick up a pencil. In his 1949 review of Wiener's *Cybernetics*, which—as we shall see—is

---

[14] It is worth noting that Hayek saw himself not just as drawing on Craik's book but also as remedying a lacuna in it by providing a more detailed account of the neurological mechanisms underpinning the production and use of models than was offered by Craik (Hayek 1947: paragraphs 353-59, 371, and p. 151 note to paragraph 354a; Hayek 1952: paragraphs 8.48-8.56). This issue is discussed in more detail in Lewis (2016b).

[15] One of the earliest, and most important, examples of cybernetic modelling to understand human beings is provided by the work of McCulloch and Pitts, who in a path-breaking paper sought to illuminate the neural mechanisms underlying mental phenomena (McCulloch and Pitts 1943). However, Hayek does not refer to their paper in *The Sensory Order*. Moreover, while a reference to one of McCulloch and Pitts' later papers, in which they show how their approach can be used to explain how people can recognise and respond to patterns of stimuli or *gestalten*, appears in the bibliography of Hayek's book, he does not refer to it in the main text, so it is unclear what, if any, influence it exerted on his thinking. Consequently, it will not be discussed further in what follows (Pitts and McCulloch 1947; Hayek 1952: p. 200).

cited by Hayek, Ashby also argues that the process is 'hierarchical' in the sense that while the higher order centres of consciousness in the brain set the overall goal of picking up the pencil, they do not specify detailed behaviour for all the muscles that must be used:

> The higher centre acts so as to make the sub-system A [governing the motion of the hand] goal-seeking for the state at which the distance between hand and pencil is zero. The activity in A is proportional to the 'error' that is, to the amount by which the pencil is not yet reached, to the difference between the two representations of position in the sensory cortex. It will be noticed that the higher centre determines only the goal … the details [that is, the movements of the shoulder, elbow, and fingers whereby the goal is achieved] are left to be determined by A itself [i.e. hierarchically]. (Ashby 1949: 722.)

Once the goal of picking up the pencil has been set, the detailed movements of the relevant muscle groups are governed below the level of full consciousness by a series of nested feedback loops, which successively control the movement of the shoulder, arm, hand, and fingers, so that the overall goal is achieved (Ashby 1949: 722-23).

3.7 The Influence of Cybernetics, Part II: Hayek on the Role of Negative Feedback and Hierarchical Control in Voluntary Human Action

Hayek draws on the notions of negative feedback and hierarchical control in Section 6 of chapter 4 of *The Sensory Order* in order to explain the "manner in which … separate motor impulses are coordinated so as to produce complex patterns of behaviour consisting of many simultaneous and successive movements" (Hayek 1952: paragraph 4.48). Significantly, the motor signals sent out from the higher centres of the nervous system in the course of generating purposeful behaviour "will be for the execution not of one particular pattern of co-ordinated movements but for any one of a class of such patterns" (1952: paragraph 4.49). The process of selection whereby, out of all the members of the class of patterns appropriate for achieving the person's goal, one specific pattern of movement will be selected "must not be conceived as taking place in one act" but rather as involving an iterative process of "continued control, modification, and adjustment" (1952: paragraph 4.53). And it is in conceptualising this process of adjustment that Hayek draws on cybernetics and invokes the notion of negative feedback:

> In connexion with these continuous adjustments, made while the movement proceeds, the interaction between the exteroceptive and the proprioceptive impulses and the operation of the 'feedback principle' become of special significance. In the first instance, the sensory representation of the environment, and of the possible goal to be achieved in that environment, will evoke a movement pattern generally aimed at the achievement of the goal. But at first the pattern of movement initiated will not be fully successful. The current sensory reports about what is happening will be checked against expectations [generated by the person's model of the situation], and the difference between the two will act as a further stimulus. The result of every step in the course of the actions will, as it were, be evaluated against the expected results, and any difference will serve as an indicator of the corrections required. (Hayek 1952: paragraph 4.54.)

In a footnote inserted immediately after the reference to the "feed-back principle", Hayek cites the work of Wiener (1948a, 1948b), McCulloch (1948) and Ashby (1947, 1948, 1949).[16] These are works in which, as noted above, the first generation of cyberneticians not only set out the concept of negative feedback but

[16] In the footnote in question, the surname of Ross Ashby is incorrectly recorded as "Ashley", as indeed it also is in the index of Hayek's book (Hayek 1952: pp. 95, n. 1, 205). The correct name is, however, used in the Bibliography (Hayek 1952, p. 195). Also, the footnote refers simply to "Ashley (1947)", while the bibliography contains references to two works published by Ashby in that year, namely Ashby (1947a, 1947b).

also applied it to voluntary human action. The fact that Hayek did not refer to negative feedback in *WIM'47*, but only in the final version of *The Sensory Order*, indicates that he added this material only after, and indeed because of, reading the work of Wiener, McCulloch and Ashby.

Immediately following the paragraph quoted above, Hayek also notes that the process through which negative feedback controls people's motor processes is hierarchical, in the same sense in which that idea was outlined by Ashby, writing:

> In this process the intervention of the highest centres is probably needed only to give the general directions, while the execution and current adjustment is left to the guidance of the lower centres. Once the 'course is set', the deviations will be automatically corrected by the differences between the expected and the effective stimuli acting as the signs which produce the correction. (Hayek 1952: paragraph 4.54.)

Given the reference to Ashby's work in the paragraph preceding this, it seems reasonable to conclude that it is from Ashby that Hayek drew this idea.

Herein, then, lies the second major influence of cybernetics on the development of Hayek's theoretical psychology: the writings of the cyberneticians afforded Hayek the idea that it is through a hierarchical process of negative feedback that the interplay between the activity of the sensory and motor neurons leads to purposeful human conduct, thereby helping him to conceptualise the psychological foundations of goal-oriented action.

4. CYBERNETICS AND HAYEK'S SOCIAL THEORY

4.1 Hayek's references to cybernetics: An inventory

Scrutiny of Hayek's postwar writings reveals that there are 12 occasions on which he refers explicitly either to cybernetics, or to its correlates such as negative feedback, in the context of his discussions of economic issues.[17]

Hayek's first references to cybernetics in the context of the spontaneous order of the market can be found in two papers first published in 1968, namely "Competition as a Discovery Procedure" (Hayek [1968] 2014) and "The Confusion of Language in Political Thought" (Hayek [1968] 1978). In the introduction to the first of those papers, Hayek states that he intends to "consider competition as a procedure for the discovery of such facts as, without resort to it, would not be known to anyone, or at least would not be utilised" (Hayek [1968] 2014: 304):

> [C]ompetition is valuable *only* because, and so far as, its results are unpredictable and on the whole different from those which anyone has, or could have, deliberately aimed at. Further, that the generally beneficial effects of competition must include *disappointing or defeating some particular expectations or intentions*. (Hayek [1968] 2014: 305; the second set of italics has been added.)

It is the italicised sentence at the end of this quotation, where Hayek states that one of the central themes of his essay is that the benefits of competition are invariably accompanied by the disappointment of some people's expectations, that is important for understanding the contribution made by cybernetics to Hayek's thought. When Hayek develops this point later in his paper, he does so in the context of a

[17] The references are as follows: Hayek ([1968] 2014: 309); Hayek ([1968] 1978: 74); Hayek ([1970] 2014: 345); Hayek (1973: 36-37, 104); Hayek (1976: 71, 94, 124-25); Hayek (1978: 63-64); Hayek (1979: xii, 158); and Hayek (1988: 2). Hayek also briefly mentions cybernetics in three other places, whilst discussing various issues in the philosophy of science (Hayek [1955a] 2014: 205 n. 11, 211 and Hayek [1964] 2014: 261).

discussion of cybernetics. Having just defined social order in terms of plan coordination, Hayek writes that:

> This mutual adjustment of individual plans is brought about by what, since the physical sciences have also begun to concern themselves with spontaneous orders, or 'self-organising systems', we have learnt to call 'negative feedback.' Indeed, as intelligent biologists acknowledge, "long before" Claude Bernard, Clerk Maxwell, Walter B. Cannon, or Norbert Wiener developed cybernetics, "Adam Smith had just as clearly used the idea in his *The Wealth of Nations* (1776). (Hayek [1968] 2014, p. 309; also see Hayek 1979: 158.)

Immediately following this reference to negative feedback, which in this context involves people adjusting their plans—about, for example, what wage or price to charge, or to pay, for some good or service—in order in reduce the gap between actual and expected or target outcomes, Hayek states that, "the fact that a high degree of coincidence of expectations is brought about by the systematic disappointment of some kinds of expectations is of crucial importance for an understanding of the functioning of the market order." In particular, as Hayek makes clear in the next section of the essay, the kinds of expectations that will tend to be disappointed by the operation of the market process will be those arising from people's views about "what is called 'social justice'" (that is, about what prices and wages "appear to be just", where for the people in question justice "means in practice preservation of the traditional structure of incomes and prices") (Hayek [1968] 2014: 309, 311-12).

This reading of Hayek's 1968 paper suggests that he used the concept of negative feedback to help make explicit, to articulate more fully than he had hitherto done, and to draw out the consequences for his broader social philosophy, of one aspect of his account of the market process, namely that it involves the *systematic* disappointment of people's expectations about the wages and prices they will be paid and the violation of their notions of 'social justice'. Hayek develops this theme in the first and second volumes of *Law, Legislation, and Liberty*, arguing that "the steering mechanism of the market" (1976: 71) operates via negative feedback, because it is only through "the systematic disappointment of some expectations that that on the whole expectations are as effectively met as they are" and plan coordination achieved (1973: 104). What this implies, Hayek argues, is that the price mechanism necessarily involves the systematic disappointment of people's expectations about how their efforts ought to be rewarded.

In the two passages just mentioned, Hayek (1973: 104, 1976: 71) directs the reader to a fuller discussion of the fact that the operation of negative feedback involves the systematic disappointment of some kinds of expectations, found in chapter 10 of the second volume of *Law, Legislation and Liberty*. There, in the section entitled, "The correspondence of expectations is brought about by a disappointment of some expectations", Hayek writes as follows:

> The correspondence of expectations that makes it possible for all parties to achieve what they are striving for is in fact brought about by a process of learning by trial and error which must involve a constant disappointment of some expectations. The process of adaptation operates, as do the adjustments of any self-ordering system, by what cybernetics has taught us to call negative feedback: responses to the differences between the expected and the actual results of those actions so that differences will be reduced. This will produce an increased correspondence of expectations of the different persons so long as current prices provide some indications of what future prices will be, that is, so long as, in a fairly constant framework of known facts, always only a few of them change; and so long as the price mechanism operates as a medium of communicating knowledge which brings it about that the facts which become known to some, through the effects of their actions on prices, are made to influence the decisions of others. (Hayek 1976: 124-25.)

As Hayek notes elsewhere in the same section, people change their plans in response to the divergences between their expectations about the market value of their labour or the goods they are trying to sell and what that labour, or those goods, can actually command on the market. And, as already noted, the disappointment of those expectations will also involve outcomes failing to satisfy widely held notions of social justice (also see Hayek 1978: 63).

4.2 The influence of cybernetics on Hayek's economics and social philosophy: An assessment

It is time to take stock. In particular, if it is the case, as Hayek seems to suggest when he directs readers of *Law, Legislation and Liberty* to it, that the paragraph just quoted is a full statement of his views on the significance of cybernetics for economics—and it is the longest passage in his writings on economics where cybernetics is mentioned—then the question arises: what, if anything, do the references to cybernetics and to negative feedback add to Hayek's argument? Do they indicate that his use of concepts drawn from cybernetics facilitated an advance of some kind in his analysis of the working of the price mechanism, enabling him to develop his theory of the market process beyond where it previously stood? Or are they more cosmetic in nature, being used principally to highlight the similarities between the principles of cybernetics and Hayek own, pre-existing theoretical approach, and to re-express Hayek's prior insights in a different idiom without advancing those ideas in a significant way?

The two key, substantive points made in the paragraph about how social order is possible in decentralised market economies are, as well shall see, both long held by Hayek. Moreover, they were arguably derived independently of his knowledge of cybernetics. First, the fact that the price mechanism acts as a means of communication whereby, through the impact of their actions on market prices, people convey to others hints about the significance of their local knowledge about the scarcity of different kinds of resource, was of course outlined by Hayek in his famous 1945 paper on "The Use of Knowledge in Society" (Hayek [1945] 2014: 98-102), which was written and published before Hayek became acquainted with cybernetics. Second, the claim that the rivalrous market process will be effective in coordinating people's plans only if it takes place against the background of a fairly constant array of known facts, most notably in the form of the "regularities of conduct" engendered by a stable set of legal rules, is also a point long made by Hayek. It appears, for example in *The Road to Serfdom*, where Hayek describes abstract legal rules "as a kind of instrument of production, helping people to predict the behaviour of others with whom they must collaborate" (Hayek [1944] 2007: 113). It can also be found in Hayek's Cairo lectures on 'The Political Ideal of the Rule of Law', where he refers to how "if a multitude of individual elements obey certain general laws, this may … produce a definite order of the whole mass without the interference of an outside force" and states that this "applies to the laws obeyed by men no less than to the laws of nature" (Hayek [1955b] 2014: 160; also see Hayek [1945] 2014: 101-02 and Hayek [1960] 2011: 123). Similarly, the insight that people are encouraged to change their plans in ways that make them more compatible by divergences between the expected and the actual results of their actions—as for example when businessmen change their plans because they have led not to the expected profits but to losses—was, of course, one well known to Hayek well before his acquaintance with cybernetics, not least thanks to his acquaintance with Ludwig von Mises' famous essay on economic calculation in a socialist commonwealth (Mises [1920] 1935). Indeed, Hayek himself had long emphasised this divergence between the actual and expected results of his action as a cause of plan revision in his own contributions to the socialist calculation debate of the 1930s and 1940s (Hayek [1935] 1997: 107-11; Hayek [1937] 2014: 63-64).

It appears, then, that the key elements of Hayek's finished account of the working of the market were well known to him, and had been marshalled by him into a coherent account of how social order is possible in decentralised market economies, independent of his acquaintance with cybernetics. Indeed, Hayek himself, both in the 1968 passage from 'Competition as a Discovery Procedure' reproduced above, and also in one of the other passages where he mentions cybernetics in the early 1970s, makes that very point, arguing that economic theory had developed these insights "long before" or "200 years before cybernetics" (Hayek [1970] 2014: 345). Perhaps most starkly, in the second volume of *Law, Legislation*

and Liberty, Hayek refers to the way in which, "What was familiar to Adam Smith has *belatedly* been *re*discovered by scientific fashion under the name of 'self-organising systems' (Hayek 1976: 178, n. 11; emphasis added; also see Hayek 1973: 37-38, 1988: 146). Far from acknowledging that cybernetics had added anything to his thinking, Hayek's concern in these passages is to establish the priority of economics over cybernetics in developing these insights. When coupled with the evidence presented above that the key elements of Hayek's account of social order were present in his work independent of cybernetics, they support the view that when Hayek did eventually mention cybernetics in the context of his discussion of economic issues, its contribution did not lead to a significant advance in his analysis of social order, but rather provided a presentational gloss on ideas Hayek had already developed.

What remains, however, is the claim, outlined above, that ideas drawn from cybernetics did add something to Hayek's broader social philosophy, by enabling him to draw out and articulate more clearly and explicitly than he had done hitherto the fact that the market process works through the systematic disappointment of people's expectations. And while it is hard to view that insight as anything more than a marginal contribution to Hayek's economics, narrowly understood, it is arguably of greater significance for Hayek's broader social philosophy; the notion of negative feedback, and the idea that such feedback manifests itself in the market via the systematic disappointment of people's expectations, was an important conceptual bridge through which Hayek sought to connect his account of the market order or catallaxy with his critique of social justice. It is perhaps for this reason more than any other that Hayek was able to write in the Preface to the final volume of *Law, Legislation and Liberty* that his use of terminology drawn from cybernetics was intended to make his ideas "more precise and unambiguous" and therefore "more readily comprehensible to the contemporary reader" (Hayek 1979: xii).

5. COMMENTARY AND CONCLUSION

In this Section, we summarise the findings of the paper, and also consider their significance for the coherence of Hayek's thought.

It was argued above that ideas drawn from cybernetics had two main influences on Hayek's theoretical psychology, both of which concerned Hayek's attempt to extend his account of the brain as an instrument of classification to encompass purposeful human action. First, the work of Kenneth Craik enabled Hayek to argue that the pattern of nerve impulses in a person's brain constitutes a model that enables the person to form expectations about the consequences of their actions, on the basis of which they are able to decide how to behave. Second, Hayek's efforts to develop an account of the neurological underpinnings of action were also facilitated by his reading of the work of other prominent cyberneticians, most notably as Norbert Wiener, Warren McCulloch, and Ross Ashby, from whom he took the idea that human action is governed by a hierarchical process of motor control acting in accordance with the principle of 'negative feedback'.

As noted in the introduction to this paper, concern has been expressed about the extent to which cybernetics is able to do justice to importance of meaning and to the subjective aspects of purposeful conduct (cf. Caldwell 2004a: 300). If well grounded, such concerns would naturally be troublesome for someone like Hayek, one of whose goals in *The Sensory Order* was of course to develop a psychological foundation for his preferred subjectivist approach to economics. To the extent that Hayek relied on the cybernetics of Wiener, McCulloch and Ashby to ground his theoretical psychology, there appear to be legitimate grounds for concern. Their approach reduced purposive behaviour to the notion of negative feedback, in the sense that it was portrayed as involving nothing more than action taken to reduce the gap between the actual and target outputs of the system. What this excludes is the way that purposeful behaviour involves people formulating an image or symbolic representation of the goal to be pursued, and of their circumstances, which guides their actions. As the historian of cognitive science Jean-Pierre Dupuy (2009: 6, 7) has put it, the process whereby the cybernetic systems described by Wiener, Rosenbluth and McCulloch—but not, significantly, Craik—determine their actions "is not symbolic

computation; that is, computation involving representations. It is purely 'mechanical', devoid of meaning … either with respect to its objects or to its aims":

> The science of mind that cybernetics wished to construct was … resolutely 'eliminativist.' The mental states invoked by ordinary or 'folk' psychology to account for behaviour—beliefs, desires, will, intentions—were to be banned from scientific explanation. (Dupuy 2009: 49.)[18]

In particular, as Dupuy has indicated, what the cybernetics of Wiener, Rosenblueth and McCulloch eliminates is the idea that purposeful human action involves people (having the subjective experience of) developing meaningful symbolic representations of the goal they are pursuing, and of the circumstances in which they act, that guide their subsequent behaviour (also see Heims 1991: 42 and Boden 2006: 220).

Another way of putting this point is to consider the concept of 'information' on which the cyberneticians relied. When the early cyberneticians explored how a system responded to the information it received, the 'information' in question concerned only the physical properties of the signal or message in question. It had no propositional content or meaning. As used by the cyberneticians, 'information' was a technical term referring to a statistical measure of the predictability of a signal, not a semantic notion, pertaining to the conveyance of meaningful ideas (Shannon and Weaver 1949: 3). This purely formal concept of 'information' therefore tells us little about how the information in the feedback loop influences the behaviour of the system. This "flight from meaning and subjectivity", as Boden (2006: 205) described it, reinforces the view that cybernetics is a distinctly an uncomfortable bedfellow for a subjectivist like Hayek (Boden 2006: 204-05, 235-36; Dupuy 2009: 118; also see Taylor 1950).

At least some of these limitations of the cybernetic approach advanced by Wiener and his colleagues were recognised early on by some commentators, perhaps most notably—for our present purposes—by the biologist and founder of system theory Ludwig von Bertalanffy. In an article published in 1950, and read by Hayek, Bertalanffy distinguished between different kinds of goal-oriented or teleological behaviour. One variety, which he describes as "Directiveness based upon structure", means that "an arrangement of structures leads the process in such a way that a certain result is achieved" (Bertalanffy 1950: 159). Bertalanffy notes that this kind of behaviour is governed by negative feedback and refers to examples of cybernetic systems drawn both from the physical world and also from biology. However, Bertalanffy also describes two other kinds of goal-oriented behaviour, both of which he seeks to differentiate from the kind of action studied by cybernetics. Bertalanffy refers in particular to what he terms "true finality or purposiveness, meaning that the actual behaviour is determined by foresight of the goal … It presupposes that the future goal is already present in thought, and directs present action. True purposiveness is characteristic of human behaviour, and it is connected with the evolution of the symbolism of language and concepts" (1950: 160) (cf. Boden 2006: 201, 220). Like the critics described above, therefore, Bertalanffy is sceptical about the capacity of cybernetics to do justice to the way in which truly purposeful human behaviour is guided by meaningful symbolic representations of the goal being pursued.

However, it is less clear that Hayek himself is similarly vulnerable. The reason is that in developing his account of purposeful action, Hayek drew not only on the work of Wiener, McCulloch and Pitts, but also on the writings of other authors, most notably—as we have seen—Kenneth Craik. As noted above, one of the key ideas that Hayek took from Craik was that thought consists in the capacity to form and manipulate symbolic models that represent both the present state of the external world and also the

[18] See, for example, Ashby (1949: 721), who claims that cybernetics involves the notion of 'purpose' being replaced by physical concepts: "Cybernetics … treats these actions [by people on the external world] … by methods which rest not on the metaphysical concepts of purpose, instinct, and libido, but on the physical concepts of structure, energy and dynamics." Similar views were expressed by McCulloch and Walter Pitts, according to whom "both the formal and the final aspects of that activity which we are wont to call *mental* are rigorously deducible from present neurophysiology … [so that] 'Mind' no longer 'goes more ghostly than the ghost'" (McCulloch and Pitts 1943: 132).

possible outcomes of the various courses actions open to people. As Hayek put it in his account of how such models make it possible to "try out on the model the effects to be expected from alternative courses of action" and thereby facilitate purposeful conduct, it is these "symbolic representations of the consequences to be expected from a given representation of events which we must conceive as those 'symbolic processes in the brain' which physiological psychology has been led to postulate in order to account for the complex adaptive responses [of the person to his or her environment]" (Hayek 1952: paragraphs 5.61-62). Hence, by drawing on Craik's notion of the model, Hayek was able to conceptualise how people are able to engage in what Boden felicitously terms "symbolic foresight"—that is, to form meaningful symbolic representations of their goals that mediate and guide their actions—in a way denied to Wiener *et al*. (Boden 2006: 201; also see pp. 218, 220). Arguably, therefore, Hayek's theoretical psychology is less vulnerable than the work of Wiener *et al*. to the charge that it eliminates the possibility of meaningful, subjective action.[19]

Further support for this claim is provided by the fact that, in articulating the philosophical significance of his theoretical psychology, Hayek drew on Bertalanffy's organicist biology in order to argue that mental phenomena are emergent properties of the structured arrangement of neurons found in the brain and, as such, are irreducible to the (physical) properties of the individual neurons of which the brain is comprised. On this account, as Hayek makes clear, the mind's capacity to imbue events with meaning and to initiate purposeful action, as well as its ability to generate the phenomenal world of sense experience, are all emergent phenomena that supervene on, but are irreducible to, mechanical, neuronal activity (Lewis 2016a, 2016c, 2016d).

Bertalanffy also identified what he saw as a second notable shortcoming of cybernetics, namely that the notions of negative feedback and homeostasis were insufficient to account for the phenomena of change, evolution, and creativity characteristic of living systems (Bertalanffy 1952: 132-46, 1969: 15-27, 44, 149-50, 163; also see Dupuy 2009: 74-75, 131-32, 126-36, 156-58; Hammond 2003: 120-21, 125-29; Lewis 2016a: 143-45, 2016b n. 27). However, once again, Hayek is arguably less vulnerable to this criticism than the early cyberneticians. For one of the key features of Hayek's postwar work in economics and social theory is his portrayal of the economy as a complex system that develops through a multi-level evolutionary process in which the system of rules that structures people's interactions changes over time. Given that, as I have argued elsewhere, Bertalanffy's writings provided some of the key building blocks of Hayek's account, and given also Bertalanffy's emphasis on accounting for the origins of social structures rather than—as did cybernetics—treating them as fixed, it is unsurprising that Hayek did not make much use of cybernetics in developing his account of the economy as a complex adaptive system, precisely because it did not provide him with the conceptual resources to do justice to the evolution of the rules that give the economy its structure (Lewis 2016a: 143-45, 2016c n. 27). Indeed, as Olivia (2016: section 4.3) has perceptively observed, Hayek explicitly distanced himself from cybernetics on these very grounds ([1967] 2014: 282-83).

It should come as no surprise, therefore, that, as noted above, the impact of cybernetics on Hayek's economics was distinctly limited. As we have seen, the principal contribution made by cybernetics to Hayek's later work in economics and social theory lay in the way that the notion of negative feedback—that is, the idea that a system adjusts its behaviour in order to reduce the difference between its actual and expected or target outcome—afforded Hayek a conceptual vehicle that enabled him to articulate his idea that a *systematic* disappointment of people's expectations was an integral part of the market process. This can hardly be described as a major advance in Hayek's economics. But it is of greater significance for Hayek's broader social philosophy because, as we have seen, it helped him to connect his analysis of the market process with his critique of the notion of social justice and to argue,

[19] There is nothing intrinsic to the notion of negative feedback that renders it necessarily incompatible with the idea that people's actions are guided by representations of the goal for which they are striving. On the contrary, as Boden (2006: 220) observes, it is perfectly possible to develop a coherent theoretical account of purposeful action that marries an emphasis on negative feedback with an acknowledgement of the importance of symbolic representations, as shown for example in the late 1950s by the work of Newell *et al*. (1959).

more specifically, that because the working of the market invariably involves the disappointment of the expectations held by some (groups of) people about what remuneration they will receive, it imposes a constraint on our ability to achieve certain notions of (social) justice. In this way, the analysis presented above sheds light on one of the conceptual influences on Hayek's broader political economy.

In the light of this, we end by returning to the other remarks, reported in the introduction to this paper, that various commentators have made about the coherence and consistency of Hayek's use of cybernetics. As noted above, Hodgson has argued that Hayek's use of ideas drawn from cybernetics betrays a shortcoming in his efforts to theorise the notion, central to his account of the market process, of 'spontaneous order'. The analysis presented above gives some ground for thinking that Hayek is less vulnerable to the charge of a lack of clarity and consistency than his critics have suggested. First, as we have just seen, Hayek was aware of the shortcomings of some of the approaches and concepts listed by Hodgson, perhaps most notably those drawn from cybernetics (including the notion of homeostasis), as vehicles for expressing the insights of the evolutionary or Scottish Enlightenment tradition. This is arguably why he relied more on ideas drawn from Bertalanffy's system theory than from cybernetics in developing his account of the development and operation of the market system. Second, as also argued above, to the extent that Hayek did use terminology drawn from cybernetics in his account of spontaneous market order, he did so principally as a presentational device to render more intelligible to his readers an explanation of the spontaneous order of the market that he had developed without significant influence from cybernetics, not in order to develop substantively his account of the market order. The fact that Hayek's use of cybernetics was principally presentational rather than substantive suggests that his account of the market order is not fundamentally compromised by the shortcomings of the cybernetics approach, or its inconsistency with some of the other ideas upon which he drew. What does remains open to question, however, is the merits of Hayek's choice of cybernetics as a presentational vehicle. Put slightly differently, if there is an inconsistency here, it is less a fundamental inconsistency in Hayek's substantive account of how a catallaxy arises and more an inconsistency between that underlying theory and the terminology he sometimes uses to put a gloss on it.[20]

Finally, Witt (2013) has observed that Hayek's use of ideas drawn from cybernetics is incomplete because it has been limited to the notion of negative feedback and did not involve Hayek drawing on the notion of positive feedback. In light of the analysis presented above, two remarks are in order. First, Witt's criticism is fair in that the evidence suggests Hayek would have been aware of the notion of positive feedback, that concept having been discussed in readings cited by Hayek (see footnote 6 above). Second, Hayek's incomplete appropriation of the ideas developed by the cyberneticians is perhaps rendered more intelligible once it is realised that his principal goal in drawing on their work for his economics was to add a presentational gloss—but not significant substantive content—to his pre-existing account of the market order. Viewed in that light, his incomplete appropriation of ideas drawn from cybernetics is unsurprising, simply because he was drawing only upon those ideas that gave voice to his already-developed theory of the market order, not seeking out new ideas that would augment, and possibly call into question, his extant theory. This is not necessarily to defend the limited scope of Hayek's use of concepts drawn from cybernetics in his economics, but rather to make his position understandable. On this view, to adapt the words of Philip Mirowski, it is perhaps better to think of Hayek as someone who "filtered various cyborg themes into economics … in a merely ceremonial manner … motivated to search them out by his prior commitment to the metaphor of the market as a [spontaneous order]" (cf. Mirowski 2002: 238, 548).

[20] For more on the notion of order to which Hayek subscribes, see Fleetwood (1995), Vaughn (1999), and Lewis (2014).

REFERENCES

Ashby, W.R. (1945). "The Physical Origin of Adaptation by Trial and Error.' *The Journal of General Psychology,* 32: 13-25.

Ashby, W.R. (1946). 'Dynamics of the Cerebral Cortex: The Behavioural Properties of Systems in Equilibrium.' *The American Journal of Psychology*, 59: 682: 86.

Ashby, W.R. (1947a). 'Dynamics of the Cerebral Cortex. Automatic Development of Equilibrium in Self-Organising Systems." *Psychometrica*, 12: 135-40.

Ashby, W. R. (1947b). 'Principles of the Self-Organizing Dynamic System.' *The Journal of General Psychology,* 37: 125–8.

Ashby, W. R. (1948). 'Design for a Brain.' *Electronic Engineering*, 20: 379-383.

Ashby, W.R. (1949). 'Critical Review: The Facts and Methods of Cybernetics.' *Journal of Mental Sciences,* 95: 716-24.

Bertalanffy, L. v. (1950). 'An Outline of General System Theory.' *British Journal for the Philosophy of Science*, 1: 134-65.

Boden, M. (2006). *Mind as Machine: A History of Cognitive Science, Volume I*. Oxford: Oxford University Press.

Burczak, T. (2001). 'Response to Butos & Koppl: Expectations, Exogeneity, and Evolution.' *Review of Political Economy*, 13: 87-90.

Butos, W. and R. Koppl (1993). 'Hayekian Expectations: Theory and Empirical Applications.' *Constitutional Political Economy*, 4: 303-29.

Butos, W. and R. Koppl (1997). 'The Varieties of Subjectivism: Keynes and Hayek on Expectations.' *History of Political Economy*, 29: 327-59.

Butos, W. and R. Koppl (2001). 'Confidence in Keynes and Hayek: Reply to Burczak.' *Review of Political Economy*, 13: 81-86.

Caldwell, B. (2004a). *Hayek's Challenge: An Intellectual Biography of F.A. Hayek*. Chicago and London: The University of Chicago Press.

Caldwell, B. (2004b). 'Some Reflections on F.A. Hayek's *The Sensory Order*.' *Journal of Bionomics,* 6: 239-54.

Cannon, W. (1932). *The Wisdom of the Body*. New York: W. W. Norton.

Carabelli, A. and N. De Vecchi (2001). 'Hayek and Keynes: From a Common Critique of Economic Method to Different Theories of Expectations.' *Review of Political Economy,* 13: 269-85.

Craik, K.J. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press.

Dupuy, J.-P. (2009). *On the Origins of Cognitive Science: The Mechanization of Mind*. Cambridge, MA.: MIT Press.

Festré, A. and P. Garrouste (2016). 'Hayek on Expectations: The Interplay between Two Complex Systems.' GREDEG Working Paper No. 2016-13. Available online at: http://www.gredeg.cnrs.fr/working-papers/GREDEG-WP-2016-13.pdf. Accessed 19th May 2016.

Fleetwood, S. (1995). *Hayek's Political Economy: The Socio-Economics of Order*. London New York: Routledge.

Frank, L. (1948). 'Forward.' *Annals: New York Academy of Sciences*, 50: 189-220.

Gaus, G. (2006). 'Hayek on the Evolution of Society and Mind.' In E. Feser (ed.), *The Cambridge Companion to Hayek*. Cambridge: Cambridge University Press.

Hammond, D. (2003). *The Science of Synthesis: Exploring the Social Implications of General Systems Theory*. Boulder: University Press of Colorado.

Hardin, G. (1961). *Nature and Man's Fate*. New York: A Mentor Book.

Hayek, Friedrich A. 1920. "Beiträge zur Theorie der Entwicklung des Bewusstseins." Unpublished manuscript. Friedrich Hayek Papers, Hoover Institution Archives, box 93, folder 1.

Hayek, F. A. ([1920] Forthcoming). "Contributions to a Theory of How Consciousness Develops." Translation by Grete Heinz of "Beiträge zur Theorie der Entwicklung des Bewusstseins." In *The Sensory Order and Other Essays*, edited by Viktor Vanberg. Vol. 14 of *The Collected Works of F. A. Hayek*. Chicago: University of Chicago Press.

Hayek, F. A. ([1935] 1997) 'The Present State of the Debate.' In F.A. Hayek, *The Collected Works of F.A. Hayek, Volume 10: Socialism and War: Essays, Documents, Reviews*. Edited by B. Caldwell. Chicago: The University of Chicago Press.

Hayek, F.A. (1945). *What is Mind? A Working Hypothesis on the Origin of Sensory Qualities and Abstract Concepts*. Unpublished paper. Hayek archives.

Hayek, F.A. (1947). *What is Mind? An Essay in Theoretical Psychology*. Unpublished paper. Hayek archives.

Hayek, F.A. ([1955a] 2014). 'Degrees of Explanation.' In F.A. Hayek (2014).

Hayek, F.A. ([1955b] 2014). 'The Political Ideal of the Rule of Law.' In F.A. Hayek (2014).

Hayek, F.A. ([1960] 2011). *The Constitution of Liberty*. Edited by R. Hamowy, vol. 17 (2010) of *The Collected Works of F. A. Hayek*. Chicago: The University of Chicago Press.

Hayek, F.A. ([1964] 2014). 'The Theory of Complex Phenomena.' In Hayek (2014).

Hayek, F.A. ([1967] 2014). 'Notes on the Evolution of Systems of Rules of Conduct.' In F.A. Hayek (2014).

Hayek, F.A. ([1968] 1978). 'The Confusion of Language in Political Thought.' In F.A. Hayek (1968), *Studies in Philosophy, Politics and Economics*. London: Routledge and Kegan Paul.

Hayek, F.A. ([1968] 2014). 'Competition as a Discovery Procedure.' In Hayek (2014).

Hayek, F.A. ([1970] 2014). 'The Errors of Constructivism.' In F.A. Hayek (2014).

Hayek, F.A. (1973). *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy. Volume I: Rules and Order*. London and New York: Routledge.

Hayek, F.A. (1976). *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy. Volume II: The Mirage of Social Justice*. London and New York: Routledge.

Hayek, F.A. (1979). *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy. Volume III: The Political Order of a Free People*. London and New York: Routledge.

Hayek, F.A. (1988). *The Fatal Conceit: The Errors of Socialism*. London and New York: Routledge.

Hayek, F.A. (2014). *The Collected Works of F.A. Hayek, Volume 15: The Market and other Orders*. Edited by B. Caldwell. Chicago: The University of Chicago Press.

Hayek, F.A. (n.d.). 'Within Systems and about Systems: A Statement of Some Problems of a Theory of Communication.' Typescript. Hoover Institution, Hayek Archives, box 104, folder 22.

Heims, S. (1991). *The Cybernetics Group*. Cambridge, MA: MIT Press.

Lewis, P.A. (2014). 'Hayek: From Economics as Equilibrium Analysis to Economics as Social Theory.' In R. Garrison and N. Barry, eds., *Elgar Companion to Hayekian Economics*. Cheltenham: Edward Elgar.

Lewis, P.A. (2016a). 'The Emergence of "Emergence" in the Work of F. A. Hayek.' *History of Political Economy*, 48: 111-50.

Lewis, P.A. (2016b). "The Origins of Hayek's 'Model': The Significance of Kenneth Craik." Unpublished paper, Department of Political Economy, King's College London.

Lewis, P.A. (2016c). 'Systems, Levels of Organisation and Structural Properties: The Influence of Ludwig von Bertalanffy on the Development of the Work of F.A. Hayek.' Forthcoming in *Research in the History of Economic Thought and Methodology*.

Lewis, P. (2016d). 'Ontology and the History of Economic Thought: The Case of Anti-reductionism in the Work of Friedrich Hayek.' Unpublished paper, Department of Political Economy, King's College London.

Lewis, P. and P. Lewin (2015). 'Orders, Orders, Everywhere: On Hayek's *The Market and Other Orders*.' *Cosmos and Taxis*, 2(2): 1-17.

McCulloch, W. (1948). 'A Recapitulation of the Theory, with a Forecast of Several Extensions.' *Annals: New York Academy of Sciences*, 50: 259-77.

Mirowski, P. (2002). *Machine Dreams: How Economics became a Cyborg Science*. Cambridge: Cambridge University Press.

Mirowski, P. (2007). 'Natuaralizing the Market on the Road to Revisionism: Bruce Caldwell's *Hayek's Challenge* and the Challenge of Hayek Interpretation.' *Journal of Institutional Economics*, 3: 351-72.

Mises, L. von ([1920] 1935). 'Economic Calculation in the Socialist Commonwealth.' In F. A. Hayek (ed.), *Collectivist Economic Planning*. London: Routledge and Kegan Paul Ltd.

Newell, A., J. Shaw, and H. Simon (1959). *Report on a General Problem-Solving Program*. The Rand Corporation.

Olivia, G. (2016). 'The Road to Servo-Mechanisms: The Influence of Cybernetics on Hayek from *The Sensory Order* to the Social Order.' Available online at: http://www.ineteconomics.org/uploads/papers/Apr-Oliva_The-Road-to-Servomechanisms.pdf.

Pitts, W. and W. McCulloch (1947). 'How we Know Universals: The Perception of Auditory and Visual Forms.' *Bulletin of Mathematical Biophysics*, 9: 127-47.

Rosenbluth, A. and N. Wiener (1943). 'The Role of Models in Science.' *Philosophy of Science*, 10: 18-24.

Rosser, J.B. (2015). "Complexity and Austrian Economics.' In P. Boettke and C. Coyne (eds.), *The Oxford Handbook of Austrian Economics*. Oxford: Oxford University Press.

Shannon, C. and W. Weaver (1949). *The Mathematical Theory of Communication*. Urbanna, IL.: University of Illinois Press.

Taylor, R. (1950). 'Comments on A Mechanistic Conception of Purposefulness.' *Philosophy of Science*, 17: 310-17.

Vaughn, K.I. (1999). 'Hayek's Implicit Economics: Rules and the Problem of Order.' *Review of Austrian Economics*, 11: 128-44.

Wiener, N. (1948a). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Paris: Hermann.

Wiener, N. (1948b). "Time, Communication, and the Nervous System." *Annals: New York Academy of Sciences*, 50: 197-220.

Witt, U. (2013). "Competition as an Ambiguous Discovery Procedure.' *Economics and Philosophy*, 29: 121-38.