Explaining Modern Management Approaches by Cybernetic Principles and Some Implications

by

Dan Trietsch
MSIS Department
University of Auckland
Private Bag 92019
Auckland, New Zealand
d.trietsch@auckland.ac.nz

Abstract

Managerial cybernetics, using Stafford Beer's cybernetic model (which dates back to the early seventies), more recently known as the viable system model (VSM), calls for designing nested hierarchical organizations where every viable unit in the hierarchy is autonomous as much as possible, but is also subject to some controls from its upper management metasystem. Said metasystem represents a more general level of the hierarchy. The functions of this control include damping of oscillations between the subunits (operational control), coordinating activities of subunits to achieve synergy for the whole (management control), strategic planning (i.e., deciding how the organization should evolve and address changes in the outside environment), and overall coordination and balancing of management control decisions with strategic ones. Beer argues that such a structure is necessary for survival. Indeed, it turns out that some modern management techniques and philosophies can be interpreted as partial applications of the VSM. For example, small *JIT* (i.e., *kanban*), *CONWIP* and *Supply Chain Management* can be viewed as damping oscillations and providing some management control functions. *Management by Constraints* (MBC) and *PERT/CPM* -- MBC's older isomorph -- can be viewed as variety attenuators and amplifiers -- concepts that are key to the VSM -- and their function, again, is to support operational decisions. The function of strategic planning is closely associated with the modern concept of *learning organizations.* Similarly, *participative management* supports both functions and provides proper amplification of upper hierarchy decisions. Since Beer's assertion that these functions are vital is supported by modern practice, it may behoove us to study in some detail parts of his model which have not yet been implemented satisfactorily: they may provide significant opportunities. Emphatically, this includes the correct role of information systems.

Introduction

Cybernetics, Greek for "steersmanship," was defined by Norbert Wiener as "the science of control and communications in the animal and the machine." On the one hand, as Wiener argued, for our purpose we can think about animals (including humans) as machines (created by god or by nature), albeit more advanced than the artificial machines we fabricated to date (Wiener, 1954). This implies that it's conceivable for (some) artificial machines to behave in what appears to an observer to be a purposeful manner. For example, Ashby's Homeostat may be said to do so (Ashby, 1960). It also implies that (some) artificial machines can be designed to build other artificial machines, including models more advanced than themselves; this claim, Wiener reports, was proved mathematically by Von Neumann. On the other hand, *systems* of machines, e.g., populations and organizations, are also [compound] machines, and it was with this insight that Stafford Beer pioneered the application of cybernetics -- which already had wide application in engineering -- to management. Hence, *managerial cybernetics* is the science of control and communications in human organizations (Beer simply called it the *science of organization*). "Control" does not imply coercion: rather it includes any method by which the system maintains its viability. "Communications" includes any transfer of data, information, directives, etc. Thus, managerial cybernetics spans management (control), communications, and information systems. Particularly, it provides important connections between operations management and information systems (IS). It also has major implications for accounting and for human behavior issues, but those are outside our focus.

That Wiener chose to include both control and communications in his framework was not accidental. Effective control invariably includes feedback loops whereby information about the results of former actions is fed back to the controller. For example, the difference between the desired result and the actual one may trigger a correction. Since the correction is in the opposite direction to the error, this is also known as *negative feedback* (whereas *positive feedback* involves self-reinforcing cycles, both vicious and virtuous). The first known example of such control in engineering is Watt's governor, which regulated the speed of steam engines regardless of the load to which they were subjected (and constituted a vital part of the industrial revolution). As for animals, scientists have discovered myriad examples of negative feedback mechanisms in living organisms (as well as positive feedbacks). For instance, there is a complex system of negative feedbacks that controls the blood temperature of warm-blooded animals within very tight limits, and does so successfully within a very wide range of conditions. The close relationship between control and feedback suggests that cybernetics must address both the controlling actions and the information flows concerning them. Feedback is also abundant in organizations, but note that not every communication constitutes feedback: to say, "I got some feedback about the lecture" is *not* a correct usage of the word, unless the information was used to change the lecture while it was given. Thus, comments after the fact do not qualify. In essence, feedback implies feeding the signal *back* into the process, and changing the process thereby. Arguably, information that will be used in the future can count as part of a slower feedback loop of a larger process, but merely collecting data does not guarantee that it will ever be looked at, much less used.

In the remainder of this paper we first briefly review some general cybernetics laws and implications leading to the viable system model (VSM) that Beer developed (1979; 1981; 1985). We then discuss how kanban (and similar inventory control methods), PERT/MBC, big-JIT (competing against time, including concurrent engineering), total quality (TQ), participative management, and learning organizations fit within the VSM. Finally, we address implications for IS. Connections between big JIT and the VSM have already been noted by Duimering et al. (1993), but the coverage here is much more explicit.

Requisite Variety (Ashby's Law)

In essence, the first role of cybernetics is to maintain a system in a desired stable state regardless of its inputs (i.e., its environmental condition). To do this the system must be capable of responding to the potential variety of these inputs. This potential variety is typically huge. Ross Ashby (1956) showed that if the system is required to maintain perfect stability in its output, then it must have *requisite variety* equal to that of the input. In his words, "only variety can kill variety." If the system cannot match the environmental variety fully, it may still be capable of achieving homeostasis, which can be defined as *stability within limits*, and this is practically the best we can hope for. As long as the system can survive within the limits that it can achieve, it's operationally viable. Following Ashby, we treat variety as a threat to homeostasis, i.e., it has a somewhat negative connotation, often called *complexity*. Nonetheless, we must recall that variety also provides selection, which is less negative. Balancing the selection we offer (a positive) and the complexity we get as a result (a negative) is up to management.

Autonomy (freedom)

Ashby's Law has profound technical implications for the issue of autonomy. Putting any value judgements aside, central management cannot provide requisite variety effectively for anything but the smallest organization; and even then it would probably fail as a result of environmental changes that are likely to go undetected by a single authoritarian manager. Furthermore, Ashby (1960) demonstrates that systems that are richly connected are practically not manageable: they take exponential time to achieve homeostasis, and death can easily come first. Death, of course, is a type of relative stability that should not be confused with viability of the original system. Unfortunately, modern reality is that our systems become more and more richly connected, and therefore it becomes much more difficult to manage them. For example, the environment was not a major issue as early as one century ago, but today, when poor nations destroy their rain-forests, the Japanese consume too much seafood, or the Americans burn too much oil, the whole world is concerned, thus creating interactions where once there were none. Modern transportation and communications media, e.g., the internet, also create new interactions: to wit the ongoing attempts to reach agreement on what constitutes freedom of speech and what exceeds it in cyberspace are very difficult due to the international implications. All this simply adds variety, so it falls under Ashby's law. But it does show that we must seek more effective ways to manage. Furthermore, it implies that we should seek to design into our systems connections that are only rarely invoked, and thus avoid the pitfalls of the richly connected system, and yet they must be invoked when necessary. Thus, we must identify what's important enough to justify communication (suggesting that there is a crucial role for statistical analysis; e.g., by using control charts and reporting special causes only). To recap, adequate autonomy is vital to prevent the deadly symptoms of richly connected systems. With today's exponentially growing rate of change, and the need for more and more coordination within the global village, these arguments are much stronger than ever before.

Indeed, one can hypothesize that authoritarian management is in decline exactly because of the new "smallness" of the world and the exponential rate of change. Organizations that did not learn to avoid authoritarian management were much more likely to disappear and much less likely to serve as role models for new organizations. In conclusion, we must provide as much autonomy as possible to the lower levels of the organization, down to the line workers level. This will deploy requisite variety effectively, since many brains will be involved. But we must stop short of complete autonomy, or the organization will simply fall apart. Thus, autonomy should be limited as necessary to ensure viability of the whole and make possible synergies that individual autonomous parts may fail to realize when left to their own devices, but no more. The balance, then, is achieved at the level at which the parts are coordinated appropriately. Any decisions that do not bear on this coordination should be left to the discretion of those who have to implement them and who have the knowledge about the necessary details.

When viewed in temporal terms, high autonomy can be interpreted as *rarely* having to comply with direct orders and other interventions, and low autonomy implies having to seek permissions and to report in detail routinely. As a generic limit to autonomy, within organizations, large monetary expenditures *always* bear on the whole, so some budgetary control is inherently appropriate. A free market economy, subject to constraints relating to common goods, provides the same function between organizations. Splitting organizations to true profit centers, as advocated by Ackoff (1994), is aimed to achieve the same free market mechanism within organizations; but Ackoff explicitly allows upper management to make decrees when necessary, as long as it is willing to make up the difference, thus curbing autonomy for the good of the whole.

Beer’s Viable System Model

Although we discuss Beer’s VSM, we’ll employ terminology introduced independently by Anthony (1965). He distinguishes three levels of planning and control: *Strategic planning* "is the process of deciding on objectives of the organization, on changes in these objectives, on the resources used to attain these objectives, and on the policies that are to govern the acquisition, use, and disposition of these resources" (ibid, p. 16). *Management control* "is the process by which managers assure that resources are obtained and used effectively and efficiently in the accomplishment of the organization’s objectives" (ibid, p. 17). *Operational control* "is the process of assuring that specific tasks are carried out effectively and efficiently" (ibid, p. 18). The boundaries between these are somewhat blurred, but Anthony argues that it is impossible to distinguish them completely without sacrificing their usefulness.

Beer’s framework (see Chart 1) is a nested (i.e., recursive) model of viable parts, each enjoying maximal but limited autonomy. In this chart circles denote operations and squares denote management. The environment -- depicted by the ameboid shape on the left -- is defined as the set of entities which influence the system or are influenced by it. In reality, operations are part of the environment and management is part of operations, so the square should reside within the circle and the circle within the ameboid, but the chart separates them to clarify the relationships between them. In more detail, the big circle (dotted line) surrounds the operations of the *system-in-focus* (i.e., the system depicted by the whole figure). The operations compose our *system 1*; the big (management) square at the top of the figure is a *metasystem* that manages system 1. The metasystem comprises four components, namely *systems 2, 3, 4,* and *5* (as marked in the chart). Dropping a recursion from the system-in-focus to subsystems that it coordinates, two such subsystems are embedded inside the dotted circle, each complete with its own operations and management box. There can be more than two subsystems, but if there are more than about nine, then it may be that a level of recursion was missed (or should be added). The system-in-focus itself, along with one or more systems at the same level, resides within a yet larger circle (not depicted) of the higher

The VSM (Source: Beer, 1985, by generic permission.
Dotted circle and numerals added.)

recursion. Thus, it is itself a system 1.[1] Likewise, at the other end, one should envisage sub-subsystems inside the smallest circles in the figure, and so on. As for the environment, the higher up in the recursion we go, the environment becomes wider: the environment of each subsystem must be contained in the environment of the system-in-focus, but the latter may be larger than the union of the former.

Finally, a system may belong to more than one metasystem. For example, a business school may be considered a subsystem of a particular university as well as a subsystem of a peers group of business schools. Indeed, it receives some of its guidance from the university (both formally and informally) and some from its peers group (mostly informally, but such groups do perform formal audits for accreditation periodically). In a similar way an accountant belongs to an office and to a profession (not to mention a family, a country, etc). If we choose to depict the business school as part of the university, the other business schools are subsumed in the environment of the business school and therefore also of the university; the university, of course, is part of the environment of the system depicting all business schools. A similar observation holds for the accountant. Using a particular depiction relates to the particular interactions we wish to study: interactions within the environment, although not ignored by Beer, are outside the focus of his model.

As Beer put it, the parts that compose system 1 are those that *produce the system.* What they do *is* the system. For example, if a university is an organization that teaches and conducts research, then teaching and research produce it, and are thus its system 1. In contrast, services to the system, however vital they may be, do not produce the system and they are not part of system 1; such services are part of the metasystem that includes systems 2 through 5. System 2 is charged with one key task only: prevent excessive fluctuations as a result of lack of coordination between parts. Thus, system 2 is a coordinator of the parts, but its objective is limited to balance, rather than achieving synergies or directing the activities of parts for some higher objectives. In Anthony's terminology, system 2 is limited to operational control functions. An example of a system 2 function within a university is the teaching schedule that prevents two classes from being assigned to the same room at the same time; another example is a booking system for scarce resources such as departmental laptop computers. The coordinating functions beyond balance are the charge of higher systems. System 3, of which system 2 may be considered a subsystem, is in charge of directing the activities of the parts, and achieving synergies between them. One can say that it is charged with management control. To maintain autonomy at the system 1 level, this intervention should be sporadic, and Beer refers to it as an audit function (denoted by 3*). Nonetheless, Beer includes other sporadic activities in this function, e.g., operational research studies. System 3 (and system 3*) can be viewed as dealing with the "inside and now," and system 2, acting as a subsystem of system 3, supports it in this endeavor. This leaves out the strategic planning function, which can be said to deal with the "outside and then." This activity is performed by system 4. To this end, system 4 is served by sensory organs that systems 1 through 3 do not have. System 4, then, can observe the outside world without direct support from other systems, while system 3 is not concerned with the outside world at all: its communication and information channels are strictly within the organization. System 5 has

---

[1] Beer is a bit ambiguous about the exact definition of system 1. In some places he refers to the **set** of operations as "system 1" while in others he refers to each of them as a "system 1." This may have to do with the level of the recursion. In the chart I marked the big circle as system 1 of the system-in-focus, but the small squares (representing the subsystems) are also systems 1 (as in Beer, 1985, Figure 37, which is Chart 1 with Beer's own annotations). My interpretation is that a system 1 is both the whole and the set of the nested operations that produce the system. Thus, the metasystem resides *in* system 1 (as the management square really resides in the operations circle).

one simple function: to achieve the correct balance between system 3 and system 4, none of which alone is capable of running the whole show.

In the body -- which inspired the VSM and is certainly an example of a viable system -- we can think of systems 3, 4, and 5 as functions provided by various parts of the brain; system 2 is the autonomous nervous system; and the system 1 parts are various organs (Beer, 1981). Functionally, each subsystem does a job that is essential for the larger system of which it is a part (e.g., the legs provide locomotion), and it has its own internal management function that controls most of its activities. Coordinating these functions on an ongoing basis is a charge of the metasystem, in particular, system 3 and, with it, systems 2 and 3*. System 3 is informed through the spinal cord (the central connections in the chart), and it controls activities largely by chemical signals through the system 3* channel, while system 2 uses the nervous (electrical) system. These chemical signals (e.g., hormones) may act by changing the operating parameters of the electrical nervous system. One can view these chemical signals as whips and carrots, motivating the parts to serve the whole. System 4 is equipped with dedicated sensory organs, such as the eyes and the ears, which it needs to study the outside. Thus, eyes and ears are not system 1 parts in their own right: rather, they provide important services. The system 5 function is provided by the cerebellum. While we are considering the body as an example, note that the cerebellum is often confused with the whole system. That is, we tend to believe that we are our consciousness. Nonetheless, the probable reason why animals evolved ever larger brains is simply because brains enhance survival. Thus brains can be viewed as nothing more than survival tools that serve the body. While such a claim is likely to meet with a lot of resistance, perhaps most of us would be far more willing to agree with the analogous observation that top managers and boards of directors (organizational system fives) are merely tools to enhance the survival of their organizations.

Considering Ashby’s Law, the majority of variety should be dealt with at the sub-system level, which is why it has its own local autonomous management function. The metasystem, including systems 2 through 5, should not be inundated with level 1 information, since it cannot handle much variety. Thus, on the one hand, information should be filtered on its way up. Such filtering constitutes *variety attenuation.* System 3 should only be exposed to information about "inside and now" variety that systems 1 cannot deal with autonomously, and that system 2 cannot then resolve. To this end, information must be attenuated as it goes up the ladder, and only exceptional signals should be sent up, and only if the lower level cannot handle it on its own. For example, if we use a control chart to do the filtering, only special causes should be reported. Systems 3, 4, and 5 may ask for more information specifically (e.g., as part of an audit), but that should not happen for routine management of operations. Furthermore, to reduce the load on system 5 to manageable proportions, it should only deal with the balance of coordination that systems 3 and 4 cannot resolve. On the other hand, when the metasystem (e.g., system 3) sends instructions to system 1, they are necessarily lacking in requisite variety: in other words, they are not detailed enough to address the full variety of the situation. Therefore, they require *variety amplification* on the part of system 1. For example, when the captain of a ship asks the engineer to increase the speed, there is a myriad of adjustments that the engineer then proceeds to perform autonomously: these start with selecting and adjusting the exact throttle position, and include adjusting the coolant flow, monitoring various measurements such as oil pressure, etc. Note that in modern ships many, if not all, of these controls can be performed by artificial machines, but such machines still amplify the variety of the simple directive to increase speed.

In general, the only known successful response to Ashby’s Law is hierarchy. As in the speed control function discussed above, even if we automate much of the decision making that is done by humans today the hierarchical approach will remain intact: the automation just assumes the role of a hierarchical level. In the VSM hierarchy is depicted by the recursive nesting. This is quite different, however, from traditional organizational trees, since here we observe a richly connected system, capable of using feedback for effective steering. In this connection, each line in the chart connecting two black dots should be interpreted as *two* directed arrows that can transmit information and directives between the two parts it connects. The dots themselves denote *transducers,* whose role is to translate between the connected parts (which need not "speak the same language").

Small JIT (kanban), CONWIP, and Supply Chain Management

Kanban, its generalization CONWIP (Hopp and Spearman, 1996), and similar applications to the management of supply chains, have two things in common: (i) they are all pull control systems; (ii) as such, they limit the maximal amount of inventory in the system. Here, the end customer generates a demand that is transmitted backwards to the initial production/supply steps. In contrast, push systems don't require an actual sale to initiate production. While pull systems include simple but strict inventory control, push production typically does not monitor inventory on a real-time basis; as a result, strong inventory fluctuations are the rule under push. Viewed in this light, "pull" simply implies more immediate feedback from sales. And when talking about feedback, a major difference between the use of kanbans between all stations (as in basic small JIT applications) and authorizing production by feedback from final sales (as in CONWIP) -- is that the feedback in the latter case is more immediate (it does not have to travel back through all stations). But this is achieved at a cost: *internal* fluctuations are more likely under CONWIP. Be that as it may, all these techniques are simply system 2 applications. This is especially remarkable when one notes two key issues: (i) according to Beer (writing mostly in the seventies), system 2 was typically very informal, and thus inefficient and less effective than it should have been; (ii) these pull techniques invoke powerful variety amplification: they are simple to impose from the top, and yet they drive a myriad of decisions at the floor level. This is similar to the powerful effect of rules of the road, e.g., agreeing to always drive on the right side. Without such rules, complexity would paralyze anything but the slightest trickle of traffic. Yet, drivers must make a tremendous amount of small decisions while following the rules, so the rules help delegate power. Of course, excessive rules (bureaucracy) stifle activity: they are really attempts to make all the decisions automatically, without paying attention to the actual variety in the field. Hence, good cybernetics may involve rules, but invoking rules is not necessarily good cybernetics. One must always retain the ability to respond to actual conditions, i.e., use feedback at all levels.

The implementation of limited inventory systems such as these also involves ongoing reviews of the exact amount of inventory that is to be allowed. This amount is subject to change over time, but such change is already part of the system 3* function (akin to changing the operating parameters of the electrical nervous system chemically). Thus, these methods include both system 2 and system 3 functions. Furthermore, such changes are often a response to improvements, e.g., reducing setup times or the frequency of breakdowns allows a reduction of inventory. But improvements go beyond the "here and now" management: it is system 4 that is in charge of system improvements (since they are geared towards the future). Thus, these techniques are also associated with system 4. (We'll return to the issue of continuous adjustment of inventory levels later.)

MBC and PERT/CPM

Management by constraints is an isomorph of the PERT/CPM project management technique, but it was originally aimed at increasing throughput rather than decreasing completion time. Under PERT/CPM the first step is to identify the critical path, which dictates the minimal completion time; under MBC the first step is to identify the system's binding constraints. (When the objective is to minimize completion time, the system's binding constraints are exactly the durations of the critical path activities; and when the objective is to maximize throughput, the binding constraints are in terms of capacity.) The second step of PERT/CPM is to schedule the critical path activities as soon as possible; under MBC the second step is to decide how to exploit the system's binding constraints to maximal benefit. (In both cases this optimizes the objective function subject to the constraints as they are.) The third step of PERT/CPM is to subjugate all other activities to the needs of the critical path, i.e., they are each assigned a completion time such that they will not interfere with the critical path needs; under MBC the third step is to subjugate all resources to the needs of the constrained resources, to serve them in maximizing throughput. The fourth step of PERT/CPM is to consider reducing the durations of critical activities (crashing); the fourth step of MBC is to alleviate the constraints. The fifth step in both cases is to reevaluate which constraints are currently binding: more and more inert constraints become binding as the initial constraints are progressively relaxed. This signals the start of a new iteration.

In our context, the importance of these techniques is that they facilitate effective allocation of authority to a large number of people. They are essentially *focusing techniques* that allow managers of the system-in-focus to determine where best to spend their limited resources. Necessarily, this implies that non-binding-constraint activities are to be delegated, usually to subsystems, but at the same time it also provides objectives to these activities. In the case of PERT/CPM the objective is to safely meet the timetable dictated by the critical path; in the case of MBC the objective is to best serve the needs of the binding constraints. Thus these techniques attenuate the variety with which managers have to deal, and at the same time support the variety amplification function of those to whom authority is delegated. Meanwhile, the needs of the system are met synergistically. Therefore, one can say that institutionalizing these techniques creates an effective system 3. Perhaps even more important is that they make possible designing organizations with proper hierarchy without resorting to the (arguably grossly inadequate) traditional functional organization tree structure. Finally, these techniques have strong interaction with the concept of participative management, as discussed in the next section. Their continuous improvement aspect gives them system 4 functionality as well. Furthermore, if we organize and delegate power based on our current constraints, then as the constraints change the de-facto organization automatically changes with them! Beer refers to such capability for *self-organizing* as a highly desirable viability trait. Such changing of the organization is, again, a system 4 function.

The VSM applies equally well to small and huge organizations. Beer reports experience in applications to whole economies, e.g., in Allende's Chile. MBC can also be applied to very large systems. Trietsch (1992) reports a case where such an analysis of the US naval system led sequentially to highly leveraged improvements at a machine shop in a naval shipyard. Similarly, MBC can be applied to running whole economies and making taxation policy decisions. Specifically, a large economy, generally based on free market principles, should tax consumption of common goods. Such consumption includes the use of the infrastructure, pollution, and, arguably, the use of land and scarce natural resources (which god gave us all but current laws usually treat as private property). Once common goods are defined, we identify the subset of them that are in short supply (e.g., clean air). These are our most legitimate binding system constraints, and they should be utilized as effectively as possible. To do that, we set a price, i.e., a user tax, for the consumption of each unit of such a constraint. The price itself is subject to feedback correction in such a manner that the total societal consumption of each common good falls within the tolerance society wants.

Individuals and organizations then proceed to make their own economic decisions to maximize their own utilities, but as a group they do it in a manner that is responsible for the whole. Note that the arrangement employs effective variety engineering typical of MBC applications: the variety of enforcing this tax is much lower than the variety of using the resources, leaving the major variety amplification function to the users. Thus, autonomy is not curbed more than absolutely necessary. To further clarify this idea, the MBC approach implies that such user taxes should be the only form of taxation society imposes. Unless a taxpayer uses resources that society must budget, there is no reason to tax him. Administratively, most of this taxation will take the form of property tax and inheritance tax on private property that is really common goods (e.g., on land), sales tax on common goods purchased (e.g., a fuel tax can cover the value of oil reserves to society, and pay for air pollution and for the use of public roads -- which are part of the infrastructure that limits us). Such taxes are much easier to collect than individual income tax. Compare this simple idea to our present feeble and exorbitantly costly attempts to control the quality of common goods by regulating exactly how they should be used, e.g., by specifying which techniques a utility should use at various places in its power stations. The latter approach, as practiced by the United States environmental protection agency, exemplifies a bureaucracy that tries to control variety from the top by an authoritative approach. Howard (1994) provides myriad examples of this type, and demonstrates how much waste they cause and how little protection they achieve.

Participative Management, Big-JIT, Competing Against Time, and TQ

Communications and feedback are crucial to cybernetics. In the VSM the bulk of this communications is within and between subsystems, especially where they have a supplier-customer relationship. Noting that the model is nested hierarchically, one necessarily has to conclude that it implies adequate communications between individuals (e.g., internal customers and suppliers), since such individuals are subsystems or represent subsystems. Similarly, such communications are important between teams, departments, etc. All this, however, is identical to the TQ approach, where everybody is encouraged to identify their customers and suppliers, with a view to request and utilize feedback from the former, and provide it to the latter. The cybernetic approach differs, perhaps, only in stressing the need to limit formal communications to what is essential.

Similarly, under JIT, large work in progress (WIP) inventories that were traditionally used to separate internal suppliers are trimmed as much as possible, and direct communications are thus facilitated and encouraged. Furthermore, JIT promotes worker participation in decision making, specifically at the level appropriate to their own activities. This may require some clarification: our objective should *not* be to share power in some "just" way, but rather to delegate the power to those who know what to do and how. Thus, participative management does not necessarily imply that line workers should be involved in managerial decisions at all levels, but rather at their own level. (Ackoff's circular organization [1994], which he promotes as democratic management, involves representatives of line workers participating in the board of directors. Similarly, German boards have 50% of their seats reserved to labor, and labor has veto power over the selection of the chairperson. Laudable as these ideas may be, they do not stem directly from the VSM. Neither do they negate the VSM, since it allows the same person to wear more than one hat, as long as any conflicts between the roles are addressed explicitly and as long as the person is capable of "speaking all the languages" required. By this we mean that persons at various levels of the recursion see issues differently, and so they should. For example, a production worker is not typically too concerned with the balance sheet while a typical manager is not truly proficient in the technology her organization employs: so these two "speak different languages.")

To see even more directly how JIT relates to cybernetics, consider how it provides requisite variety. One of the key techniques of modern JIT is the use of 100% source inspections by gadgets and by workers, in their role as responsible suppliers of their internal customers. Thus, the production of defective items due to defective inputs is prevented. This includes inspecting the product and the setup of the process. It is also known as *zero quality control* (Shingo, 1986). The gadgets involved are fail-safe devices, known as poka-yoke (error-proofing), which are designed to capture defect causes before they cause damage, and do so at a very low cost. Experience shows that in the presence of myriad potential human errors (variety), quality control by sampling simply cannot provide requisite response variety to capture anything near all the defects, but if we find repetitive errors we can typically design poka-yoke against them. Poka-yoke provides the one-to-one response variety that is required to capture these errors. Note that such poka-yoke must be cheap to be cost-effective, which is why the major application is in repetitive manufacturing.

A poka-yoke may simply be an improved procedure. For example, in the assembly of electrical switches workers often forgot to insert one of two springs, and it was expensive to detect this error by 100% inspection after the fact. The springs were picked from a large bin; an improved procedure specified preparing the two springs in a small dish prior to assembly. Now, a worker who forgot a spring would find one remaining in the dish, so the defect could be corrected cheaply and immediately (at the source). The error was thus completely eliminated (ibid). Other poka-yoke devices involve design changes. For example, parts that are too symmetrical may be assembled in the wrong orientation (creating unnecessary variety), so a design change that removes symmetry (e.g., by punching a hole in one corner) can make possible the use of jigs that do not accept a part in the wrong orientation (e.g., a pin is located so that it must go through the hole). But poka-yoke often involves the use of cheap sensors, placed in such a manner that typical error causes are detected. Toyota has an average of 12 such gadgets on each machine, e.g., sensors that detect whether all nuts are present before tightening a wheel. This is an example of providing requisite variety by gadgets. It is also an example of low-level automation (called *autonomation* at Toyota) that avoids much of the complexity (variety) associated with full-fledged automation. And, again, catching errors at the source removes much variety downstream, and thus it not only matches existing variety but also prevents potential variety downstream. Finally, this arrangement treats the worker as a VSM, responsible for his own outputs, and not as a cog in a wheel.

Choosing to concentrate on repetitive manufacturing, and providing requisite variety there, is a strategy the Japanese employed successfully during their ascent to economic victory. Focused plants speak to the same general idea: reduce variety. Henry Ford I had a good reason why he refused to provide wide selection to his customers: problems proliferate with variety. Unfortunately, as Ford himself already had to learn, modern consumers limit the ability of producers to use this strategy today. This strategy, if you will, does not have requisite variety as far as the market is concerned. Therefore, the use of repetitive manufacturing is generally in decline. Furthermore, services become more and more important. The use of poka-yoke for non-repetitive production and for services is much less straightforward (since they have to address much more variety), but some examples of successful application to repetitive *elements* of these operations already exist. For example, parity checks in computers, and similar checks against errors and fraud in credit card numbers (where not every set of digits constitutes a legitimate potential number), are successful poka-yoke. Chase and Stewart (1994) provide other examples, including the design of robust service procedures.

One of the major techniques of JIT is the systematic reduction of setup times. The method is called SMED (single minute exchange of die; Shingo, 1985). This is absolutely necessary to make possible small batches, and thus achieve short lead times. SMED also eliminates many sources of variation (variety). For example, infinite adjustments are avoided in favor of "clicking into place" and using modular dimensions for setup parts. Thus the capability of most general machines to be infinitely adjusted is an unnecessary source of variety that SMED eliminates. It also involves the use of implicit check sheets that provide the requisite variety against all the possible ways of forgetting important ingredients. A fundamental way in which SMED provides requisite variety is by carefully allocating some setup activities to be done in relative tranquility while the machine is still working: this implies that workers have time to respond, and having enough time is equivalent to having more response variety (a limited capacity resource can provide requisite variety by taking enough time to do it, but it’s useless if during this time the situation changes and the response is no longer adequate). Finally, SMED utilizes quick fasteners, but beyond saving time this particular aspect does not seem to be highly relevant to the VSM. SMED is also geared towards repetitive manufacturing, but Trietsch (1992) reports a highly successful implementation of SMED in a custom job shop.

JIT is also associated with the notion of *competing against time*, which should perhaps be really called *combatting against waste of time,* whereby lead times are continuously reduced. It can be shown that this has direct benefits, by reducing many waste sources (e.g., defects also cause waste of time, so they are often reduced to combat waste of time). But it also has an indirect benefit through reducing dynamic complexities of the type identified by Forrester and exemplified by the Beer Game. These benefits can be viewed as the reduction of variety. In this connection, we mentioned that system 3 may change the inventory allowance over time. Such adjustments are certainly called for following process improvements that, say, reduce the lead time or the consumption variance. But in practice such changes are also associated with demand fluctuations, as per the EOQ formula. And such demand-related changes increase the Beer-game type fluctuations. Therefore, we should be very careful in changing the order size based on temporary changes in demand. Furthermore, when the inventory is financed by investment, and we wish to maximize our return on the total investment (including the inventory itself and other assets), then Trietsch (1995) shows that the best response is often to keep the order size constant, and adjust for demand fluctuations by the frequency of ordering. This is tantamount to using kanbans for purchasing without changing their number as a function of demand. As a side benefit, this reduces the dynamic fluctuations.

Concurrent Engineering

Concurrent engineering is a big-JIT approach to the design of new products. Trietsch (1997) suggested that as such it could also be applied to the renewed design of organizations, but this cannot be reported as a known practical application. Be that as it may, concurrent engineering involves participants from the whole system, starting with suppliers and ending with customers (and thus extending even beyond the boundaries of the organization that does the design itself).

To the extent that designing new models of existing products is concerned (e.g., the design of the Ford Taurus), concurrent engineering is a method for providing requisite variety to system 3 functions (since such design is really an ongoing management control function rather than dealing with the outside and then; the latter have more to do with totally new markets and changes in the organization). But concurrent engineering principles apply equally to designing the future, i.e., to strategy. As such it is applicable, at least in theory, to system 4 activities.

Concurrent engineering is associated with modern design principles that have a cybernetic role in and as of themselves. One of them is *modular design*, where the use of common modules is encouraged. This, obviously, reduces variety considerably. Another is the minimization of the number of parts and, in the same vein, promoting the use of off-the-shelf items whenever possible; this has the same variety reducing effect. A third principle involves limiting the introduction of new features in each design cycles to as few as possible, ideally one. When more than one new features are introduced at the same time, the number of potential problems proliferates exponentially: it is not only necessary to check the interaction of the new features with all the existing ones, and their combinations, but rather every combination of new features has to be checked with the existing ones. It is much easier to divide and conquer, i.e., to add features one by one. This suggests that cycles must be shorter, or we'll fall behind the competition. But short cycle times and the principle of ongoing design *not shackled to budget cycles* are fundamental to concurrent engineering. The result is that major breakthrough design changes are broken down to a series of much smaller ones that are much easier to deal with. Another result is that the income and competitive advantage from the improved designs start during the design of the additional features, in contrast with longer design cycles that are typically associated with trying to incorporate many changes at once.

Learning Organizations

Using Anthony's framework, we identified system 4 with strategic planning. This is concerned with changing the organization, typically in an evolutionary manner, to adapt to changing realities. Some reflection reveals that such evolution is basically identical to organizational learning (Senge, 1990). Learning organizations have been hailed as the wave of the future. The argument -- identical to Beer's claim that cybernetics is more important now than it ever was -- is that in today's rapidly changing environment it is imperative that organizations become adaptive, i.e., capable of learning. Thus, the need for organizational learning is the need for system 4. Notably Beer considered this function lacking in most organizations; Anthony -- at roughly the same period -- commented that only a small minority of firms engaged in true and ongoing strategic planning (much of what goes under the name "strategic planning" is not). But merely noting the importance of accelerating organizational learning is not enough to make it happen. The chief promoters of organizational learning did not invent it: they just gave a name to a phenomenon that viable organizations cannot fail to have to some degree at least (or they would not be viable). Nor did these promoters, yet, come up with clear and proven suggestions how to make organizations learn better than they already do (arguably, TQ principles are conducive to this purpose). Be that as it may, there seems to be consensus that organizational learning is important, and as such this supports the VSM as well. Basically, *all* viable systems must be flexible and adapting: in most animals this is chiefly relegated to adaptations across generations (evolution), but in humans and other high animals the ability to learn and adapt is already an *individual* survival trait. One can say that such learning covers all but the instinctive innate knowledge, and it is often argued that higher animals rely on such learning as their chief survival mechanism to an increasing degree.

Note now that if we view the production plant as a machine, adding poka-yoke devices should count as learning: the plant learns to detect and control more and more variety causes. Furthermore, when people learn the use of poka-yoke, SMED, etc, the organization as a whole learns: it is now capable of designing better production processes. But organizational learning via individual learning is not limited to these specific JIT applications.

Some Information Systems Implications

Typical management information systems have been called *misinformation systems* by Ackoff (or should it be *MIS*information systems?), reflecting a view, shared by Beer, that they do not perform their systemic role effectively. Beer considers the computer as the linchpin of modern cybernetic applications in large organizations, but emphatically not in the way it is done today. Looking at the VSM it becomes apparent that information flowing up should be filtered (attenuated), unless it is solicited explicitly as part of a system 3* audit function. But traditional IS do not include filtration as an important design objective. Instead, IS designers typically attempt to capture all the conceivably pertinent information about an organization and its environment, push it to the in-baskets of managers, and also archive it (forever). The sheer volume of information that is thus generated is guaranteed to prevent its effective use even by the most conscientious managers. Thus, on the one hand, it is necessary to design IS by the pull principle, where managers request information before it is generated and provided. On the other hand, managers need to know about any unforeseen developments that may have a bearing on their operations, and such information must be *pushed* at the initiative of the level at which it is identified. One of Beer's suggestions about this issue is that such information should be automatically pushed up, but only after giving the lower level the right amount of time to solve the problem on their own. Once pushed up, the issue becomes the responsibility of the upper level. This is based on the assumption that if the lower level could not solve the problem it must be because it must be done from the more general level. Thus, information may be pushed sequentially to ever-higher recursion levels. Be that as it may, to determine what is special enough to justify this treatment, filters (e.g., statistical) must be built into the system, and any signals that are strong enough to go through the filter (exceptions) should be addressed (while other data are ignored). The amount of this filtration should be determined in such a way that upper management would not become inundated with too much information, or it won't be able to act at all. For example, if we use control charts as our filter, the question whether to use 3σ limits, 2σ limits, or, perhaps, 3.75σ limits should be determined for each level such that the people receiving the resulting exception reports will be able to act on them. It may make sense to inform the level directly responsible for these results at a, say, 2σ level (i.e., many false signals should be expected), and only relay upwards signals beyond, say, 3.3σ.

Upon demand, e.g., for an audit, IS have to provide detailed information to upper levels. But there is a question whether this should imply ever-increasing data bases, and ever-increasing data collection efforts. Some (e.g., Ackoff, Anthony and Beer) contend that this is impossible, and certainly inefficient. Much better would be to collect most information when needed, especially information that can actually be either collected from outside sources or be generated only when needed (don't store a table of costs of all the combinations of machines and materials: store formulae for its calculation instead). Furthermore, IS should have the capability to "forget" old information that is not used. Models of the brain memory function point to two interesting traits that are useful for IS as well. First, there is short-term and long-term memory, and this is imitated by computers where short-term memory is distinguished from long-term (permanent) storage. Second, with regard to long-term memory, in the brain it is not permanent, but rather it is modeled as subject to exponential decay with time, if unused. But whenever information is retrieved, the relevant memory is renewed and reinforced. Thus, active storage is effective, while information that is not used is allowed to decay away. In this connection note that the use of exponential smoothing gives a decaying weight to historical serial numerical data. It also requires minute current storage. Thus, proper numerical information about ongoing processes can be reduced considerably. Nonetheless, the information managers get is often too attenuated: it may be the average of averages of averages. What it lacks too often is a clear indication of variation. Techniques for the graphical presentation of distributions have been promoted by Tukey (1977), and with the capacity of typical personal computers today to display and manipulate graphics, we should be able to provide such information routinely.

Executive information systems (EIS) constitute one of the areas in which IS today fall short of the mark all too often (Bussen and Myers, 1997). This speaks to the needs of system 4. Beer suggested that the "battle for Britain" war room was the most successful EIS prototype, and, accordingly, he proposed the use of a computerized control room that emulates that war room. Executives will be spending much of their time there, in the company of each other, exposed to appropriately filtered and graphically presented information. They can also request information, and, specifically, they can ask "what-if" questions. The answers would be provided graphically by a computer (as late as 1981, Beer actually recommended analogue computers for this purpose, which was certainly much more cost-effective than digital computers with the necessary graphic interface in the seventies; today, there is no need to consider analogue computers here). Note the similarity of the war room approach to concurrent engineering: in both cases teams of decision makers are housed together in a rather small area that includes their chief information sources, and they are expected to work out solutions together.

Such a computerized control room is especially important for system 4, but it would also be helpful in the execution of system 3 functions. Actually, since system 3 and system 4 should interact strongly, they can do so within this control room (system 5 is also welcome). Nonetheless, with today’s proliferation of computerized networks and electronic communications, an argument could be made that the information should also be available on each desk top. Perhaps we need a control room club, where executives are encouraged to spend time together as in the VSM, but also provide electronic links to desktops for those who cannot be physically present. This question relates to the issue of electronic meetings and group decision support systems in general, and, perhaps, it should not be resolved based on the VSM alone. Then again, the existence of such networks itself is a step in the direction of providing necessary informational links among systems and between players at various systems. One can say that these networks provide a skeleton of a nervous system for the organization, but only humans who decide to use them, both to send messages and to respond to them, can make them act as real nervous systems. And, in this connection, it is doubly important to consider the role of proper filters carefully: when every seminar organizer from every department may choose to send three reminders for each seminar, a typical user may decide to ignore messages altogether. A similar problem on the internet is when a user sends multiple copies of the same message to multiple boards. Moderated boards are really simply filters that avoid repetition and get rid of the frivolous. This function, of course, has always been provided by editors of various outlets, and board moderators are editors.

Acknowledgement: I am grateful to John Buzacott and to David Robb for providing useful feedback to the process of writing this paper.

References

Russell L. Ackoff (1994), *The Democratic Corporation: A Radical Prescription for Recreating Corporate America and Rediscovering Success,* Oxford University Press, New York, NY.

Robert N. Anthony (1965), *Planning and Control Systems: A Framework for Analysis,* Harvard University, Boston, MA.

W. Ross Ashby (1956), *Introduction to Cybernetics,* Wiley, New York, NY.

W. Ross Ashby (1960), *Design for a Brain: The Origin of Adaptive Behavior,* Chapman and Hall, London.

Stafford Beer (1979), *The Heart of Enterprize,* Wiley, Chichester.

Stafford Beer (1981), *Brain of the Firm: The Managerial Cybernetics of Organization,* 2nd Edition, Wiley, Chichester.

Stafford Beer (1985), *Diagnosing the System for Organizations,* Wiley, Chichester.

Wendy Bussen and Michael D. Myers (1997), Executive Information Systems Failure: A New Zealand Case Study, *Journal of Information Technology,* 12(2), June (forthcoming).

Richard B. Chase and Douglas M. Stewart (1994), Make Your Service Fail-Safe, *Sloan Management Review,* Spring, 35-44.

P. Robert Duimering, Frank Safayeni, and Lyn Purdy (1993), Integrated Manufacturing: Redesign the Organization before Implementing Flexible Technology, *Sloan Management Review,* Summer, 47-56.

Wallace J. Hopp and Mark L. Spearman (1996), *Factory Physics: Foundations of Manufacturing Management*, Irwin, Chicago, IL.

Peter M. Senge (1990), *The Fifth Discipline*, Doubleday Currency.

Shigeo Shingo (1985), *A Revolution in Manufacturing: The SMED System,* Productivity Press, Cambridge, MA.

Shigeo Shingo (1986), *Zero Quality Control: Source Inspection and the Poka-yoke System,* Productivity Press, Cambridge, MA.

Dan Trietsch (1992), Focused TQM and Synergy: A Case Study, *AS* Working Paper 92-06, September, Naval Postgraduate School.

Dan Trietsch (1995), Revisiting ROQ: EOQ for Maximizing Company-Wide Return on Investment, *Journal of Operational Research Society*, 46, 507-515.

Dan Trietsch (1997), *Total Quality: Leadership, Focusing, Analysis, and Engineering,* textbook, 2nd draft (available from the author).

John W. Tukey (1977), *Exploratory Data Analysis,* Addison-Wesley, Reading, MA.

Norbert Wiener (1954), *The Human Use of Human Beings: Cybernetics and Society,* Eyre and Spottiswoode, London.

Philip K. Howard (1994), *The Death of Common Sense,* Warner Books, New York, NY.