Untold Essence of the Toyota Production System (TPS)

Dr. Ushio Sumita

Productivity Insights

Vol. 4–1

Asian Productivity Organization
The Asian Productivity Organization (APO) is an intergovernmental organization that promotes productivity as a key enabler for socioeconomic development and organizational and enterprise growth. It promotes productivity improvement tools, techniques, and methodologies; supports the National Productivity Organizations of its members; conducts research on productivity trends; and disseminates productivity information, analyses, and data. The APO was established in 1961 and comprises 21 members.

APO Members
Bangladesh, Cambodia, Republic of China, Fiji, Hong Kong, India, Indonesia, Islamic Republic of Iran, Japan, Republic of Korea, Lao PDR, Malaysia, Mongolia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Turkiye, and Vietnam.
Untold Essence of the Toyota Production System
The P-Insights, short for “Productivity Insights,” is an extension of the Productivity Talk (P-Talk) series, which is a flagship program under the APO Secretariat’s digital information initiative. Born out of both necessity and creativity under the prolonged COVID-19 pandemic, the interactive, livestreamed P-Talks bring practitioners, experts, policymakers, and ordinary citizens from all walks of life with a passion for productivity to share their experience, views, and practical tips on productivity improvement.

With speakers from every corner of the world, the P-Talks effectively convey productivity information to APO member countries and beyond. However, it was recognized that many of the P-Talk speakers had much more to offer beyond the 60-minute presentations and Q&A sessions that are the hallmarks of the series. To take full advantage of their broad knowledge and expertise, some were invited to elaborate on their P-Talks, resulting in this publication. It is hoped that the P-Insights will give readers a deeper understanding of the practices and applications of productivity as they are evolving during the pandemic and being adapted to meet different needs in the anticipated new normal.
INTRODUCTION

The Toyota Production System (TPS) is symbolized by the slogan “Produce only needed things at exactly needed times in exactly needed quantities” and expressions such as “kanban” and “just-in-time (JIT).” After the first oil crisis, many corporations outside the automobile industry adopted the TPS, first in the apparel sector and then in the electronics and other sectors. The power of the TPS expanded during the 1980s, led by a research project organized by MIT to study the lean production approach, soliciting a variety of contributions from scholars and practitioners from all over the world. As a result of that research project, the book entitled The Machine That Changed the World was published [1], which firmly established the reputation of the TPS in manufacturing.

Most of the TPS literature has focused on the structure and mechanisms of lean production in manufacturing environments. Naturally, such studies have been described by terms used in manufacturing. While conclusions derived from the literature could help many manufacturing corporations, it may be difficult for corporations outside that sector to grasp the relevance of the TPS to their own businesses. The purpose of this report is to show that the TPS contains hidden management principles, which explain not only why it has been widely and successfully implemented in manufacturing but also how one may adopt them in the more general context of corporate management well beyond manufacturing. More specifically, the untold essence of the TPS is discussed from a new perspective by focusing on the analysis of organizational functionalities achieved by it. Since the organizational functionalities can be transferrable from one industry to others more easily, this approach provides a better vantage point for understanding what can be learned from the TPS beyond manufacturing.
What Is Missing in the TPS Literature: Importance of Functional Analysis, Not Mechanisms, of the TPS

The value chain in manufacturing consists of all business processes along the line of procurement/outsourcing, inbound logistics, production/inter-logistics, outbound logistics, sales, aftersales services, and reverse logistics. The TPS aims at eliminating all kinds of waste from the value chain since they propagate from one to other types. An excess of production, for example, results in excess inventory, which increases the inventory cost as well as the inventory management cost and requires expanded inventory space. Abolishing the chain effect of this sort throughout the value chain is the essence of the TPS.

The slogan “Produce only needed things at exactly needed times in exactly needed quantities” can be easily extended to other industries in manufacturing beyond the automotive sector, as actually happened. However, the question of how far the slogan can be stretched beyond manufacturing remains unanswered. Would the slogan enable law firms, accounting firms, and the like to devise actual implementation plans to enhance their performance? This task will remain difficult as long as only the mechanical aspects of the TPS are focused on. To answer this question, it is necessary to examine the organizational functionalities achieved by the TPS, which can be transferable more easily from one industry to others.

Material Requirement Planning and the TPS: Push vs. Pull System

It is useful to briefly review the traditional manufacturing management approach called material requirement planning (MRP) and to highlight its
differences compared with the TPS. MRP is a production planning, scheduling, and inventory control system used to manage manufacturing processes. Based on the demand estimate and bills of materials (BOM) for the underlying product, it specifies what items are required in what quantity and when, and then schedules the entire production process by cascading them back from the delivery of the final products to the procurement of necessary materials and receiving outsourced, partially processed components. Once the schedule is determined, however, both information and materials flow from the upper stream to the lower stream. In this sense, this approach is called a “push system,” as depicted in Figure 1.

In contrast, the TPS determines only the timing and the speed of flow of the underlying products into the market based on the demand estimate and confirmed orders from the network of dealers as they occur. Once this outflow of the product into the market is determined, each work unit informs all adjacent upper stream work units of when and in what quantity to produce and deliver their partially processed components to the work unit through the use of kanban. Accordingly, information flows from the lower stream to the upper stream, while the materials flow from the upper stream to the lower stream. Because of this, the TPS is called a “pull system,” as shown in Figure 2.

It should be noted that this control mechanism of the TPS is applied to the entire value chain including outside suppliers. The core of Toyota’s strength is
that the same disciplines surrounding the TPS are shared across the global network of value chains involving R&D, procurement, production, all logistics, sales, and after-sales service.

**Organizational Functionality Achieved by the TPS: Incidental Problems vs. Institutional Problems**

One of the key concepts of the TPS, which is the most difficult to explain to non-Japanese, is “genchi-genbutsu.” The concept can be translated into English as “Any problem should be dealt with at the site of the problem, not through deskbound discussions.” This translation, of course, can be readily understood by businesspeople. However, the hidden implications of the concept go far beyond those words. This section reflects the concept of genchi-genbutsu in a more general context from the viewpoint of “incidental problems vs. institutional problems.”

Incidental problems result from accidental reasons, such as careless mistakes, sudden power failures, and natural disasters, to name only a few. On the other hand, institutional problems are caused systemically, such as how human, material, and monetary resources are allocated; how information is controlled; what type of decision structure is employed; etc. Institutional problems are peculiar in that they can occur repeatedly unless the causes are identified and removed, although the magnitude of the resulting effects of the problems may differ depending on the persons involved.
When a problem occurs, it is necessary to identify the cause, sometimes based on hypotheses, to deal with it. Accordingly, one has to make a decision on whether the problem is incidental or institutional, possibly without realizing it. Here, there exist two types of risks similar to those found in statistical hypothesis testing (Figure 3). If an incidental problem is judged to be institutional, such a mistake is called a type I risk; judging an institutional problem to be an incidental problem corresponds to a type II risk.

If an incidental problem is judged to be an institutional problem, it may seem that the time and energy spent in trying to detect systematic reasons for the problem are in vain. However, this review process provides an opportunity to improve management quality and the efforts are not totally wasted. On the other hand, if an institutional problem is considered as an incidental problem, the opportunity of finding systemic causes of the problem would be lost and the problem might occur repeatedly in the future. It is now clear that type II risk is much more hazardous to the organization than type I risk.

From this perspective, genchi-genbutsu can be interpreted in the following manner.

1. Every problem should be considered as an institutional problem, at least at the beginning of dealing with it. In other words, take type I risk if you have to, but make sure to avoid type II risk.
2. Detect and remove institutional causes at the site of the problem, not through deskbound discussions.

As an example of how genchi-genbutsu may be employed, consider the case of an imaginary Corporation X, which contracts a third-party logistics company to transport new cars from a factory in Kentucky to the west coast of the USA. The train system was used for such transportation by loading automobiles onto double-decker train cars before about 15% of new autos were damaged during shipment.

If this problem were considered to be incidental, the logistics company would simply be blamed for the damage. The contract would be terminated, and the matter would be taken to court to recover the monetary losses. A new contract would be made with a new logistics company. However, the causes of the problem would be hidden through this approach, and the same problem could occur repeatedly even with a new logistics company.

As soon as Corporation X heard the news, it decided to detect systemic reasons involved and dispatched a team of six logistics experts from the headquarters in Japan to the factory in Kentucky. The team observed the logistics operations at key points between Kentucky and the west coast for two months and summarized kaizen ideas to avoid any damage to new automobiles during the logistics operations. A summary report was given and explained to the logistics company free of charge. The contract was continued without any penalties, and kaizen ideas were implemented in close collaboration between Corporation X and the logistics company. Through this approach, the causes of the problem were eliminated, and the same problem would not recur. Naturally, the loyalty of the logistics company to Corporation X would be enhanced substantially.

Organizational Functionality Achieved by the TPS (2): Production at the Speed of Sales in the Market

One of the most important wastes to be eliminated in the TPS is inventory cost. It is an opportunity cost in that the corporate capital resting in inventory without producing any profit might otherwise be used for different purposes with positive returns. Since the most valuable inventory item would be the final product, its inventory cost is the highest and the TPS begins with its minimization. This objective can be achieved by producing at the speed of sales on the market.
Figure 4 illustrates the concept of production at the speed of sales on the market. Here, the blue, pink, and green arrows represent the speed of sales on the market for three types of cars, A, B, and C, respectively (for example, 8 cars per minute for A, 4 cars per minute for B, and 2 cars per minute for C). Assuming 8 working hours per day, this means that 3,840 A cars, 1,920 B cars, and 960 C cars must be produced each day. If production takes place with large lot sizes of 3,840, 1,920, and 960 for types A, B, and C, respectively, the final product inventories are necessary to match the sales speeds of cars A, B, and C appropriately. In contrast, in the TPS, the production speed is synchronized with the sales speed for each type with smaller lot sizes of 8, 4, and 2 for A, B, and C, respectively. The three types of cars are then shipped from the production facility in accordance with the speeds of sales of A, B, and C. Hence, theoretically, it would not be necessary to keep the inventory of the completed cars. In reality, of course, the transportation times would fluctuate, and some “safety (reserve) stock” would be needed for ensuring customer satisfaction. In the remainder of this report, however, we conceptually assume the simple idealistic situation so that the philosophy of the TPS is communicated clearly.

Organizational Functionality Achieved by the TPS (3): Conversion of Every Segment of the Production Process into Part of a Critical Path

The critical path method (CPM) is a project modeling technique, originally developed by Kelley and Walker (1959) [2]. In brief, activities are on a critical path, and if any activity on the path is delayed by Δ, then the completion of the
entire project would be also delayed by $\Delta$ (see [3] for further details). One possible interpretation of the TPS may be: “Convert every segment of the production process into a part of a critical path by removing extra resources whenever possible.”

As discussed in the section above, the TPS synchronizes the speed of production with the speed of sales on the market for each product by producing a variety of products in small lot sizes through the same final assembly line. In order to sustain production operations on the final assembly line in this manner, necessary parts and materials are fed into it based on the JIT approach, that is, only the needed things at exactly the needed times in exactly the needed quantities are delivered to predetermined spots on the final assembly line by exchanging kanban information. The effects of these processes cascade back upstream including outside suppliers.

The spirit of “Convert every segment of production process into a part of a critical path” means that any excess resources ought to be identified and removed from every production process, provided that the flows of materials toward the final assembly line are sustained intact and sound working environments are maintained. If a work unit could reduce the number of kanban actions employed without altering the output flow of the work unit, it would do so, thereby reducing the work-in-progress (WIP) inventory there. If a work unit could find a way to sustain the same number of kanban actions with one less person while keeping the output flow intact, it would do so, and the removed person would be reassigned to a busier work unit, possibly in another factory.

Traditionally, sustaining stable throughput was the most important goal in manufacturing process management. Following the principles in reliability theory, a back-up machine would be prepared for an often-breakable production machine. If demand increased suddenly, substantial safety stocks of parts and final products would be kept to avoid stockouts. Against this traditional approach, the TPS brought about a Copernican revolution in manufacturing process management, where the traditional principles were disregarded as sources of generating unnecessary waste. Admittedly, the TPS might be fragile and vulnerable in that one breakdown in one production segment could stop the entire production process. This vulnerability has been overcome by developing excellent human resources and various management devices, thereby
establishing a production system that achieves the minimum cost and rarely stops even when every production path is set on a critical path.

**Organizational Functionality Achieved by the TPS (4): Explicit Expression of System State through Daily Operations**

In manufacturing process management, one encounters a variety of different problems, such as deterioration of quality, low yield rates, machine breakdowns, stockouts of materials and parts, etc. Such problems may occur in different work units, and it is often not easy to detect the sources of the problems, demanding many hours of meetings and substantial analytical efforts across different business sections. Accordingly, many problems cannot be dealt with promptly.

In the TPS, by making every segment of the production process a critical path, manufacturing operations are organized in such a way that daily production activities themselves provide explicit expressions of system state. For example, a problem of one work unit would propagate to other work units through stagnation of kanban operations and the location of the problem can be identified immediately. Then the causes of the problem would be investigated, even when the entire production line might be stopped. Removing the cause of a problem now at the sacrifice of short-run productivity results in better long-run productivity by avoiding repeated occurrences of the problem.

Sustained by well-trained management teams, foremen, and line workers who rigorously comply with operational and engineering standards, the TPS has converted the potential vulnerability of very thinly allocated resources so that locations and causes of problems can be identified easily. This in turn enables dealing with them promptly by removing the causes.

**Organizational Functionality Achieved by the TPS (5): Direct-sum Decomposition of Production Processes**

By converting every segment of the production process into part of a critical path, the TPS decomposes the entire production process into a set of mutually exclusive and exhaustive segments. Such decomposition is called a “direct-sum decomposition.” Following the example given in section 2.3, Figure 5 illustrates Work Unit I (WU-I), which has two subassembly segments I-1 and
I-2, as indicated by three yellow circles. At the end of WU-1, partially processed parts for three types of cars A, B, and C in lot sizes of 8, 4, and 2, respectively, are delivered to the final assembly line, where it is assumed that each car requires one part.

Here, the two subassembly segments I-1 and I-2 are independent of each other in that one is not affected by how the other operates, e.g., what types of production machines are used, how many numbers of kanban actions are circulated, how many workers are involved, etc., as long as both keep up with their output flows in accordance with the production operations of WU-I. Because of this separable structure, individual work units enjoy lowered decision rights for managing manufacturing processes under their responsibility and organizing kaizen activities. This decomposition structure is propagated upstream, covering the entire supply chain with external suppliers. In this manner, through the use of kanban, the problem of how to manage the entire production process is decomposed into a set of management problems associated with individual production process segments. Clearly, such local management problems are mutually exclusive and exhaustive, constituting a direct-sum decomposition.

The direct-sum decomposition makes the contributions of cost reductions visible and additive. For example, suppose that group I-1 found a new way of processing, by which the number of kanban actions could be reduced by one while keeping the same output flow. This kaizen idea resulted in the reduction
of the WIP inventory cost of $X per month. If group I-2 achieved the kaizen effect of $Y per month in a similar manner, the overall production cost reduction would be $X + $Y. In this manner, the TPS makes local contributions of kaizen to the overall enhancement of productivity visible. At the same time, such local contributions are directly summed up to overall productivity enhancement. This strongly motivates workers and management teams to organize kaizen activities positively. This is the key reason why quality circle activities have been yielding remarkable production cost reductions involving many different work units every year over many years.

In the TPS, the apparatus of the organization determines the flow of the final assembly line as a set of constraints to be followed by various work units. Based on the direct-sum decomposition of the whole production process, individual work units then function independently, enjoying lowered decision rights of their own for satisfying the constraints and organizing kaizen activities for further cost reductions.
CONDITIONS UNDER WHICH THE JIT APPROACH WOULD WORK WELL

If a management approach works well under certain circumstances, there are reasons why it works well. As discussed in the previous sections, such reasons for the TPS can be best understood by analyzing it from the viewpoint of organizational functionality. The validity of the TPS from this perspective goes well beyond manufacturing, and many corporations outside manufacturing, including various service industries, could benefit by learning from it. However, if one focuses on the structure and the mechanism of the TPS, represented by the term “the JIT approach,” it is not almighty. To my best knowledge, no literature exists for analyzing the conditions under which the JIT approach would work well in a systematic manner. The purpose of this section is to examine the conditions under which the JIT approach would be effective in manufacturing management and when it might be irrelevant.

Manufacturing Products for which the JIT Approach Would Work Well

As explained above, the essence of the TPS could be interpreted as “production at the speed of sales on the market.” This interpretation implies that if the sales speed cannot be approximated as a flow, the JIT approach may not necessarily be effective. In this regard, depending on when procurement and production take place, we first classify manufacturing products into three categories, as shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>THREE TYPES OF MANUFACTURING PRODUCTS.</td>
</tr>
<tr>
<td>Product type</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Type I</td>
</tr>
<tr>
<td>Type II</td>
</tr>
<tr>
<td>Type III</td>
</tr>
</tbody>
</table>
For type I products such as airplanes and bullet trains, both procurement and production take place only after receiving confirmed orders. Manufacturing products with dependent demand constitute type II products, represented by automobile parts, semiconductor chips, sensors, etc., which are devices and components used in consumer products. In general, production occurs after receiving confirmed orders from manufacturers that produce the final products to be sold in consumer markets. Type III products consist of various consumer items sold in corresponding consumer markets. Because of severe competition in shortening market lead times, type III producers are forced to rely upon demand estimates for both procurement and production.

In light of “production at the speed of sales on the market,” the JIT approach is clearly irrelevant for type I products, since the sales are occasional and cannot be approximated as a flow. For type I products, contracts are typically obtained through open bidding. Once a company wins a contract, it receives payments through multiple installments starting before the delivery of the final products. Accordingly, it is not so important to reduce the WIP inventory through the JIT approach. It is much more important to establish an efficient knowledge management system to promptly prepare a competitive proposal ready for bidding and win the bid.

The JIT approach is relevant to type II companies in that confirmed orders often come along with corresponding kanban activity requests from manufacturers based on the TPS. Shipments from type II companies to their manufacturers are typically made only several times a day, and the outflows of products from type II companies cannot be approximated as a flow. However, they may also have to adopt the JIT approach for their own production processes in order to sustain total quality management (TQM) activities on a daily basis in close collaboration with manufacturers. In addition, they are not strong enough to impose the JIT approach on their own suppliers. In order to keep up with rigorous due dates imposed by manufacturers, type II companies are often forced to preserve materials and partially processed parts as safety stock in large volumes. In practice, many type II companies exercise the JIT approach in a limited manner without enjoying its full potential.

In contrast, the JIT approach is most effective for type III manufacturers, where they often enjoy large sales volumes with good sales speeds. In the case of automobiles, for example, individual dealers may sell several cars a week,
which would not constitute a flow of sales. However, a typical automobile maker has many dealers all over Japan and the accumulated nationwide sales can be approximated as a flow. Based on the JIT approach, where only necessary parts and materials are fed into the final assembly line at exactly needed times in exactly needed quantities from the subassembly lines by means of kanban, type III manufacturers organize their production processes to produce at the speed of sales. Other upstream production processes are organized similarly, with the chain of JIT operations cascading backward including outside suppliers, thereby minimizing the cost for WIP inventories.

**Manufacturing Products with Short Total Lead Time**

In manufacturing, the total lead time is the time required for a product to reach its market after the beginning of the manufacturing operation and essentially consists of three key components: procurement lead time; production lead time; and delivery lead time. If the total lead time is X days, the current production planning would aim at the market after X days from now. Accordingly, the demand of X days from now has to be estimated for organizing the current production plan. Even when the JIT approach is employed, this situation is the same and the current production line has to produce at the estimated speed of sales in the market after X days from now. It is clear that the longer X is, the less accurate demand estimates would be. Hence, X must be short enough for the JIT approach to be effective.

**Manufacturing Products with Small Setup Cost**

Based on the concept of “produce at the speed of sales on the market,” the JIT approach requires producing a variety of products in small lot sizes using the same production line. This inevitably results in increasing the number of necessary setups. To illustrate this point clearly, consider the example where the speed of sales on the market for three car types A, B, and C are given as 8, 4, and 2 cars per minute, respectively. Assuming 8 working hours per day, this means that 3,840 type A cars, 1,920 type B cars, and 960 type C cars must be produced each day. If the production takes place in large lot sizes of 3,840, 1,920, and 960 for A, B, and C, respectively, only three setups are needed including the initial setup. In the JIT approach, setups have to be repeated 480×3=1,440 times.
Hence, we can conclude that the JIT approach is effective if and only if the cost decrease due to reduction of WIP inventories throughout the value chain is greater than the cost increase resulting from repeated setups.

**Manufacturing Products Sustained by Multiskilled Workers**

Through the spirit of “convert every segment of the production process into a part of a critical path,” the JIT approach identifies any excess resources and removes them from every production segment, provided that the flows of materials toward the final assembly line can be sustained intact and sound working environments can be maintained. For example, a work unit may find a way to sustain the same number of kanban actions with one less person while keeping the output flow intact. Unless the removed personnel can be utilized to enhance productivity elsewhere, however, the improvement achieved at the work unit would not contribute to the overall improvement of the company as a whole. This means that the JIT approach is effective if and only if the entire manufacturing process is sustained by experienced multiskilled workers to ensure the enhancement of overall productivity.
CONCLUSION

This report analyzes the organizational functionalities achieved by the TPS, providing important steps to expand the applicability of the TPS approach beyond manufacturing. The key findings are summarized as follows:

1) If a problem occurs, deal with it as an institutional problem, at least at the beginning, rather than as an incidental problem.

2) If the speed of sales of products or services can be approximated as a flow, produce such products or services at a synchronized speed.

3) Structure a set of business processes in such a way that every segment becomes a part of a critical path by removing extra resources whenever possible. At the same time, recognize the potential vulnerability of the approach and make sure that people involved understand this vulnerability and behave accordingly.

4) Organize any business project in such a way that daily business activities themselves provide explicit expressions of system state for the progress of the project. Through this approach, make the cause of any problem explicit and visible at the time of occurrence.

5) Structure any decision problem in two layers. An upper-level decisionmaker should concentrate on providing a set of constraints to lower-level managers by clearly stating decision authority, responsibility, and how to control information for the problem. Such constraints should be presented so that cross-divisional effects are well reflected and decision problems at lower levels approximately constitute a direct-sum decomposition, as long as the constraints are firmly met.

6) Lower-level managers should freely exercise their own initiatives for achieving their goals within the given constraints.
Any management approach could be “almighty,” and the JIT approach, which is the core of the TPS, is no exception. In manufacturing, the JIT approach is effective under the following conditions:

1) The speed of sales on the market should be approximated as a flow.

2) The total lead time should be short enough to contain the demand estimation error within a certain range.

3) The cost decrease due to reduction of WIP inventories throughout the value chain should exceed the cost increase resulting from repeated setups.

4) Multiskilled workers should be developed constantly for enhancing overall productivity through flexible work assignments.
REFERENCES


LIST OF TABLES

TABLE 1  Three types of manufacturing products .......................................................... 12

LIST OF FIGURES

FIGURE 1  MRP as a push system .................................................................................. 3
FIGURE 2  The TPS as a pull system ............................................................................ 4
FIGURE 3  Incidental vs. institutional: Two types of risk ............................................. 5
FIGURE 4  Produce at the speed of sales ..................................................................... 7
FIGURE 5  Direct-sum decomposition of production processes .................................. 10
ABOUT THE AUTHOR

Dr. Ushio Sumita
Professor Emeritus
Graduate School of Science and Technology
University of Tsukuba
Japan

Watch the Productivity Talk on
Untold Essence of the Toyota Production System (TPS)