

<http://revistas.ucr.ac.cr/index.php/ingenieria>  
ISSN 2215-2652

# Ingeniería

Revista de la Universidad de Costa Rica

AGOSTO/DICIEMBRE 2013 - VOLUMEN 23 Número (2)



# LISTAS DE ESPERA EN LOS SERVICIOS PÚBLICOS DESDE UNA PERSPECTIVA DE LA LOGÍSTICA DE MANUFACTURA

*Ronny Pacheco Segura*

## Resumen

El uso de las técnicas de reducción de tiempos de entrega en la manufactura moderna, puede ser de gran utilidad para la gestión de las líneas de espera en los servicios públicos, sin embargo, la difusión de conceptos, aun los más sencillos, no se percibe que influyan en la administración pública costarricense y la forma como se administran los tiempos de espera en este país, pese a su éxito a nivel mundial.

**Palabras Clave:** servicios públicos, listas de espera, SMED, tiempos de entrega, secuenciamiento.

## Summary

The manufacturing lead time reduction techniques can be useful in the public services lines reduction, however, these concepts are not impacting the public administration and the management of public services lead time in Costa Rica, regardless the global success of similar applications.

**Keywords:** public services, waiting lines, SMED, lead time, production sequencing.

**Recibido:** 3 de Junio del 2013 • **Aprobado:** 16 de Setiembre del 2013

## 1 INTRODUCCIÓN

Los beneficios de la aplicación de los avances tecnológicos provenientes de la manufactura moderna en los servicios públicos, se confirma con la gran cantidad de casos de éxito que se encuentran en la literatura técnica referente a este tema. Administradores de servicios de salud, universidades y muchos profesionales en campos relacionados han puesto especial atención a esta interacción y el cómo realizar la adopción de estos conocimientos.

La tendencia industrial japonesa orientada a la eliminación de desperdicios y que fue forzada por las condiciones de un país destruido durante la segunda guerra mundial, ha marcado la pauta de la manufactura mundial durante años y los demás países, especialmente los occidentales, han tratado de imitar la bella sencillez de los

conceptos de manufactura japoneses, pero sobre todo lograr sus índices de productividad, así como la tiempos de entrega y eficacia con la que llevan sus productos al mercado.

Cuando se discuten estos temas con industriales poco familiarizados a los sistemas de manufactura modernos en Costa Rica, es difícil transmitir los conceptos, principalmente porque muchos de estos conocimientos son divergentes con algunos de las bases administrativas que han sido aceptadas tradicionalmente. Más aun cuando se trata de aplicar estos principios a los servicios, no solo sus responsables, muchos de ellos con conocimientos básicos de administración o ingeniería, sino que incluso estudiantes universitarios avanzados de ingeniería industrial enfrentan dificultades tratando de trasladar las ideas de la fábrica a la atención de personas.

**Cuadro 1.** Planta de Framingham de General Motors comparada con la planta de Takaoka de Toyota en 1987

	GM Framingham	Toyota Takaoka
Horas de ensamble por automóvil	31	16
Defectos de ensamble por cada 100 automóviles	135	45
Inventario de partes (Promedio)	2 sem	2 hrs

Fuente: Hill & Jones, 2005.

Este esfuerzo requerido para asimilar el concepto y tratar de aplicarlo en un ambiente muy diferente, vale la pena. Son muchos los ejemplos que pueden mencionarse donde se produce más con menos recursos, en menos tiempo. Pero no solo esto es lo impresionante, también lo son las brechas correspondientes, así cuando la productividad de algún competidor en el mercado cuadruplica a la de su oponente es porque se están haciendo las cosas significativamente mejor.

Diferentes autores han documentado el caso Toyota y su evolución:

Las consecuencias del sistema de producción de Toyota incluían un aumento en la productividad de la mano de obra y una disminución en el número de defectos por automóvil (Hill & Jones, 2005).

En el mismo documento, los autores citan a J.P. Womack, D.T. Jones y D. Roos, de su libro "La máquina que cambió el mundo".

En el caso de los servicios también es común encontrar líneas de espera, similares a las de las fábricas de manufactura: líneas de espera de pacientes por servicios de salud, de barcos por ser atendidos en los puertos, de productos por sus registros sanitarios, ciudadanos por trámites municipales o en los juzgados a la espera de resolución, son solo algunos ejemplos. Pero, ¿Cómo ayuda la mejora tecnológica de la línea de fabricación de automóviles japonesa, a la lista de espera por la sala de operaciones? La respuesta está relacionada con el concepto: "tiempo de entrega". En el ejemplo mencionado anteriormente, Womack (2005) muestra cómo

el inventario promedio de partes en la fábrica de Framingham era de 2 semanas (lo que significa una línea de espera de dos semanas para cualquier pedido nuevo), mientras que en la planta de Toyota en Takaoka, de apenas 2 horas (ver Cuadro 1).

Las oportunidades que ofrecen algunas herramientas de reducción de tiempos de entrega en manufactura y de sencilla implementación, para ayudar a reducir las listas de espera en los servicios públicos de Costa Rica, es el tema que ocupa este ensayo.

## 2. ¿CÓMO REDUCIR LOS TIEMPOS DE ESPERA?

Administrar manufactura es muy diferente a administrar servicios, en los segundos en vez de procesar materias primas, se "procesan" personas, lo que incluye una característica de inmediatez en la prestación. Los servicios no se pueden almacenar:

La interacción con el cliente a menudo afecta de forma negativa el desarrollo del proceso. Pero un servicio, por su misma naturaleza, implica que es necesaria alguna interacción y personalización (Heizer & Render, 2001, p. 260).

Los trabajos que involucran reducción en los tiempos de entrega no son nuevos y la industria de manufactura ha trabajado este tema desde hace muchos años, por esta razón, la base de la estructura en la mejora de los tiempos de entrega se sustenta en algunos conceptos enunciados

hace ya varios años. Los autores Hopp, Spearman y Woodruff (1990) publicaron un artículo llamado "Estrategias prácticas para la reducción del tiempo de entrega", donde definen cinco métodos para reducir estos tiempos. Estas se van a utilizar cómo base de la estructura de este artículo:

- Vigile el inventario en el proceso.
- Mantenga las cosas en movimiento.
- Sincronice la producción.
- Suavice el flujo de producción.
- Elimine la variabilidad.

A continuación se desarrollan estas metodologías con más detalle.

### 2.1 Vigile el inventario en proceso

En el tema de servicios y listas de espera el inventario es el equivalente a la cantidad de personas esperando por ser atendidas. Por esa razón, debería resultar fácil entender que cuantas más personas hay en la fila, más tiempo

va a tener que esperar el último en llegar para recibir el servicio. En manufactura se diría que cuanto más inventario, mayor va a ser el tiempo de entrega. Este apartado establece que es de vital importancia tener permanente conocimiento de la cantidad de personas en la fila, pero sobre todo, la toma de decisiones que debe estar relacionada.

Resulta en exceso común apreciar en los servicios públicos cómo existe una fila para un servicio o grupo de servicios específicos, mientras la ventanilla o grupo de ventanillas de al lado están desocupadas, o casi, que para el caso es lo mismo. Y sin embargo, no hay quién tome la decisión de balancear la carga. En manufactura moderna se podría implementar un sistema de alertas visuales, con una toma de decisión básica pre establecida. Se va a discutir esta propuesta con el ejemplo que se muestra en la Figura 1.

Se presentan dos líneas de espera para recibir dos servicios diferentes. Las posiciones de servicio uno, dos y tres atienden a los clientes en la línea uno y las posiciones de servicio cuatro

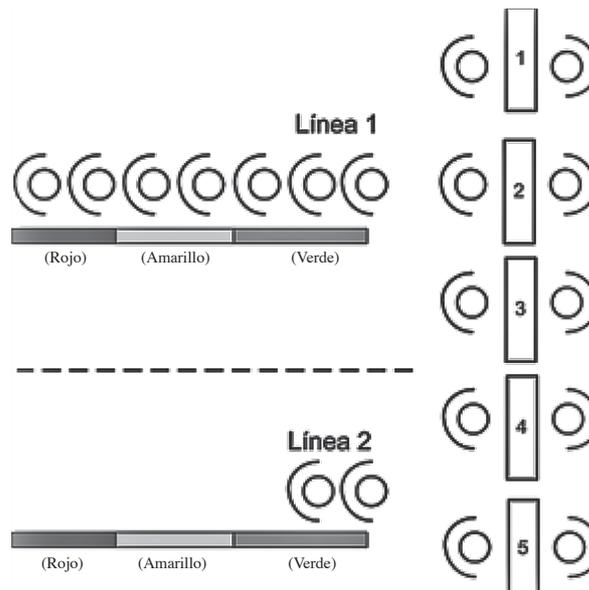


Figura 1. Ejemplo de líneas de espera con centros de atención diferenciados.

Fuente: Elaboración propia, 2013

y cinco atienden a los de la línea dos. Se puede observar, con base en el semáforo (Colores verde, amarillo y rojo colocados debajo de las líneas de espera), que existe un desbalance.

La identificación del desbalance sería mucho más difícil si no estuviera definido el semáforo, porque se podrían apelar situaciones como por ejemplo: que los trámites que se solicitan en la línea dos son más complicados y por lo tanto, tardan mucho más, o que la tasa de llegada a esa línea es mucho mayor y por lo tanto, se debe mantener una baja ocupación de la línea. Inclusive se escuchan explicaciones más livianas, como por ejemplo: "esta división ya estaba hecha, y mala suerte que les tocó tanto trabajo". En conclusión, el semáforo se convierte en la forma de visualizar la cantidad de inventario y por lo tanto, de poner en evidencia la situación imperante en cada momento.

En este sentido, la vigilancia del inventario consta de dos partes:

- Visualizar el inventario. Con base en las estimaciones de tiempo de atención y el espacio ocupado por una persona en la fila (puede ser un expediente, portafolio, un grupo de máquinas para ser reparadas u otra materialización que represente a la persona en la fila), establezca límites que representen tiempo, por ejemplo, si la tasa de atención es de diez minutos por persona, seis personas representan una hora, por lo tanto, la séptima persona en la fila, deberá esperar aproximadamente una hora antes de ser atendido.
- Establecer medidas y tomar decisiones. Con base en las políticas internas y la estrategia de la organización, o en muchos casos con base en lo que establecen reglamentos o legislación de la administración pública, establezca señales de alerta y límites, que le permitan tomar decisiones oportunas y evitar situaciones indeseables. Estas medidas pueden incluir:
  - a- Reforzar servicios críticos. Esto se logra estableciendo reglas para reforzar servicios reasignando personal automáticamente, extendiendo la jornada de atención con horas extra o más turnos. En el caso de que sea un problema per-

manente, se debe evaluar la posibilidad de simplificar las operaciones y automatizar. En el ejemplo de las dos líneas presentado anteriormente, el puesto de atención cuatro puede tener la directriz de que en caso de que se llegue al nivel rojo en la fila uno, se cambie de línea y en vez de atender la línea dos, atienda la que está sobrecargada, hasta que la cola llegue de nuevo al color verde, entonces, puede retomar su labor original.

- b- Trate de orientar la demanda. De igual manera se debe hacer con base en planificación e información. Podría generar promociones, publicidad, o asignación de citas, para orientar la demanda y que las personas soliciten los servicios en momentos del año, la semana o del día, en que se conoce con anterioridad, existe una menor ocupación del servicio debido a la estacionalidad de la demanda. Se pretende que de esta manera se logre inducir al cliente a solicitar el servicio en momentos en que hay menos demanda y se "descargue" el servicio en horarios o momentos de mayor ocupación, por la estacionalidad de la demanda.

Estas soluciones pueden aplicarse de igual forma cuando en lugar de dos tipos de ventanillas se tiene una ventanilla única. En este caso en vez de reforzar pasándose personal entre ventanillas, se haría con otro personal, quizá de oficina o las mismas jefaturas. Este personal de refuerzo puede encargarse de los casos más sencillos en líneas diferenciadas, si no poseen la experiencia y habilidad necesaria.

Para realizar las actividades aquí descritas se requiere que la organización desarrolle diferentes habilidades, como por ejemplo, establecimiento de estándares y políticas de servicio, se requiere conocer la demanda y lograr anticiparla de alguna manera. También sería necesario que los funcionarios tengan el conocimiento y disposición necesarios para poder ejecutar con éxito varios puestos de trabajo.

Todas estas son habilidades de la organización que deben formar parte de los

objetivos de crecimiento organizacional de las compañías o instituciones que aun no las poseen.

## 2.2 Mantenga las cosas en movimiento

El concepto implica que siempre la lista de espera debería estar siendo trabajada y en movimiento, reduciendo o eliminando aquellas actividades, decisiones o políticas que inducen a retrasos o demoras, en conclusión, obstáculos que aletargan la consecución de los objetivos.

Es frecuente escuchar en conversaciones referentes a reducción de tiempos de espera, argumentos como:

- Los funcionarios ya no pueden hacer el trabajo más rápido porque se afectaría la calidad y están haciendo su máximo esfuerzo, por lo tanto para producir más necesitamos más personal. (El personal)
- No podemos contratar más personal, el personal que tenemos debería ser suficiente para salir adelante con lo requerido. (La administración)

Esta discusión no lleva a un final que resuelva la situación presentada. Primero, porque los participantes no se están enfocando en el verdadero tema del problema; segundo, porque casi nunca tienen datos fehacientes de lo que está pasando, por lo tanto se convierte en un juego de poder y no en un ejercicio de resolución de problemas. En el mejor de los casos, se llega a un acuerdo de mejora tecnológica, que en muchos casos viene a ser una medida paliativa, que no ataca realmente la causa raíz del problema y automatiza malos hábitos de trabajo.

La mejora de procesos tradicional se enfocó en mejorar el método en la operación, en este caso se reducen los tiempos de las actividades productivas y la consecuencia de esa mejora siempre es requerir del operario una mayor producción. Esta práctica reduce el tiempo de actividades que agregan valor al producto o servicio.

La orientación japonesa de manufactura moderna es: "eliminar o reducir las actividades que NO agregan valor al producto". Es cada vez más frecuente leer autores que nos señalan este enfoque:

De los cuatro elementos del tiempo de entrega (operación, preparación, movimiento y tiempo en fila), los últimos dos pueden ser comprimidos en un buen diseño y práctica de control de actividades de producción (Volman, Berry, Whybark, & Jacobs, 2005, pp. 398 - 399).

El tiempo de operación se refiere al tiempo empleado propiamente en realizar la actividad que se requiere: en un banco, realizar la transacción bancaria, en un departamento de consulta externa, la revisión del médico. La preparación se refiere al tiempo requerido para preparar la estación de trabajo y generalmente se contabiliza desde que se finalizó la atención de un cliente, hasta que se inicia la atención del siguiente. Movimiento, es el tiempo que transcurre para el traslado de una estación de trabajo a la siguiente, en servicios pueden ser personas, documentos, o ambas. Algunas veces este tiempo de movimiento puede incluir una "espera para ser movido" que se puede estudiar de manera independiente. Por último, está el tiempo en fila, que evidentemente es el tiempo que espera el cliente

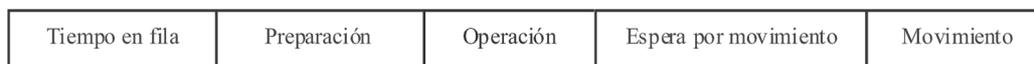


Figura 2. Ejemplo de líneas de espera con centros de atención diferenciados.

Fuente: Elaboración propia, 2013

ante una estación de trabajo previo a ser atendido, como se observa en la Figura 2.

En la Figura 2 no se pretende representar una proporcionalidad del tiempo que implica cada una de estas etapas, sin embargo para diferentes servicios podemos, con base en experiencias, suponer las proporcionalidades. Por ejemplo, en el banco, ¿cuántos minutos de fila se deben hacer en relación a los minutos de atención? En la cita médica, ¿cuántas horas o días se debe esperar entre la solicitud de la cita y la atención del médico, y cuál es su proporción con respecto al tiempo de atención?

Este apartado se preocupa por los momentos en que no hay movimiento y que afectan a la fila de espera: la preparación y la espera previa al movimiento.

2.2.1 *Tiempos de preparación*

En la mayoría de servicios pasar de un cliente al siguiente no representa mayor demora, se realiza casi de inmediato y sin mucho ajuste. Sin embargo, existen servicios dónde el cambio de una persona a otra ocurre considerando en el intermedio, un proceso de ajuste tardado y cuidadoso.

En los servicios hospitalarios esta condición es familiar, sucede con frecuencia. Así por ejemplo, las salas de cirugía suelen ser un recurso crítico junto con todo su personal, costoso y altamente calificado, por esto interesa cada minuto de utilización de éstas. Pero además, en muchos de los casos las personas que están haciendo fila para ser atendidos, se debaten en situaciones de vida o muerte, donde el tiempo de espera por la cirugía puede marcar la diferencia. En este caso,

los pacientes no están en movimiento, sino que están en espera.

Parte de la espera a la que se somete el paciente corresponde a la preparación que se debe dar de la sala de cirugía y los equipos, entre una cirugía y otra. Esta es la parte del tiempo de entrega al que se refieren Volman y sus co autores, cuando mencionan "preparación". Este tiempo ha sido estudiado en manufactura con una técnica conocida como SMED:

SMED (Single Minute Exchange of Dies) -

Un conjunto de técnicas ideadas por Shigeo Shingo para realizar en menos de diez minutos la preparación de una máquina para otro tipo de fabricación. Preparación al primer toque (One-touch set up) es la expresión empleada cuando la modificación se hace en menos de un minuto. Evidentemente el objetivo a largo plazo es reducir estos plazos de tiempo a cero, de forma que los cambios no obstruyan el flujo continuo. (Womack & Jones, 2005, p. 476)

Como se puede apreciar en esta definición, se refiere a una serie de técnicas para reducir los tiempos de preparación. En su mayoría son muy sencillas, sin embargo sumamente poderosas. La evolución de la mejora a través del tiempo, ha llevado a los sistemas productivos a tiempos de preparación de minutos, cuando originalmente se hablaba de horas. Básicamente el efecto que se quiere se muestra en la Figura 3.

Esta mejora en los tiempos de preparación se obtiene por medio de un proceso de mejora continuo, que inicia con algunas técnicas relativamente sencillas. La mayoría de los procesos de mejora relacionados con las



**Figura 3.** Cambio esperado con técnicas para mejorar los tiempos de preparación.

Fuente: Elaboración propia, 2013

técnicas japonesas de manufactura inician con la implementación de un programa de 5s: “5s es utilizada en organizaciones relacionadas con Lean 6 sigma, Justo a tiempo (JIT), Mantenimiento productivo total (TPM) y Administración de la calidad total (TQM), El principio es ampliamente aplicable no solo al piso de producción, sino en las oficinas también”. (Basu, 2009)

Sobre esta base y siguiendo con la conceptualización SMED, se presenta a continuación, una versión simplificada de los pasos de implementación de esta técnica propuestos por Basu (2009):

- Estudie y mida las operaciones de la línea de producción para discriminar:
  - a- “Set up interno” la operación que debe ser realizada mientras la máquina está detenida.
  - b- “Set up externo” la operación que puede realizarse mientras la máquina aun está corriendo.
- Elimine las actividades que no agregan valor y convierta operaciones de “Set up interno” en operaciones de “Set up externo”.
- Simplifique el diseño de la máquina, especialmente mecanismos de llenado y ajuste.
- Logre balance en el trabajo que debe realizarse y asegure el trabajo en equipo.
- Minimice pruebas y controles

### 2.2.2 División de lotes

La división de lotes está enfocada en disminuir la espera previa al movimiento de las personas u objetos que van de una estación de trabajo a otra.

En un servicio público de nivel nacional, la recepción de documentos estaba en el primer piso, mientras que la revisión de los documentos se encontraba en el segundo. La recepción de documentos tardaba aproximadamente una hora, mientras que la revisión podía tomar cuatro horas, por lo que había más revisores que receptores de los trámites. El trasiego de los expedientes estaba a cargo de una persona, quien realizaba un viaje a la semana, es decir, como su puesto se encontraba en el primer piso, una vez a la semana subía los expedientes recibidos y bajaba los expedientes revisados.

La consecuencia de esta decisión de hacer un viaje semanal, de la cual no tenían conocimiento las autoridades en la organización, causaba por si misma un atraso de dos semanas en el tiempo de entrega del servicio, tal y cómo se muestra en la Figura 4.

En este ejemplo, que era solamente parte de un problema más complejo que enfrentaba esta organización, se muestra cómo decisiones sobre el tamaño de lote a transferir, puede impactar el tiempo de entrega de los servicios. En este caso,



Figura 4. Ejemplo de tiempo de entrega delimitado por los lotes de traslado.

Fuente: Elaboración propia, 2013

los expedientes no se mantenían en movimiento, y una decisión al parecer sin importancia, generaba que el tiempo de entrega que perfectamente podía ser de dos días, tuviera que ser de al menos dos semanas y media.

### 2.3 Sincronice la producción

Probablemente la metodología de sincronización de la producción sea la que más se relacione con el concepto de justo a tiempo en ambientes no fabriles. Este concepto se refiere a tener lo que se necesita en el tiempo y en la cantidad que se necesita.

Ejemplos de falta de sincronía que afectan el tiempo de entrega se pueden apreciar constantemente en los servicios. Por ejemplo, en una institución como la Caja Costarricense de Seguro Social, se acostumbra programar grupos de citas a una hora en la mañana. Esto genera que las personas que son atendidas al final del período tengan que esperar. Así por ejemplo, si para el servicio de consulta externa se tiene un acuerdo de atención de 15 minutos por paciente, y en determinado centro médico se cuenta con 5 consultorios, donde se atienden a 4 personas cada hora (una persona cada 15 minutos), se asignarían 20 personas para el horario de las 7 de la mañana, si se atiende en bloques de una hora.

De esta manera, si el proceso se cumple según lo esperado, 5 pacientes harán fila por 45 minutos de manera innecesaria, porque fueron citados mucho antes de que fuera necesario.

Es importante cuando se trabaja con filas, asegurarse que se cuenta con la capacidad suficiente para satisfacer la demanda, y que no se está frente a un problema de recursos restringidos por la demanda, es decir un "Cuello de botella":

Entendemos que un cuello de botella constituye todo recurso que tiene una capacidad inferior a la demanda que se le impone (Chase, Jacobs, & Aquilano, 2005, p. 754).

Es evidente que si se tiene un recurso que no tiene la capacidad suficiente para satisfacer la demanda, generará una fila.

Algunas observaciones hechas en la industria han demostrado que la mayoría de las plantas tienen muy pocas operaciones que formen un cuello de botella (Chase, Jacobs, & Aquilano, 2005, p. 754).

Por esta razón es importante que la administración de los servicios identifiquen si tienen cuellos de botella dentro de su organización y se aseguren de poder administrarlos adecuadamente. Entre ellos procurar que se tenga un flujo continuo hacia el recurso restringido por la capacidad y evitar desperdicios de tiempo en ese recurso. De esta forma procurar que exista sincronía entre el resto de la organización y el recurso restricción. Además de que se debe centrar en este recurso los esfuerzos por aumentar la capacidad.

Muchos de los problemas de colas están relacionados con el orden en el que se programan los trabajos. En manufactura esta es una decisión que se toma diariamente, y se hace bajo conceptos relacionados a técnicas de control de actividades de producción. Dentro de esta serie de técnicas se encuentran las reglas de secuencia por prioridades. Una de ellas es la famosa "primero en entrar, primero en salir", que es la que tradicionalmente se aplica en la gran mayoría de servicios, cómo por ejemplo bancos, juzgados, y algunos servicios de salud. Generalmente se considera esta regla justa, incluso motivo de queja si no se asigna prioridad de esta manera, sin embargo, ¿esta es la mejor?

Los autores Chase, Jacobs y Aquilano (2005) indican que la regla del asignar primero los trabajos que requieren menos tiempo de operación es la mejor opción para disminuir el tiempo de espera y el tiempo promedio esperado para recibir atención, cuando el trabajo es realizado por un único centro de trabajo.

Podemos demostrar, matemáticamente, que la regla del SOT produce una solución óptima para el caso  $n/1$ , con otros criterios de evaluación como la media del tiempo de espera y la media del tiempo para terminar. De hecho, esta simple regla es tan potente que ha sido llamada "el concepto más

importante para el tema de secuenciado".  
(Chase, Jacobs y Aquilano, 2005, pag 697)

Bajo esta premisa se hace evidente que si se tiene una fila de espera para un puesto de trabajo específico, atender primero aquellos casos que requieren menos tiempo de atención, es una opción para reducir los tiempos de espera y por lo tanto la fila.

Dado que el concepto de primero en salir, primero en entrar es muy difícil de cambiar culturalmente, la recomendación en este caso es hacer filas separadas que diferencien la complejidad del servicio.

## 2.4 Suavice el flujo de producción

Dentro de los conceptos de manufactura moderna que más se discuten hoy en día esta el de suavizar el flujo de producción. Es difícil de lograr debido a que previo se requiere del aprendizaje y crecimiento de la organización en muchos aspectos, sin embargo, aquí se discutirán algunos temas relacionados.

El mejor reflejo del flujo continuo es procurar una salida continua y constante de clientes del servicio, y que esta sea consistente con la demanda en el momento específico. Esto hace necesario que las personas que trabajan en los servicios tengan un enfoque al logro y un proceso establecido para lograr el crecimiento cómo organización. Los responsables de los servicios deben procurar salidas frecuentes, para esto han de disminuir los lotes de trabajo hasta que se logre un avance de uno a la vez, y reducir al mínimo los defectos y reprocesos. Estos son temas que se discuten en otros apartados.

Para lograr este avance es importante que los responsables de los servicios tengan un buen conocimiento de la demanda de sus servicios y de la capacidad que debe tener la organización para responder rápidamente ante cambios relacionados. Este concepto está relacionado con cargas de trabajo y planificación de la capacidad. Si evaluar un crédito le toma a un analista 30 minutos, medido cómo esfuerzo del funcionario no como tiempo de entrega, este tiempo debe ser tomado en cuenta para planificar la capacidad de este puesto de trabajo que evalúa créditos.

Si se estima que durante el próximo periodo de un mes se va a requerir evaluar 100 créditos,

significa que se requerirán  $100 \times 0,5$  horas= 50 horas de analistas de crédito para esta función. Si se cuenta con un analista de crédito y este tiene 160 horas disponibles durante este periodo, significa que el funcionario tendrá una saturación esperada de 31,25% ( $(50/160) \times 100$ ).

Las cargas de trabajo deben planificarse, la capacidad que tendrá el servicio no debe dejarse al azar. Por razones de eficiencia, se quisiera una utilización elevada de los recursos, sin embargo, esto podría repercutir en clientes mal atendidos o inclusive no recibidos del todo. Por otra parte, personal sobrecargado, puede enfrentarse a situaciones de estrés físico o mental excesivo:

Para los empleados de tiempo completo, las cargas de trabajo es uno de los factores estresantes más citados (Macdonald, 2006) .

Según diferentes autores, por ejemplo Haywood – Farmer y La Nollet (1991), mencionados por Chase, Jacobs y Aquilano (2005), los recursos que se encargan de atender al cliente en los servicios deben planificarse a una saturación aproximada del 70% para lograr una adecuada calidad de atención. Sin embargo, esta decisión no se debe tomar de manera tan determinista, principalmente porque responde a diferentes condiciones estocásticas. Algunas de estas variables, como por ejemplo los tiempos de atención, corresponden a una variable aleatoria, de igual manera la demanda específica para un determinado momento del año, mes, semana, inclusive del día, puede variar considerablemente y corresponden a diferentes variables aleatorias. Por esta razón, se debe planificar con mucho cuidado la carga de trabajo que se espera tener en los servicios, para que responda a la variabilidad inherente, atienda los requerimientos del negocio y sobretodo, satisfaga las expectativas de los clientes al considerar la flexibilidad necesaria para lograrlo.

De esta manera, se debe tener control de las cargas de trabajo del personal y lograr nivelar el trabajo. Esta nivelación se refiere a que los costos de mantener los recursos se nivelen con la cantidad de trabajo que se espera que realicen estos mismos recursos. Una manera de realizar este balance son las técnicas de modelaje de costo basado en

actividades guiado por el tiempo. (Time driven activity based costing).

Ejecutivos pueden utilizar sus modelos de costeo basados en actividades guiados por el tiempo como el centro del análisis para pronosticar la demanda de recursos. Este pronóstico brinda a los ejecutivos la información para ajustar el suministro futuro de recursos, y además, los costos relacionados, para satisfacer sus demandas (Kaplan & Anderson, 2007, p. 86).

## 2.5 Elimine la variabilidad

En el caso de los servicios, al igual que en manufactura, la variabilidad es el centro de los problemas de calidad.

Se puede decir que la calidad abarca todas las cualidades con las que cuenta un producto o un servicio para ser de utilidad a quien se sirve de él (Cantú, 2006, p. 5).

Para lograr la calidad apropiada en cualquier producto o servicio es indispensable asegurarse que los procesos relacionados funcionen adecuadamente. Básicamente estamos hablando de procurar que la cantidad de errores y re trabajos se reduzcan al mínimo.

En esta materia, se estudia el estado de un proceso para estudiar su capacidad y estabilidad. Un proceso capaz es:

Proceso que cumple con las especificaciones de tal forma que el nivel de disconformidades es suficientemente bajo para garantizar que no habrá esfuerzos inmediatos para tratar de bajarlas y mejorar así su capacidad (Gutiérrez & De la Vara, 2004, p. 308).

Por su parte, la estabilidad de un proceso

"implica estudiar la variabilidad de un proceso a través del tiempo. Un proceso tiene estabilidad si su desempeño es predecible en el futuro inmediato y se dice que está en control (Gutiérrez & De la Vara, 2004, pag 310).

Se debe procurar que la gestión de los procesos ayude a disminuir la variabilidad. El presidente de la empresa Telefónica S.A. César Alierta Izuel, indica en su prólogo del libro *Las Claves de Seis Sigma*:

Los valores medios deben conocerse pero, para acoplarse a la psicología del cliente, es más importante obtener datos sobre la variabilidad de nuestros procesos: La variabilidad es el enemigo. A un cliente no le importa cómo de buenos somos de media, sino cómo de bien atendemos su caso particular (Pande, Neuman, & Cavanagh, 2002, p. 8 prólogo).

Existen diferentes sistemas y metodologías para diseñar, implementar y administrar proyectos de calidad, sin embargo, cuando se habla de calidad, siempre es bueno hablar de cultura y procesos de mejora sostenibles en el tiempo. Por esta razón es importante conforme se avance en el crecimiento de la organización, acercarse a metodologías robustas de mejora, por ejemplo, las conocidas como seis sigma.

Un sistema completo y flexible para conseguir, mantener y maximizar el éxito en los negocios. Seis sigma funciona especialmente gracias a una comprensión total de las necesidades del cliente, del uso disciplinado del análisis de los hechos y datos, y de la atención constante a la gestión, mejora y reinversión de los procesos empresariales (Pande, Neuman, & Cavanagh, 2002, p. 23 prefacio).

Existe otro tipo de variabilidad que tiene que ver con la falta de estandarización en los procesos y la toma de decisiones. Si la respuesta de los procesos no se logra estandarizar para obtener operaciones con tiempos sin mucha variación, los tiempos totales de respuesta del proceso se van a extender innecesariamente.

Para exponer este caso, se presenta un ejemplo de tiempos de servicio de dos líneas de producción que van a procesar dos servicios o productos de manera consecutiva, que se pueden apreciar en los Cuadros 2 y 3.

Para el procesamiento de cada uno de los productos se presentan diferentes tiempos, el

**Cuadro 2.** Ejemplo de línea de producción con alta variabilidad en la duración de las operaciones.

Línea 1	Duración (horas)	Producto 1		Producto 2		Tiempos muertos
		Tiempo inicio	Tiempo final	Tiempo inicio	Tiempo final	
Ope 1	3	0	3	3	6	
Ope 2	6	3	9	9	15	3
Ope 3	3	9	12	15	18	0
Ope 4	8	12	20	20	28	2
Ope 5	5	20	25	28	33	0
	25					5

Fuente: Elaboración propia, 2013

**Cuadro 3.** Ejemplo de línea de producción con menor variabilidad en la duración de las operaciones.

Línea 2	Duración (horas)	Producto 1		Producto 2		Tiempos muertos
		Tiempo inicio	Tiempo final	Tiempo inicio	Tiempo final	
Ope 1	4	0	4	4	8	
Ope 2	5	4	9	9	14	1
Ope 3	5	9	14	14	19	0
Ope 4	6	14	20	20	26	1
Ope 5	5	20	25	26	31	0
	25					2

Fuente: Elaboración propia, 2013

primero es “Duración” este representa el tiempo estándar de la operación. Luego se presenta el tiempo de inicio y final de cada operación. La diferencia entre estos dos tiene que ser igual a la “Duración”. Se supone que el producto pasa de una operación a la siguiente apenas termina su procesamiento en la operación anterior, siempre y cuando la operación destino esté desocupada. Si la operación destino no está desocupada, el producto tiene que esperar a que se desocupe, y se generan

así lo que denominaremos “Tiempos muertos”. Se puede apreciar que los tiempos estándar utilizados en cada línea son iguales, veinticinco horas cada uno, sin embargo, los tiempos muertos entre operaciones y la duración total para terminar los dos productos no son los mismos.

Los tiempos de atención de la primer línea se presenta en el Cuadro 2.

Se puede apreciar como en el producto dos, la operación dos inicia hasta la hora nueve,

pese a que la operación uno del producto dos terminó en el momento seis. Esto generaría un tiempo muerto de tres horas desde que termina la operación uno del producto dos e inicia la operación dos del mismo.

En este caso se tiene un total de cinco horas muertas, es decir un desperdicio de cinco horas en el producto dos. El efecto se incrementa cuando hay más productos, por esta razón si se extrapola a líneas de espera de cientos de personas o expedientes por ser procesados, se tendrán más horas desperdiciadas de las que se quisiera. Si el proceso tuviera menos variación, se tendría una situación como la que se observa en el Cuadro 3.

En este caso se puede apreciar que los tiempos muertos son solo dos horas, esto debido a que existe menos variación entre las operaciones. Además el producto dos en este segundo caso está terminado en la hora treinta y uno y no en la hora treinta y tres como en el primer caso, esto reduce en dos horas la conclusión con el producto dos y por lo tanto la espera por el mismo. Al terminarse antes, se puede atender al siguiente cliente antes, reduciendo la línea de espera, utilizando la misma cantidad de horas del personal.

Este ejemplo nos ilustra en una pequeña escala los efectos de la variación sobre los tiempos de entrega y por consiguiente su repercusión sobre las líneas de espera.

### 3. CONSIDERACIONES FINALES

Las líneas de espera son una realidad y en los servicios son de especial importancia. Sin embargo y pese a haber mucha literatura al respecto, la aplicación de herramientas de manufactura para mejorar el desempeño de los servicios no ha sido suficientemente utilizada en Costa Rica.

Se puede reducir la línea de espera de los servicios utilizando los mismos conceptos que se utilizan en manufactura para reducir los tiempos de entrega. Entre ellos se han mencionado cinco metodologías traducidas a servicios:

- Vigile constantemente el largo de la fila.
- Mantenga las cosas en movimiento.

- Sincronice la capacidad, las operaciones y la capacidad.
- Suavice el flujo de salida, cuide la capacidad de respuesta ante cambios en la demanda y reduzca los defectos y reprocesos.
- Reduzca la variabilidad.

A pesar de esto, todos los esfuerzos solo podrán implementarse y desarrollarse mediante el involucramiento gerencial. La convicción de que se puede mejorar y el esfuerzo sistemático para conseguirlo serán las claves del éxito de una mejora que impacte sosteniblemente, los indicadores de éxito de nuestros servicios.

### REFERENCIAS BIBLIOGRÁFICAS

- Basu, R. (2009). *Implementing Six Sigma and Lean. A practical Guide to Tools and techniques*. Victoria: Taylor & Francis.
- Cantú, H. (2006). *Desarrollo de una cultura de calidad*. Corea: Mc Graw Hill.
- Chase, R. B., Jacobs, F. R., & Aquilano, N. J. (2005). *Administración de la producción y operaciones para una ventaja competitiva* (10a Edición ed.). México: Mc Graw Hill Interamericana.
- Gutiérrez, H., & De la Vara, R. (2004). *Control estadístico de la calidad y seis sigma*. México D.F.: Mc Graw Hill.
- Heizer, J., & Render, B. (2001). *Dirección de la Producción, decisiones estratégicas* (6ta ed.). Madrid: Pearson educación.
- Hill, C., & Jones, G. (2005). *Administración estratégica*. México: Mc Graw Hill.
- Hopp, W., Spearman, M. L., & Woodruff, D. L. (1990). Practical strategies for lead time reduction. (A. S. Engineers, Ed.) *Manufacturing Review*, 3 (2), 78-85.
- Kaplan, R., & Anderson, S. R. (2007). *Time Driven Activity Based Costing*. Massachusetts, Boston, USA: Harvard Business School Press.
- Macdonald, W. A. (2006). Managing Workloads to Optimize Performance, Health and Well-Being. En W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors*. Victoria: Taylor and Francis Group, LLC.

- Pande, P., Neuman, R., & Cavanagh, R. (2002). *Las claves de seis sigma*. Madrid, España: Mc Graw Hill.
- Volman, T., Berry, W., Whybark, C., & Jacobs, R. (2005). *Planeación y control de la producción. Administración de la cadena de suministros*. México, DF, México: Mc Graw Hill.
- Womack, J. P., & Jones, D. T. (2005). *Lean Thinking*. Barcelona: Gestión 2000.

## **SOBRE EL AUTOR**

### **Ronny Pacheco Segura.**

Universidad de Costa Rica.

Máster en Ingeniería Industrial.

Docente en la Escuela de Ingeniería Industrial.

Correo electrónico: ronny.pacheco@ucr.ac.cr

Teléfono: 8827-9742

