8 Gordon Pask and His Maverick Machines

Jon Bird and Ezequiel Di Paolo

A computer that issues a rate demand for nil dollars and nil cents (and a notice to appear in court if you do not pay immediately) is not a maverick machine. It is a respectable and badly programmed computer. . . . Mavericks are machines that embody theoretical principles or technical inventions which deviate from the mainstream of computer development, but are nevertheless of value.

—Gordon Pask (1982a, p. 133)

Gordon Pask (1928–1996) is perhaps most widely remembered for his technical inventions in the field of automated teaching. Less widely appreciated are the theoretical principles embodied in Pask's maverick machines. He described himself as a ''mechanic philosopher'' (Scott 1980), and building machines played a central role in the development of a conceptual framework that resulted in two theories later in his career: Conversation Theory (CT) (Pask 1975) and Interaction of Actors Theory (de Zeeuw 2001). Even adherents of these theories concede that they are difficult to understand. Pask wrote over two hundred fifty papers and six books and his prose can be hard to follow and his diagrams difficult to untangle. B. Scott (1980, p. 328), who collaborated with Pask on CT, characterizes some of Pask's writing as ''esoteric, pedantic, obscurantist.'' R. Glanville (1996), who wrote his doctorate under Pask's supervision, admits that CT is ''in many parts very hard to understand, because of a tendency to present it all, all the time, in its full complexity.'' Pask's presentations were dramatic and furiously paced and often left the audience baffled. Consequently, ''some dismissed him, almost with resentment because of their inability to come to terms with him, but others recognised something both intriguing and important in what he said and the way that he said it. I myself often found I had lost the thread of what Gordon was saying, yet strangely he was triggering thoughts and insights'' (Elstob 2001, p. 592). The psychologist Richard Gregory, who was a contemporary of Pask's at Cambridge, remembers

Figure 8.1
Gordon Pask (c. 1963). Printed with permission of Amanda Heitler.

(2001), "A conversation with Gordon is (perhaps too frankly) memorable now as being extraordinarily hard to understand at the time. Or is this just my inadequacy? He would come out with an oracular statement, such as 'Life is fire,' and would defend it against all objection. No doubt it had a certain truth, but I for one was never quite clear whether he was dealing in poetry, science, or humour. This ambiguous mixture was a large part of his charm" (p. 686). However, Gregory acknowledges that "without doubt, Gordon was driven by genuine insight" (p. 685). Heinz von Foerster and Stafford Beer, who both collaborated closely with Pask, also rated his intellect very highly, describing him as a genius (von Foerster 2001, p. 630; Beer 2001, p. 551).

In this chapter we focus on the early period of Pask's life, tracing the development of his research from his days as a Cambridge undergraduate to the period in the late 1950s when his work started to have an impact internationally. We describe three of his maverick machines: Musicolour, a sound-actuated interactive light show; SAKI, a keyboard-skill training machine; and an electrochemical device that grew an "ear." We assess the value of these machines, fifty years after they were built, in particular, the maverick ideas that they embody. We hope this will not only provide a way in to the challenging Paskian literature for the interested reader, but also demonstrate that many of Pask's ideas remain highly relevant for many current research areas.

School and University

What do we mean by conflict? Basically, that two or more time sequences of computation, which may have been proceeding in parallel, interact. Instead of remaining parallel and (by the definition of parallel) separate, they converge in a head-on collision from which there is no logical-deductive retreat.
—Gordon Pask (1982a, p. 62)

School Years

Pask stood out at Rydal, a Methodist public school in North Wales, where he was a boarder during the Second World War.[1] It was fairly liberal, but the headmaster, a prominent churchman, had a reputation for severity and would beat pupils (a common practice in public schools at the time). Pask's dress sense distinguished him from his fellow pupils and made him seem older than he was; he wore: double-breasted business suits and bow ties, compared to the blazers and gray flannel trousers of his contemporaries. It was a style that he kept for the rest of his life (adding an Edwardian cape once he had left school). He was a small and sickly child and did not excel on the sports field—a very important part of Rydal culture (the school's two most famous alumni distinguished themselves as international rugby players). He spent his spare time building machines, for example, a device to detect rare metals that he tested out in nearby mines. A story about another one of his inventions, possibly fantasy, circulated through the school and contributed to Pask's reputation as a "mad professor." It was said that at the beginning of the Second World War he sent the War Office a design for a weapon. After a few months he received a reply stating that his proposal had been considered and it was thought it would work, but its effect was too dreadful to be employed against a human enemy.

Although his were not the usual preoccupations of teenage boys, he was not disliked, as he had a sense of fun and mischief. As a prank he would deflate large numbers of rugby balls that the sports master had inflated and left outside his room ready for the next day's sports activities. Pask also demonstrated his independence by slipping away from school some evenings, catching the train to Liverpool, and returning in the early hours. He said he was involved in producing stage shows in the city.[2] One day the whole school was summoned to a general assembly, as was always the case when the headmaster wanted to single out somebody in front of the school for disciplinary offenses. Nobody knew who the offender was until his name was announced. Pask's absence had been discovered the previous evening and the headmaster publicly berated him. Pask was not cowed and in fact took offense at his treatment: he stood up and stormed out of the hall, telling the headmaster, "I shall speak to my solicitor about this." Apparently he escaped a beating.

Pask did not do national service after Rydal, perhaps because of ill health. Instead he went to Liverpool Technical College, where he studied geology and mining. In 1949 he went to Downing College, Cambridge University, to study medicine. He had a vivid impact on his contemporaries, just as he had done at school.

Cambridge

At Cambridge Pask read Norbert Wiener's *Cybernetics*, which had an "emotional impact" on him (Pask 1966). He had found a field of study that was broad enough to accommodate his wide range of interests and also combined theory and practice: "As pure scientists we are concerned with brain-like artifacts, with evolution, growth and development; with the process of thinking and getting to know about the world. Wearing the hat of applied science, we aim to create . . . the instruments of a new industrial revolution—control mechanisms that lay their own plans" (Pask 1961, p. 11). Pask met Robin McKinnon-Wood, a physicist, at Cambridge, and they began to build machines together. It was a relationship that continued for the rest of their lives. When they graduated they set up System Research Ltd., a company that sold versions of the machines that they had first started developing as undergraduates.

Pask also began to investigate statistical phenomena. Cedric Price, the architect, knew him as an undergraduate and was roped into some statistical experiments: "'It's simple, just throw these wooden curtain rings as quickly as possible into the numbered box—which I shall call out. Then do it backwards with a mirror, then blindfolded.' He took my arm and led me into Jordan's Yard. I could see that he was not to be trifled with" (Price 2001,

p. 819). This strange-sounding experiment was Pask's way of generating different probability distributions in order to predict the enlistment numbers for the RAF in the year 2000.

Stationary and Nonstationary Systems

A broad distinction that can be drawn about the statistics of a series of events is whether they are *stationary* or *nonstationary*. A scientist observing the behavior of a system over time might identify some regularities, for example, if the system is in state A, it goes to state B 80 percent of the time and to state C 20 percent of the time. If this behavior sequence is invariant over a large number of observations of the same system, or an ensemble of similar systems, then an observer can infer that statistically the system is stationary. The observed properties are *time-independent*, that is, various statistical measures, such as the mean and standard deviation, remain invariant over time. Therefore, given the occurrence of A we can be confident about the probability of B or C following, irrespective of what time we observe A.

Nonstationary systems do not display this statistical invariance; there are time-dependent changes in their statistical properties, and the relationship between A, B, and C can change. Human behavior, for example, is often nonstationary, as was dramatically demonstrated by Pask when he was studying medicine. He would get through anatomy tests by memorizing footnotes from Gray's *Anatomy*; by dazzling on some arcane anatomical details he usually managed to cast shadows over the holes in his knowledge. But on occasion he got found out. Gregory (2001) recalls an anatomy exam where Pask was asked to dissect an arm. One might predict, having observed the behavior of other anatomy students, that he would use a scalpel. Instead, he used a fire axe, smashing a glass dissecting table in the process. Unsurprisingly, Pask graduated from Cambridge in physiology, rather than medicine.

Learning provides less dramatic examples of nonstationary behavior. We can measure the skill of a novice at performing some skill, for example, typing, by recording the person's average response time and error rate. As the novice practices, his skills will improve, and although his performance might be stationary for periods of time, it will also show discontinuities as it improves. Dealing with nonstationary systems is a challenge, as their behavior is difficult to characterise. Pask started developing two learning machines while he was an undergraduate and developed a mechanical and theoretical approach to dealing with nonstationary systems. In the next two sections we describe these machines in detail.

Musicolour

> Man is prone to seek novelty in his environment and, having found a novel situation, to learn how to control it.
> —Gordon Pask (1971, p. 76)

Pask built the first Musicolour system, a sound-actuated interactive light show, in 1953. Over the next four years, Pask, McKinnon-Wood, their wives, and a number of other individuals were involved in its development (Pask 1971). Pask's initial motivation for building the system was an interest in synesthesia and the question of whether a machine could learn relations between sounds and visual patterns and in doing so enhance a musical performance. From the outset, Musicolour was designed to cooperate with human performers, rather than autonomously generate ''aesthetically valuable output'' (Pask 1962, p. 135). The way musicians interacted with the system quickly became the main focus of research and development: the performer ''trained the machine and it played a game with him. In this sense, the system acted as an extension of the performer with which he could co-operate to achieve effects that he could not achieve on his own. Consequently, the learning mechanism was extended and the machine itself became reformulated as a game player capable of habituating at several levels to the performer's gambits'' (Pask 1971, p. 78).

How Does Musicolour Work? The sounds made by the musicians are relayed to the system via a microphone and amplifier. A bank of filters then analyze various aspects of the sound (see figure 8.2; the system had up to eight filters, but five are shown). An early system just used band-pass filters, but in later systems there were also filters that analyzed attack and rhythm. Each of the filters has a parameter that can take one of eight prespecified values. These values determine the frequency range of the band-pass filters and delays in the attack and rhythm filters.

The output from each filter is averaged over a short period, rectified, and passed through an associated adaptive threshold device (figure 8.2). If the input exceeds a threshold value, the output is 1, otherwise it is 0. These devices adapt their threshold to the mean value of the input, habituating to repetitive input, for example a continuous sound in a particular pitch band, and outputting 0. The outputs from the adaptive threshold devices

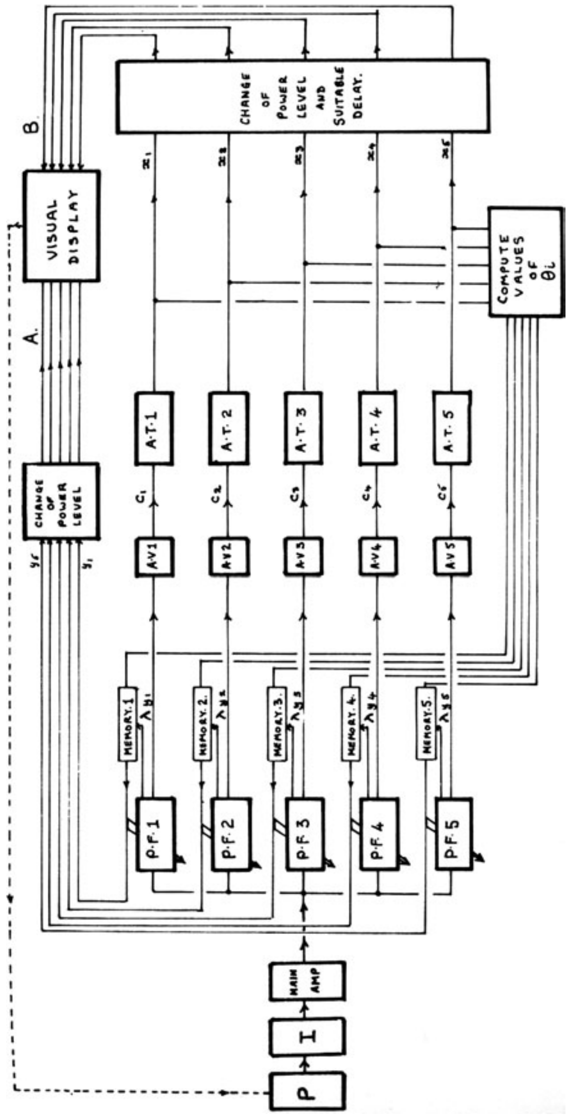

Figure 8.2
Diagram of a typical Musicolour system. P = performer; I = instrument and microphone; AT = adaptive threshold device; A = inputs to the visual display that determine *what* patterns are projected; B = inputs to the visual display that determine *when* the patterns are projected. From Pask (1971). Reprinted with permission of Jasia Reichardt.

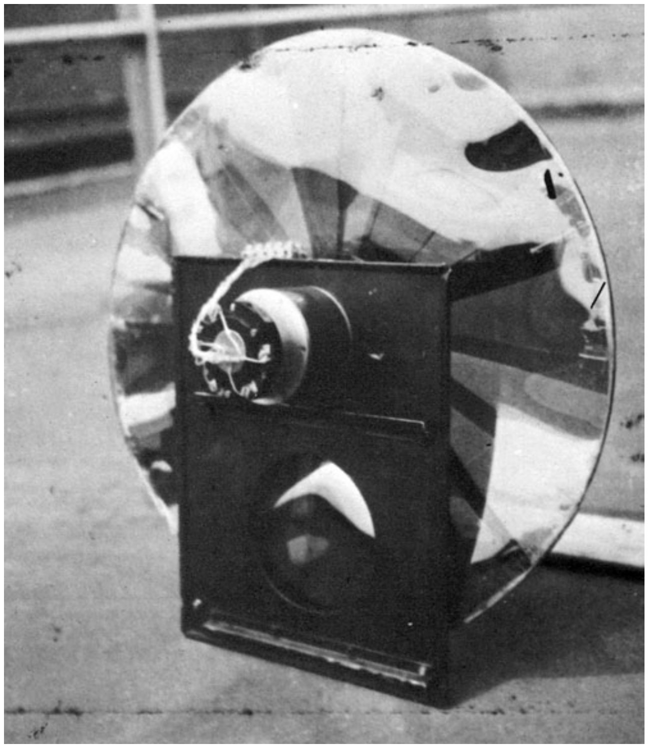

Figure 8.3
A servo-positioned pattern wheel used in Musicolour. From Pask (1971). Reprinted with permission of Jasia Reichardt.

determine *when* a selection is made from the available visual patterns by controlling dimmers connected to the lights.

The values of the filter parameters determine *what* visual pattern is selected by controlling a servo-positioned pattern or color wheel (see figure 8.3). The particular parameter values are selected on the basis of how different the output of the filter's associated adaptive threshold device is, compared to the other filter's thresholded outputs, and how long it is since a particular value has been selected. The selection strategy aims to increase the novelty of the filter outputs and to ensure that all of the parameter values are sampled.[3]

If the input to Musicolour is repetitive, it habituates and adjusts its filter parameter values in an attempt to generate more variety in the light patterns. If there is no input, the system becomes increasingly sensitive to any sound in the environment and a gain control prevents this from disrupting the system too much.

Was It a Success? Musicolour was found to be ''eminently trainable'' (Pask 1971, p. 80). Performers were able to accentuate properties of the music and reinforce audio-visual correlations that they liked (for example, high notes with a particular visual pattern). Once performers became familiar with the filter-value selection strategy of the machine, they were able to establish time-dependent patterns in the system and reinforce correlations between groups of musical properties. It is important to note that there was no fixed mappings between sounds and lights: these were developed through the interaction of the musicians with Musicolour. There is reciprocal feedback between Musicolour and the performers: ''The machine is designed to entrain the performer and to couple him into the system'' (Pask 1971, p. 80). From the performer's perspective, ''training becomes a matter of persuading the machine to adopt a visual style that fits the mood of his performance,'' and when the interaction has developed to this level ''the performer conceives the machine as an extension of himself'' (p. 86). Pask did some ''rough and ready'' studies of how visual patterns affect performance, finding that short sequences of visual events acted as releaser stimuli (p. 86).[4] It was also found that once a stable coordinated interaction had been established, it was robust to a certain level of arbitrary disturbances.

Musicolour developed from a small prototype machine that was tested at parties and in small venues to a large system that toured larger venues in the north of England and required two vans to transport the equipment and five people to set it up. After this tour, Musicolour was used in a theatrical performance at the Boltons Theatre in 1955, where it was combined with marionettes in a show called *Moon Music*. Musicolour and puppets were ''unhappy bedfellows,'' and after a week of technical problems, the stage manager left and the show closed (Pask 1971, p. 81). With Jone Parry, the music director for Musicolour, Pask and McKinnon-Wood then used the month's paid-up rental on the theater to develop the musical potential of the system, and the show became a concert performance. Subsequently, Pask developed a work, *Nocturne*, in which he attempted to get dancers interacting with Musicolour. This was technically challenging, but Pask thought it showed some artistic potential.

The Musicolour project began to fall into debt and Pask explored different ways of generating income, ranging from adapting it for juke boxes (then at the height of their popularity) to marketing it as an art form. Bankruptcy was avoided by a regular gig at Churchill's Club in London (and by Cecil Landau becoming a partner in the business). People participated in the system by dancing, responding to the music and light show. After a year Musicolour moved to another club, the Locarno in Streatham, London, a large ballroom with a capacity of several thousand as well as a huge lighting rig (120 kW), which Musicolour modulated. This cavernous environment was not conducive to audience participation as there were too many other visual elements, such as exit signs, that distracted dancers from the visual display. Churchill's Club had been more intimate, and Musicolour had integrated with the space. Pask (1971) says that Landau "was prone to regard an archway across the middle of the night-club as a surrogate proscenium and everything beyond it a stage" (pp. 87–88). In larger, commercially viable spaces, Musicolour became just "another fancy lighting effect" and it "was difficult or impossible to make genuine use of the system" (p. 88). In 1957, after a final performance at a ball in London, Musicolour was shelved and Pask and McKinnon-Wood concentrated on the commercial development of their teaching machines.

SAKI

Teaching is control over the acquisition of a skill.
—Gordon Pask (1961, p. 88)

In 1956, Pask, his wife, Elizabeth, and Robin McKinnon-Wood applied for a patent for an "Apparatus for Assisting an Operator in Performing a Skill."[5] This patent covers a wide range of teaching machines built along cybernetic principles, including SAKI (self-adaptive keyboard instructor), which Stafford Beer (1959, p. 123) described as "possibly the first truly cybernetic device (in the full sense) to rise above the status of a 'toy' and reach the market as a useful machine." SAKI trains people to operate a Hollerith key punch (see figure 8.4), a device that punches holes in cards used for data processing.[6] By pressing keys the operator makes holes in selected columns on the cards to encode data in a form that can be read by a card reader and stored in computer memory. The Hollerith keyboard was designed to be operated with one hand and had twelve keys: 0 to 9, an X, and a top key. One digit can be entered per column by pressing the corresponding key. Alphabetic characters are entered by punching two holes in the same column: the top key and 1 to 9 for A to I, the X key and 1 to 9 for J to R, and 0 and 1 to 9 for S to Z. Up until the 1970s the key punch was

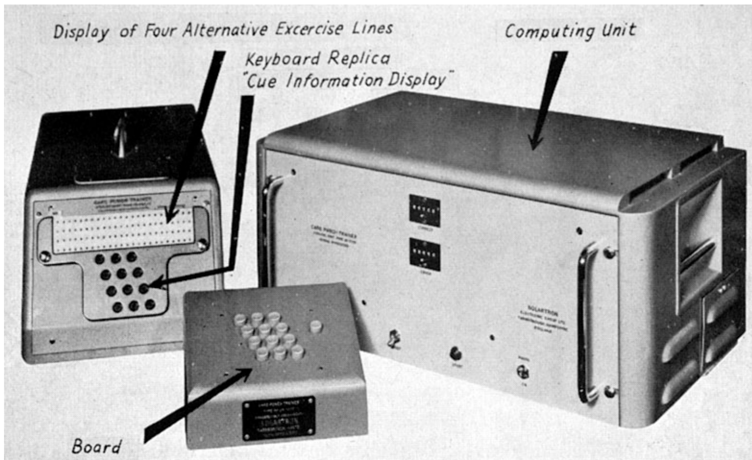

Figure 8.4
SAKI (self-adaptive keyboard instructor). Image taken from Plate II, Gordan Pask (1960) *An Approach to Cybernetics*, Harper and Brothers, with kind permission of Springer Science and Business Media.

a common form of data entry and there was a large demand for skilled operators.

One challenge in automating teaching is to ensure that a student's interest is sustained: ''Ideally the task he is set at each stage should be sufficiently difficult to maintain his interest and to create a competitive situation yet never so complex that it becomes incomprehensible. A private tutor in conversation with his pupil seeks, in fact, to maintain this state which is not unlike a game situation'' (Pask, McKinnon-Wood, and Pask 1961, p. 32). It requires that the tutor respond to the particular characteristics of a pupil. A multitude of factors determine a person's skill level (previous experience, motor coordination, level of tiredness) and some of these factors will change as a result of the learning process. Pask's novel approach was to build teaching machines that construct a continuously changing probabilistic model of how a particular operator performs a skill. Furthermore, the machines do not force an operator to perform in a particular way; operators are ''minimally constrained by corrective information'' in order to provide the ''growth maximising conditions which allow the human operator as much freedom to adopt his own preferred conceptual structure'' (p. 33). By adapting the task on the basis of a dynamic, probabilistic model of the operator, SAKI teaches in a way that responds to students' (non-stationary) individual characteristics and holds their interest.

How Does SAKI Work? The operator sits in front of a display unit (see figure 8.4) that presents the exercise material (four lines of twenty-four alphanumeric characters to be punched) and cueing lights, arranged in the same spatial layout as the keyboard, that indicate which key, or key sequence, to press on the key punch. Initially the operator works through all four exercise lines. Starting with the first line, items are randomly presented at a slow, uniform rate and the cueing lights are bright and stay on for a relatively long period of time. The operator's response time for each item is stored in the "computing unit." This consists of a series of capacitors that are charged from the moment an operator makes a correct response until the next item is presented: the faster a correct response is, the higher the charge stored. When all four exercise lines have been completed correctly, SAKI has a preliminary analogue "model" of the operator's key-punch skills for every item in the four exercise lines, stored as charges on the series of capacitors.

The exercise line for which the operator has the slowest average response time is then repeated. The capacitors drive valves, which determine how the individual items in this exercise are presented to the operator—specifically, the available response time and the clarity of the cueing lights (their brightness and duration). In a prototype design, Pask uniformly varied the difficulty of the items according to average performance on an exercise line. However, it was found that uniformly increasing the difficulty of *all* the items in the exercise results in oscillations in an operator's performance—the task alternating between being too difficult and being too easy (Pask, McKinnon-Wood, and Pask 1961). The computing unit therefore *individually* varies the difficulty of each item in an exercise line so as to better match the performance of the operator. For example, it increases the difficulty of items where the operator has performed relatively successfully by reducing the cue information as well as the available response time. The reduction in available response time also reduces the maximum charge that can be stored on the associated capacitor. As the operator's skill on an item increases, the cueing information reduces, until finally there is only an indication of the alphanumeric character that has to be punched. This reduction in cueing information initially increases the likelihood that the operator will make a mistake. SAKI responds by reintroducing the visual cues and extending the available response time. Operators using SAKI show plateaus in their learning curves, but can ultimately reach a final stable state where there is no visual cueing information and an equal distribution of available response times for all items in an exercise line (Pask 1961). That is, they punch each key with equal proficiency. To maintain this level, the operator has to consistently perform a sequence of key punches at or below predetermined error and response rates.

Beer (1959) describes his experience of using a version of SAKI in *Cybernetics and Management* (pp. 124–25):

> You are confronted with a punch: it has blank keys, for this is a "touch typing" skill. Before you, connected to the punch, is Pask's machine. Visible on it is a little window, and an array of red lights arranged like the punch's keyboard. The figure "7" appears in the window. This is an instruction to you to press the "7" key. But you do not know which it is. Look at the array of lights. One is shining brightly: it gives you the position of the "7" key, which you now find and press. Another number appears in the window, another red light shines and so on. Gradually you become aware of the position of the figures on the keyboard, and therefore you become faster in your reactions. Meanwhile, the machine is measuring your responses, and building its own probabilistic model of your learning process. That "7," for instance, you now go to straight away. But the "3," for some obscure reason, always seems to elude you. The machine has detected this, and has built the facts into its model. And now, the outcome is being fed back to you. Numbers with which you have difficulty come up with increasing frequency in the otherwise random presentation of digits. They come up more slowly, too, as if to say: "Now take your time." The numbers you find easy, on the contrary, come up much faster: the speed with which each number is thrown at you is a function of the state of your learning. So also is the red-light system. For as you learn where the "7" is, so does the red-light clue gradually fade. The teacher gives you less and less prompting. Before long, if you continue to improve on "7," the clue light for "7" will not come on at all. It was getting fainter on "5," for you were getting to know that position. But now you have had a relapse: "5" is eluding you altogether. Your teacher notes your fresh mistakes. "5" is put before you with renewed deliberation, slowly; and the red light comes back again, brightly. . . . So the teaching continues. You pay little intellectual attention: you relax. The information circuit of this system of you-plus-machine flows through the diodes and condensers of the machine, through the punch, through your sensory nerves and back through your motor nerves, the punch, the machine. Feedback is constantly adjusting all the variables to reach a desired goal. In short, you are being conditioned. Soon the machine will abandon single digits as the target, and substitute short runs of digits, then longer runs. You know where all the keys are now; what you have to learn next are the patterns of successive keys, the rhythms of your own fingers.

Was It a Success? Beer began as a complete novice and within forty-five minutes he was punching at the rate of eight keys per second.[7] It seems likely that he was just doing single-key exercises, rather than key combinations. Generally, SAKI could train a novice key-punch operator to expert level (between seven thousand and ten thousand key depressions per hour) in four to six weeks if they completed two thirty-five-minute training sessions every working day. A conservative estimate of the reduction in training time, compared to other methods, was between 30 and 50 percent (Pask 1982b).

SAKI deals with incomplete knowledge about the characteristics of individual operators and how they learn by taking the cybernetic approach of treating them as a "black box"—a nonstationary system about which we have limited knowledge. In order to match the characteristics of the operator, the computing unit is also treated as a black box that builds a probabilistic, nonstationary analogue of the relation between itself and the operator through a process of interaction. The overall goal is to find a stable relation between the user and SAKI, with the additional constraint that the operator meets a prespecified performance level defined in terms of speed and accuracy of key punching. Pask summarizes this design methodology: "a pair of inherently unmeasurable, non-stationary systems, are coupled to produce an inherently measurable stationary system" (Pask 1961, p. 98). SAKI found the appropriate balance between challenging exercises and boredom: "Interest is maintained, and an almost hypnotic relationship has been observed, even with quite simple jobs" (Pask, McKinnon-Wood, and Pask 1961, p. 36). In 1961 the rights to sell SAKI were bought by Cybernetic Developments and fifty machines were leased or sold, although one unforeseen difficulty was getting purchasers to use SAKI as a training machine, rather than as a status symbol (Pask 1982b). SAKI was a very effective keypunch trainer but a limited financial success.

Summary of Musicolour and SAKI

Pask described Musicolour as "the first coherence-based hybrid control computer" where a nonstationary environment was tightly coupled with a nonstationary controller and the goal was to reach stability, or coherence, through reciprocal feedback (Pask 1982a, p. 144). He describes it as "hybrid" because rather than executing a program, it adapted on a trial-and-error basis. SAKI differs from Musicolour in that for commercial reasons there was also a performance constraint driving the activity. There were no such constraints on how Musicolour and musicians reached stable cycles of activity, the search for stability being an end in itself. Interestingly, having observed people interacting with both systems, Pask concluded (1961) that people are motivated by the desire to reach a stable interaction with the machines, rather than to reach any particular performance goal: "After looking at the way people behave, I believe they aim for the non-numerical payoff of achieving some desired stable relationship with the machine'' (p. 94).

Both Musicolour and SAKI are constructed from conventional hardware components (capacitors, valves, and so forth), but it is difficult to *functionally* separate the machines from their environments, as they are so tightly coupled. However, Pask wanted to develop *organic* machines that were built from materials that *develop* their functions over time, rather than being specified by a design. An organic controller differs from Musicolour and SAKI by not being limited to interacting with the environment through designer-specified channels (such as keyboards and microphones): it ''determines its relation to the surroundings. It determines an appropriate mode of interaction, for example, it learns the best and not necessarily invariant sensory inputs to accept as being events'' (Pask 1959, p. 162). The next sections describe the collaboration between Pask and Stafford Beer as they explored how to build such radically unconventional machines.

Pask as an Independent Cybernetic Researcher

Stafford Beer (1926–2002) and Pask met in the early 1950s and they collaborated for the rest of the decade. They were ''both extremely conscious of the pioneering work being done in the USA in the emerging topic that Norbert Wiener had named cybernetics, and knew of everyone in the UK who was interested as well'' (Beer 2001, p. 551). Both men were ambitious and wanted to make an impact in the field of cybernetics. They were particularly interested in W. Ross Ashby's work on ultrastability (Ashby 1952) and the question of how machines could adapt to disturbances that had not been envisaged by their designer. Beer was working for United Steel, doing operations research, and had persuaded the company to set up a cybernetics research group in Sheffield. Pask was developing learning machines and trying to market them commercially. They grew close as they both faced similar challenges in trying to persuade the business world of the value of their cybernetic approach. They also shared a deep interest in investigating the suitability of different ''fabrics,'' or media, as substrates for building self-organizing machines:

> If systems of this kind are to be used for amplifying intelligence, or for 'breeding' other systems more highly developed than they are themselves, a fixed circuitry is a liability. Instead, we seek a fabric that is *inherently* self-organizing, on which to superimpose (as a signal on a carrier wave) the particular cybernetic functions that we seek to model. Or, to take another image, we seek to *constrain* a high-variety fabric rather than to fabricate one by blueprint (Beer 1994, p. 25).

The ''high-variety'' criterion came from Ashby's argument that a controller can only control an environment if it has variety in its states greater or equal to the variety in the disturbances on its inputs.[8] Another requirement for a suitable fabric was that its behavior could be effectively coupled to another system.

The Search for a Fabric

Both Beer and Pask investigated a wide range of media for their suitability as high-variety fabrics. From the outset, Beer rejected electrical and electronic systems as they had to be designed in detail and their functions, well specified, and this inevitably constrained their variety. Instead, he turned to animals.

In 1956 Beer had set up games that enabled children to solve simultaneous equations, even though they were not aware they were doing so. Their moves in the game generated feedback in the form of colored lights that guided their future moves. He then tried using groups of mice, with cheese as the reward, and even tried to develop a simple mouse language. Beer considered the theoretical potential of other vertebrates (rats and pigeons) and, with Pask, social insects, but no experiments were carried out using these animals.

Beer then investigated groups of *Daphnia*, a freshwater crustacean. He added iron filings to the tank, which were eaten by the animals. Electromagnets were used to couple the tank with the environment (the experimenter). Beer could change the properties of magnetic fields, which in turn effected changes in the electrical characteristics of the colony. Initially this approach seemed to have potential, as the colony ''retains stochastic freedom within the pattern generally imposed—a necessary condition in this kind of evolving machine; it is also self-perpetuating, and self-repairing, as a good fabric should be'' (Beer 1994, p. 29). However, not all of the iron filings were ingested by the crustaceans and eventually the behavior of the colony was disrupted by an excess of magnets in the water.

Beer then tried using a protozoan, *Euglena*, keeping millions of them in a tank of water, which he likened to a ''biological gas'' (Beer 1994, p. 30). These amoebae photosynthesize in water and are sensitive to light, their phototropism reversing when light levels reach a critical value. If there is sufficient light they reproduce by binary fission; if there is a prolonged absence of light they lose chlorophyll and live off organic matter. The amoebae interact with each other by competing for nutrients, blocking light and generating waste products. Although the green water was a ''staggering source of high variety'' and it was possible to couple to the system (using a point source of light as an input and a photoreceptor to measure the behavioral output), the amoebae had "a distressing tendency to lie doggo, and attempts to isolate a more motile strain failed" (Beer 1994, p. 31). Beer started to experiment with pond ecosystems kept in large tanks, thinking that his single-species experiments were not ecologically stable. He coupled the tank and the wider world in the same way as he had done in the *Euglena* experiments, using a light and photoreceptors. However, it proved difficult to get this system to work as a control system—the feedback to the environment was too ambiguous. "The state of the research at the moment is that I tinker with this tank from time to time in the middle of the night. My main obsession at the moment is at the level of the philosophy of science. All this thinking is, perhaps, some kind of breakthrough; but what about an equivalent breakthrough in experimental method? Do we really know how to experiment with black boxes of abnormally high varieties?" (Beer 1994, p. 31). The first experimental breakthrough came during one of his visits to Pask.

Growing an Ear

Although based in Sheffield, Beer would regularly go down to London and work most of the night with Pask.[9] In 1956 or '57, he had "the most important and indeed exciting of my personal recollections of working with Gordon" (Beer 2001, p. 553): the night they grew an electrochemical ear. Pask had been experimenting with electrochemical systems consisting of a number of small platinum electrodes inserted in a dish of ferrous sulphate solution and connected to a current-limited electrical source. Metallic iron threads tend to form between electrodes where maximum lines of current are flowing. These metallic threads have a low resistance relative to the solution and so current will tend to flow down them if the electrical activation is repeated. Consequently, the potentials at the electrodes are modified by the formation of threads. If no current passes through a thread, then it tends to dissolve back into the acidic solution. Metallic threads develop as the result of two opposing processes: one that builds threads out of ions on relatively negative electrodes; and one that dissolves threads back into ions. The trial-and-error process of thread development is also constrained by the concurrent development of neighboring threads and also by previously developed structures. Slender branches extend from a thread in many directions and most of these dissolve, except for the one following the path of maximum current. If there is an ambiguous path then a thread can bifurcate. As the total current entering the system is restricted, threads compete for resources. However, when there are a number of neighboring unstable structures, the threads can amalgamate and form one cooperative structure. Over time a network of threads literally grows dynamically stable structures.

These electrochemical systems display an elementary form of learning. If a stable network of threads is grown and then the current to the electrodes is redistributed, a new network will slowly start to form. If the current is then set to the original distribution, the network tends to regrow its initial structure. The longer a network has been stably growing, the slower it breaks down when the current distribution changes, and the quicker it returns to its original structure when the current distribution is reset.

Beer vividly remembers the night that he and Pask carried out the electrochemical experiments that resulted in an ear (Beer 2001, pp. 554–55). They were discussing Ashby's concept of ultrastability and the ability of machines to adapt to unexpected changes—changes that had not been specified by their designer. Pask had recently been placing barriers in the electrochemical dishes and the threads had grown over them—they had adapted to unexpected changes in their environment. That night they did some experiments to see how the threads would respond to damage by chopping out sections of some of the threads. When current was applied to the system the threads regrew, the gap moving from the anode to the cathode until it was gone.

Although excited by this result, they thought that these were relatively trivial disturbances. They wanted to perform an experiment to investigate whether a thread network could adapt to more radical, unexpected disruption. "We fell to discussing the limiting framework of ultrastability. Suddenly Gordon said something like, 'Suppose that it were a survival requirement that this thing should learn to respond to sound? If there were no way in which this 'meant' anything, it would be equivalent to your being shot. But this cell is liquid, and in principle sound waves could affect it. It's like your being able to accommodate to a slap, rather than a bullet. We need to see whether the cell can learn to reinforce successful behaviour by responding to the volume of sound.' . . . It sounded like an ideal critical experiment" (Beer 2001, p. 555).

Beer cannot remember the exact details of how they rewarded the system.[10] However, it did not require any major changes to the experimental setup. They basically connected one, or more, of the electrodes with output devices that enabled them to measure the electrical response of the electrochemical system to sound. The reward consisted of an increase in the current supply, a form of positive reinforcement. Regardless of how the electrodes are configured, the electrochemical system will tend to develop a thread structure that leads to current flowing in such a way that it is rewarded further. Importantly, the reward is simply an increased capacity for growth—there is no specification of what form the growth should take.

The electrochemical system is not just electrically connected to the external world: threads are also sensitive to environmental perturbations such as vibrations, temperature, chemical environment, and magnetic fields. Any of these arbitrary disturbances can be characterized as a stimulus for the system, especially if they cause a change in current supply. "And so it was that two very tired young men trailed a microphone down into Baker Street from the upstairs window, and picked up the random noise of dawn traffic in the street. I was leaning out of the window, while Gordon studied the cell. 'It's growing an ear,' he said solemnly (*ipsissima verba* [the very words])" (Beer 2001, p. 555).

Pask (1959) describes further experiments that were carried out where a thread network was grown that initially responded to 50 Hz and then, with further training, could discriminate between this tone and 100 Hz. He was also able to grow a system that could detect magnetism and one that was sensitive to pH differences. In each case the electrochemical system responded to positive reinforcement by growing a sensor that he had not specified in advance. Beer is clear why he and Pask thought this experiment was significant: "This was the first demonstration either of us had seen of an artificial system's potential to recognize a filter which would be conducive to its own survival and to incorporate that filter into its own organization. It could well have been the first device ever to do this, and no-one has ever mentioned another in my hearing" (Beer 2001, p. 555).

Pask (1959, p. 262) argues that the electrochemical ear is a maverick device, as it shows the distinction between

> the sort of machine that is made out of known bits and pieces, such as a computer . . . and a machine which consists of a possibly unlimited number of components such that the function of these components is not defined beforehand. In other words, these 'components' are simply 'building material' which can be assembled in a variety of ways to make different entities. In particular the designer need not specify the set of possible entities.

Importantly, electrochemical systems, although finite, "are rendered nonbounded by the interesting condition that they can alter their own relevance criteria, and in particular, by the expedient of building sense organs, can alter their relationship to the environment according to whether or not a trial relationship is rewarded" (p. 262).

The Value of Gordon Pask

Ideas that were dear to Gordon all that time ago, on interactive circuits with dynamic growth, are coming back in the form of neural nets, with parallel processing in digital computers and also analogue systems. My bet is that analogue self-adapting nets will take over as models of brain function—because this is very likely how the brain works—though AI may continue on its course of number crunching and digital computing. Surely this is alien to the brain. So we would fail the Turing Test, being too good at pattern recognition, and much too poor at arithmetic compared with digital computers. In short, the kind of philosophy that Gordon nurtured does seem to be returning. Perhaps his learning machines have lessons for us now.
—Richard Gregory (2001, pp. 686–87)

The naive picture of scientific knowledge acquisition is one of posing increasingly sophisticated questions to nature. But, of course, such questions, and therefore the knowledge obtained from them, are never pure, unaffected by the questioners' ulterior motives, or unconstrained by technological and conceptual barriers. Science manifests itself as a social and cultural activity through subtle factors such as concept management, theory creation, and choice of what problems to focus on. It is far from being passive observation followed by rational reflection. It is *active*. But even in this picture, experimental data, the source of scientific information to a community of researchers, is still seen as the detached, passive observation of nature at work. Observer intervention (today most apparent in quantum measurement or the behavioral and cognitive sciences) is often treated as a *problem* we would wish to minimize if we cannot eliminate.

Pask's approach goes against this view. For him, not only can we gain new understanding by actively constructing artefacts instead of just observing nature, we can also increase our knowledge by engaging in an interaction with them. Pask's design methodology can be characterized as "meeting nature half way": accepting that we have limited, incomplete knowledge about many systems we want to understand and treating them as black boxes. By interacting with these systems we can constrain them, and ourselves, and develop a stable interaction that *is* amenable to analysis. For him, both the construction and the interaction become a necessity if we wish to understand complex phenomena such as life, autonomy, and intelligence.

Let us consider construction. The first thing that must be clarified is that Pask, and nowadays some of current research in AI and robotics, is not simply proposing that technology and science interact, often in a positive, mutually enhancing manner. The construction he refers to is not that of more sophisticated artefacts for measuring natural phenomena or the construction of a device that models natural phenomena by proxy, but the construction of a proper object of study, in other words, the synthesis of *a scientific problem in itself*. This idea is radical—fraught with pitfalls and subject to immediate objections. Why create problems deliberately? Are we not just using our existing knowledge to guide the creation of an artefact? Then, how do we expect to gain any new knowledge out of it?

Indeed, the idea seems not just a minefield of methodological issues; it seems absurd and a nonstarter, at most a recipe for useful pedagogical devices, toy problems for scientific training, but not the stuff of proper science. To answer these criticisms it is necessary to demonstrate not only that interesting artefacts can be constructed that will grasp the attention of scientists but also that we can do science with them, that they can advance our understanding of a problem.

It is clear, in relation to one of the objections above, that if by construction we mean the full specification of every aspect of our artefact and every aspect of its relation to its environment, then little new knowledge can be expected from it, except perhaps the knowledge that confirms that our ideas about how to build such an artefact were or weren't correct. This is traditional engineering, which of course is a source for that kind of knowledge. But what if the construction proceeds not by a full specification of the artefact but by the design of some broad constraints on processes that lead to increased organization, the result of which—with some good probability—is the artefact we are after? Now, if we succeed in this task, the workings of such a system are not fully known to us. It may surprise us. It may challenge our preconceptions by instantiating a new way of solving a problem. Or, more subtly, it may make us revise the meaning of our scientific terms and the coherence of our theories.

Is such an underspecified synthesis possible? Yes, it is. It was for Pask, as he demonstrated with his maverick machines (most dramatically with the electrochemical "ear") and it is common currency in biologically inspired AI (self-organizing optimization algorithms, evolutionary robotics, stochastic search, and so forth). Hardware evolution, which uses genetic algorithms to constrain reconfigurable devices such as field-programmable gate arrays (FPGAs), also provides striking examples of how relaxing conventional engineering constraints (such as a central clock) can lead to the invention of novel circuits—or should that be "discovered"?[11] Pask's research also provides a valuable reminder of the constraints that conventional computer architectures impose on machines. Although digital computers have an invaluable "number-crunching" role, they are not necessarily the best medium for building controllers that have to interact with dynamic environments. Similarly, conventional computer architectures might not be the best models of adaptive systems.

Pask provides a methodology for developing controllers that can deal with nonstationary environments about which we have limited knowledge. His cybernetic approach of coupling two nonstationary, unmeasurable systems in order to generate a stable, measurable relation will probably not appeal to conventional engineers. For example, Beer (2001, p. 552), discussing SAKI, lamented, "The engineers somehow took the cybernetic invention away. I suspect that they saw themselves as designing a machine to achieve the content-objective (learn to type), instead of building a Paskian machine to achieve the cybernetic objective itself—to integrate the observer and the machine into a homeostatic whole. Machines such as these are not available to this day, because they are contra-paradigmatic to engineers and psychologists alike."

Even if we can successfully synthesize an artefact that would indeed be a source for furthering our understanding about a given problem, what makes us think that such a device would be easier to understand than nature? Valentino Braitenberg (1984), a proponent of a related synthetic approach, convincingly pointed out a curious fact that he dubbed the "law of downhill synthesis and uphill analysis": it is rather easy to build things that look very complex and are hard to understand. This is true particularly if we specify lower-level mechanistic building blocks and leave as unspecified higher-level and interactive aspects of the system. If we now present the latter as the easily observable variables, the system can seem devilishly complex. W. Grey Walter had already demonstrated this with his robotic tortoises (1950, 1951, 1953). Simple combinations of a very few basic mechanisms could interact in sophisticated and surprising manners with a complex environment, giving the illusion of sophisticated cognitive performances (such as decision making, adaptive goal constancy, self-sustenance, and others).

This "law" of downhill synthesis, uphill analysis, is on the one hand quite interesting in itself, and often the source of entertaining explorations. It is also a stark reminder that we need not theorize complex mechanisms when we are faced with complex systemic behavior—a much unheeded warning. It is, in this sense, a powerful positive idea. On the other hand, though, it points to a major problem with the proposal of furthering scientific understanding by construction. Yes, we may be successful in constructing our artefact, but how shall we understand it? We seem to be at an advantage over understanding similarly complex phenomena in nature. We may have more access to data, we know many things, if not everything, about how the system is built, we can restart it and do experiments that would be impossible in nature. But will these advantages always suffice? Have we given our process of synthesis too much freedom and the result is now an intractably complex system?

Two answers can be given to this problem. One suggests that our greater gain is by proceeding in a more or less controlled manner in exploring increasingly complex systems. This answer advocates minimalism as a methodological heuristic (Harvey et al. 2005). By building systems that are underdetermined but in a controlled fashion (which sounds like a paradox, but simply means that we should carefully control the constraints to the process of automatic synthesis), we stand our highest chance of creating new knowledge because we advance minimally over our previous understanding, which can largely be deployed on the analysis of our current system. There is a sense in which such a minimalism will provide us with the simplest cases that instantiate a phenomenon of interest, for instance learning or decision making, and allow us to develop the right kind of "mental gymnastics" to deal with more complex cases (Beer 2003).

But Pask proposes a different, more radical solution that has, paradoxically, been in use in dealing successfully with nature since the advent of culture, much before anything like science ever existed. Pask proposes that we should base our understanding of a complex system on our interactions with it and the regularities that emerge from such interaction. We should approach complex systems, even those we synthesize ourselves, as a natural historian would (perhaps even as an animal trainer, a psychotherapist, or an artist would). This interactive method for understanding complex systems is still a hard pill to swallow in many areas of science.

Pask's machines and philosophy often seem so maverick that they are hard to evaluate. Interacting with his work, one can struggle to achieve a stable understanding because of the demands he places on the reader. However, we have found it a worthwhile struggle and we hope that others will be encouraged to interact with his ideas: although fifty years old, they are highly relevant for every discipline that is attempting to understand adaptive (nonstationary) behavior.

Acknowledgments

We would like to thank Michael Renshall, CBE, and Paul Pangaro for agreeing to be interviewed about their memories of Gordon Pask and for giving valuable feedback on an earlier draft. It was through the research of Peter Cariani that Jon Bird first became aware of Gordon Pask and he has enjoyed and benefited from conversations with Cariani over the last few years. Many thanks to Philip Husbands for his support for Paskian research and for first introducing Jon Bird to the cybernetic literature.

Notes

1. Michael Renshall, CBE, who was at Rydal from 1941 to 1948 and also was a contemporary of Pask's at Cambridge, provided all of the information about Gordon's school days.

2. The radical theater director Joan Littlewood was certainly aware of Pask by 1946 (Littlewood 2001). Pask wrote shows for his Musicolour system in the early 1950s.

3. A clear description of the strategy for selecting filter parameter values is given in Pask (1971, p. 80).

4. Ethologists coined the term "releaser stimulus" to refer to a simple perceptual feature of a complex stimulus that elicits a pattern of behavior. Niko Tinbergen (1951) had shown that crude models of a stickleback could elicit behavior patterns in the real fish—they attack red-bellied models and court swollen-bellied models.

5. The complete patent specification was published in 1961 (GB866279).

6. A later version of SAKI was developed to train operators in the use of key punches with larger numbers of keys (see Pask 1982, p. 71, figure 2). Herman Hollerith developed the first automatic data-processing system to count the 1890 U.S. census. A key punch was used to record the data by making holes in dollar-bill-sized cards. A tabulating machine contained a pin for each potential hole in a card. A card was passed into the reader and if a pin passed through a hole a current was passed, incrementing a counter. On the basis of these counters the card was automatically dropped into the appropriate section of a sorting box. It took just three years to tabulate the 62 million citizens the census counted. Building on this success Hollerith set up the Tabulating Machine Company, which eventually, after a series of mergers, became IBM in 1924. Some key-punch devices continued to be marketed as Hollerith machines, for example, the IBM 032 Printing punch produced in 1933 and the keyboard used in the first versions of SAKI.

7. This does seem a remarkably fast rate—the average response time for pressing a key after training on SAKI was about 0.2 seconds (Pask 1961a, p. 96).

8. For details of Ashby's "Law of Requisite Variety," see Ashby (1956, pp. 202–18).

9. Both Pask and Beer worked eccentric hours. Pask would regularly stay awake for thirty-six hours and then sleep for twelve hours, regulating the cycle with pills

(Elstob 2001). His wife thought that he was often at his best at the end of these marathon work sessions (Paul Pangaro, personal communication; Pangaro, who earned his doctorate with Pask and maintains an on-line archive of Pask's work at http://www.pangaro.com/Pask-Archive/Pask-Archive.html).

10. We lack clear information about the experimental details, even though Pask continued in-depth investigations into electrochemical systems at the University of Illinois under Heinz von Foerster. There has not, to our knowledge, ever been an independent replication of these experiments.

11. Adrian Thompson (1997) evolved a circuit on a small corner of a Xilinx XC6216 field-programmable gate array (FPGA) that was able to discriminate between two square wave inputs of 1 kHz and 10 kHz without using any of the counters–timers or RC networks that conventional design would require for this task. Layzell (2001) developed his own reconfigurable device, the Evolvable Motherboard, for carrying out hardware evolution experiments. One experiment resulted in the "evolved radio," probably the first device since Pask's electrochemical "ear" that configured a novel sensor (Bird and Layzell 2001).

References

Ashby, W. Ross. 1952. *Design for a Brain: The Origin of Adaptive Behaviour*. London: Chapman & Hall.

———. 1956. Introduction to Cybernetics. London: Methuen.

Beer, R. D. 2003. "The Dynamics of Active Categorical Perception in an Evolved Model Agent." *Adaptive Behavior* 11, no. 4: 209–43.

Beer, Stafford. 1959. *Cybernetics and Management*. London: English Universities Press.

———. 1994. "A Progress Note on Research into a Cybernetic Analogue of Fabric." In *How Many Grapes Went into the Wine: Stafford Beer on the Art and Science of Holistic Management*, edited by R. Harnden and A. Leonard. New York: Wiley.

———. 2001. "A Filigree Friendship." *Kybernetes* 30, no. 5–6: 551–59.

Bird, Jon, and P. Layzell. 2001. "The Evolved Radio and Its Implications for Modelling the Evolution of Novel Sensors." *Proceedings of the 2002 Congress on Evolutionary Computation*. Washington, D.C.: IEEE Computer Society.

Braitenberg, Valentino. 1984. *Vehicles: Experiments in Synthetic Psychology*. Cambridge, Mass.: MIT Press.

Elstob, M. 2001. "Working with Gordon Pask: Some Personal Impressions." *Kybernetes* 30, no. 5–6: 588–92.

Foerster, Heinz von. 2001. "On Gordon Pask." *Kybernetes* 30, no. 5–6: 630–35.

Glanville, R. 1996. ''Robin McKinnon-Wood and Gordon Pask: A Lifelong Conversation.'' *Cybernetics and Human Knowing* 3, no. 4: n.p.

Gregory, Richard. 2001. ''Memories of Gordon.'' *Kybernetes* 30, no. 5–6: 685–87.

Grey Walter, W. 1953. *The Living Brain*. London: Duckworth.

———. 1950. ''An Imitation of Life.'' *Scientific American* 182, no. 5: 42–45.

———. 1951. ''A Machine That Learns.'' *Scientific American* 185, no. 2: 60–63.

Harvey, I., E. A. Di Paolo, E. Tuci, R. Wood, and M. Quinn. 2005. ''Evolutionary Robotics: A New Scientific Tool for Studying Cognition.'' *Artificial Life* 11: 79–98.

Layzell, P. 2001. ''Hardware Evolution: On the Nature of Artificially Evolved Electronic Circuits.'' Ph.D. diss., University of Sussex.

Littlewood, Joan. 2001. ''Gordon Pask.'' *Kybernetes* 30, no. 5–6: 760–61.

Moore, H. 2001. ''Some Memories of Gordon.'' *Kybernetes* 30, no. 5–6: 768–70.

Pask, Gordon. 1959. ''Physical Analogues to the Growth of a Concept.'' *Proceedings of Symposium No. 10 of the National Physical Laboratory: Mechanisation of Thought Processes, National Physical Laboratory, 24–27 November 1958*. Volume 2. London: Her Majesty's Stationery Office.

———. 1961. *An Approach to Cybernetics*. New York: Harper & Brothers.

———. 1962. ''Musicolour.'' In *A Scientist Speculates*, edited by I. M. Good. London: Heinemann.

———. 1966. ''Comments on the Cybernetics of Ethical Psychological and Sociological Systems.'' In *Progress in Biocybernetics*, edited by J. P. Schade. Volume 3. Amsterdam: Elsevier.

———. 1971. ''A Comment, a Case History and a Plan.'' In *Cybernetics, Art and Ideas*, edited by J. Reichardt. London: Studio Vista: London.

———. 1975. *Conversation, Cognition and Learning: A Cybernetic Theory and Methodology*. Amsterdam: Elsevier.

———. 1982b. ''SAKI: Twenty-five Years of Adaptive Training into the Microprocessor Era.'' *International Journal of Man-Machine Studies* 17: 69–74.

Pask, Gordon, and S. Curran. 1982a. *Microman: Computers and the Evolution of Consciousness*. New York: Macmillan.

Pask, Gordon, Robin McKinnon-Wood, and E. Pask. 1961. ''Patent Specification (866279) for Apparatus for Assisting an Operator in a Skill.'' Available at http://v3.espacenet.com/textdoc?DB=EPODOC&IDX=CA624585&F=0&QPN=CA624585.

Price, C. 2001. Gordon Pask, *Kybernetes* 30, no. 5–6: 819–20.

Scott, B. 1980. ''The Cybernetics of Gordon Pask, Part 1.'' *International Cybernetics Newsletter* 17: 327–36.

Thompson, Adrian. 1997. ''An Evolved Circuit, Intrinsic in Silicon, Entwined with Physics.'' *International Conference on Evolvable Systems (ICES96)*, LNCS 1259, Berlin: Springer.

Tinbergen, Nikolaas. 1951. *The Study of Instinct*. Oxford: Oxford University Press.

Zeeuw, G. de. 2001. ''Interaction of Actors Theory.'' *Kybernetes* 30, no. 5–6: 971–83.