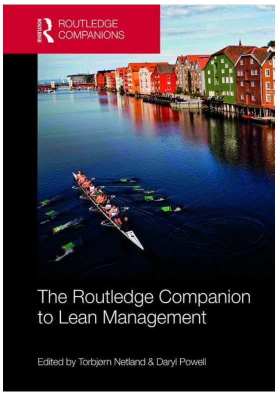

The Routledge Companion to Lean Management

Torbjørn H. Netland, Daryl J. Powell

Lean Engineer-to-Order Manufacturing

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9781315686899.ch25
Daryl J. Powell, Aldert van der Stoel

How to cite :- Daryl J. Powell, Aldert van der Stoel. 28 Dec 2016, *Lean Engineer-to-Order Manufacturing from:* The Routledge Companion to Lean Management Routledge
Accessed on: 01 Apr 2023
https://test.routledgehandbooks.com/doi/10.4324/9781315686899.ch25

25

LEAN ENGINEER-TO-ORDER MANUFACTURING

Daryl J. Powell and Aldert van der Stoel

Introduction

Engineer-to-order (ETO) is a type of manufacturing that is inherently different from mass production. It refers to a strategy by which design, engineering, and production do not commence until after a customer order has been received. ETO as a manufacturing approach is typically characterized by a high variety of customized products produced in low volumes (often one-of-a-kind), a series of complex, non-repetitive, labor-intensive processes (often demanding highly skilled labor), and long lead times due to the additional elements of engineering lead time and procurement lead time in addition to production and delivery lead times.

ETO can be described using a concept called the customer order decoupling point (CODP), which is often used to distinguish between ETO and other manufacturing approaches, namely make-to-order (MTO), assemble-to-order (ATO), and make-to-stock (MTS). As illustrated in Figure 25.1, the CODP separates the part of the supply chain that responds directly to customer demand from the part that relies on forecasts and speculation.

Due to the uncertainties in both the supply and demand perspectives of ETO, ETO manufacturers can neither produce to stock nor purchase materials based on forecast, which significantly pushes the CODP upstream, lengthening the supply lead time. Though MTO producers can also experience the problems associated with delivering high-variety products in low volume, this chapter primarily focuses on ETO environments, where lengthy purchasing, production, and delivery lead times are also accompanied by significant engineering lead times.

Where lean production has traditionally been very effective in eliminating variability through standardization of both product and process, in ETO the very core of the value proposition is based on designing and producing a customer-specific product, which inherently introduces a degree of variability into the production system. The result is that ETO producers struggle to improve their performance adequately using the standard Toyota Production System (TPS) lean toolbox.

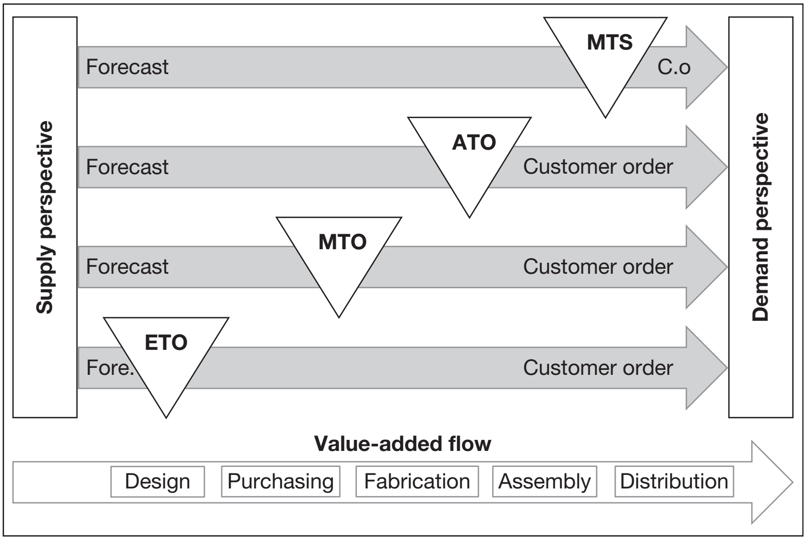

Figure 25.1 The CODP

Source: Adapted from Rudberg and Wikner (2004).

What is Lean ETO?

Taiichi Ohno, one of the chief architects of TPS, is often quoted as saying:

> All we are doing is looking at the time line, from the moment the customer gives us an order to the point when we collect the cash. And we are reducing the time line by reducing the non-value adding wastes.
>
> *(Ohno, 1988, p. ix)*

He goes on to say that one of the primary goals of TPS, or indeed the phenomenon we now know today as lean production, is to reduce the customer delivery lead time. Due to the inherently long lead times in ETO, the principles and practices of lean production have therefore recently become a very interesting concept for increasing the competitiveness of the sector, where the importance of shortening lead times is a stringent requirement for remaining competitive. However, Ohno (1988) also describes lean production as an alternative way of organizing mass production, suggesting that large volumes are intrinsically associated with the lean production paradigm, at least in the traditional sense. This leads to subsequent problems when attempting to apply lean production to ETO. For example, the basic principles of mass and flow production state that a) mass production demands mass consumption, and b) flow production requires continuity of demand (Woollard, 1954). ETO manufacturers exhibit neither of these traits; thus an alternative approach to "traditional" lean production is required for lean ETO.

Pavnaskar et al. (2003) suggest that a headlong rush into becoming lean has resulted in many misapplications of lean tools, often due to an inadequate understanding of them. After all, the traditional lean production tools were initially developed in order to pursue lean in a given context—the Toyota Production System. In the case of manufacturers in other types of industry, particularly in ETO, the relevance of the original toolbox must be reconsidered. Matt and Rauch

(2014) support this notion by suggesting that the suitability of certain lean methods, such as value stream mapping or kanban, is limited in ETO, while the more basic lean practices such as 5S and continuous improvement can indeed still generate significant improvement in any given situation. Though basic lean tools such as 5S are seemingly universally applicable, alternative methods are required to realize the greatest benefits from a lean transformation in ETO. This requires a fundamental re-examination of lean as both a management philosophy and a set of guiding principles, rather than simply attempting to make the existing tools fit in the new application area by adopting a "square peg, round hole" approach. Forcing a square peg into a round hole will only lead to inferior performance.

We suggest the following as a formal definition of lean ETO: *the adoption and deployment of lean principles in high-variety, low-volume manufacturing and project-based production environments*. We suggest that the ideal starting point for evaluating lean principles in the context of ETO producers is Womack and Jones' (1996) five lean principles: precisely specify *value* by specific product; identify the *value stream* for each product; make value *flow* without interruptions; let the customer *pull* value from the producer; and pursue *perfection*.

Value

One of the major reasons for the requirement of an alternative approach for lean in ETO is the way in which customer value is defined, with particular reference to the idea of variability. In traditional lean production, one of the goals is to eliminate all forms of variability through rigid standardization of product and process such that continuous flow production is achieved. This has often been accomplished through demand and production leveling, a concept known as "*heijunka*" (a Japanese term for "steady wave"). However, heijunka has itself been criticized as making lean production too inflexible and not applicable in more volatile markets, such as in ETO (Holweg, 2005). For example, where it is always desirable to reduce and eliminate dysfunctional variability in ETO, customer value propositions are built on the very idea of exploiting strategic variability. As such, customization and bespoke product offerings are recognized as a strategic advantage in ETO manufacturers; thus the elimination of this type of variability is not desirable.

Value Stream

The second lean principle in Womack and Jones (1996) is the identification of all value streams for products/product families. A value stream consists of all actions (both value added and non-value added) that are currently required to bring a product from raw material into the arms of the customer, from concept to launch (Rother and Shook, 1999). The logical starting point in identifying value streams is value stream mapping (VSM). VSM is a very useful tool, but, like many of the lean tools, should not be treated as a one-size-fits-all solution. For example, systems without a highly linear material and information flow are unlikely to benefit from this approach. ETO environments are dynamic environments in that they encounter many product variants, often with non-linear material flows with iterative cycles through similar process stages. This presents difficulties with regard to the use of VSM. The requirement to deliver one-of-a-kind products also makes VSM difficult to apply, due to several underlying circumstances (see also Alves et al., 2005):

- *Takt time*: The definition of *takt* time is contingent upon a certain and consistent demand. Thus, if demand is not known or constant, neither is the takt time.

- *Every part, every. . . (EPE)*: Again, where there is no constancy of demand, it is difficult to achieve a steady-state EPE. The actual loading of the production system as well as the demand for each product/product family typically varies from one period to the next in ETO.
- *Batch sizes*: Though batch sizes in ETO can be in the order of one, production orders in an ETO environment do have the likelihood of varying largely in terms of number of units (e.g. anything from one-offs upwards) and work content (from minutes to days). Therefore, the definition of a batch size that keeps a smooth flow of work and matches the takt time is a challenge in ETO.
- *Pacemaker*: The absence of a takt time and existence of complex production routings makes the implementation of a pacemaker hard to achieve in ETO.
- *FIFO*: Establishing FIFO (first-in-first-out) lanes is hard to achieve in ETO, for example where products must return to previous operations as part of the defined production routing. These items should not go to the back of the FIFO queue, as this simply lengthens the throughput time in the system.

As an alternative to VSM, swim lane diagrams are a visual management tool that can be used to visualize value streams in ETO and take into account non-linear and returning material and information flows.

Besides these challenges, ETO manufacturers must still attempt to segregate products based on similar characteristics as well as frequency of occurrence, for example, a “runners, repeaters and strangers” type classification. In this respect it may be possible to establish flow lines for the “runners” (the more regular, higher volume articles) and flexible cells for repeaters/strangers with similar characteristics. This type of classification often leads to a trade-off between a focus on the power of time (Suri, 2010) and increasing flow efficiency (Modig and Åhlström, 2012), and a focus on high capacity utilization through economies-of-scale thinking. Suri (2010) suggests that the huge impact time reduction can have on a manufacturing operation is unknown to most managers. Suri’s quick response manufacturing (QRM) concept focuses on reducing the lead times for all tasks in the entire enterprise through emphasizing the power of time. As an alternative to (or more precisely an enhancement of) value stream mapping, Suri suggests manufacturing critical-path time (MCT) mapping as a simplified tool for understanding high-variety, low-volume environments. MCT mapping distinguishes touch time (value-added time) from white space (non-value-added time), where “touch time typically accounts for less than 5% of total lead time” (Suri, 2010, p. 9).

Flow

A central tenet of lean production has been improving the flow of material and information throughout a product’s value stream. In their discussion of *flow efficiency*, Modig and Åhlström (2012) suggest three laws that should be adhered to in order to realize efficient flow:

1. *Little’s law*: Basic queueing theory states that a reduction in work-in-process (WIP), for example a reduction in the number of flow units in the system, will result in a reduction of throughput time, given a constant cycle time (Little and Graves, 2008). Hence the recommendation for reducing the number of flow units and enforcing a WIP cap on the system (Hopp and Spearman, 2004).
2. *Law of effect of utilization and variability on lead time*: Theory of queueing systems also states that the greater the variation in the process, the longer the resulting throughput time, especially

with a high level of resource utilization (e.g. Kingman, 1961). Thus the recommendation of planning to operate at a maximum of 80 percent or even 70 percent capacity in ETO environments, particularly on critical resources (Suri, 1998).

3 *Law of bottlenecks*: Basic queueing theory also states that the throughput time in a process is primarily affected by the stage in the process that has the longest cycle time, therefore the recommendation is to subordinate the non-bottleneck (non-constraint) resources in order to protect the bottleneck (Cox and Goldratt, 1984).

Central to flow thinking is the choice of buffer that an enterprise uses to manage day-to-day variations in demand. The choices available are time, capacity, and inventory, or indeed a mixture of all of these. In high-volume environments where economies of scale and a high level of capacity utilization is desired, buffers are often found in the form of inventories, e.g. semi-finished and finished goods. However, in an ETO situation, where it makes little sense to buffer against variation with inventory due to the customized nature of products, and where it makes little sense to add extra time buffers to already lengthy lead times, capacity buffers are the only real solution (see the second law, above). This means that the traditional focus on achieving 100 percent capacity utilization and keeping machines running as much as possible must be replaced with a focus on achieving the greatest flow efficiency through reducing batch sizes, WIP, and throughput time. Where flow efficiency becomes the new goal, production operations and those responsible for them must not be measured on resource utilization, at least in the immediate term. Spare capacity becomes essential in order to achieve a satisfactory level of flow in light of high product and process variation (Kingman, 1962).

Pull

Once the value stream has been identified and products begin to flow, Womack and Jones (1996) suggest that the next step is to let the customer pull value from the producer. This is achieved by deploying pull systems throughout the value stream. The fundamental difference between pull systems such as the kanban system (e.g. Sugimori et al., 1977) and push systems such as material requirements planning (MRP) is that push systems schedule releases, while pull systems authorize them (Hopp and Roof, 1998). Rother and Shook (1999) suggest that where continuous process flow cannot be achieved, pull systems should be established to help support and maintain the flow. However, the concept of supermarket-based pull systems raises some issues in the context of ETO. Powell and Arica (2014) suggest that although the term pull has become a cornerstone of modern manufacturing operations, there seem to be mixed views and interpretations of the pull concept across different contexts. Exploring the various interpretations, they go on to propose three context-dependent definitions of pull (Table 25.1).

In the case of ETO, one could say that production would always be in response to demand pull, as production occurs after the CODP. However, production pull, for example the traditional kanban system with its associated supermarkets and product-specific kanban cards as described in Sugimori et al. (1977), is difficult to apply in ETO due to the requirement of min-max inventories of standardized components, sub-assemblies, and finished goods (e.g. the two-bin system). On the other hand, plan pull would almost always be a suitable pull mechanism for production control in lean ETO, whereby the number of flow items (or items of work in process) are limited and a system of constraint management is used to control and reduce flow time, regardless of product type or configuration. *Constant work-in-process* (CONWIP) or *paired-cell overlapping loops of cards with authorization* (POLCA) (e.g. Spearman et al., 1990; Riezebos, 2010) are two common examples of this type of hybrid push-pull system. We also consider *Scrum*

Table 25.1 Context-dependent definitions of pull

Context	*Definition*
Demand pull	Value-adding activities only take place in response to real customer demand (however, production can still be either pull-based or push-based).
Production pull	Value-adding activities take place in response to a specific withdrawal from an explicitly limited inventory buffer or supermarket. The direction of information flow is the reverse direction of material flow, and production takes place in order to replenish an exact amount of consumed products and/or components.
Plan pull	Value-adding activities take place based on a priority rule such as earliest due date (EDD) and constraint management.

Source: Powell and Arica (2014).

(Sutherland, 2014) to fit this classification. And indeed, when we consider kanban in the literal sense (kanban in fact means "to look closely at the wooden board"), kanban "boards" are also a suitable workflow control mechanism (see, for example, Kniberg and Skarin, 2010). Therefore, in this chapter, we adopt the terminology kanban to mean visual board, and maintain that this is just one part of Toyota's kanban system, which also uses supermarkets and product-specific kanban cards as a material control system (as described in Sugimori et al., 1977).

Perfection

Continuous incremental improvement, or *kaizen*, is perhaps the most universally applicable of the original lean production principles in terms of the ways in which it is applied, for example through the use of what one might refer to as traditional lean production practices. If continuous process flow is the goal of the lean system, the involvement of everybody in daily continuous improvement activities and team-based problem solving is the substance by which this can be accomplished. Here, continuous experimentation and reflection is the key, through the adoption of a plan-do-check-act (PDCA) problem-solving approach. PDCA is of course just as applicable in one-of-a-kind ETO as it is in high-volume production environments. This approach to continuous improvement, otherwise known as the scientific method (Spear and Bowen, 1999), helps to create common understanding among all involved parties. The A3 process, as used by Toyota and other lean exemplars, is a supporting visual approach that helps strengthen the focus on continuous improvement throughout the enterprise.

Challenges and Opportunities of Lean ETO

The obvious challenge in lean ETO lies in creating a common understanding among employees in this sector that lean thinking applies just as much in ETO as it does in mass production ATO/MTS environments. The "we're different" mentality is a common obstacle for lean transformations in ETO. A major contributor to this issue has been the all-too-common focus on tools and techniques rather than adopting a focus on purpose, process, and people. For example, Shook (2014) suggests that starting with the question "*what is our value-driven purpose*?" allows for a situational approach to lean transformation, where every given situation is different, and in many cases the result will be a rather different set of countermeasures, tools, and techniques. There is much more to lean ETO than simply dwelling on an already established set of tools and techniques that have been made famous by the overwhelming interest in the practices employed by the Toyota Production System.

A copy-and-paste, cookie-cutter approach to lean transformation should be avoided at all costs. A simplistic view of transferring proven tools and techniques from one context to another often leads to a number of challenges and essentially failure of the lean transformation. One issue in particular is that of the kanban system. As we know, kanban systems, with supermarkets of standard parts and product-specific kanban cards, have been proven to work very well as the production control mechanisms in the context of high-volume production of similar products and with fairly stable demand, reducing work-in-process (WIP), smoothing production, and improving quality (Hopp and Spearman, 2004). However, there are three fundamental issues that prevent the application of the traditional kanban system as a suitable control mechanism in ETO manufacturing (Suri 2010):

1. *The kanban system's reliance on takt time*: Though the kanban system creates a balanced flow through the shop floor given stable demand and regular takt, high-variety, low-volume project-based production environments encounter a high level of instability in processes, particularly processing times and capacity utilization. This renders the traditional kanban system ineffective in ETO manufacturing.
2. *Kanban system as a pull system*: With a kanban system, parts are replenished—pulled through the factory—as they are consumed in downstream operations. However, given a high variety of products with low annual demand, it does not make a lot of sense to have supermarkets of finished goods waiting to be pulled by the customer and semi-finished goods waiting to be pulled from an upstream operation.
3. *Difficulties with completely custom/bespoke products*: Finally, kanban systems are unable to manage custom-engineered parts and components. The success of the kanban system is highly dependent on the use of standard components and parts.

Even though the traditional kanban system is not necessarily suitable for ETO environments, this doesn't mean that the pull principle that first motivated the development of such a simple workload control system cannot be adopted in ETO (see Powell et al., 2013b for a description of workload control: job entry, job release, priority dispatching/WIP control). After all, an ETO producer is always acting in response to customer demand (i.e. demand pull). In this respect, no value-adding activity is carried out based on speculation, as little or no engineering, procurement, production, or assembly takes place before a customer order is received. This makes it necessary for an ETO producer to pay close attention to the way in which materials flow through the value-adding activities that make up the value stream. The real benefits of adopting the pull principle in ETO can only be achieved when we realize that increased customer service levels and reduced throughput times are achieved by adopting a focus on flow and explicitly limiting the amount of work-in-process (WIP) that can be released into the system (Hopp and Spearman, 2004). Some ways in which the pull principle can be deployed in ETO include the implementation of alternative pull systems such as CONWIP, POLCA, and Cobacabana. We suggest that Scrum and kanban (in the form of production control boards) can also be applied for workload control in order to reduce throughput times and improve flow in ETO environments.

CONWIP

The CONWIP system has been described in some detail in Spearman et al. (1990). In general, CONWIP reduces throughput time by explicitly limiting WIP levels on the shop floor. This is achieved by limiting the number of jobs that are released into a system, i.e. by defining a constant (max.) level of WIP. Slomp et al. (2009) identify CONWIP as one of three elements (CONWIP, FIFO, and takt time control) for lean production control in make-to-order (MTO) job shops, where volumes are typically low and variety is high. Thus a CONWIP system of production control is a likely possibility for establishing pull also in ETO.

POLCA

Paired-cell overlapping loops of cards with authorization (POLCA) is a material control system designed for both make-to-order and engineer-to-order companies (Riezebos, 2010). It is the predominant production control system in quick response manufacturing (QRM). POLCA is a material control system that regulates the authorization of order progress on the shop floor in a cellular manufacturing environment (Powell et al., 2013b). As the name suggests, POLCA is dependent on the establishment of a cell-based layout as a prerequisite for its successful application. Again, by strictly limiting the amount of WIP between and in cells, POLCA aims to increase the flow efficiency of the system by speeding up job transfer and reducing imbalances in the system. Before work can begin on a production order, upstream capacity must be reserved through taking a POLCA card from each of the pair of cells next in sequence in the production routing. If either of the cards is not available, this is a sign that there is not sufficient available capacity at that particular cell, so the next job on the prioritized release list (backlog) for which there can be found an available pair of required cards should be released into production. Each time a production operation is complete, capacity must again be reserved at the next pair of cells, again by attaching the next available card to the work order/routing document. The previous card can be returned to the POLCA board to visualize available capacity at the respective operation. A detailed description of designing a POLCA system can be found in Riezebos (2010).

Cobacabana

As an alternative to both CONWIP and POLCA, *Control of balance by card-based navigation* (Cobacabana) is a simple card-based pull system for job-shop control (Land, 2009). Land suggests that the Cobacabana system has been specifically organized around the order acceptance and order release decisions in job-shop control. The system uses card loops consisting of color-coded cards between the planner performing the releases and all critical workstations (see Land, 2009 for a more detailed description). In essence, available cards authorize the planner to release a new order to a specific work center. The planner attaches the cards to a work order, and, on completion of the work, the cards are returned to the planner from the workstation to signal additional available capacity at the station. Each card represents a certain percentage of available capacity at a workstation, with the total number of cards adding up to 100 percent of a station's available capacity. Thus each order will normally require multiple cards.

Scrum and Kanban

Scrum was developed as an alternative to using the Waterfall method (Gantt charts) for managing projects, particularly in software development (Sutherland, 2014). It is often considered as an agile software development approach. This also goes for kanban, which in the agile development world is primarily defined in terms of its visual representation of work. When we consider the two Japanese characters that make up the word kanban, the visual nature of this approach becomes highly apparent (Figure 25.2).

The first character—Kan—is made up of the symbols for *hand* and *eye*. It represents a man shielding his brow in order to see clearly, and means "*to look at closely*." The second character—

Figure 25.2 Japanese characters for kanban

Ban—is made of the symbols for *tree*, *wood*, and *wall*. It represents a wooden board leaning against a wall, and literally means "*wooden board*." In essence, then, kanban means "*to look at the wooden board closely*." As such, kanban can simply be translated as board, and provides an effective form of visual management for development activities, project management, and production control.

Both Scrum and kanban use visual boards to show available work, work-in-process, and completed work (often under the headings "To-do," "Doing," and "Done"). Scrum and kanban encourage and promote team-based problem solving throughout the completion of prioritized tasks, and can be effectively adopted in both software and hardware development, as well as in project-based production environments, such as in ETO manufacturing.

In Scrum, tasks are selected from a prioritized backlog and posted under "To-do." The tasks that are selected are expected to be completed by the Scrum team during the *Sprint* that follows, which typically takes place over the next seven days to four weeks (Sprint duration is normally constant from one sprint to the next). Each day during the Sprint, all team members gather around the Scrum board in what's known as a stand-up meeting (typically maximum 15 minutes long). During the stand-up, the participants discuss what was achieved on the previous day, the plan for today, and the problems that were addressed/remain to be addressed such that the Sprint can be completed on time. On the first stand-up meeting, individuals will select and move "To-do" items to "Doing." Upon completion, these jobs will be moved to "Done." On the last day of the Sprint, team members should reflect on the Sprint in order to improve during the subsequent Sprint. In essence, Scrum constitutes a pull mechanism as the quantity of tasks that is released into each Sprint remains fixed for the duration of the Sprint. The quantity of tasks released is based on the amount of required capacity that is anticipated to complete each task during the Sprint. No further tasks will be released into the current Sprint unless there is a plausible exception, and only then by approval of the product owner.

Kanban is very similar to Scrum in that a visual board and daily stand-up team meetings are the mechanisms for success. However, the major difference is that a kanban board explicitly limits the amount of tasks that can be assigned to each stage of the operation. In the case of kanban, "Doing" is often segregated into separate operations, each with a visually defined limit of tasks that can be processed simultaneously in that operation. Tasks are still selected from a prioritized release list, as in Scrum (and indeed POLCA and CONWIP). Yet the fundamental difference between kanban and Scrum is that kanban can be more reactive to changes in requirements. This is due to the fact that tasks are not locked in for the duration of the Sprint (e.g. 7 to 28 days); the response time of a kanban team is the time it takes for capacity to become available after having completed a current task in process. This follows the general principle of one-out-one-in, as in CONWIP.

For example, in Figure 25.3, all items selected in the current iteration or Sprint (A–H) are locked for the duration of the Sprint, and must be completed before the next set of new jobs from the release list/backlog can be released. On the other hand, in Figure 25.4, as soon as one of the jobs C or D is completed, capacity becomes available to process one of the "To-do" jobs E–H, or, in kanban, an alternative job from the backlog replaces any one of the jobs E–H and subsequently becomes the next in the queue.

Besides the more common challenge of overcoming the "we are different" perception of ETO manufacturing personnel and the issues that result due to misplaced attempts to simply transfer lean tools and techniques from high-volume, low-variety production environments into ETO manufacturing, there are some more general challenges that should be highlighted when it comes to applying lean in ETO. These include:

- how to address the complex information flows and frequent delays (e.g. during the request-for-quotation (RFQ) phase right up to order confirmation),
- how to simplify the complex and unpredictable flow of materials,
- how to secure the supply of long-lead items,
- how to overcome frequent engineering change orders (ECOs), and
- how to prevent penalties for late delivery.

Backlog		To-do	Doing	Done
I	J	E	C	A
K	L	F	D	B
M	N	G		
O	P	H		

Figure 25.3 Scrum board (jobs A–H will be completed during current Sprint)

Backlog		To-do (4)	Doing (2)	Done
I	J	E	C	A
K	L	F	D	B
M	N	G		
O	P	H		

Figure 25.4 Kanban (notice maximum levels 4 (To-do) and 2 (Doing) respectively)

We suggest that such challenges can be addressed through the application of concepts such as "quick response office cell" (QROC), cellular manufacturing, pull production, and Scrum.

For example, QROCs are formed to cut through functional boundaries such that information flows can be simplified and accelerated. A QROC (pronounced queue-rock) is a closed loop, co-located, dedicated multifunctional, cross-trained team that is responsible for the office processing of all jobs belonging to a particular product family or value stream (Suri, 2010). QROCs use the same principles as those applied in the cellular manufacturing approach that is used to simplify material flows on the shop floor, only here we drastically simplify information flows in the supporting administrative tasks in addition. QROC would be a potential solution to the issue raised by Hicks et al. (2000), who suggest that the variety of work that is involved in ETO projects, e.g. designing and producing customized products while simultaneously addressing the underlying uncertainties of markets, indicates that procurement and marketing must be integrated with other processes, particularly tendering and design.

In an attempt to manage the supply of long-lead items, an ETO producer may well have to hold a minimum level of safety stock of these items in order to eliminate such lengthy procurement lead times. This of course encounters a cost that must be considered along with the cost of holding the inventory, in particular the cost of risk of obsolescence, seeing as we are perhaps dealing with one-of-a-kind products. However, as the success of lean ETO really is about realizing the power of time, there are few alternative options other than holding a small supermarket for long-lead items, if supplier agreements cannot be put in place to reduce these lead times in the first place. We suggest that best-practice supplier development and supplier collaboration are fundamental to addressing the problem of long-lead items in lean ETO.

A satisfactory system for managing ECOs must also be established in those ETO environments where regular, often unexpected, ECOs are the norm. A satisfactory way of following up and closing out ECOs could be the adoption of Scrum, for example, with regular week-long Sprints to address and deliver effective engineering changes through regular collaboration with the customer.

The Future of Lean ETO

The complexities associated with ETO make it very much a form of *advanced manufacturing*. Advanced manufacturing is not a static entity. Washington (2015) suggests that what was once considered advanced decades ago has now become traditional, and what is advanced today will be considered mainstream in the future. While the traditional production of high-volume, low-variety products with low profit margins may well have migrated to low-cost countries, the knowledge-based engineering and technological development involved in delivering customer-specific products and solutions to turbulent markets can be considered as the future of competitive manufacturing in high-cost regions. Most simply, Washington (2015) states that advanced manufacturing is about innovating products and processes in the manufacturing cycle to operate more productively. The future of ETO will therefore include the complete set of activities from the engineering and design stages all the way through production and assembly processes to the development of innovative after-sales services. We suggest that this makes lean ETO a strategic factor for the success of the future of manufacturing in high-cost countries. Furthermore, we also suggest that the real triumph of a lean transformation in ETO will depend on the adoption of ETO-specific tools and techniques in order to successfully deploy the fundamental lean principles. Powell et al. (2013a) suggest that a lean implementation process typically consists of three phases: a basic lean phase, an advanced lean phase, and a continuous improvement phase. Such a view on lean implementation is perhaps just as applicable in ETO manufacturing as it is in mass production. First, the basic lean phase involves the application of the more universally applicable lean practices such as 5S workplace organization and visual management. This prepares the workplace for what is to come in the second phase—the advanced phase. In this phase, ETO-specific lean tools are deployed, such as the alternative pull systems discussed previously. Finally, for the results to be sustainable, the third phase is continuous improvement, involving all personnel in ongoing incremental improvement activities, every day.

A final thought comes from Rajan Suri (2011, p. 3), who encourages us to realize that success in this area may have major impact on the way we look at our current business models:

> Many employees in developed countries live in fear of their operation being outsourced to low-wage countries, such as China. However, for a typical product made in a developed nation, direct labor accounts for only 10% of its cost. Moreover, in terms of the selling price of a product, the number is lower: less than 7% of the price to the customer is attributable to direct labor. Thus, if you use QRM methods to reduce cost by 25%, you wipe out the labor-cost advantage of low-wage countries. When you consider that overseas competitors need considerable lead-time for shipping, your short response time makes it impossible for them to compete on the same terms. You can compete against anyone, making products anywhere.

Case Study: Lean Engineer-to-order in Bosch Hinges and Metal

Bosch Hinges and Metal designs and manufactures high-quality metal bespoke hinges for industrial applications. At the start of 2004, the company was challenged with the need to adapt its strategy in order to survive the onset of increased competition. At that time, in light of low-cost competition, the company was in fact operating at loss. Today, with its 30 staff in the east of the Netherlands, Bosch Hinges and Metal is an example of a very successful lean engineer-to-order (ETO) manufacturer. The company has over 4,000 orders for approximately 600 different customers annually, with batch sizes ranging from 1 to 1,000 pieces and product sizes from just a few centimeters to 4 meters in length. On average, a production operator at Bosch Hinges and Metal (otherwise known as a hinge maker) is working on one particular order no longer than 30 minutes before starting on the next. Some of the operations needed to produce the company's customer-specific hinges include laser cutting, punching, sawing, rolling, folding, carving, bending, drilling, welding, grinding, and brushing. Products flow through these operations in many different routings. But even before production can begin, 10 employees in the office at Bosch Hinges and Metal carry out tasks for quoting (acquisition, calculation, and quotation), engineering, order preparation, and production and material planning.

Challenge

The primary reason for a change of strategy in 2004 was the fact that Bosch Hinges and Metal was confronted with a market that was demanding ever-smaller series and shorter lead times. This had led the company into the vicious circle of prioritizing and re-prioritizing "rush" orders, which led to an increased order backlog, increased work-in-process, and increased lead times. More and more new rush orders simply escalated this problem. A typical quoted delivery lead time in 2004 was six weeks, with actual delivery lead times of around eight weeks or more. At that time, planning, scheduling, and rescheduling was a time-consuming effort. The company typically used a one-week planning horizon, with detailed day-to-day planning and a separate backlog order planning. This demanded the full attention of one full-time employee. Delivery performance was poor and customer satisfaction suffered, which caused the company to lose some of its customers.

Journey

Bosch Hinges and Metal first attempted to apply traditional lean techniques but, like many ETO companies, failed to realize the expected results. Despite some initial success with 5S and single-minute exchange of dies (SMED), the company kept struggling with vast amounts of backorders, high amounts of WIP, and poor delivery performance (around 60 percent on-time delivery). It was not until 2007, when Godfried Kaanen, director and owner of the company, heard of Rajan Suri's work on *quick response manufacturing* (QRM), that he realized the importance of lead time reduction. From this moment, the company managed to realize breakthrough improvements. Examples of these include a reduction in delivery lead time from eight weeks to three weeks and an increased delivery performance from 60 percent to around 90 percent. Production planning, which was previously a full-time job, now takes no more than 1.5 hours per day. Quotation lead time was improved from 3–5 days to same-day delivery. Together, these improvements transformed the company from being unprofitable to turning a healthy profit. In the remainder of this case study, some of the specific practices that helped transform Bosch Hinges and Metal into a lean ETO exemplar are described.

Focused Sales

On a strategic level, Bosch Hinges and Metal adopted a new business model. Realizing that its primary customer value proposition was flexibility to offer unique hinges fast, the company began to focus on the delivery of bespoke hinge solutions. This meant refusing the occasional large order, as large orders lead to substantial disruption in the flow, consume large amounts of capacity, and offer relatively low margins compared with specialized, low-volume orders.

POLCA Cells in Production

Originally, the shop floor at Bosch Hinges and Metal was functionally organized in a traditional job-shop type layout. This was also the case in the office. MRP was used for production planning and control, which led to high WIP, low visualization, poor quality control, long lead times, and a large backlog of orders and the occurrence of many rush orders. In order to regain control of its production processes and reduce throughput times, the company implemented a POLCA system to improve and control order release and material flow. With the introduction of POLCA the shop floor was transformed into just a few production cells, which were also color coded (within the cells, floor marking, equipment, and operators' overalls correspond to the cell color). After an analysis of machine capacity, investments had to be made for spare capacity on some of the more highly utilized machines. The POLCA system creates an improved material flow by using POLCA cards to authorize production, in combination with a high-level MRP release list. The high-level MRP system back-schedules using fixed throughput times per cell to calculate the earliest release dates. The release list indicates the earliest allowed start dates of all orders in queue and prevents overproduction. The POLCA cards limit overall WIP and prioritize orders based on the next operation with the least work in the queue. Within-cell planning is carried out autonomously by the cell operators, using the predetermined conditions of the POLCA system. The prioritization of orders and strictly controlled WIP limits, together with the strategic choice to have spare capacity, creates a robust system with stable and predictable throughput times.

Quick Response Office Cells

Once the main problems in production were controlled and corrected, further opportunities for the reduction of lead times were found in the office processes. Three *quick response office cells* (QROCs) were formed in the office environment. The "commercial cell" consists of four full-time employees (FTEs) and is responsible for acquisition, calculation, and quotation. As a rule, potential customers now receive their quotation within 24 hours after first inquiry. The "engineering cell" (3 FTEs) is responsible for all required engineering, for order preparation, and for customer confirmation (which is needed before production is allowed to start). The third QROC is called "general affairs" (3.5 FTEs) and is responsible for daily management, administration, and human resources. Employees within the cells are all cross-trained and jointly responsible for continuously improving and reducing total lead time.

Future Outlook/Further Improvements

While the introduction of the POLCA system led to the needed breakthrough improvements, the company struggled with the day-to-day handling of the system. Additionally, since the introduction of the POLCA system the company grew, leading to larger distances between the cells. This hampered the effective use of the traditional POLCA card system. Also the physical nature of the cards lacked the capability to monitor the shop floor status in real-time. To overcome these issues, Bosch Hinges and Metal engaged the services of an external software development entrepreneur and developed the *production and POLCA operating system* (PROPOS), a company-specific manufacturing execution system (MES). In PROPOS, the physical cards and POLCA boards are replaced by touch screens at all cells. The PROPOS system shows the release list in order of earliest starting date and shows the availability of materials and cards for the next cells. The system gives the operators insight into which order to start first, what is to come, and what the current lead time performance is. To the engineering cell, PROPOS provides real-time insight into production flow, and highlights where issues may arise such that they can be anticipated as early as possible.

Godfried Kaanen, Director and Owner of Bosch Hinges and Metal, concludes that despite the huge success of the implementation so far, the company has only just begun—"In order to stay in front of the competition, we must understand that today's lead times are still too long for tomorrow." Time is on their side though; a relentless focus on lead time reduction gives Bosch Hinges and Metal a bright future in lean ETO.

References

Alves, T. C. L., Tommelein, I. D. and Ballard, G. (2005). Value stream mapping for make-to-order products in a job shop environment. *Construction Research Congress 2005*, San Diego, CA.

Cox, J. and Goldratt, E. M. (1984). *The Goal: A Process of Ongoing Improvement*, Croton-on-Hudson, NY, North River Press.

Hicks, C., McGovern, T. and Earl, C. F. (2000). Supply chain management: A strategic issue in engineer to order manufacturing. *International Journal of Production Economics*, *65*(2), 179–190.

Holweg, M. (2005). The three dimensions of responsiveness. *International Journal of Operations and Production Management*, *25*, 7–20.

Hopp, W. J. and Roof, M. L. (1998). Setting WIP levels with statistical throughput control (STC) in CONWIP production lines, *International Journal of Production Research*, *36*(4), 867–882.

Hopp, W. J. and Spearman, M. L (2004). To pull or not to pull: What is the question? *Manufacturing and Service Operations Management*, *6*(2), 133–148.

Kingman, J. (1961). The single server queue in heavy traffic. *Mathematical Proceedings of the Cambridge Philosophical Society*, *57*(4), 902.

Kingman, J. F. C. (1962) On queues in heavy traffic. *Journal of the Royal Statistical Society, Series B (Methodological)*, *24*(2), 383–392.

Kniberg, H. and Skarin, M. (2010). *Kanban and Scrum: Making the Most out of Both*, C4Media.

Land, M. J. (2009). Cobacabana (control of balance by card-based navigation): A card-based system for job shop control. *International Journal of Production Economics*, *117*(1), 97–103.

Little, J. D. C. and Graves, S. C. (2008). Little's law. In: Chhajed, D. and Lowe, T. J. (eds) *Building Intuition: Insights from Basic Operations Management Models and Principles*, Springer Science+Business Media.

Matt, D. T. and Rauch, E. (2014). Implementing lean in engineer-to-order manufacturing: Experiences from an ETO manufacturer. In: Modrak, V. and Semanco, P. (eds), *Design and Management of Lean Production Systems*, Hershey, PA, IGI Global, pp. 148–172.

Modig, N. and Åhlström, P. (2012). *This is Lean: Resolving the Efficiency Paradox*, Stockholm, Rhealogica Publishing.

Ohno, T. (1988). *Toyota Production System: Beyond Large-Scale Production*, Boca Raton, FL, CRC Press.

Pavnaskar, S., Gershenson, J. and Jambekar, A. (2003). A classification scheme for lean manufacturing tools. *International Journal of Production Research*, *41*, 13–15.

Powell, D. J. and Arica, E. (2014). To pull or not to pull: A concept lost in translation? *American Journal of Management*, *15*(2), 64–73.

Powell, D. J., Alfnes, E., Dreyer, H. C. and Strandhagen, J. O. (2013a). The concurrent application of lean production and ERP: Towards an ERP-based lean implementation process. *Computers in Industry*, *64*(3), 324–335.

Powell, D. J., Riezebos, J. and Strandhagen, J. O. (2013b). Lean production and ERP systems in small- and medium-sized enterprises: ERP support for pull production. *International Journal of Production Research*, *51*(2), 395–409.

Riezebos, J. (2010). Design of POLCA material control systems. *International Journal of Production Research*, *48*(5), 1455–1477.

Rother, M. and Shook, J. (1999). *Learning to See: Value Stream Mapping to Add Value and Eliminate MUDA*. Cambridge, MA, Lean Enterprise Institute.

Rudberg, M. and Wikner, J. (2004). Mass customization in terms of the customer order decoupling point. *Production Planning and Control*, *15*, 445–458.

Shook, J. (2014). *Transforming Transformation*. Available at: www.lean.org/shook/DisplayObject.cfm?o=2533 (accessed February 2016).

Slomp, J., Bokhorst, J. A. and Germs, R. (2009). A lean production control system for high-variety/low-volume environments: A case study implementation. *Production Planning and Control*, *20*(7), 586–595.

Spear, S. J. and Bowen, H. K. (1999). Decoding the DNA of the Toyota Production System. *Harvard Business Review*, 77, 96–108.

Spearman, M. L., Woodruff, D. L. and Hopp, W. J. (1990). CONWIP: A pull alternative to kanban. *International Journal of Production Research*, *28*(5), 879–894.

Sugimori, Y., Kusunoki, K., Cho, F. and Uchikawa, S. (1977), Toyota Production System and kanban system: Materialization of just-in-time and respect-for-human system. *International Journal of Production Research*, *15*(6), 553–564.

Suri, R. (1998). *Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times*, Boca Raton, FL, CRC Press.

Suri, R. (2010). *It's About Time: The Competitive Advantage of Quick Response Manufacturing*, Boca Raton, FL, CRC Press.

Suri, R. (2011). *Beyond Lean: It's About Time!* Available at: http://quickresponse-enterprise.com/wp-content/uploads/2013/10/Beyond-Lean-it-is-about-time-eng.pdf (accessed February 2016).

Sutherland, J. (2014). *SCRUM: The Art of Doing Twice the Work in Half the Time*, London, Random House.

Washington, L. (2015). The importance of advanced manufacturing to the changing global economy. Available at: www.cincom.com/blog/advanced-manufacturing/the-importance-of-advanced-manufacturing-to-the-changing-global-economy/ (accessed July 2016).

Womack, J. P. and Jones, D. T. (1996). *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, New York, Simon & Schuster.

Woollard, F. G. (1954). *Principles of Mass and Flow Production*, London, Iliffe and Sons.