

Chapter 19ⁱ

Lowlands Sociotechnical Design Theory and Lean Production

Jac Christis en Erik Soepenbergh

Introduction

Lean Production (LP) can be regarded as a design approach in search of a theoretical foundation. In this paper we show that Lowlands' Sociotechnical Design Theory (STSL) could function as such a foundation. To reach this goal, we first describe STSL as a system theoretical reformulation of Original Sociotechnical Theory (OSTS). Then, we introduce the Toyota Production System as the origin of LP and the challenge it poses for the academic field of organization design. This academic field should (1) assess LP's success, (2) generalize it by embedding it in more abstract concepts and theories in order to be able to (3) re-specify it for different manufacturing and non-manufacturing contexts. Next, we give an exposition of STSL as a structural design approach based on developments in system theory. At last, we reformulate lean production in STSL terms and so show that LP is a subcase within the more general theory of STSL. We discuss the merits of both approaches and clarify some misunderstandings of lean both outside and inside the lean community. Embedding LP in the more general language of STSL should enable us to discover similarities and differences, to start a process of mutual learning, to integrate diverse design approaches in a theory of organizational design and to add content to redesign proposals of for example the health care system as proposed by Porter and Teisberg (2006) and Christensen et al. (2009). We quote extensively from the lean literature (to convince our sociotechnical friends) and embed both STSL and LP in the broader literature on organization design. We hope this adds a new perspective to the one given in the Operations Management literature on LP. Again, mutual learning is the goal.

19.1 Original and Lowlands Sociotechnical Design Theory

Original Sociotechnical Theory

STS originated in the coalmines of Durham. The story is a familiar one. In the pre-mechanized phase, cross-functional teams worked at the coal-face. The introduction of coal cutting machines and mechanical conveyers made it possible to work on a single long face instead of a series of shorter faces. In order to reap economies of scale, management organized work in three shifts. Each shift was responsible for one phase in the process of coal-getting: the cutting, the ripping or the filling shift. This way of organizing is called the long wall method. By placing interdependent activities in separate shifts, it wrought havoc on both productivity and morale. When, after an accident, it was impossible to work at a long face, the workers proposed to reintroduce the old way of working: cross-functional teams working at short faces. This is called the composite or short wall method. By placing interdependent activities back again in the same organizational units (shifts), this way of working improved both productivity and morale.

The theory: joint optimization of technical and social subsystems

Reflecting on the Durham case, OSTS proposed to define an organization as a system with two subsystems, a social and technical subsystem. Design should be aimed at the joint optimization of both subsystems. An important design tool was the variance matrix, in which the distance between disturbances and their control was measured.

The conceptualization of an organization as a sociotechnical system and the notion of joint optimization were questioned from the start. What are, for example, the elements (and structures) of both subsystems? A transformation process or primary process is a system of interrelated operations that are partly carried out with the help of artefacts (tools, machines, and so on). If you

take the artefacts from this system and put them together, you do not create a technical subsystem but an aggregate without an internal structure. It is just a whole that is composed of similar things (these things are all artefacts). In this way you lose insight in the role these artefacts play in the transformation process. The same applies to the so-called social subsystem. Both are at best nominal systems of similar things (artefacts and people) and not real systems with an internal structure (and external environment).

And in the Durham case no joint optimization took place. Instead, in the short wall method the same technology was used as in the long wall method, but now combined with a different work organization. That means that the Durham case was an example, not of joint optimization, but of organizational choice (the title of the book the Tavistock researchers subsequently wrote). With the Durham case, technological determinism was refuted and the existence of organizational choice proven (a mayor feat indeed).

The origin of STSL

STSL was developed in the Netherlands by Ulbo de Sitter (see chapter 6). In the opinion of De Sitter, OSTS was correct in its practice. Instead of adapting workers to existing, Tayloristic structures, it aimed at the transformation of that structure itself. Because the structure of the division of labour is at the same time a structure of power relations, it stroke at the heart of the organization, its power structure. However, De Sitter was dissatisfied with both the concepts and design tools of OSTS. So, he set himself the task of a system theoretical reformulation of OSTS. On the conceptual level he considered the distinction between technical and social subsystems to be a reification of what in fact are interconnected aspects of the same system (or aspect systems). An organization as a system has many aspects such as costs, quality, time, safety, health and so on. Those diverse aspects could be grouped roughly in 'technical' and 'social' aspects. However, by reifying them into separate subsystems, you tear apart what in fact is interrelated and so, make the actual relations between those aspects within systems and subsystems invisible. Those aspects are emergent system properties and the way they are interconnected is determined, as we will see, by the structure of the system. Redesign should be aimed, not at the joint optimization of so-called social and technical subsystems, but at integral design. The organizational structure should be designed in such a way that all aspects are improved simultaneously, both the quality of the organization (in terms of costs, quality and time), the quality of work (in terms of stress risks and learning opportunities) and the quality of labour relations (in terms of cooperation and shared decision-making).

Practical critique of original theory: the causes of uncontrolled variances

With respect to the variance matrix, De Sitter stresses the need for a theory that can explain the level of disturbances and the lack of local control. Without such a theory, organization redesign lacks a point of direction. The basic idea is this. Because all organizations are characterized by a certain level of functional or horizontal and hierarchical or vertical division of labour, their primary process can be defined as a network of functional interdependencies with workplaces as their nodes. When disturbances occur, these are either absorbed at the node (which requires local control opportunities) or passed on to the next node and propagated through the whole network which, then, will be out of equilibrium. The network can be analysed at the level of the nodes and at the level of the network as a whole (figure 19. 1).

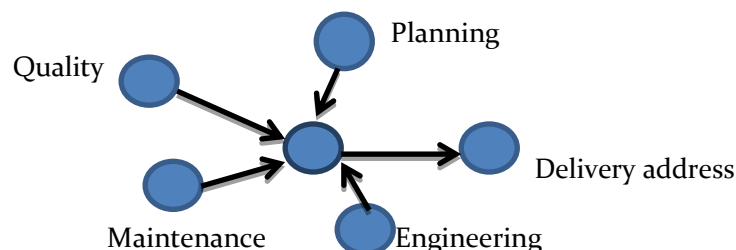


Figure 19.1. Primary process: a network of functional interdependencies. Source: De Sitter 1994: 82.

An inventory of uncontrolled disturbances at the nodes gives a measure of both work stress and efficiency problems at the job level. Uncontrolled problems at the job level cause work-life-frustrations that can be used to start a process of improvement. Lean concurs: Frustration relief is a potent avenue for process improvement. If you think about it, on-the-job frustrations are likely also to be serious performance shortcomings for the company. ... Such frustrations add up and sink morale. A work-centre team's priority-ordered list of frustrations tends also to be a rather sound priority list of opportunities for improvement for the company (Schonberger 2008: 63-64). The level of the network as a whole is used to explain the occurrence of uncontrolled disturbances. It is the complexity of the network (the number, variability and predictability of its interfaces) that explains both the probability of disturbances and the lack of local control opportunities. Redesign of the network should, therefore, be aimed at the reduction of the complexity of the production structure of the network (in order to decrease the probability of disturbances) as a precondition for the decentralization of its control structure (in order to enable local control). Without these structural insights, no proper use can be made of the variance matrix. And, with these structural insights, you do not need the variance matrix (or the network analysis), although it can be helpful from a change perspective.

STSL: a system theoretical reformulation of OSTs

So, De Sitter set himself the task of a system theoretical reformulation of OSTs. He did so in different publications. Most famous is his densely written book 'Synergetisch produceren (De Sitter 1994), a book of 418 pages in length. The book contains a set of conceptual tools, design rules and tools that enable designers (including the frontline workers themselves) to analyse existing structures in an integral way (that is, on all aspects) and to redesign them in an integral way (in order to improve the organization simultaneously on all aspects). It does so in a detailed way and differentiates for example between high volume/low variety situations (to which lean is applicable) and low volume/high variety situations (to which Quick Response Manufacturing is applicable). So, STSL is not a quality of work approach, although it encompasses quality of work as an important aspect of organizational analysis and redesign. In fact, it did so in an innovative way (see Christis, 1998).

Some preliminary definitions

The object of STSL is 'the organization of the organization (Luhmann).' An organization (in the institutional meaning of the word 'organization') has an organization (in the instrumental meaning of the word) or a work organization. So, STSL studies the work organization of organizations and as such it does not belong to the field of Organization Theory, but to the field of Organization Design. It has as its object the way organizations organize their primary process, that is, their work organization. The work organization is an instrument, used by organizations to reach their different and changing goals. If they do so in the wrong way, they will meet with difficulties in reaching their goals. The meta-question that is answered by STSL is 'what do work organizational structures have to look like in order to enable organizations to reach different and changing goals simultaneously?' Because goals are system-environment relations, this is of course a re-formulation of the general notion of adaptive systems, applied to organizations.

De Sitter defines the primary process of an organization (its manufacturing process, educational process, care delivery process, service delivery process, and so on) as a norm-based transformation process with performance and control operations as its elements. Note that operations or actions, not persons, are defined as the elements of the system. As already pointed out by Barnard (1938) and Simon (1997), if persons are defined as the elements of an organization, each shift change would amount to a total change of the organization. That would be nonsensical. With each shift change, the organization remains the same in so far as its operations (elements) and their interrelations (structure) remain the same.

In a next step, De Sitter replaces the distinction between social and technical subsystems by the distinction between subsystems and aspect systems. A subsystem is a selection or subset out of all operations with preservation of all interrelated aspects. Organizational units such as a quality department would be an example of an organizational subsystem. An aspect system is a selection

of one out of all aspects, applied to all elements of the system. A quality system would be an example of such an aspect system. In it, all elements of the primary process are analysed from only a quality perspective. The same would be true of a cost accounting system. Obviously, all elements of the quality department could be analysed from both a quality and cost perspective. But these aspects are interrelated and therefore we need an integral assessment of the performance of the quality department. This requires insight in the structure of the quality department and its place in the overall structure of the organization. It is the structure of the system that explains the way different aspects are related.

The primary process has a production structure (the grouping and coupling of performance operations) and a control structure (the grouping and coupling of control operations). De Sitter developed a set of design parameters to analyse and redesign the production and control structures of organizations. We will use those parameters to re-describe lean in system theoretical terms and in this way show that Lean is a subcase of STSL as the more general theory. But, first, we will introduce Lean Production.

19.2 Lean Production

The origin of Lean: the Toyota Production System

The Toyota Production System (TPS), now known as Lean Production (LP) or Lean Thinking (LT) was developed by the Toyota Motor Company as an answer to two problems it met after WW II. First, a small home market for different types of cars necessitated a flexible way of producing cars: “Toyota did not have the resources or the market to support many plants, and the product mix was too eclectic to justify dedicated plants” (Standard, Davis 1999: 60). Second, because of shortages on the capital market, it needed short cycle times (as the sum of processing time and waiting time). The time between purchasing raw materials and being paid by the customer had to be as short as possible. The result of years of experimenting with solutions to those problems was the TPS, a system that differed in essential ways from the mass production systems of Ford and General Motors: *So, why didn't Toyoda [after his visit of the Ford Rouge plant] take the Ford production system from the Rouge plant back to Toyota? There was too much material, floor space, time, and investment tied up for too long. Instead, the Toyota executives developed a completely different way of thinking about manufacturing* (Standard & Davis, 1999:61).

According to Hopp and Spearman, what was so revolutionary about the TPS was the fact that they did not take the existing production system with its functional structure, large batches, long setup times and high levels of inventories as given. Instead, they simplified the production system by introducing a flow system of just-in-time (JIT) production that necessitated small batches, short set-up times and continuous process improvement. This, in its turn, enabled them to install greatly simplified planning systems (pull systems such as KANBAN) and cost accounting systems (known as lean accounting): *The JIT ideals suggest an aspect of the Japanese production techniques that is truly revolutionary: the extent to which the Japanese have regarded the production environment as a control. Rather than simply reacting to such things as machine setup times, vendor deliveries, quality problems, production schedules, and so forth, they have worked proactively to shape the environment. By doing this, they have consciously made their manufacturing systems easier to manage* (Hopp, Spearman 2008:158).

Lean and organization design

According to Standard and Davis, lean is a design approach in search of a scientific foundation. The science of organization design should (1) assess its success, (2) generalize it by embedding it in more abstract concepts and theories in order to be able to (3) re-specify it for different manufacturing and non-manufacturing contexts. Lean is a success story (Schonberger 2008) and we will concentrate in this paper on embedding Lean in a more general theory. Within the field of operations management Hopp and Spearman's theory of Factory Physics (2008 third edition) has been called the science of lean (as in Standard, Davis 1999). It explains in a scientific way the success of lean and clears up some (self-)misunderstandings of Lean. You could say that lean

kicked operations management out of its dogmatic slumbers and that Factory Physics is the, in our opinion, outstanding result of the awakening process. However, it concentrates on manufacturing and within manufacturing on discrete parts production on disconnected flow lines (Hopp & Spearman, 2008:11). Others try to embed lean in Goldratt's theory of constraints (Levinson & Rerick 2002; Levinson, 2007). In this chapter we offer STSL as a non-competing but complementary scientific foundation of Lean. It offers a general system language that encompasses applications of lean in other contexts such as high variety/low volume manufacturing (Suri, 1998; 2010), service organizations (Seddon, 2005), construction (Ballard, 2008) and public organizations (Seddon, 2008). It enables us to discover both general similarities and context specific differences, to start a process of mutual learning, to integrate all these insights in a theory of organizational design and to add content to redesign proposals of for example the health care system as proposed by Porter and Teisberg (2006) and Christensen et al. (2009).

19.3 STSL as a structural design theory

STSL is a design theory, not a change theory. And it is a theory that focuses on work organizational structures. So, first, we will look at lean from a design perspective, not from a change management perspective. We are interested in the 'what' of the change, not in the 'how' and just note that the 'what' and 'how' should be aligned. And, second, we look at lean from a structural perspective. All primary processes have a structure. To contrast structure and process is committing a logical fallacy. A process is a non-arbitrary sequence of events. What makes the sequence non-arbitrary is its structure: without a structure there would not be a process. Compare this with a melody. A melody is a non-arbitrary, that is, structured sequences of tones. Without such a structure, it would be just noise. So, we look at the way lean structures or organizes its primary process and we do so in the language of STSL.

The language of STSL is a system theoretical language based on the work of Ashby, Beer, Simon and Thompson (see for an introduction Achterbergh, Vriens 2009). STSL is based on Ashby (1958) and his law of requisite variety: the variety of the regulator of a system should equal the variety of its disturbances. According to Beer (1979) that means that the management of variety combines two strategies: the reduction or attenuation of the variety of disturbances and amplification of the variety of the regulator. According to Simon (1996), you reduce the variety of disturbances by decomposing your system into subsystems with high internal and low external interdependencies. The high frequency interactions should be contained within the subsystems, leaving the low frequency interactions to higher system levels. In such a modular, hierarchic or 'nearly decomposable system,' you can make the modules self-regulating within higher level system constraints. According to Thompson (1976), you design such a modular structure by placing sequential or reciprocal interdependent positions into the same organizational unit, as was done in the composite wall method in the Durham case. In this way you create 'local, conditionally autonomous units' and substitute intra-unit coordination for costly inter-unit coordination. When interdependencies cannot be contained at this level, you move at the next higher level: *When reciprocal interdependencies cannot be confined to intragroup activities, organizations subject to rationality norms seek to link the groups involved into a second-order group, as localized and conditionally autonomous as possible... We have now introduced the first step in a hierarchy ... Each level ... is a more inclusive clustering, or combination of interdependent groups, to handle those aspects of coordination which are beyond the scope of any of its components* (Thompson 1976:59).

According to De Sitter (1994), you do so, not bottom up (as suggested by Thompson) but top down, starting at the macro level. If you design a system, whether a product, organization or paper, you don't start at the micro level (which words do belong to which sentences, working upwards to paragraphs and chapters and in the end discovering what your paper is about). Instead you start at the macro level (what is the topic of the paper, what should be its chapters) and work your way downwards. Of course, if you meet with difficulties in the process, that can be a reason to restructure your paper at higher levels: writing a paper is an iteration between top down and bottom up processes. It is a process of reshuffling until you feel satisfied with the structure of your

paper. The same applies to the design of organizational structures. So, in order to get units at the micro level with high internal interdependencies, you have to start at the macro level with higher order units that are characterized by pooled interdependencies (figure 19.2).

	Macro	Meso	Micro
Pooled	X	X	
Sequential		X	
Reciprocal			X

Figure 19. 2. Levels of design. Source: De Sitter 1981:122

A note on hierarchy and heterarchy

A ‘nearly decomposable system’ is called a hierarchical system: a system that is decomposed in subsystems with high internal and low external interdependencies. We call this ‘hierarchy’ in the technical meaning of the word. In this sense, our body is a hierarchical system and many products are designed as hierarchical or modular systems. That is why Simon in his famous paper on hierarchical systems referred mainly to examples of natural (physical and biological) systems in order to defend his thesis that all complex systems are hierarchical. Now, all organizations are decomposed in subsystems (divisions, business units, departments, teams) and all those units have someone who is in charge, that is, someone who has authority over his or her subordinates. These organizations are also called hierarchies, but now because authority is organized in a hierarchical way. We can call this hierarchies in the social meaning of the word. But note that not all organizations are hierarchical in the technical meaning. Take the functional structure in its most simple form as an example. At the first level, the organization is decomposed in three functional subsystems: preparation (such as sales, engineering, planning work preparation and procurement), production and support units (such as quality, maintenance, safety and health and personnel). At the second level, production is decomposed in functional subsystems such as cutting, bending, welding, painting and assembly. As a result, you get the neat boxes of the organogram with a manager within each box. But such a hierarchy in the social sense (hierarchy of authority) is not a hierarchy in the technical sense (a nearly decomposable system). If you fill in the interdependencies between the units, its structure is shown to be that of a heterarchy or network, in which all subsystems and their operations are connected to each other (figure 19.3).

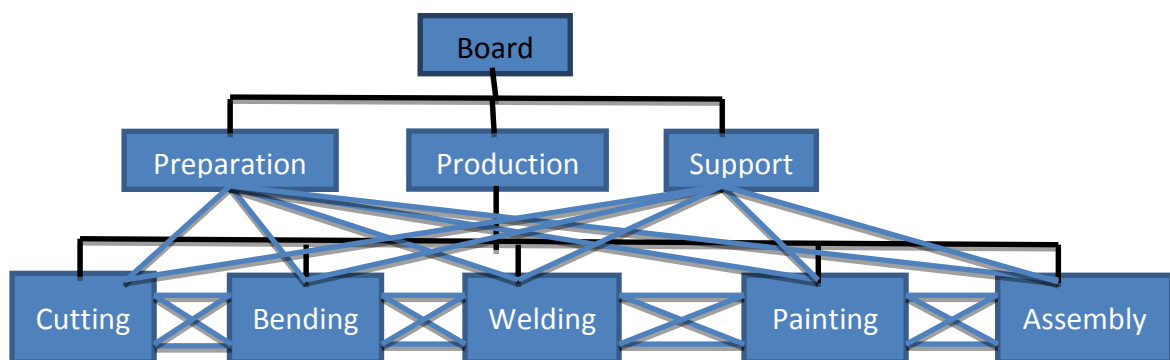


Figure 19.3. The functional structure as a heterarchy

In this structure, all units are interconnected and interdependent, because all functionally specialized operations are – potentially or actually – coupled to all customer orders. According to both LP and STSL, in a complex and dynamic environment, such a heterarchical structure should be transformed in a hierarchic, that is, nearly decomposable or modular structure that consists of subsystems in which different and interdependent operations are carried out on a subset of similar

order types.

STSL: an overview

STSL is a design theory that can be summarized until so far in the following way:

- Design object: the primary process as an instrument to reach different and changing organizational goals.
- Design goal: requisite variety or controllability as a generic structural capacity that enables organizations to reach different and changing goals simultaneously.
- Functional design requirements or criteria: quality of organization (cost, quality and time), of work (stress risks and learning opportunities) and of labor relations (cooperation and shared decision-making).
- Design strategy: reduction of the complexity of the production structure as a precondition for the decentralization of the control structure.
- Sequence rules: design your production structure top down (and then your control structure bottom up) and start design of the production structure with the making function and then align your preparation and support functions.
- Design parameters: as Mintzberg says, these are the knobs a designer turns to reach design goals and criteria. These design parameters will be discussed in the next section.

19.4 Design parameters

In this paragraph we shall describe the design parameters that will then be used to reformulate Lean in STSL concepts. To enable understanding of the STSL design parameters by non-Lowlands readers, we start with the way Mintzberg introduced the concept of design parameters.

Mintzberg on design parameters

Mintzberg (1983) introduced four groups of design parameters in his book 'Structure in fives': design of positions, design of superstructure, design of lateral linkages and design of decision-making system. It is not difficult to recognize the first two groups as the design of the production structure and the last two groups as the design of the control structure (and Mintzberg correctly notes that the design of the decision-making structure should ante cede the design of lateral linkages). According to Mintzberg, unit grouping (a parameter of the super structure) is the strongest parameter of all. As we will see, it is an enabling constraint both 'downwards' with respect to the design of positions (job design) and 'upwards' with respect to the design of the control structure (decision-making system and lateral linkages).

Unit grouping: functional and market-based

We owe Mintzberg the insight that all bases or criteria for grouping can be reduced to two: functional and market grouping. To quote him extensively: *In fact, we shall comprise all the bases for grouping discussed above to two essential ones: market grouping, comprising the bases of output, client and place, and functional grouping, comprising the basis of knowledge, skill, work process, and function. ... In effect, we have the fundamental distinction between grouping activities by ends, by the characteristics of the ultimate markets served by the organization – the products and services it markets, the customers it supplies, the places where it supplies them – or by the means, the functions (including work processes, skills, and knowledge) it uses to produce its products and services* (Mintzberg, 1983: 53-54).

The same distinction is used in STSL, although in a different terminology. A primary process has as its end the transformation of a requested order (a customer, client or patient with a wish) in a delivered order. And the activities or operations carried out in the primary process are the means to reach that end. So, you can group on means. You then look inside the organization in search of similar activities/operations that are grouped together into the same functionally specialized unit. In STSL this is called operations-based grouping. In an operations based structure all operations are potentially or actually coupled to all customer orders. That is what makes them heterarchical or network structures. And that is why these structures are so complex and prone to disturbances.

You can also group on ends. You then look outside your organization at the market environment in search of similar order types. The different and interdependent operations that are needed for the production of a restricted set of similar orders are then placed together in the same organizational unit. In this way you create independent or pooled streams around similar orders. That is why this is called stream-based grouping in STSL and value streams in lean. According to Mintzberg, you will find market grouping at the higher and functional grouping at the lower levels of organizations as in the divisionalized form. Both STSL and lean propose to apply market grouping 'all the way down' till you reach the level of cross-functional teams or cells as the lowest level building blocks of your organization. This is one of the reasons you will not find this kind of organizations in Mintzberg's configurations (see also Sabel, 2006). In general, by stream-based grouping, you try to reduce variability in input, process or output. Because an order is a customer/client with a wish (for a product or service), stream-based or market grouping can be:

- Product-based: similar wishes for different customers (as often in manufacturing);
- Customer-based: similar customers with different wishes (as often in services).¹;
- Project-based: similar projects in which unique wishes of customers are handled (as in architect bureaus).

The functional and market structures have different 'downward' and 'upward' effects. A functional structure enables job specialization or function differentiation at the job level and constrains the possibilities of job enlargement, job enrichment and cross-functional teamwork. The opposite is the case with market grouping. Also the functional structure leads to both a centralization of decision-making and a proliferation of lateral linkages. The opposite is the case with market grouping: *In effect, the functional structure lacks a built-in mechanism for coordinating the work flow. Unlike the market structures that contain the work-flow interdependencies within single units, functional structures impede both mutual adjustment among different specialists and direct supervision at the unit level by the management. The structure is incomplete; additional means of coordination must be found. The natural tendency is to let coordination problems rise to higher-level units in the hierarchy, until they arrive at a level where the different functions in question meet. The trouble with this, however, is that the level may be too far removed from the problem (Mintzberg, 1983: 59).*

The design parameters of STSL

In this paragraph we present the parameters that will be used in the next paragraph to reformulate lean in terms of STSL's parameters. The parameters refer to the production structure (the grouping and coupling of performance operations), the separation of performance and control and the control structure (the grouping and coupling of control operations). The performance operations are subdivided in preparatory, making and support operations. The preparatory functions include quoting, engineering, order processing and procurement. The support functions include quality control, maintenance, logistics, accounting and control and personnel). For reasons of space, we concentrate on the parameters of the production structure including the parameter of separation/integration of performance and control:

- Functional concentration versus de-concentration. This parameter refers to the way of grouping of making functions: functional versus market grouping;
- Functional specialization versus de-specialization. This parameter refers to the centralization versus decentralization of preparatory and support functions;
- Functional differentiation versus integration. This parameter refers to the level of division of labour or job specialization within the preparatory, making and support functions;
- Separation versus integration of performance and control. This parameter refers to the separation versus integration of conception and execution.

Note that the three parameters of the production structure refer to the same phenomenon

¹ The local cross-functional district teams of Buurtzorg Nederland and the local branches of Svenska Handelsbanken are examples of customer-based groupings (see the respective case studies).

(functional or market grouping) applied both to different parts and to different levels of the primary process. The parameters are used in both a descriptive and a normative way: in complex and dynamic environments the combination of functional de-concentration, de-specialization and integration enhances both the quality of the organization, work and labour relations.

Functional structure: economies and diseconomies of scale

A typical functionally concentrated, specialized and differentiated organization with separation of performance and control would look like figure 19.4 (and note that in Mintzberg's divisionalized form this is the internal structure of the divisions and/or business units it is composed of).

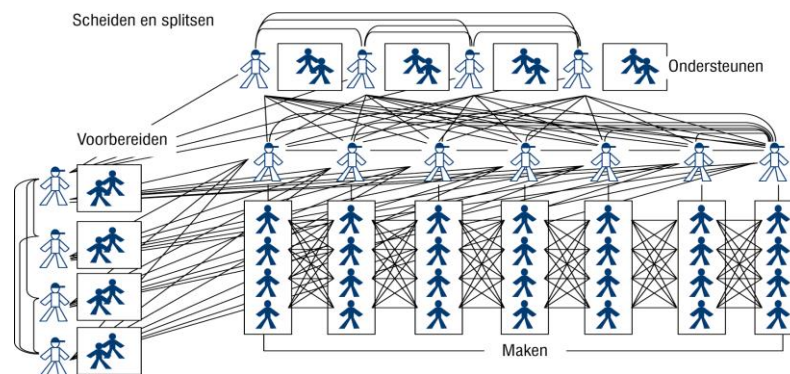


Figure 19.4. A complex organization with simple jobs. Source: De Sitter 1983: 138).

The underlying idea of such an organizational structure is that cost reductions can be best achieved by exploiting economies of scale. At the same time the diseconomies of scale that are created in this way are accepted and even made invisible. The idea of cost reduction by economies of scale not only informs the design of organizational structure, but also the design of incentive and reward structures. As such, the functional structure is entrenched and difficult to change: you have to change both taken for granted ways of thinking (cultural change) and of doing (structural change).

In a nutshell, the idea tells you that in order to reduce costs, you should group in a functional way and you should optimize the functionally specialized parts or segments of the organization by aiming at job specialization (in order to economize on wage costs), maximal capacity utilization (in order to avoid idle capacity) and large batch production (in order to amortize set-up times). Paradoxically, this leads to a sub-optimization of the whole: it disrupts the order flow, creates excessive inventories with negative effects on costs, quality, cycle time and flexibility and increases coordination costs immensely which necessitates high staff levels and the instalment of complex planning and cost accounting systems. Overhead costs will rise and the paradox result of cost reduction in this way is a worsening of performance in both quality, time/speed **and costs**. The system dynamics are shown in figure 19.5.

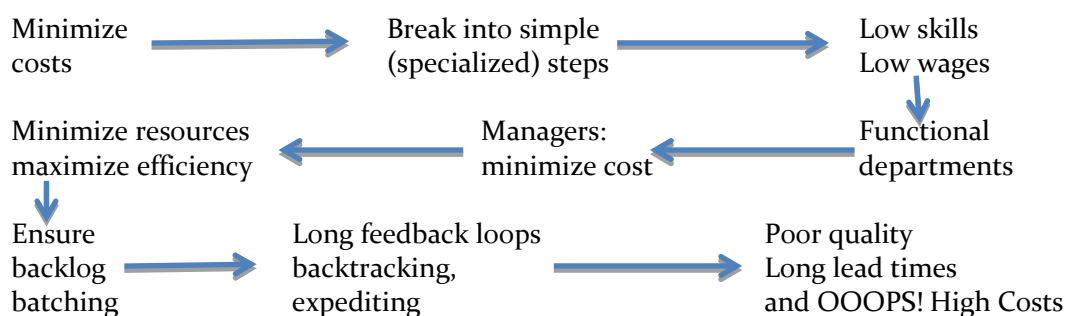


Figure 19.5. Diseconomies of scale. Adapted from Suri, 2010:45.

19.5 A system theoretical reformulation of lean

How not to characterize lean

For some, lean is about waste reduction (the seven forms of muda) and/or continuous process improvement. This is a misleading characterization of lean. First, waste reduction is neither a goal nor a manufacturing strategy: *Is losing weight the definition of a good diet? No; it is better health, increased capabilities, longer life. So it must be for lean. More fittingly, lean employs a large set of concepts and tools to reduce delays and quicken response in all processes. That is fundamental lean, with time compression as its main focus* (Schonberger 2008:45). Standard and Davis concur: *If eliminating waste is the central theme of an improvement effort, the benefits will be superficial. ... System improvement focusses on improving how material and information flow through the plant. Its objective is to minimize cycle time* (Standard, Davis 1999:134). So does Suri (2010): it's about time. Second, if all important forms of waste are caused by functional batch-and-queue production, it isn't helpful to try to reduce waste within that structure. Improvement within those structures is called kamikaze kaizen by Womack and paintball kaizen by Standard and Davis (1999: 134). The most important gain is in substituting 'organizing with the flow' for 'organizing across the flow.'

The lean strategy: flow production with capacity buffers

According to Hopp and Spearman (2008), in an ideal world without variability, demand and transformation can be perfectly aligned. However, the real world with variability necessitates buffers. Variability comes in two forms: demand and process or transformation variability. And buffers can only take three forms: time (when demand waits for products/parts), inventory (when product is finished before demand) and capacity (which reduces the need for the other two buffers). For traditional bureaucratic organizations, idle capacity is the main waste. The guiding idea for organizing production is maximum resource utilization which idea is best realized in a functional structure (Modig & Ahlstrom, 2012). But such a manufacturing strategy necessarily produces waste in the form of time and/or inventory buffers (and, according to De Sitter, adds internal variability to demand variability). In contrast, the lean strategy is aimed at flow efficiency (Modig, Ahlstrom 2012). The most important wastes are time and inventory buffers. To reduce those buffers, lean introduces a flow structure, installs a pull planning system and then starts a process of continuous improvement to further reduce process variability. But such a manufacturing strategy **necessitates capacity buffers**: *Toyota exploited its understanding of the science of operations by using a 30 percent capacity buffer to support its strategy to drive consistent, low cycle times. Most Lean practitioners would label such a capacity buffer as non-value-added and try to eliminate it ... Toyota chose to pay for inventory reduction, low cycle times, and continuous improvement efforts with its capacity buffer. The cost of the capacity buffer was outweighed by the ability it provided Toyota for buffering against variability to achieve lower inventories, reduced scrap, and better response time. This was the right choice for Toyota and was reflected in its financial statements* (Pound, Bell & Spearman, 2014: 175-176).

Flow production: functional de-concentration or market-based grouping

The evil-doer for both lean and STSL is functional batch-and-queue production and both turn over the production structure from a functional to a market-based one (functional de-concentration). In lean, market-based groupings are called value streams: *The best way to think about a value stream is as a business segment focused on a product family, or sometimes, customer family. There is probably nothing more effective, in process improvement, than breaking up the functional silo's and realigning the processes by the work flow in a product family. The work cell is a microcosm of this realignment. The focused factory and plants-in-a-plant are enlarged variants. Linking a focused factory to a supply chain or customer chain extends the scheme further* (Schonberger, 2008: 106).

Lean is based on both a different way of thinking (the idea of economies of flow or flow efficiency) and of doing (designing organizational structures and reward and incentive systems that are informed by this idea). To create a flow, you take a restricted set of similar orders (for example all titanium bicycles, as in Womack & Jones, 2003), determine the operations/machines you need to produce this subset of orders, unfasten them from the floor and place them together in a cross-

functional value stream and/or manufacturing cell. The advantages are many. First, you reduce the complexity of the production structure by placing interdependent activities in the same unit. This reduces coordination needs and enables you to install immensely simplified planning and cost accounting systems. Both planning and cost accounting are directed at the higher levels of value streams and manufacturing cells and not at individual levels of machines and operations as in traditional planning and cost accounting systems. In terms of costs, you reduce overhead costs, decrease the number of lateral linkages and introduce performance criteria that are targeted at the optimization of the process as a whole.

Second, by introducing flow production, you lower inventories with positive effects on costs (less capital tied up in inventories, smaller storage space, less material handling, less risk of obsolescence), quality (early discovery of defects, less scrap and rework and improvement of root cause analysis), cycle time (shorter cycle times and less variability of cycle times) and flexibility (shorter cycle times postpones committing resources to production and so enables adaptation to changes in customer orders). Instead of pitting costs against quality and time, lean stresses the cost aspects of quality and time. By concentrating on improvements in quality and speed, costs will be reduced as a consequence and new business will be generated.

Note that the same idea ('hospitals within hospitals') is applied by Porter and Christensen to the organization of the health care system. According to Porter: *In health care, the traditional functional structure must radically shift to a structure that medically integrates the care of patients with particular medical conditions. We term such a structure the integrated practice unit [IPU] structure. IPU's are defined around medical conditions, not particular services, treatments, or tests. ... Ultimately, the IPU model lends itself to the notion of hospitals within hospitals and practices within practices, rather than monolithic, functionally structured entities* (Porter, Teisenberg 2004:168, 171). And according to Christensen, the all-purpose, general hospital should be reorganized along three different business models: the value added chain, the facilitated network and the solution shop. Each business model is characterized by a different value proposition, different technological and human resources, a different process organization and a different profit formula. A star case would be the introduction of team-based medicine in Kaiser Permanente (see Lawrence 2002).

Flow production: functional de-specialization or decentralization of preparation and support functions

Functional de-specialization is the same as functional de-concentration but now applied to the preparatory and support functions. It means that formerly centralized preparatory and support functions are decentralized to the different value streams (as in the Bromont case, in which each cross-functional product unit has its own support unit). The same in lean/QRM: *Notice that each cluster (focused factory) has its own staff of engineering (manufacturing, quality, design), maintenance and material support* (Nicholas, Soni 2006: 195). In QRM this is called an office cell. In QRM, shop floor and office cells are always designed around a Focused Target Market Segment (FTMS). An office cell is based on the same design principles as applied to a shop floor cell and it is defined as "a closed-loop, collocated, dedicated, multifunctional, cross-trained team responsible for the office processing of all jobs belonging to a specific FTMS. The team has complete ownership of the cell's operation and the primary goal of the team is reduction of the cell's [cycle time]" (Suri, 2010: 14). Former overhead is now contained within the value streams and/or cells which greatly simplifies both planning and cost accounting.

Functional de-specialization corresponds to Mintzberg's horizontal and vertical decentralization or decentralization of staff and line functions as in the divisionalized form. According to Mintzberg, decentralization can be selective and parallel. This corresponds to STSL's distinction between aspectual and integral control. Parallel decentralization or integral control is logically associated with market-based grouping: *Each unit or division is decoupled from the others and given the power necessary to make all those decisions that affect its own products, services, or geographical areas. In other words, parallel vertical decentralization is the only way to grant market-based units the power they need to function in a quasi-autonomous manner* (Mintzberg, 1983: 102). Note that in Mintzberg's divisionalized form, market-based grouping and parallel decentralization stop at the

divisional level, while in both STSL and lean they are applied ‘all the way down’ to the level of cross-functional teams or cells. That is one of the reasons why sociotechnical and lean organizations do not fit any of Mintzberg’s five configurations.²

Flow production: functional integration and cross training

Independent streams take the environment as given and segments or cells within the stream take the stream as given. Designing segments is the same as determining the external structure of the segments (interdependencies between the segments within a stream should be pooled or sequential). Functional de-concentration can now be applied to the internal structure of the segments. In each segment, different and interrelated operations are carried out on a restricted set of similar orders. The result is cross-functional teams that contrast with the functional teams of the functional structure. Application of cross training makes the team members multi-skill and the team a flexible one. To record the level of cross training, the flex matrix is used by both STSL and lean. The vertical axis of the matrix contains all direct and indirect team tasks, the horizontal axis all team members and within the matrix you can see who is able to carry out which task. You can replace the indications within the matrix with skill or competence levels (beginner, competent and expert). To this, two rules can be added. The first one states that work is not top sport: you cannot be an expert in all team tasks (see Benner, 1984). Therefore, the aspiration is to make as many team members competent in as many team tasks as possible. The next rule states that professionalization should be in the hands of the professionals themselves. Therefore, the experts at a task coaches the beginners at the same task to the level of competent task performance. In this way you have created a simple competence management tool that enables workers to help each other in getting better at their work. It does so without having to define competences and to operationalize them in behavioural criteria as in standard forms of competence management.³ According to Mintzberg, there is a trade-off between efficiency and quality of work: *job enlargement pays to the extent that the gains from better-motivated workers in a particular job offset the losses from less than optimal technical specialization* (Mintzberg, 1983: 31). Against this, both lean and STSL can explain why organization and job design have direct and simultaneous effects on both efficiency and quality of work. They both point to the necessary macro and meso level preconditions for job enlargement and enrichment in conditionally autonomous teams at the job level. In this way they create simple organizations with complex jobs and so improve both the quality of the organization and the quality of work.

Flow production: integration of performance and control

According to Lazonick (2005), functional and hierarchical integration is one of the characteristics of the innovative enterprise (the other two being strategic control and financial commitment). All organizations have a certain level of functional or horizontal and hierarchical or vertical specialization (fs and hs respectively in the triangle):

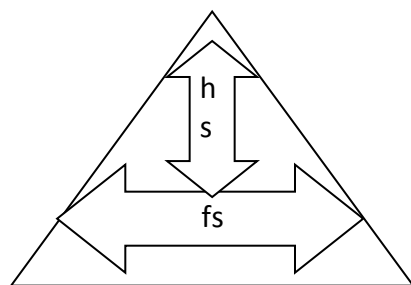


Figure 19.5. Functional and hierarchical specialization. Adapted from Lazonick 2005.

Most organizations apply a deep level of division of labour. By functional differentiation and

² And because he thinks that the distinction between functional and market grouping is irrelevant for professional bureaucracies he misses the innovations that actually take place in health care and education, see Christis, 2011.

³ For the tool, see Christis et al. (2013) and for the theoretical background Christis, Fruytier (2006).

segmentation, they stretch the triangle in a horizontal direction. This increases coordination needs and necessitates a stretching of the triangle in the vertical direction (hierarchical differentiation and segmentation), resulting in one huge triangle that has all the marks of a bureaucratic, complex organizations with simple jobs (as in figures 3 and 4). A small vanguard of innovative organizations introduce functional and hierarchical integration in their organization. They do so by designing an organization that is composed of smaller triangles (the principle of modularity). They, then, make these triangles or modules conditionally autonomous by delegating control tasks to those modules: *Operators in work cells typically have autonomy to make decisions and perform their own basic equipment maintenance, changeover, quality control, and job-scheduling (and) also engage in continuous improvement efforts, data collection and performance management, and even materials procurement from vendors* (Nicholas & Soni, 2006: 79). See figure 19.6.

	Cost-based	Time-based
Organization:	Functional	Cellular (cross-functional manufacturing and office cells)
Management:	Top-down control	Team ownership
Team members:	Narrow, specialized	Cross-trained
Mind-set:	Efficiency and utilization goals	Relentless focus on cycle time reduction

Figure 19.6. Key transformations. Adapted from Suri, 2010: 46

Continuous improvement

Most of the advantages of LP are a consequence of its JIT flow system with shorter cycle times and lower inventory levels. At the same time, this makes them extremely vulnerable to disruptions of the flow. First, inventories act as safety buffers. Lowering them means lowering the safety buffers. Second, by limiting the number of possible routings the system becomes more rigid. To elucidate the situation, the metaphor of the river is often used. The water level of the river stands for the inventory level, the boats on the river represent the orders and the rocks at the bottom symbolize possible disruptions of the progress of the boats/orders. High inventory levels buffer against the rocks, but as a consequence more boats/orders are on the river which will result in congestion problems. LP lowers the inventory level and by this reduces the number of boats/orders on the river and concomitant congestion problems. But, now, the risk of boats hitting the rocks at the bottom is increased. If the rocks are not removed, the progress of the boats will be disrupted. So, introducing LP, at the same time makes the rocks visible (by lowering inventory levels), creates the urgency to remove them (to prevent disruption of the flow) and creates the possibilities to do so (by multi-skilled workers who have an overview of the interrelated operations of the process).

Note that the design sequence is this: first install flow production (value streams and cells) and pull planning systems and only then start the process of continuous improvement. Without continuous improvement on the process level, the system would fall apart. And without the introduction of flow and pull at the system level, process improvement makes little sense: *Process improvement alone cannot produce system wide advantages and system improvement requires that specific processes within the system be modified* (Standard & Davis, 1999: 127). Process improvement without changes at the system level is also called kamikaze kaizen (Womack) or paint ball kaizen (Standard & Davis). This is the top down design approached that is proposed in for example Womack & Jones (2003), Nicholas & Soni (2006), Suri (2010) and Standard & Davis (1999). Top down in design (from macro to micro) should not be confused with top down in change (without worker participation): you can and should design top down in a participative way.

A note on routines: lean organizations as high reliability organizations

Routines and standard work are an essential part of the process of continuous improvement. In this, lean organizations resemble so-called High Reliability Organizations (HRO's). HRO's have a primary process that is a high risk system with a high catastrophe potential: if something goes

wrong, the consequences for people and environment can be severe. Yet, in HRO's, no or less accidents occur than in comparable high risk organizations. Their success in preventing accidents is explained by the introduction of a set of mindful practices (Weick & Sutcliff, 2007). Central to those mindful practices is the ambivalent attitude to routines and standards. In HRO's everything is formalized. If the best way to do something is discovered, it is turned into a standard, routine or standard operating procedure. At the same time, these routines are distrusted. Using them in slightly different contexts could have disastrous effects. These routines are, therefore, critically reviewed and updated permanently by the frontline workers themselves. The level of formalization explains why HRO's don't fit Mintzberg's adhocracies. The described ambivalent attitude to routines applies literally to the lean process of continuous improvement (Christis, 2010). First, by lowering inventory buffers, LP creates an internal catastrophe potential and so creates the urgency and motivating force to install those mindful practices (see Sabel 2005). It shows that 'normal' organizations can learn from HRO's (pace Roberts, 1991), but only by creating an internal catastrophe potential (pace Weick & Sutcliff, 2007). Second, in the lean literature a distinction is made between work standards and standard work. The former refers to the standards formulated by Mintzberg's techno-structure and imposed on the workers. It is an example of the separation of conception and execution. Standard work refers to standards that are developed, reviewed and updated by the workers themselves. It is an example of the integration of conception and execution: *Whereas the former [standard work] relies mostly on the efforts of shop floor teams to develop standards, the latter [work standards] imposes standards that are developed by staff specialists and engineers. ... Another difference between traditional work standards and standard work is that the former are considered semi-permanent; the latter, provisional. Work standards, first conceived by Frederick Taylor around 100 years ago, represent the "one best way" to do something; they leave no opening or incentive for improvement. Standard work, however, represents the best way known at the moment; if a problem occurs while the standard work is being performed, the standard work is considered a contributor to the problem and in need of revision to prevent the problem from reoccurring. ... Standard work represents the best ideas a team can generate at the time; it is the current 'gold standard'* (Nicholas & Soni, 2006: 163-164). This practice was introduced by Ohno early in the development of the TPS: *"One of Taiichi Ohno's first assignments at Toyota was to prepare job descriptions, and he found the experience invaluable. Only through preparing their own job descriptions, he concluded, could workers comprehend the details of their work and know why they should have to do things that way, and only then would they be capable of pondering other, better ways to do it – the basis for kaizen. Ever since, shop-floor teams at Toyota have prepared their own work instructions and standard work descriptions for their work areas"* (Nicholas & Soni, 2006: 163).

Discussion and conclusion

Commonalities

We showed that lean corresponds to the prescriptive STSL parameters: lean combines functional de-concentration, functional de-specialization, functional integration and integration of performance and control functions. And it does so for the same reasons: to reduce internal variability and cope with the remaining variability. In the language of Ashby: to reduce the variety of disturbances and to enhance the variety of the regulator. Both STSL and lean regard cross-functional teams or cells as the basic building blocks of the organization. The advantages of the shift from functional to cross-functional units as the building blocks of the organization are many:

- Cells eliminates inventory (and so frees cash flow);
- Cells shortens cycle time (and so creates cash flow);
- Cells draws interrelated processes together in time and place (and so enables continuous improvement of quality, speed and costs);
- Cells eliminates overhead costs (by simplifying planning and cost accounting systems);
- Because cells are the first customer of product engineering, they enable design for manufacturing.

However, designing value streams and/or cells can be a difficult affair: there are many ways of

grouping similar orders and finding the right one can be a hard nut to crack. Designing streams and/or segments within streams can take different forms according to the situation you start with. To simplify matters, De Sitter distinguishes three different design contexts or start situations: crisscross or spaghetti streams, latent streams and one single stream (figure 19.7). These are called heterogeneous, semi-homogeneous and homogeneous streams in Kuipers, Van Amelsfoort (1990).

Crisscross stream	Latent streams	One stream
Many different products/variants	Mainly product-variants	One product with different designs
Many differences in operational combinations and sequences	Differences in operational combinations but less in sequences	Identical operations and fixed sequences
Hundreds of routings	Tens of routings	One routing
Example: part supplier	Example: producer of furniture	Example: assembly of mass product

Figure 19.7. Design situations. Source: De Sitter 1994:245

These correspond to Hopp and Spearman's distinction (based on Hayes & Wheelwright 1979) between the jumbled flow (job shop), the disconnected line flow (batch production), and the connected line flow (assembly line).

Design principles and contexts in manufacturing

We can first use these different design contexts to resolve a debate that is raging in the lean community: is QRM the same as or totally different from lean? Lean in its classical, Toyota, form was developed for high volume/low variety production. It then moved on to the different design situations of semi-homogeneous and heterogeneous streams. In the way it had to invent new tools because tools such as value stream mapping, KANBAN and levelled production make no sense in the high variety/low volume situation (crisscross streams). QRM was developed by Suri to meet the high variety/low volume situation. In this way, he developed tools that are applicable in the other design situations. So, both apply the same higher level principles to different lower level design contexts. And both are extensively covered in De Sitter 1987 and 1994. De Sitter offers a set of analytical tools for introducing flow production in these different situations that are broader than and partly overlap with the tools offered in the lean/QRM literature. At the other hand, STSL was never as creative as lean in developing simplified planning and cost accounting systems and visual control systems.

The discussion between STSL and lean seems to focus on the design of the assembly line. The general idea of De Sitter is to redesign the assembly line into parallel flows with less stations in which less people carry out more interdependent tasks in longer 'takt' or work cycle times. This can take the form of phase groups, mini-lines, dock groups, pre-assembly in module groups and can cumulate in assembly groups in which the complete product is made in a few phases by one or a few groups (as in the Volvo Uddevalla plant in Sweden). Apart from the effect on the quality of work this would drastically reduce system losses: takt time is inversely related to system losses of the line structure such as stochastic losses, balancing losses, and so on. In Womack, Jones and Roos (1990) the assembly groups of Uddevalla are presented as a naïve return to a lost time of craft production. At the other hand, the average Japanese takt time was 90 seconds which is long when compared to European firms (32 seconds for the Fiesta in Belgium and 40 seconds for the Peugeot in France, according to De Sitter). This explains at least part of the Japanese success. Moreover, when confronted with a tight labour market, Toyota started to experiment with 'Scandinavian'

forms of work restructuring at the assembly line (Pill & Fujimoto, 2007).

Design principles and contexts in health care

Functional structures, based on a flawed notion of cost reduction by resource efficiency, still dominate the health care sector. In manufacturing this leads to fragmented production and in health care to fragmented care. So, the first step in reaching integrated care is the introduction of so-called clinical or care path's. In such a path the routing of similar patients through the different functional silo's is traced. This is the same as value stream mapping in manufacturing. And the results are similar: waiting time as percentage of cycle time is at least 90%, quality and safety are poor and costs are high. Of course, care delivery as a primary process differs fundamentally from manufacturing. But the question then is 'if care delivery is so different, why do we organize it as a functional manufacturing plant?

So, in a next step, clinical or care path managers are appointed to optimize the care path from the viewpoint of the patients. We now have the nightmare of a matrix organization in which unit managers, responsible for resource efficiency fight with care path managers, responsible for flow efficiency. The way out of this situation is to organize healthcare along care paths for similar patients (as proposed, among others by Porter, Christensen and Lawrence). These can be:

- Product-based: grouped around patients/clients with the same medical/care conditions as in a migraine centre, asthma centre, cancer centre, hip street, cataract street, and so on.
- Customer-based: grouped around similar patients/clients with different medical/care conditions as in the district teams of Buurtzorg Nederland and the social district teams many Dutch municipalities are experimenting with.

Evidence is growing that these forms of integrated care deliver higher quality care against lower costs as was to be expected from the manufacturing experience. To be successful, health care organizations with team-based medicine should substitute capacity buffers for time buffers (in health care there are no inventory buffers; instead, patients have to wait). To start with, they should follow Toyota and work with a capacity buffer of 30% and use it to absorb fluctuations in the patient flow and to give frontline workers the time for continuous improvement in order to further reduce process variability and enhance quality and safety of care. Next, they should install pull planning systems to put a cap on waiting time. And of course, they will meet the same problems as in manufacturing: how to make cells within a value stream, what to do with shared capacities (such as operating theatres), and so on.

Lean and mean

At last, the sociotechnical community should shed its prejudices on lean. The first one is that lean concentrates on process improvement and neglects organizational design. As we showed, this is an absurd accusation. A second one is that lean neglects quality of work. Bromont is a star case of lean/six sigma within General Electric. Those who visited the plant and spoke with the workers and team leaders know, again, that this is an absurd accusation. Its structure, with cross-functional manufacturing and support cells can be called sociotechnical and we all admired the way frontline workers are involved in the continuous process of improvement and innovation of the plant. Long ago, lean added the under-utilization of human capacities as a form of waste and the spirit of lean is well captured by Schonberger: *lean is hard on processes and soft on people*.

References

- Achterbergh, J., D. Vriens (2009). *Social systems conducting experiments*. Berlin: Springer.
- Ashby, W. (1956). *Introduction to Cybernetics*. London: Chapman & Hall.
- Ballard, G. (2008). The Lean Project Delivery System: An Update. *Lean Construction Journal* 2008, 1-19.
- Barnard, C. (1938). *The functions of the executive*. Cambridge: Harvard University Press.
- Beer, S. (1979). *The heart of enterprise*. New York: Wiley.
- Black, J, S. Hunter (2003). *Lean manufacturing systems and cell design*. Dearborn: Society of Manufacturing Engineers.
- Christensen, C. et al. (2009). *The Innovator's Prescription. A disruptive Solution for Health Care*. New York: Mc Graw Hill.
- Christis, J. (1998). *Arbeid, organisatie en stress*. Amsterdam: Spinhuis.
- Christis, J. (2010). Organization and job design: what is smart organizing? In: H. van Lieshout, L. Polstra, J. Christis, B. Emans, *Managment of labour. Societal and managerial perspectives*. Groningen: KCA.
- Christis, J. (2013). *Handleiding competentiemanagement instrument*. Groningen: KCA.
- Christis, J. (2014). *Netwerkanalyse*. Groningen: KCA.
- Christis, J., B. Fruytier (2006). Competentiemanagement : Een kritiek en een alternatief. *Tijdschrift Voor HRM*, 4, 6-16.
- Christis, J. (2009). *Wat is slim organiseren?* Groningen: Hanzehogeschool.
- Hopp, W., M. Spearman (2008). *Factory Physics*. New York: McGraw Hill.
- Hayes, R, S. Wheelwright (1979). Link manufacturing process and product life cycles. *Harvard Business Review* (January-February), 133-140.
- Kuipers, H., P van Amelsvoort (1990). *Slagvaardig organiseren*. Deventer: Kluwer.
- Lazonick, W. (2007). *Varieties of capitalism and innovative enterprise*. Comparative Social Research, 24, 21-69.
- Lawrence, David (2002) *From chaos to care. The promise of team-based medicine*. Cambridge: Da Capo Press.
- Levinson, W. (2007). *Beyond the theory of constraints*. New York: Productivity Press.
- Levinson, W., R. Rerick (2002). *Lean manufacturing: a synergistic approach to minimizing waste*. Milwaukee: ASQ Quality Press.
- Mintzberg, H. (1983). *Structure in fives*. New Jersey: Prentice-Hall.
- Pound, E., J. Bell, M. Spearman (2014). *Factory physics for managers*. New York: Mc Graw Hill.
- Porter, M., E. Teisberg (2006). Redefining healthcare. Boston: *Harvard Business School Press*.
- Nicholas, J. A. Soni (2006). *The portal to Lean production*. New York: Auerbach Publications.
- Porter, M. (2008). Value-based health care delivery. *Annals of Surgery*, 248 (4), 503-510.
- Roberts, K. (1990). Some characteristics of one type of high reliability organization. *Organization Science*, 1(2), 160-176.
- Sabel, C. (2006). A real-time revolution in routines. In: C. Heckscher, P. Adler (eds.) *The firm as a collaborative community*. Oxford: Oxford University Press.
- Schonberger, R. (2008). *Best practices in Lean six sigma process improvement*. New York: Wiley.
- Seddon, J. (2005). *Freedom from command & control. Rethinking management for lean service*. New York: Productivity Press.
- Seddon, J. (2008). *Systems Thinking in the public sector*. Axminster: Triarchy Press.
- Simon, H. (1996). The architecture of complexity. In: Simon *The sciences of the artificial*. Cambridge: MIT Press.
- Simon, H. (1997). *Administrative behavior*. New York: Free Press.
- Sitter, L. U. de et al. (1987). *Het flexibele bedrijf*. Deventer: Kluwer.
- Sitter, L. U. de (1994). *Synergetisch produceren*. Assen: Van Gorkum.
- Standard, C., D. Davies (1999). *Running today's factory*. Cincinnati: Hanser Gardner Publications.
- Suri, R (1998). *Quick Response Manufacturing: a companywide approach to reducing lead times*. New York: Productivity Press.
- Suri, R. (2010.) *It's about time*. New York: Productivity Press.

Takahiro, F., F. Pil (2007). Lean and reflexive production. *International Journal of Production Research*, vol. 45 (16), 3741-3762.

Thompson, J. (1967). *Organizations in action*. New Brunswick: Transaction Publishers.

Weick, K. & Sutcliffe, K. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty*. Second edition. San Fransisco: John Wiley & Sons, Inc.

Womack, J. & Jones, D. (2003). *Lean thinking*. New York: Free Press.

Womack, J., Jones, D. & Roos, D. (1990). *The machine that changed the world: The story of lean production*. New York: Rawson.

ⁱ A shorter version will be published in:

Mohr, Bernard J., Pierre van Amelsvoort ed. (2015) Co-creating humane and innovative organisations: evolutions in the practice of socio-technical system design, (to be published Fall 2015 by Amazon.com 'print-on-demand')