This document should be cited as:
Garcia-Sabater, Jose P. ; Maheut, J. (2023)
Aligning operations with the market and priorities. Technical Note
RIUNET Repository UPV
http://hdl.handle.net/10251/199965

Aligning operations with the market and priorities

1. Introduction

An operations system uses resources to **transform inputs into outputs via processes**. Unfortunately, the term "process" can mean many things. It can refer to the work done by a machine, by office workers, by a computer or by all of these together.

Many "process engineers" focus only on getting machines up and running. Moreover, many "process industries" have never defined a process.

The definition of operations should be aligned with the market requirements to be satisfied: more or less volume, more or less variety, more or less variation, more or less visibility, more or less speed.

Once clear about what we are going to do, we can define various strategies to serve this market, such as setting the decoupling point, organising resources, outsourcing, automation and, above all, defining the excess capacity to be installed. All of this is done based on the information available at any given time.

At the micro level, information is less ambiguous but more complex. The situation is conditioned by the resources acquired.

Defining a process requires defining inputs, outputs, the rules that govern the process, the resources it uses and also the method (the set of ordered tasks that will transform inputs into outputs). Moreover, defining a process is a recursive activity. Processes are usually made up of other processes.

Many processes require tools, in some cases physical tools, in others, logical tools. In any case, the tool should support the process, and not the other way around. In other words, the process comes first and then the tool.

Unfortunately, companies often design (or select) tools or machines that are very costly in terms of money or effort to implement and then have to use other machines to make them cost-effective. This is particularly evident in the Spanish public service, where the following tends to occur: the legislator publishes a law or regulation; a person carrying out their duties, or perhaps even usurping these duties, implements an IT tool (partially related to the issue); and suddenly the IT system (usually defective) rules the aforementioned service and its users (Graeber, 2015).

Even without the possibility of modifying the resources, there can still be room for manoeuvre if the processes are redesigned and adapted to the organisation's priorities (which may be shifting): cost, reliability, flexibility, personalisation, speed, security, beauty, etc.

These priorities may change, as may the performance of tools. This will have to be achieved with the resources available, which often requires the redefinition of processes. Whatever else is said, there are always alternatives. This chapter attempts to create a mental framework that facilitates the continuous review of processes, whether they be production, logistics or business processes.

It then goes on to reflect briefly on the lack of information that designers inevitably face. Subsequently, an estimate is made of the resources to be selected and strategies are designed to help define the necessary resources. Even when the available resources have been defined, there are still ways to align the processes with the company's priorities. Eight lines of action are proposed, potentially offering alternative ways of implementation (even where no new resources are allocated). The chapter ends with reflections on the difficulty involved in making the change.

2. Define the acceptable level of uncertainty

If the best is the enemy of the good, then precision is the enemy of action.

At times, product design should be fully defined before production starts. However, this would mean producing "test products" for months, during which time the specifications would keep changing, meaning no new product would ever reach the market, since nothing but test products would be produced.

A newly graduated engineer tasked with designing a system expects to receive the data first. The engineer's boss expects the system to be designed with the data available (i.e. with no data). The boss is always right, probably unlike the data that the graduate considers necessary.

2.1. Ambiguity and uncertainty in process design

Usually, the data required are not known until the design begins. Fortunately, very few solutions are sensitive to input data.

It is also fortunate that operations systems are complex.

In the design process it is important to distinguish between scenarios and alternatives. Scenarios are sets of uncertain data over which the designer has no influence. Alternatives are sets of uncertain data which the designer defines.

A designer who lacks reliable data (by definition any designer) would do well to set about clarifying ambiguities, defining alternatives and estimating how they behave in different scenarios.

In this way, the designer can seek data that will have an impact on the solution rather than struggling to find information that is not necessary.

2.2. Uncertainty in process management

The previous section discussed process design. This section considers process management (which also has to be designed).

The stock in transit, the amount of product already included in the final product, the production rate and the amount of time required for setup will be more or less uncertain depending on the type of product and the organisation creating it.

At times we need to know exactly how much stock is available before we start production. In most cases it is enough to know that there is sufficient stock to finish the order. In others, we find out whether there was sufficient stock when, halfway through execution, we discover that the ambient heat has dried the product and we are unable continue due to lack of material (in reality the material is there, but without the water it previously contained).

Sometimes we need to know demand before starting production, but often we have to produce without knowing demand, since there will be no demand without stock in the warehouse, given that the customer expects speed.

In any case, it is always possible that the quality of the products may not be entirely adequate.

3. Selection of resources

The tasks performed by an organisation were assigned at a particular moment as the result of perhaps unstructured decision-making that responded to the needs of the decision-makers at that time. Changing the allocation of tasks to resources or even the task structure itself leads to changes in the outputs obtained, i.e. a different quantity, quality or lead time.

The resources available should be sufficient to operate at the right level of usage:

- Machines
- Workers
- Unit load components
- Energy
- Facilities

Machines (and transforming resources in general) can be faster or slower and more or less standard. Selecting such equipment will require higher (or lower) levels of investment and lead to higher or lower operating costs. Some resources will lead to bottlenecks requiring special protection in the form of buffers or requiring production scheduling to be built around them.

In some companies, machines are so important that they become monuments. Large and impressive, they are a place of pilgrimage for visiting tourists, whereas locals avoid them and often wonder why it was here that pharaoh erected his pyramid.

A machine with high changeover times and costs will be unproductive, making coordination with other activities difficult, delaying delivery of the product to the next stage and determining the size of production batches, all of which will lead to blockages in warehouse systems, which will have to work above their capacity.

A machine with high or unpredictable maintenance times will require high stock levels to "protect" upstream and downstream stages and will result in longer and less reliable response times. A machine that cannot deliver products with the required quality will make it necessary to launch larger production batches than needed, at an earlier point in time than strictly required, generating stocks that will have to be managed and delaying the delivery of customer orders.

A machine that does not produce at the expected speed, i.e. with a low cycle time, or whose speed cannot be controlled, will generate stocks (upstream or downstream) which have been expensive to obtain.

The workers manning the machines should be more or less specialised and will carry out more or less repetitive tasks depending on the company's needs. However, they may also be more or less multi-skilled (worker training is usually the responsibility of the operations manager) and may be structured in teams or work groups according to their skills or some other criterion. Where workers are in short supply or require months or even years of training, they should be protected by the operations manager.

Work teams can be used to create a sense of belonging and even to incentivise worker improvement. This "bonus" cannot be achieved if the company meets its labour needs through abusive temporary contracts, temporary employment agencies, trainees operating as full-time workers or any other type of subcontracting of human resources. This strategy clearly provides flexibility, but at the expense of loyalty (which is not, however, taken into consideration by operations management).

Unit loads can (must) be used by operations managers as a resource to control the amount of stock in the system.

At times energy (its availability and cost) can affect resource efficiency as much as raw materials (their availability and cost, but also their quality).

Facilities (both the building and its ancillary systems—electrical, fire-prevention, etc.) also play a role, each of which is subject to some form of public authority licence. Storage and handling equipment will also influence the use of other resources and the conditions in which workers (line operators and middle managers) perform their duties.

4. Operations Design. Long-term

Depending on the environment for which the process is defined, the company will determine its strategy. In the long term, it is a matter of defining the type and quantity of resources required.

In the design of any process there are options, which can be listed as follows:

- Process structure (project, batches, continuous)
- Define where to fix the decoupling point (ETO, MTO, ATO, MTS)
- Automate—at various levels—both basic and ancillary operations, whether related to materials or information
- Outsource (what? how much? when?) operations or insource those previously purchased
- Define the level of resource utilisation

At first glance, it often seems that decisions are not being made, but rather that the company evolves from more "simple" to more "complex" modes as it grows and seeks to reduce prices. However, at times companies are forced to, or voluntarily, take a step backwards in response to market requirements.

4.1. Process strategies in terms of volume and variety

The Hayes-Wheelwright matrix (Hayes and Wheelwright, 1979) is one of the most classic tools for analysing process strategies. It classifies process structures (specifically the grouping of resources) according to the volume and variety of the products to be manufactured. Some suggest that rather than being a classification tool, it instead describes the likely evolution of systems.

If the aim is to offer maximum variety, it is assumed that resources will be used only when necessary, by expert operators who know when to use them. In this case, resources are "allocated or withdrawn" to the extent that the new product to be manufactured requires them. This way of arranging resources is typically used when the volume to be produced is very low (perhaps a single unit) and the variety high. It is often associated with very diverse tasks and very intermittent (or even non-existent) material flows.

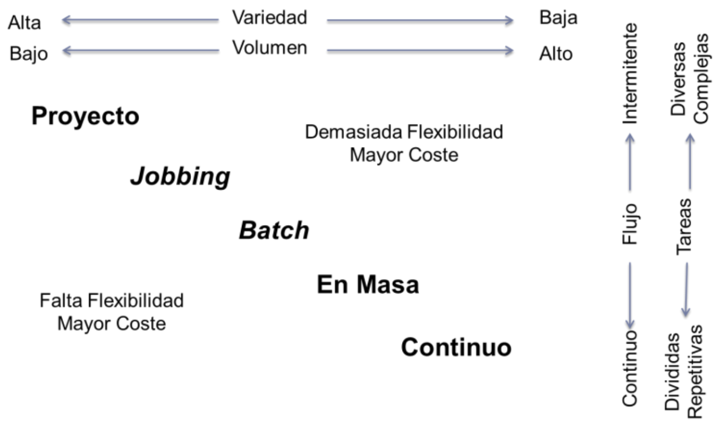

Figure 1: processes according to volume and variety (prepared by the author based on (Hayes and Wheelwright, 1979)*)*

If it is decided to allocate resources to functional units based on "what they are good at", products will have to move from one place to another in search of the capacity required at any given moment. This is called "job shop". The traffic of products and workers will be significant but resources can be managed more efficiently. This way of arranging resources is best when variety is high and volume low, and when resources cannot, should not or are difficult to move.

In environments where products are reasonably similar, resources are organised in so-called "flow shop" structures. They are grouped according to "what they are good at", but the sequence of activities is reasonably stable and therefore one section feeds into another.

Resources can be positioned to allow the product to move as efficiently as possible, finding along the way the resources it needs at any given moment. These are so-called "production lines". The structure of the process is "mass" or "continuous". This mode of organisation makes it hard to use the necessary productive resources very efficiently, but reduces the resources associated with internal logistics. This is the method used for high volume, low variety systems.

4.2. Locate the decoupling point

The location of the decoupling point should depend on the level of variety requested by the customer and the required speed of delivery. The customer order decoupling point (CODP) is the point in the value chain up to which the customer order is allowed to penetrate. In general, the decoupling point also influences the level of customer involvement in the operation.

It is generally accepted that the system acts by demand pull from the customer to the decoupling point and in push mode from the decoupling point to the supplier.

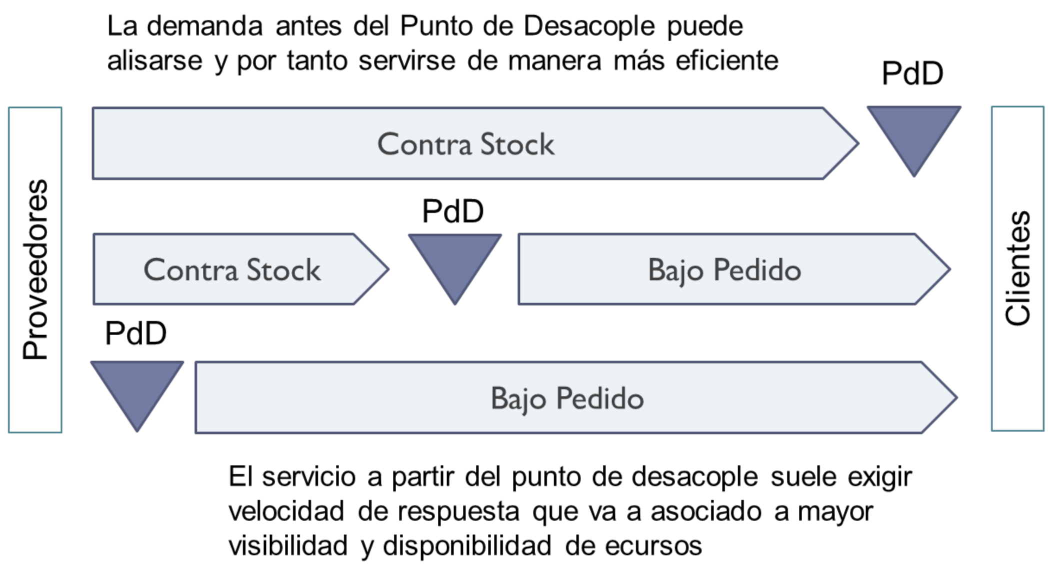

Figure 2: decoupling point

The decoupling point is the point in the system where the available stock makes it possible to separate the customer cycle from the manufacturing cycles. Locating the decoupling point close to the finished product makes it possible to manufacture independently of demand (make-to-stock) up to this point, allowing a more stable (and therefore potentially more efficient) use of resources in most of the production process.

Locating the decoupling point close to the raw material (upstream) reduces stock levels and facilitates the manufacture of products that are less standardised and closer to the customer's needs. By locating the decoupling point at a distance from the customer, the process will have to react with varying degrees of agility to changes in demand, often requiring excess installed capacity.

P>>D	P>D	P=D	P<D
Make-to-stock (MTS)	Assemble-to-order (ATO)	Make-to-order (MTO)	Engineer-to-order (ETO)
Finished products stored	Sub-assemblies stored	Raw materials stored	Some raw materials not stored
Design defined; "catalogue" product	Design defined; configurations may require specification	Some options can be redesigned	Design created when customer decides on order
Delivery promised for next day	Delivery promised subject to available capacity	Delivery promised subject to availability of capacity and material	Delivery promised subject to availability of capacity and material

Figure 3: MTS, ATO, MTO, ETO

By locating the decoupling point somewhere in between, a balance can be struck between agility and the sufficient use of resources.

Some systems are designed to move the decoupling point based on demand when demand is highly variable. As such, the CODP can move backwards at times of low demand and forwards at times of higher demand.

4.3. Outsourcing or vertical integration

Coping with high variation in the volume to be supplied is one of the circumstances where companies may outsource their activities. Further reasons are outlined below.

It is not necessary for all activities to be carried out on company premises or by company staff. The process of allowing others the pleasure of doing our work is called "outsourcing".

The operations manager may outsource part of the operations for economic reasons (reduction of investments, fixed costs or variable costs), for reasons of technology availability (the supplier may be more familiar with the required technology than the customer), or to maintain flexibility in the volume or variety of the product manufactured.

Before a decision is made to outsource, the company should analyse the risks involved in teaching the supplier its work method and in giving the supplier control over the technology, including the required materials.

On the other hand, the reduction in investments and/or costs may not be worthwhile if reversing the outsourcing process in the future will be very costly, or if any additional transaction costs in the execution of the outsourced processes are prohibitively high.

When an activity is outsourced, the supplier can become so familiar with the product and its possibilities that it would be easy for it to pass on this knowledge to competitors or to become the competitor itself. In any case, a company that is highly dependent on its suppliers has placed itself in a less favourable bargaining position.

From an operational point of view, when outsourcing part of its operations, a company runs a high risk of losing control over the activity and over the capacity to generate new opportunities from the incorporation of technologies or from cost reductions linked to the learning effect.

A variant of outsourcing is offshoring, where the external operator carrying out the activity is located in a foreign country (usually allowing a significant reduction in labour costs).

The opposite of outsourcing is vertical integration, i.e. performing in-house certain activities previously carried out by customers or suppliers. Greater vertical integration allows more control over the entire process and adds more value without the need to expand the customer base.

The decision to increase or decrease the level of outsourcing is strategic in nature. It can have serious repercussions if some of the potential risks have not been adequately protected against.

However, it is not uncommon for the decision to be made on the basis of purely operational criteria such as the cost per square metre of available facilities or quality problems the company intends to secretly pass on to the supplier.

4.4. Level of automation

High volumes of a product with relatively low variety can be achieved with high levels of automation, which in turn can help reduce costs or improve delivered quality.

As the cost of technology falls, and relative labour costs rise, the level of automation in a firm tends to increase. However, economics (labour costs or productivity) is not the only reason an organisation might choose to increase automation.

Sometimes automation is required because there is no way to reduce the danger or "toughness" inherent to certain tasks which modern Western society cannot accept. At other times, only machines can perform certain tasks with sufficient precision or strength.

Furthermore, some companies may opt for automation simply to show off.

Just as customers may be willing to pay more for a handmade product, suppliers, customers, and even workers and managers, may trust a company more because it is more automated.

Perhaps the golden rule in automation should be the old adage: "first simplify, then mechanise and, if necessary, automate".

In many automated environments the worker is left merely as a support to the machine.

In lean environments, automation follows the Jidoka approach, i.e. "respecting the human factor", which can be translated here as not eliminating work content but instead making work less cumbersome and more suited to worker dignity. Moreover, the aim is to automate that which it is feasible and practicable to automate (Baudin, 2007).

Therefore, automation usually starts with what is easiest to automate and removes the most drudgery for the worker:

- Automate unloading because you know where to take the product from and where to leave it, and it is often a repetitive and heavy job.
- Automate the return to the initial operation.
- Automate shutdowns when an abnormality is detected (and communication of the situation to the relevant people).
- Automate feeding (with associated quality controls).
- Automate the process itself.

However, automating from the outside in is not the only specific feature of the lean automation approach. Attentive readers may be interested in *Karakuri* puppets, but since they are automata, it would be better to view them on YouTube.

First simplify, then mechanise, then automate. As far as possible (although it should always be the case), this should be done while respecting the human factor.

4.5. Installed overcapacity: Level of resource utilisation

Coping with a requirement for fast delivery while maintaining a certain variety in supply usually calls for a certain amount of spare capacity. On the other hand, keeping pace with highly variable demand may require installed overcapacity that is incompatible with business survival in terms of cost.

Overcapacity always means an additional cost and reducing it is thus an obvious goal. One way to reduce total installed capacity while maintaining the system's responsiveness is by positioning a decoupling point. The operations manager is responsible for designing the physical and logical systems that make it possible to get the most out of the available resources.

To the extent that the other actors in the organisation allow it, the operations manager will have the resources (machines, equipment, facilities, staff, energy, materials, etc.) to respond to short- and long-term needs in an appropriate manner.

Basic economic theory states that the optimal use of resources in the short term does not systematically coincide with their optimal use over the long term. It makes little sense to invest tens of millions in a car manufacturing plant only to find that there is not enough capacity to meet demand.

Another decision relates to the level of subsequent use of the resources to be made available. Using everything "at full throttle" is not necessarily the best option.

Bringing operations to a utilisation level of 95% often results in uncontrollable delays unless there is a very detailed coordination of activities, which requires sophisticated planning and operations control systems, fed by highly reliable data that also comes at a high cost.

If trucks are always full (FTL or "full truckload" strategy), stock levels at origin and destination are higher than if the trucks are allowed to transport without being full (LTL or "less-than truckload" strategy).

Filling a warehouse up to the ceiling, which allows a greater number of loads to be stored, means the loads must be lifted higher, usually requiring more expensive specialised forklifts and even more forklift trucks and forklift operators.

Warehouse space (downstream and upstream) is needed in order for workers paid by the hour to always work to the best of their ability. This space will also determine work for the rest of the system.

However, having more capacity than necessary is not always a good option.

Having more unit load components available makes it easier for machines to avoid stoppages due to a lack of racks on which to leave products, but can lead to sharp increases in the amount of intermediate stock, and require additional forklift operators to move full and empty racks around the warehouse.

Having more space in a warehouse simplifies and facilitates flows. However, the human tendency towards accumulation makes teams fill available space with pallets of products (the industrial version of Diogenes syndrome) that spread like wildfire through the factory, impeding movement and strangling productivity.

The auxiliary equipment available (handling elements, handling robots, shelving and cabinets, etc.) should facilitate the use of limiting resources (so-called “bottlenecks”), as well as health and safety conditions (ergonomics) for the machinery operators.

For an operations manager, it is important to know whether the machines are reliable (whether they deliver quality products or break down uncontrollably), flexible (specifically the time it takes to go from producing one product to producing another) and efficient (whether they work using the resources at the right cost).

One way or another, abnormal behaviour from individual machines or the associated facilities will affect the system’s performance. Therefore, if these values are not well known at the design stage, it would be wise to install excess capacity, which can subsequently be eliminated if required.

However, even a system in which each resource works perfectly in isolation can have a less than optimal performance if the connections between the various stages of a process have not been properly configured.

Conversely, a system made up of elements that do not perform optimally, but which have sufficient legroom, can deliver a service that meets the customer’s requirements if the connections between the stages in the process have been properly defined.

4.6. Centralisation and decentralisation

Both physical operations and the information systems supporting logical processes can have varying degrees of centralisation.

Centralising operations clearly has an impact on economies of scale (and therefore on costs) but can also affect other equally relevant aspects such as risks or transport costs. Centralising all operations in a single location sharply increases risk levels. For high volume products, it may impact transport costs. At times, proximity to the customer (not only because of logistics costs) represents a market opportunity and promotes product customisation.

The planning and scheduling of operations can also be carried out with varying degrees of centralisation. Centralisation can lead to cost improvements if information of sufficient quality is available. Decentralisation (more or less coordinated) usually entails higher costs but allows for better functioning with more uncertain information. An intermediate step is coordination, which allows the flexibility of decentralisation but with similar results to centralisation (Rius-Sorolla *et al.*, 2020).

It can make sense to decentralise monitoring, evaluation and the communication of evaluation, but in certain environments and with certain parameters this can lead to a loss of coordination.

5. Processes and competitive priorities

When designing processes, competitive priorities will be transformed into objectives that will bring different types of benefit to the various elements of the organisation.

Prioridad Competitiva	Objetivos en el Diseño del Proceso	Beneficios
Calidad	Procesado Libre de Errores	Menos retrabajos y menos esfuerzos
Velocidad	Tiempo de tránsito mínimo Capacidad elevada	Tiempo de espera del cliente corto
Confianza	Alternativas de Ejecución Capacidad suficiente	Habilidad para hacer frente a lo inesperado.
Flexibilidad en Mezcla	Tiempo de Cambio Reducido	Costes de Cambios reducidos Incorporación rápida de nuevos productos
Flexibilidad en Volumen	Amplio Rango de Capacidad	Habilidad para cambiar de volumen de producción
Coste	Eliminación de los principales tipos de desperdicio	Costes de Procesado Bajos Costes de Inventario bajos Costes de Inversión limitados

Figure 4: competitive priorities and objectives in process design

If the competitive priority is quality, the fundamental goal is error-free processing. This is likely to lead to more robust processes being designed, perhaps with more comprehensive quality control schemes and machinery better able to fine-tune the product. An error-free system results in less rework, and in the long run is less effort-consuming and cheaper—but only in the long run.

If speed is the competitive priority, reducing transit time as much as possible will be the key goal. This requires sufficient installed capacity to avoid queues and congestion. Eliminating intermediate buffers makes the system more sensitive to failures, but minimises waiting times for the customer.

If confidence is the competitive priority, there must be implementation alternatives that will make it possible to react to any unforeseen events. Such implementation alternatives are investments that stand idle waiting to be used and are only useful when something unexpected occurs.

When the priority is to serve the customer with a sufficient variety of products (product-mix flexibility), the goal is to make newly designed products quickly available. This priority requires designing in such a way that changeovers have a low cost (in time and effort). Some suggest that the ideal approach is to have small, flexible machines rather than large ones (which generally produce at lower cost).

Sometimes the priority is to be able to cope with variation (volume flexibility), in which case the system must have the capacity to expand and contract. To this end, resources are arranged in such a way that more staff can be hired, the number of jobs can be increased, and more facilities can be made available, although this means that in times of low demand there will be much less availability. Sometimes the solution is to store large quantities of stock in times of low demand.

At other times, and much less often than the public realises, the competitive priority is cost. Here, the design will try to eliminate all kinds of waste (time, stock, movements) in search of low processing and inventory costs with the minimum possible investment.

6. Design of production processes. Short term

The set of tasks required to make a product or provide a service is more or less known. Sometimes certain tasks can be performed at various times in the work sequence. At other times, tasks have to wait for several "threads" of work to be completed. In most cases, inputs are the outputs of other processes (and vice versa).

Tasks can be performed by operators (more or less specialised) and/or machines (more or less automated). In addition to tasks, waiting times (drying, cooling, etc.) are sometimes needed.

At times the process (as a set of interrelated tasks) has been explicitly defined; more often, it is the result of decisions made following a resolution to use one resource rather than another or to locate the resource in one place rather than another.

In some organisations, the process is defined on the basis of the bottleneck(s), while in others it can be influenced by the background of its designer. In certain cases, the customer (member of the public or patient) has been taken into account when defining the process.

There is always at least one other configuration that would achieve the same output but with a different process (quicker, more efficient, safer, more flexible... or less so).

How things are being done right now is not the "only and best" way. The current way, perhaps once seen as "optimal" (seriously?) or the "only way possible" (really?), was designed, selected and accepted at some previous point in time. It is therefore unreasonable to assume there can be no alternative.

Sometimes the person who executes a process does not want to change the way he or she works. Although not "optimal" for anyone, this is reasonable.

Less reasonable, but also very common, is where the person opposing the change is the one who designed the process. They probably had good reasons to design it that way, although these are often outdated or long forgotten.

Why are you calling my solution a problem?

The worst situation is where the process was never designed and yet the "owner" does not want to change it.

In the design of any process there are options, which can be listed as follows:

- Split/group/reassign tasks so that they are performed by various actors with varying levels of differentiation or specialisation.
- Increase/replicate resources so that processes are carried out in parallel by a more or less homogeneous "battery" of actors or replace parallel resources with a line (serialise).
- Establish the capacity of buffers that protect operations from variability stemming from suppliers and customers (internal and external).

- Establish the when, how and where of the triggers that give commands to the various elements of the system.
- Arrange resources spatially (floor plan layout).
- Allocate time for resources (shifts, holidays, overtime).

Rather than a definitive classification, the above list can be used to support the generation of ideas when this process becomes stuck due to the designer's lack of experience or the resistance of those who will implement the process. Ultimately, the design (and redesign) of a system will always be limited by an elusive mindset:

"*We've always done it this way here*".

Changing the ways tasks have "always" be done means dispossessing process owners of the only certainty they have about their work: "*The unfinished work will still be there tomorrow, and so I'll still be needed, since I'm the only person who knows how to do it*".

6.1. Split/group/reassign tasks and resources

The tasks needed to realise a set of products or services were jointly assigned a long time ago, under different external conditions (volume, variety, product visibility, etc.) and internal conditions (product knowledge, worker experience, level of automation). Since then, new participants and technologies have gradually been brought in and responsibilities and tasks have been assigned to them, not according to a specific set of rules but in an attempt to alter the status quo as little as possible...

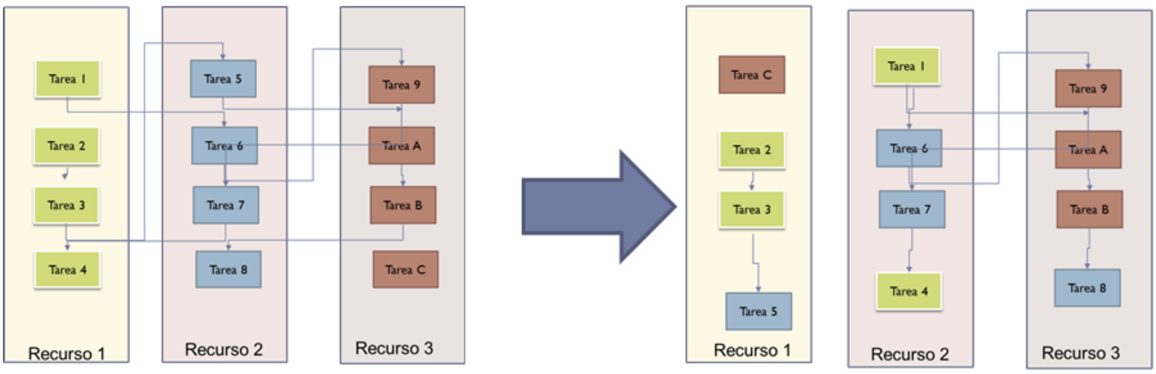

Figure 5: reorganising tasks

And it is certainly possible that the current allocation of tasks and responsibilities is not the best possible for customers, for the organisation, and even for individuals.

Intrinsic phenomena, such as the learning effect, and extrinsic phenomena, such as technological changes, mean that the workload resulting from each task in a process will change over time. Therefore, it seems logical to reorganise over time.

Perhaps the most important activity of an operations manager is reallocating tasks. However, this is by far the most difficult part of their job with the highest emotional cost. Nobody wants to change, especially those who are so well-off that they think any change will make things worse for them.

People in bureaucratic organisations (often the finger is pointed at the organisation itself, but in reality it is the people inside it) will generally fight tooth and nail against a change in the structure of activities and the resulting allocation of the workload. Since it is hard to see why someone who is overworked would not want his or her job to be rationalised, it can be assumed that someone opposing such a review has just as weak a position to defend.

A very curious and perverse phenomenon occurs in some organisations where, after detailed analysis, someone discovers that the total workload can be reduced by automating part of the process and transferring another part from one section to another. Ironically, the section that is already the most overworked (the section that has tended to receive the most tasks without complaining) is often the one that receives the additional load with the less loaded section now having an even lighter load. All this is an especially pernicious version of the Matthew effect, whereby "to him that hath shall more be given, and from him that hath not, that which he hath shall be taken away".

This particular phenomenon is very common when "management" starts to "computerise" processes. An example of this is when doctors start spending more time looking at their computers rather than at patients, eliminating the need for administrative assistants.

6.2. Parallelise or serialise

As demand for a particular operation rises, it may be necessary to decide whether or not to increase production capacity.

Such an increase can be achieved by purchasing a machine (or hiring a worker) that is the same as the existing one (parallelise) or by decomposing the process into tasks that can be executed sequentially (serialise). Serialisation taken to its extreme is the assembly line, but serialisation is also where work in a department store is divided between those who stack shelves, those who provide customer support, those who collect cash at the checkout, and so on.

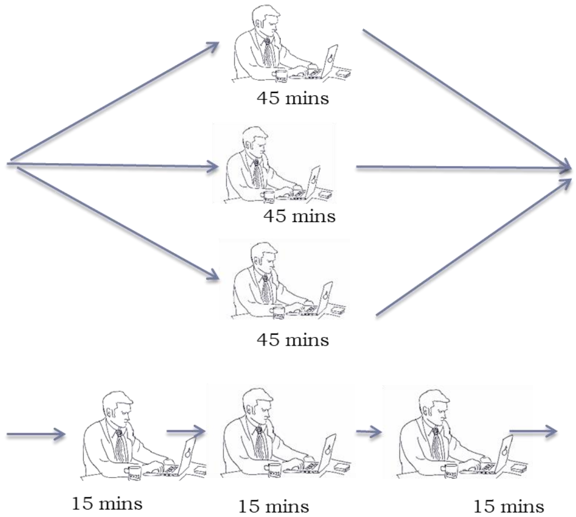

Figure 6: serialise vs. parallelise

Breaking down tasks and making execution sequential has its benefits:

- Shorter operating times mean the experience curve can take effect much more swiftly.
- The process requires less skilled workers.

- Working in sequence (if intermediate stocks are limited) makes it easier to control the total throughput.

The main disadvantage is that the reduction of intermediate stocks becomes a necessity, since a serialised system with no control over intermediate stocks can (unnecessarily) lead to sharp increases in response times.

Having parallel resources with higher cycle times has obvious advantages:

- A single type of workstation is designed and replicated.
- It makes production independent and allows for a more flexible use of production capacity.
- By having comparable systems, it is easier to set a benchmark between them.
- A parallelised system better absorbs variability (both input and service time variability), potentially making it possible to reduce required stock levels.

If it is decided to have systems working in parallel, an additional decision to be made is whether to specialise these parallel systems. If specialisation is chosen, the chances of being more efficient increases, but large variations in demand may lead to idle resources based on the "that's not my speciality" argument. This may result in the worst of both worlds.

6.3. Action triggers (pull or push)

Tasks are performed because someone decides to perform them. At this point we want to know why they decide to do something rather than nothing. These are the process triggers.

It may be that the decision to perform a task is made simply because someone or something decided that the activity should be carried out at a given time. If so, we are probably dealing with a robot, a cobot or a computer programme.

Action is typically triggered by backlogs, low stock levels or complaints about late deliveries.

Before each stage of a process there is usually a pile of unfinished work (or products to be transformed). This pile may be almost non-existent or it may form a small mountain (or several of them).

The operator chooses the next product (or set of products) to be processed based on some criterion, which may be a work schedule (more or less explicit), how long the product has been there, how much work it requires or its level of urgency.

After each stage of a process there is usually a pile of finished work waiting to be sent to the next stage. There may be no such pile or it may form a small mountain.

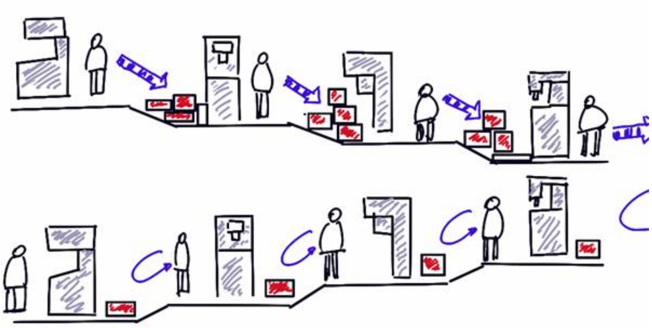

Figure 7: pull vs. push

Some people value the “one piece flow” method, but most prefer to accumulate small mountains of work before and after them as protection against others (and to protect others from themselves). In such cases, the size of the mountain will act as the trigger that sends the product to the next stage or requests more material from the previous stage.

Products (parts, services and papers) do not move from one place to another because of some cosmic energy impelling them to be completed. Someone or something has to ask that they to move to the next stage.

Much information found online implicitly associates the pull system with kanban systems and, by a not very clear derivation, with stockless production. Thus people tend to think that pull processes do not involve stock management. At the same time, push is often associated with “make to stock”, and “make to stock” with stock management.

A “pull” is where the downstream stage requests material it needs to continue working from the upstream stage.

A “push” is where the upstream stage sends the product downstream when it has finished processing it, either individually, as a whole batch, or as the production of a shift. In a push system, the order to manufacture or transport is given, in whatever way, because it is possible to do so (product is available and shipped, raw material is available and processed). Therefore, all supply chain stages that “make to stock” (i.e. produce before the customer has explicitly requested the product) work according to a push system.

In a pull system, the order to manufacture or transport is given, in whatever way, at the request of the next stage (usually due to a lack of material or work to be done). The downstream order can result from any trigger, the three most common of which are: the reorder point, periodic review or an assembly order that requires product.

The reorder point acts as follows: when a certain stock level is reached (a box is emptied or a gauge light turns on), an order is sent to the previous stage to manufacture/ship (or that it can now manufacture/ship).

If a periodic review is the trigger, the stock level is reviewed from time to time and replenished (items are manufactured or shipped) until a certain predefined level is reached. This is the usual way of operating with a kanban system, where the maximum level is set by cards. Whether this method is an authentic pull system or a controlled push system is an interesting point of discussion, but one that should be left for academia.

The third method—manufacture only on order—at industrial level only occurs in assembly lines of large and customised products (cars, aeroplanes, etc.) where the order to manufacture/ship is given based on the next product to be manufactured (e.g. customised car seats).

In short, a pull system acts like a conventional stock management system. The level of stock in the downstream pile is checked from time to time and when it falls below a certain threshold a request is made to the upstream stage to provide or process more. A push system acts the other way around, instead of the downstream stage giving the order based on its needs, the order is given by the upstream stage based on the amount of work it can do.

Pull-type systems are usually associated with lower stock levels as the stock is under control, meaning there is no need to hold more than necessary.

However, a pull system still has to plan its requirements. In fact, it needs to plan them with even greater care, so that it can define baselines more precisely and respond more swiftly.

As mentioned above, the best known pull system is the kanban system, where work orders take the form of cards that limit the amount of stock in the system.

6.4. Set the location and capacity of the required buffers

Buffers are a type of protective barrier against uncontrolled variability (stemming from external or internal factors).

A buffer is essentially a safety cushion. This metaphor illustrates the buffer's function but also reminds us that a buffer should not be rigid.

In operations management, buffers act as a protective barrier against variability and generally refer to a stock accumulated before or after a stage or resource where variability may occur.

Stock which fluctuates on the basis of factors other than variability is not strictly a buffer. Stock that does not fluctuate is not a buffer, it is merely an accumulation of work.

The queue on a server is a buffer that arises through variability. Thus, limiting the buffer (limiting the queue) reduces the server's capacity.

Similarly, a buffer that is always full is not really a buffer, since it is either not useful or cannot be used.

When a station cannot work because it has no raw material (because the previous station is not supplying it), it is said to be starved. When a station cannot work because it has nowhere to leave products, it is said to be blocked. Both situations reduce the availability of the station and, therefore, of the system. Both can be resolved by incorporating intermediate buffers.

Another reason to place buffers between consecutive stages is that variability in times (stops, setups, reduced times) without buffering can multiply stoppage times (if one stops, all stop).

There are two key decisions to make:

a) Where to put the buffer
b) The size of the buffer

The buffer is typically placed immediately before the bottleneck. Take note that the bottleneck is the resource which limits the system's ability to produce more money, not the resource which is most loved or feared.

The aim is to protect the resource from variability in previous stages so that it can continue to produce where possible. Noteworthy in this respect is the "Drum-Buffer-Rope" approach proposed by (Goldratt and Cox, 1999).

When the bottleneck is easily identifiable, the approach of protecting it seems logical. However, a production line can be balanced and even made bottleneck-free by introducing a number of intermediate buffers. To achieve this, a detailed analysis may be needed to identify and locate any bottlenecks.

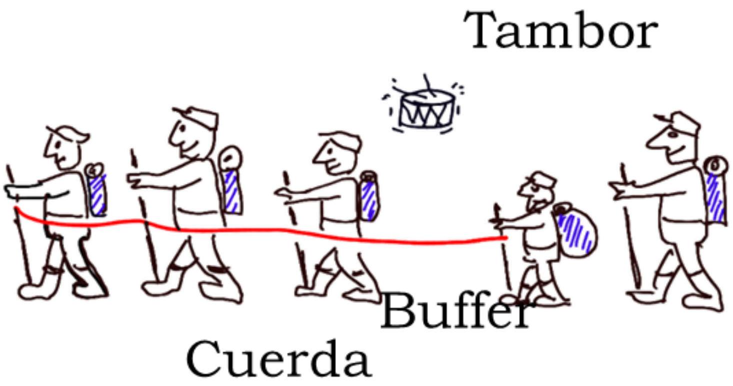

Figure 8: the "Boy Scout Hike" parable in The Goal

The buffer size required to obtain the maximum system capacity will depend on the ratio of the mean time to repair (MTTR) to the cycle time. The higher the ratio, the larger the buffer needed to absorb the variability and avoid interference between one stage and the next. With an infinite buffer the system would be interference-free.

Of course, buffers can be placed before and after each machine. They can be of any size needed to protect the system from an incident.

On the other hand, the mere existence of a buffer increases the throughput time and prevents improvement by hiding opportunities to improve.

When sizing a buffer, a balance should be struck between system throughput and delivery delay. In managing a buffer, it is important to avoid keeping it full at all times. Otherwise, rather than a buffer, it is simply a load.

Safety stock is a type of buffer introduced to protect against variability in demand or in supply lead time, i.e. against external factors.

In project management, a buffer is extra capacity (in the form of time) placed before or after a task in order to absorb the variability of preceding tasks without affecting the total time.

Therefore, in principle, time buffers make it possible to meet deadlines.

However, Parkinson's law states that: "Work expands so as to fill the time available for its completion". Similarly, the amount of required product expands so as to fill the space available for its storage. This ultimately makes the space allocated to the buffer indispensable and therefore useless.

6.5. Arranging resources in space

Having decided what resources are needed (or can be made available), another important decision is how to arrange them in space and time.

Resources must be physically available on the premises. It seems reasonable that certain resources, once installed, should then stay put; but by keeping resources mobile we have the ability to position them where they are useful.

The arrangement of resources in space will be discussed in greater detail in the chapter on floor plan layout. Here the focus is merely on whether to move resources or not, or to group them or not.

Resources can be fixed or can move from one place to another. The concept of "getting rid of anything that does not contribute" requires the ability to move resources, as well as lots of free space. The logic of moving the product rather than the resource leads us to ask whether resources should be grouped by function or to facilitate the product's journey.

It is no less important to change the location of resources, since experience shows that those who take part in a process will resist change, and by changing their work environment we make it easier to change routines.

6.6. Arranging resources in time

Working time (the number of hours worked in a year) is limited in Spain by the Spanish Workers' Statute, collective agreements and agreements between the company and its works council.

Collective agreements also regulate rest periods (hourly, daily, weekly and annual) required to ensure workers' health and performance.

Spain's climate (specifically the heat) determines the annual rest period. This makes perfect sense when there is no air conditioning. However, it is difficult to uphold when the market being supplied is a seasonal market linked to the same climate.

Excess working hours (if legal) are considered overtime, which workers may consider part of their pay agreement. Overtime (which is technically voluntary) has more recently been impacted by flexitime.

The working day can be divided into one or more shifts (shifts have increasingly fewer hours per day and days per week), since this is considered reasonable to ensure the company's productivity (more can be produced with less work).

The working day can be split or continuous, with each alternative having its advantages and disadvantages (subject to debate depending on the parties concerned).

In some organisations, particularly those where workers have a lot of power in defining shifts and where the boss is not very visible, 24-hour shifts or 84-hour weeks are acceptable. Apparently, in such scenarios rest is not entirely necessary, perhaps because it matters more to be available than actually performing a task.

There are organisations where the most important decisions of the year are taken when on-call days or days off are distributed, whereas holidays cannot be distributed and must necessarily be taken in the summer (since, of course, everyone prefers holidays in summer as no one gets sick or goes to restaurants when the day is longer, nor do they write their Master's thesis in August when their tutor is on holidays).

As almost everything is regulated, the operations manager is often left with only one final decision: whether to have employees work one, two or three shifts.

This decision is usually linked to the planning stage known as aggregate planning, where the company decides whether it prefers to store time (in the form of stock) or to use it when needed.

One shift is good for coordination but underutilises physical resources (those that require investment and cannot be replicated). It also allows for flexible hours, facilitates split shifts and has other advantages.

Two shifts (with the operations manager working eleven hours a day to meet everyone's needs) allow for slightly better use of resources (although the reader can imagine a factory at six in the morning after a freezing winter's night).

Operating night shifts means that physical resources will ostensibly be used more. However, this decision has side effects. The night shift has undesirable effects on workers' health, removes the possibility of using overtime during the day and, above all, unleashes the so-called "Toy Story effect" among machines and workers, resulting in there being less product and raw material at the end of the shift than at the beginning.

If one part of the facility works one or two shifts, while the other works two or three, space should be made available to store the product that will be manufactured but not consumed or consumed but not manufactured.

6.7. Gathering information, decision-making and communication

To properly manage any process it is necessary to gather information, transform it and make decisions, and then communicate the decisions back.

The way in which this process is structured can vary significantly from company to company, and may even be the main element shaping the company culture.

Operations managers deal with those who ultimately execute decisions in most cases.

What takes place at the plant can only be seen at the plant. It is therefore essential for the operations manager (and obviously his or her subordinates) to be present on site.

Moreover, while it is important to ask workers how the work is going, it is surely more important to inquire about useful ways to improve and to avoid asking about things that the boss should already know.

This is only possible where the quantity produced can be shown visually (computerised panels or traditional visual management boards). This means you can access the information beforehand (if computerised) or see it on a panel on arrival and simply ask why it has deviated from the standard.

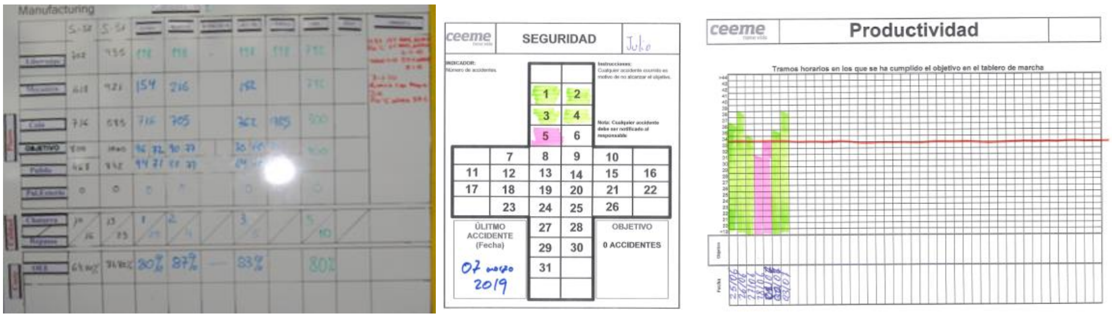

Figure 9: visual tool

6.8. *Standardising activities*

The engineer's role is to standardise. By setting a standard you supposedly lose the ability to be flexible (the most common excuse for avoiding standardisation), but setting a standard is the best possible starting point for improvement.

To paraphrase Lord Kelvin, "if you cannot define it, you cannot measure it." And if you cannot measure it, you cannot improve it. If you cannot improve it, it gets worse (even if only relative to competitors as they improve).

Tasks are standardised using a standard operations sheet (SOS).

SOSs should include both the tasks (separated into manual and machine tasks) and an estimate of the time required to execute them. Sometimes the tools needed to perform the tasks, as well as explanatory diagrams, are also included.

Figure 10: example of SOS (source: Master's degree dissertation of Cristina Fuentes)

Some companies are replacing SOSs with videos or animations that the operator must reproduce. There are even sensors and other control elements that check whether the process is being replicated as designed.

Standardisation reduces variability, thereby reducing the length of queues, which allows for smaller buffers for the same level of saturation.

One common criticism about standardisation is that it impedes improvement. This is really more of an excuse than a criticism. In fact, it is the other way around: the only way of ensuring improvement is through standardisation. Otherwise, what you have is merely change (and not necessarily for the better).

According to (Liker, 2003), standardised tasks are the basis for continuous improvement and employee empowerment.

Proper standardisation has four steps:

1. Prepare standard operating procedures.
2. Train members of the organisation to work with the documentation prepared.
3. Audit the monitoring of standards.
4. Improve established procedures to make the process more productive and robust.

The manufacturing process can be standardised (in the knowledge that standards can/should change), as can the product or its components.

Bottles sold in standard units (litres) facilitate price comparisons, while bottles of a similar shape and size facilitate the use of shelves in refrigerators. Furthermore, if the refrigerators are a standard size, they will be easier to combine with the rest of the kitchen furniture.

If components are standard, managing them is easier, since it will be easier to purchase and handle them. If equipment is standard, it can be maintained more efficiently, since the components required are standard.

It may be difficult to standardise the product since designers always seek to be different and in differentiation may lie the added value. However, by creating standard unit loads, the product can be adapted to standard means of handling and storage.

All this is achieved by standardising unit loads rather than the product. This is the concept of “standardising the interface”.

7. A general commentary on process (re)design

Except in very rare cases, the design of a process targeting a specific product is generally the redesign of a process immediately preceding it. If the purpose of the redesign is to introduce a product variant, the system will be slightly less resistant to change.

However, if the purpose of the redesign is to "improve the process", resistance will be fierce, with the person who designed the previous version leading the opposition:

"Why are you calling my solution a problem?"

The design of a new system is not only complex (many interacting components) and highly uncertain, but can often be hard to evaluate, in the sense that it is governed by vague criteria. Perhaps for this reason using the word "optimisation" is here considered sacrilege.

With so many options and such vague evaluation criteria, when faced with a blank slate the novice designer can become blocked.

One way to unblock yourself is to use the simplest solution you can think of (even if it is expensive, inflexible, time-consuming and inadequate). This solution will be the basis for a subsequent grand design when priorities start to become known.

If the priority is quality, an error-free process should be designed. If the priority is speed, the system should be given excess capacity and intermediate warehouses should be limited. If the priority is reliability, then as well as giving the system excess capacity, it is also necessary to have implementation alternatives. Flexibility in volume is achieved by having the ability to extend available capacity.

Only where reducing costs is the objective should companies seek to eliminate the main types of waste.

The various users and stakeholders will have objectives, restrictions, criteria and priorities that are both unspecific and contradictory, but these will not emerge until they have something concrete to criticise.

It is therefore very important to prepare a quick draft in order to bring discrepancies to the surface as swiftly as possible. The opinions will make it possible to put "in black and white" the selection criteria and restrictions of users and stakeholders.

Principles such as simplicity (understood as "less is more"), sustainability (understood not only in economic terms, but also socially and environmentally) and resilience (understood as the ability to withstand change) should guide the design of the process.

The designer should also take into account concepts such as the "learning effect", ergonomics and safety in the workplace, and even "Design for All".

However, in designing a process, the most important factor is the human who will ultimately interact with it. A machine's behaviour is more or less predictable, but the worst design can be fixed by a motivated employee, while the best design may fail to give positive results if the employee so chooses.

If, on paper, the design of a process is not simple and clear to all, the complexity during its implementation will make it impracticable.

Bibliography

Baudin, M. (2007) *Working with machines : the nuts and bolts of lean operations with jidoka*. Productivity Press.

Goldratt, E. M. and Cox, J. (1999) *La meta : un proceso de mejora continua*. North River Press.

Graeber, D. (2015) *La utopía de las normas David Graeber David Graeber*. Ariel.

Hayes, R. H. and Wheelwright, S. C. (1979) 'The dynamics of product-process life cycles', *Harvard Business Review*. March-April, 57(2), pp. 127–136.

Liker, J. (2003) 'Toyota Way. Blacklick, OH'. USA: McGraw-Hill Professional Publishing.

Rius-Sorolla, G. *et al.* (2020) 'Coordination mechanisms with mathematical programming models for decentralized decision-making: a literature review', *Central European Journal of Operations Research*, 28(1). doi: 10.1007/s10100-018-0594-z.

Acknowledgment

The authors received financial support for the translation of this technical note through the Program "Plan de Acciones Estratégicas para la Internacionalización de la UPV, "UPV - No Borders" of the Vicerrectorado Internacionalización y Comunicación" de la convocatoria 2021.