

Sobre Hombros de Gigantes

Los Conceptos de Producción frente a Las Aplicaciones de Producción

El Ejemplo de *Hitachi Tool Engineering Ltd*

Eliyahu Goldratt

2008

Introducción

Es fácil establecer la conexión entre la popularidad de producción Lean con el éxito de Toyota. El éxito de Toyota es innegable. Toyota ahora fabrica tantos vehículos como el líder tradicional – GM – y lo hace generando utilidades. En los últimos cinco años, la utilidad neta promedio sobre ventas de Toyota ha sido un 70% más alta que el promedio de la industria, mientras que GM está perdiendo dinero.¹ El éxito de Toyota se atribuye completamente al Sistema de Producción Toyota (TPS: Toyota Production System).² Al menos esta es la convicción de la gerencia de Toyota – el reto número uno de Toyota que han hecho explícito es transferir el TPS como el DNA de la empresa a la siguiente generación.

¹ <http://moneycentral.msn.com/investor/invsb/results/compare.asp?Page=ProfitMargins&Symbol=TM>

² El Sistema de Producción Toyota se conoció primero mundialmente bajo el nombre de “Justo A Tiempo” (JIT: Just in Time) y luego como Producción Lean. Toyota misma afirma que Producción Lean no abarca completamente el espíritu de TPS, debido a las distorsiones en las comunicaciones e implementaciones.

Dado que Toyota es la bandera de la industria japonesa, uno esperaría que LEAN estuviese ampliamente implementada en Japón. Sorprendentemente este no es el caso. Es un hecho aceptado que menos del 20% de las empresas manufactureras del Japón han implementado LEAN. ¿Cómo puede ser esta la situación?

No se debe a que no hayan intentado implementarlo. Muchas empresas en el Japón han realizado serios esfuerzos para implementar LEAN, pero han fallado. Una de esas empresas es *Hitachi Tool Engineering Ltd.* Su inhabilidad para implementar LEAN no puede explicarse por la falta de seriedad en sus esfuerzos. Esta empresa intentó repetidamente implementar LEAN, pero el deterioro en el desempeño de su producción los obligó a volver a formas más tradicionales de gerencia de la producción.

De la misma manera, el hecho de que la mayoría de la industria japonesa no haya implementado LEAN no puede atribuirse a la falta de suficiente conocimiento. Toyota ha sido más que generosa en compartir su conocimiento. Esta compañía colocó todo el conocimiento de TPS en el dominio público e incluso llegó tan lejos como a invitar a sus competidores directos a que visitaran sus plantas. Hitachi, como muchas otras empresas, estaba utilizando el conocimiento disponible y no ponía límites para contratar ayuda de los mejores expertos disponibles.

Existe una explicación aparente al fracaso de estas empresas en la implementación de LEAN; una explicación que salta a la vista para cualquier observador objetivo de una empresa como *Hitachi Tool Engineering Ltd.* El fracaso se debe a una diferencia fundamental en los ambientes de producción. Cuando Taiichi Ohno desarrolló TPS, no lo hizo en abstracto; lo desarrolló para su empresa. No es una sorpresa entonces que la aplicación poderosa desarrollada por Ohno puede que no funcione en entornos de producción fundamentalmente diferentes.

Pero eso no quiere decir que el trabajo de Ohno no sea extremadamente valioso para otros ambientes. El genio de Ohno se revela completamente cuando nos damos cuenta que él estaba enfrentando exactamente la misma situación. En esa época, el sistema de producción que revolucionó la producción era el método de ensamble en línea, desarrollado por Henry Ford. El método de Ford ya se aplicaba no solo en casi todas las fabricas de automóviles, sino también en diferentes industrias como bebidas y municiones. También, en esa época, ya se aceptaba que el ensamble en línea podía y debía ser implementado solamente en entornos donde las cantidades

requeridas justificaran la dedicación de un equipo a un solo producto. En casos donde las cantidades a producir no fueran lo suficientemente altas, nadie consideraba la posibilidad de utilizar líneas. Nadie con la excepción de Ohno.

Ohno se dio cuenta de que los conceptos detrás del sistema de Ford eran genéricos; que su aplicación se limitaba a ciertos tipos de ambientes, pero que los conceptos eran universales. Ohno tuvo la clara visión de partir desde los conceptos, la genialidad para diseñar una aplicación que se adaptara al ambiente de Toyota, donde no es viable dedicar equipo a la producción de un componente, y la tenacidad para superar obstáculos enormes que se interponían a la implementación de semejante aplicación. El resultado es TPS.

En vez de rehusarnos a utilizar los conceptos correctos, o, aún peor, tratar de forzar la aplicación a ambientes que aparentemente son muy diferentes, deberíamos seguir los pasos de Ohno.

En este artículo se presentarán

- Los conceptos básicos de las cadenas de suministro – los conceptos sobre los cuales está basado LEAN.
- Una aplicación genérica de estos conceptos que puede ser utilizada en un rango mucho más amplio de ambientes, y
- Los resultados impresionantes obtenidos por *Hitachi Tool Engineering Ltd*, con esta aplicación más amplia.

Perspectiva Histórica

La industria manufacturera ha sido moldeada por dos grandes pensadores, Henry Ford y Taiichi Ohno. Ford revolucionó la producción en masa al introducir el ensamblaje en línea. Ohno elevó las ideas de Ford al siguiente nivel con su TPS, un sistema que obligó a toda la industria cambiara su forma de entender al inventario, pasando de considerarle un activo a ser visto como un pasivo.

El punto de inicio de Ford fue que la clave para una producción efectiva es concentrarse en mejorar el flujo global de los productos a lo largo de las operaciones.

Sus esfuerzos por mejorar el flujo fueron tan exitosos que, para 1926, el tiempo de entrega desde el momento de extraer el mineral de hierro hasta tener el automóvil completo, compuesto de más de 5,000 partes, en el tren listo para entregar, era de 81 horas.³ Ochenta años después, ningún fabricante de automóviles en el mundo ha sido capaz de lograr semejante tiempo de entrega tan corto, ni siquiera llegarle cerca.

Flujo significa que los inventarios en la operación se están moviendo. Cuando el inventario no se mueve, el inventario se acumula. La acumulación del inventario consume espacio. Por lo tanto, una forma intuitiva para lograr un mejor flujo es limitar el espacio disponible para acumular el inventario. Para obtener un mejor flujo, Ford limitó el espacio asignado para trabajo en proceso entre cada dos centros de trabajo. Esta es la esencia del ensamble en línea que puede verificarse en el hecho de que las primeras líneas de ensamble no tenían ningún medio mecánico, como bandas transportadoras, para mover el inventario de un centro de trabajo al siguiente.

La naturaleza osada del método de Ford se revela cuando uno se da cuenta que una consecuencia directa de limitar el espacio es que cuando el espacio asignado está lleno, los trabajadores que lo alimentan deben parar de producir. Por lo tanto, para poder lograr que haya flujo, Ford tuvo que abolir las eficiencias locales. En otras palabras, la línea de ensamble contradice directamente a la sabiduría convencional; forma usual de pensar, que para ser efectivo, cada trabajador y cada centro de trabajo deben estar ocupados el 100% del tiempo.

Uno puede pensar que impedir que los recursos trabajen continuamente va a reducir el tróput (el rendimiento) de la operación. Este efecto indeseable pudo haber sido el resultado si Ford se hubiese quedado satisfecho con solo limitar el espacio. Pero hay otro efecto que surge de limitar la acumulación del inventario. Hace que se hagan muy visibles los problemas reales que deterioran el flujo y poderles señalar– cuando un centro de trabajo en una línea para de producir durante más de un corto tiempo, pronto toda la línea se detiene. Ford sacó ventaja de esta clara visibilidad para balancear mejor el flujo, enfrentando y eliminando las paradas aparentes.⁴ El resultado final de abolir las eficiencias locales y balancear el flujo es un incremento sustancial en el tróput. Henry Ford logró el más alto tróput por trabajador que cualquier otro fabricante de automóviles de su tiempo.

³ Ford, Henry, *Today and Tomorrow*, Productivity Press, 1988 (originally published in 1926).

⁴ Balancear el flujo no es igual a balancear la capacidad – tener la capacidad de cada centro de trabajo igual a su carga – un error común que sucede cuando se balancean las líneas de ensamble.

En resumen, el ensamble en línea de Ford se basa en los siguientes cuatro conceptos:

1. Mejorar el flujo (o su equivalente tiempo de entrega) es un objetivo primario de las operaciones.
2. Este objetivo primario debe traducirse en un mecanismo práctico que guíe a la operación cuándo no producir (impedir la sobre producción).
3. Las eficiencias locales deben ser abolidas.
4. Hay que implementar un proceso de enfoque para balancear el flujo.

Al igual que Ford, el objetivo primario de Ohno era mejorar el flujo – reducir el tiempo de respuesta – como lo muestra en su respuesta acerca de qué es lo que está haciendo Toyota:

Todo lo que estamos haciendo es observar la línea de tiempo desde el momento en que el cliente nos hace un pedido hasta el punto en el que recolectamos el efectivo. Y estamos reduciendo este tiempo...⁵

Ohno enfrentó un obstáculo casi insuperable cuando quiso aplicar el segundo concepto. Cuando la demanda para un solo producto es alta, se justifica dedicar una línea a producir cada componente, como lo hizo Ford. Sin embargo, en esa época, en Japón, la demanda del mercado era de cantidades pequeñas para una variedad de automóviles. Por lo tanto, Ohno no podía tener líneas dedicadas en Toyota. Como ya lo hemos dicho, todas las otras industrias que enfrentaron esta situación sencillamente no consideraron la utilización de líneas. Ohno, sin embargo, estaba jugando con la idea de utilizar líneas aún cuando el equipo no fuese dedicado, cuando cada centro de trabajo esta produciendo una variedad de componentes. El problema era que en este caso, utilizar el mecanismo de limitar el espacio conduciría a atascos del tráfico – no todos los componentes están disponibles para ensamble (ensamblaje no trabaja) y al mismo tiempo el espacio asignado ya está lleno (las líneas alimentadoras ya no pueden trabajar).

Ohno escribe que encontró la solución cuando escuchó acerca de los supermercados (mucho antes de que realmente haya visto un supermercado en su

⁵ Ohno, Taiichi, *Toyota Production System*, Productivity, Inc. 1988, page ix (in Publisher's forward). Vale la pena mencionar que en este y sus otros libros, Ohno le otorga todo el crédito a Ford por los conceptos básicos.

visita a los EUA en 1956). Ohno comprendió que tanto los supermercados como las líneas alimentadoras en Toyota necesitaban manejar una gran variedad de productos. En los supermercados, los productos no se amontonaban en los pasillos, sino que la mayoría de la mercancía se mantenía almacenada en la bodega trasera. En la tienda misma, a cada producto se le asigna un espacio limitado en el anaquel. Solo cuando un producto es tomado por un cliente se realiza luego el resurtido desde la bodega del almacén, para llenar el espacio en el anaquel asignado para ese producto. Lo que Ohno visualizó fue el mecanismo que le permitiría guiar a Toyota para cuándo no producir. En lugar de utilizar un espacio único limitado entre los centros de trabajo para restringir la producción de trabajo en proceso, tuvo que limitar la cantidad a acumular de cada componente específico. Basado en este entendimiento Ohno diseñó el sistema Kanban.

El sistema Kanban ha sido descrito en numerosos artículos y libros. En este artículo vamos solamente a describir la esencia, para mostrar qué tanto se apegó Ohno a los conceptos fundamentales. Entre cada uno de dos centros de trabajo⁶, y para cada componente por separado, la acumulación de inventario se limita estableciendo un número determinado de contenedores y el número de unidades por contenedor. Estos contenedores, como todo contenedor en toda industria, contienen también la documentación relevante. Pero, una de las páginas de la documentación – usualmente una tarjeta (Kanban en japonés) – es una página que especifica solo el código y el número de unidades por contenedor, que se le da un tratamiento no convencional. Cuando el centro de trabajo siguiente retira un contenedor para ser procesado, esa tarjeta no se mueve con el contenedor, sino que se le entrega al centro de trabajo anterior. Esta es la notificación a ese centro de trabajo de que un contenedor se ha retirado, que el inventario asignado no está lleno. Solo en este caso el centro de trabajo anterior se le permite producir (un contenedor de partes especificadas en la tarjeta). En esencia, el sistema Kanban guía a cada centro de trabajo a cuándo y qué se debe producir, pero lo más importante es que dirige cuándo no producir. No hay tarjeta - no hay producción. El sistema Kanban es el mecanismo práctico que guía a la operación cuándo no producir (previene la sobre producción). Ohno tuvo éxito expandiendo los conceptos de Ford al cambiar la base del mecanismo, de espacio a inventario.

⁶ Para reducir el número de lugares donde es necesario mantener contenedores, Ohno utilizó Celdas en U intensivamente, en vez de tratar de usar centros de trabajo compuestos por un solo tipo de máquinas.

Adherirse al concepto de flujo obliga a la abolición de las eficiencias locales. Ohno se refirió a este asunto una y otra vez en sus libros, insistiendo que no tiene sentido estimular a las personas a producir, si los productos no se requieren en el muy corto plazo. Este énfasis fue probablemente la razón por la cual fuera de Toyota, el TPS se conoció primero como producción Justo-A-Tiempo.⁷

Una vez que el sistema Kanban – el sistema que le guía a la operación cuándo no producir – fue implementado en el piso de la planta, la reducción inmediata en el tróput hizo obligatorio un esfuerzo colosal para balancear el flujo. El reto enfrentado por Ohno fue de órdenes de magnitud mayor al que enfrentó Ford. Para comprender qué tan grande fue el reto, es suficiente con destacar solo un aspecto de muchos. A diferencia de los ambientes con líneas dedicadas, el sistema de Ohno obligaba a un centro de trabajo a cambiar frecuentemente su producción de un componente a otro. Para la mayoría de los centros de trabajo cada uno de estos cambios requería gastar tiempo para hacer la preparación requerida. Puesto que los contenedores, por diseño, necesitaban un número relativamente pequeño de partes resultando en lotes de producción, muchas veces ridículamente pequeños, comparados con el tiempo de ajustes y cambios requeridos. Inicialmente, para muchos centros de trabajo el tiempo requerido para las preparaciones y ajustes era mayor que el tiempo requerido para la producción, lo que resultaba en una caída significativa del tróput. No sorprende entonces que Ohno enfrentara una resistencia enorme – tanta así, que Ohno escribió que se referían a su sistema como “el sistema abominable Ohno”, desde finales de 1940 hasta los inicios de los 60.⁸ Ohno (y sus superiores) ciertamente tuvieron una determinación y visión extraordinarias para continuar empujando por la implementación de un sistema que, para cualquier persona que lo observara desde una perspectiva local, como la que tendría la mayoría del personal de la planta, simplemente no tenía sentido.

Ohno tuvo que allanar un nuevo camino para superar el obstáculo de las preparaciones. En esa época, y hasta que TPS se volvió famoso en el mundo entero, la forma tradicional de manejar las preparaciones era incrementar el tamaño del lote, “el tamaño económico del lote” fue el nombre popular bajo el cual se escribieron miles de

⁷ Sin embargo, en la literatura de Lean no hay un énfasis explícito en que TPS obliga la abolición de las eficiencias locales.

⁸ Ohno, Taiichi and Setsuo Mito, *Just-In-Time For Today and Tomorrow*, Productivity Press, 1988.

artículos.⁹ Ohno ignoró todo ese cuerpo de conocimiento debido a que ceder y utilizar cantidades “económicas” hubiera llevado al fracaso su lucha para reducir los tiempos de entrega. En su lugar, insistió en que las preparaciones requeridas no estaban escritas en piedra, que los procesos pueden modificarse para reducir drásticamente el tiempo de preparación requerido. Él lideró los esfuerzos para desarrollar e implementar técnicas de reducción del tiempo de preparación que eventualmente redujeron todos los tiempos de preparación en Toyota para que fueran, a lo mucho, unos pocos minutos.¹⁰ No es de extrañar entonces que LEAN esté asociado fuertemente con lotes pequeños y técnicas para la reducción del tiempo de preparación.

Pero la necesidad de balancear el flujo requería de mucho más que solo superar el obstáculo del tiempo de preparación. El hecho de que la mayoría de centros de trabajo no estaban dedicados a un solo componente hizo casi imposible detectar por medio de la observación directa los verdaderos problemas que ponen en peligro al flujo. Ohno estaba muy al tanto que había demasiadas cosas por mejorar, que sin una forma de enfocar los esfuerzos de mejora, iba a tomar mucho tiempo balancear el flujo.

El sistema Kanban le proveyó de la forma de hacerlo. La analogía de las rocas y del agua de LEAN es útil para entender cómo hacerlo. El nivel de agua corresponde al nivel de inventario, mientras que las rocas son los problemas que perturban el flujo. Hay muchas rocas en el fondo del río y toma tiempo y esfuerzo removerlas. La pregunta es cuáles rocas son las importantes a remover. La respuesta se da al reducir el nivel del agua; aquellas rocas que emergen por encima del agua son las que deben removerse. Al inicio del sistema Kanban, para lograr un tróput razonable, Ohno tuvo que comenzar con muchos contenedores, cada uno con una cantidad no despreciable de una pieza específica. Gradualmente, Ohno redujo el número de contenedores y luego las cantidades en cada contenedor. Si el flujo no se perturbaba de forma notoria, entonces la reducción del número de contenedores y de cantidades por contenedor continuaba. Cuando el flujo se interrumpía, se utilizaba el método de los cinco Por Qué

⁹ El primer artículo fue publicado por Ford W. Harris en in *Factory, The Magazine of Management*, Volume 10, Number 2, February 1913, pp. 135-136. Desde entonces se publican más artículos sobre este tema mensualmente.

¹⁰ Por ejemplo, los cambios de moldes en Toyota pasó de dos a tres horas en los 40's a menos de una hora y hasta 15 minutos en los 50's a 3 minutos en los 60's (Ohno escribió esto en su libro, *El Sistema de Producción Toyota*).

para identificar la causa raíz. El problema raíz debía resolverse antes de poder seguir reduciendo las cantidades. Tomó tiempo, pero el resultado final fue una mejora notable en la productividad.

Es necesario resaltar que aunque en los últimos veinte años todas las demás empresas de automóviles han implementado una versión u otra del sistema Toyota y cosechado grandes beneficios, la productividad de Toyota no ha sido alcanzada por ninguna otra empresa fabricante de automóviles. Este hecho destaca la importancia que tiene escoger correctamente el proceso de enfoque para los esfuerzos locales de mejora. Desafortunadamente, los esfuerzos de mejora en las otras empresas están mal orientados, debido a que están dirigidos a lograr reducción de costos en lugar de enfocarse totalmente a mejorar el flujo.

Ohno no invirtió tanto esfuerzo en reducir los tiempos de preparación para poder obtener algunos ahorros en costos. Si el ahorro en costos hubiese sido su objetivo, no hubiese “desperdiciado” el tiempo ganado, reduciendo aún más los lotes y por tanto tener que hacer muchas más preparaciones. Ohno no trató de reducir el número de partes defectuosas para poder ahorrar algunos costos (triviales); lo hizo para eliminar las mayores perturbaciones al flujo, que resultan de tener una parte defectuosa. Ohno ni siquiera trató de exprimir mejores precios de los proveedores de Toyota, o de cortar la nómina de Toyota (los dos principales elementos del costo); más bien dedicó toda su energía a mejorar el flujo.

Lo que está destruyendo esta imagen es que el resultado final de enfocarse en el flujo e ignorar las consideraciones de costos locales es un costo por unidad mucho más bajo. Exactamente igual a como el resultado final de abolir las eficiencias locales es una eficiencia mucho mayor de la fuerza de laboral. Si suena extraño, es debido a que los gerentes aún no han internalizado la diferencia conceptual de orientar las operaciones a concentrarse en mejorar el tróput en lugar de concentrarse en reducir el costo. Una de las ramificaciones de concentrarse en la reducción de costos es que casi todas las iniciativas para promover un proceso de mejora continua rápidamente alcanzan un punto de retornos decrecientes, y como resultado, muchas de ellas se deterioran hasta terminar en simples palabras vacías. Pero ese tema es muy amplio e importante para ser metido a la fuerza en este artículo.

En resumen, tanto Ford como Ohno siguieron cuatro conceptos (a partir de ahora nos referiremos a ellos como los conceptos de la cadena de suministros):

1. Mejorar el flujo (o su equivalente tiempo de entrega) es un objetivo primario de las operaciones.
2. Este objetivo primario debe traducirse en un mecanismo práctico que guía a la operación a cuándo no producir (previene la sobreproducción). Ford utilizó espacio; Ohno utilizó inventario.
3. Las eficiencias locales deben ser abolidas.
4. Debe implementarse un proceso de enfoque para balancear el flujo. Ford utilizó la observación directa. Ohno utilizó la reducción gradual del número de contenedores y luego la reducción gradual de las partes por contenedor.

Las fronteras de TPS

El enfoque de Ohno para desarrollar LEAN demuestra una idea importante: existe una diferencia entre una aplicación y los conceptos fundamentales sobre los cuales esta basada la aplicación. Los conceptos fundamentales son genéricos; la aplicación es la traducción de los conceptos para un ambiente específico. Como ya hemos visto, la transferencia no es trivial y requiere de un número de elementos de solución. Lo que tenemos que tener en mente es que la aplicación hace supuestos (algunas veces son supuestos ocultos) acerca del entorno. No debemos esperar que una aplicación funcione en ambientes para los cuales los supuestos no son válidos. Podemos ahorrarnos mucho esfuerzo y frustraciones si nos molestamos en verbalizar explícitamente estos supuestos.

El supuesto más exigente que TPS hace del ambiente de producción es que es un ambiente estable. Y exige estabilidad en tres aspectos diferentes.

El primer aspecto se revela cuando prestamos atención al hecho de que, incluso cuando se selecciona un ambiente apropiado y los mejores expertos están supervisando la implementación, la implementación de LEAN toma un tiempo considerable. Liker señala en el libro *The Toyota Way* que la implementación de LEAN liderada por el *Toyota Supplier Support Center (TSSC)* (Centro de Apoyo Toyota al Proveedor - la organización creada por Toyota para enseñarle TPS a las empresas de EUA), tomó un mínimo de seis a nueve meses por línea de producción.¹¹ Esto no es sorpresa para cualquiera conozca el número de perturbaciones al flujo que existe casi en cualquier entorno de producción, y la sensibilidad del sistema Kanban, una vez que

¹¹ Liker, Jeffrey K., *The Toyota Way*, McGraw-Hill, 2004.

comienza a alcanzar su objetivo de bajo inventario. Debido a que el sistema Kanban toma tiempo en ser implementado, el supuesto es que el ambiente es relativamente estable - que los procesos y los productos no cambian significativamente durante un tiempo considerablemente largo.

Toyota disfruta de un entorno relativamente estable. La industria del automóvil permite cambios solamente una vez al año (el cambio anual del modelo) y usualmente, de un año al otro, la gran mayoría de los componentes son los mismos. Este no es el caso para muchas otras industrias. Por ejemplo, en grandes sectores de la industria electrónica, el horizonte de vida de la mayoría de los productos es menor a seis meses. Hasta cierto grado, la inestabilidad de los productos y procesos existe en la mayoría de otras industrias. Por ejemplo, *Hitachi Tool Engineering Ltd.* produce herramientas de corte, un tipo de producto relativamente estable, pero la competencia feroz obliga a esta empresa a lanzar nuevas herramientas de corte, que requieren de nueva tecnología, cada seis meses. Es una tarea de Sísifo implementar LEAN en un ambiente como este.

Un segundo aspecto de la estabilidad requerida por TPS es la estabilidad de la demanda en el tiempo por producto. Supongamos que el tiempo de respuesta para producir determinado producto es de dos semanas, pero la demanda para ese producto es esporádica; en promedio solo hay un pedido para este producto cada trimestre. Actualmente, este producto contribuye al trabajo-en-proceso solo durante dos semanas de cada trimestre; el resto del tiempo no se encuentra presente en el piso de la planta. Pero ese no sería el caso bajo LEAN, que obliga a mantener permanentemente contenedores para cada producto entre cada dos centros de trabajo.

En *Hitachi Tool Engineering Ltd.* están produciendo cerca de veinte mil SKUs diferentes. Para la mayoría de los SKUs la demanda es esporádica. La necesidad de mantener permanentemente, para cada SKU, inventario entre cada dos centros de trabajo conduciría, en el caso de Hitachi, a mantener un inventario de trabajo en proceso considerablemente mayor a lo que tienen hoy en día. Esto hace evidente que este no es un entorno adecuado para la aplicación de Ohno.

Pero el aspecto más exigente de la estabilidad requerida por TPS es la estabilidad en la carga total colocada por las órdenes sobre los diferentes tipos de recursos. Supongamos que, como en la mayoría de empresas, las órdenes no son

uniformes a lo largo del tiempo. Es muy probable que la carga colocada esta semana sobre un centro de trabajo específico sea muy inferior a su capacidad, mientras que la carga colocada la semana siguiente es ligeramente mayor a su capacidad. En este caso muy común, el sistema Kanban, que está para prevenir anticipar la producción – impidiendo producir antes de tiempo - va a llevar a incumplir las fechas de entrega en la segunda semana. Las órdenes de Toyota son relativamente estables y sin embargo, Toyota tuvo que establecer un mecanismo para recibir las órdenes (y promesas de entrega) que restringe el cambio en la mezcla de productos de un mes al otro. La mayoría de las empresas no son capaces de forzar a sus clientes a cumplir con condiciones semejantes tan favorables.

Es importante destacar que la estabilidad requerida está fuera del área de poder de Producción para mejorarla. Los tres aspectos de la estabilidad están relacionados directamente con la forma como la empresa diseña y vende sus productos, y no con la forma como los produce. Desafortunadamente, la mayoría de empresas sufren de al menos un tipo de inestabilidad, si no es que tienen los tres.

Lo anterior no significa que, para ambientes en donde los supuestos de LEAN no son válidos, no se pueda utilizar fragmentos de LEAN (por ejemplo, las celdas de producción en U pueden ser útiles en muchos entornos y las técnicas de reducción de los tiempos de preparación se pueden usar casi en cualquier ambiente). Pero significa que en estos ambientes, uno no debería esperar lograr la misma magnitud de resultados alcanzados por Toyota – resultados que elevaron a esta compañía al lugar donde se encuentra actualmente. Utilizar algunas técnicas específicas de LEAN, quedar satisfecho con algunos programas de ahorros de costos, no debe considerarse como implementar LEAN.

La importancia del flujo en entornos relativamente inestables

Ford y Ohno abrieron nuestros ojos al hecho de que un mejor flujo – reducir el tiempo de respuesta – conduce a unas operaciones mucho más efectivas. Ellos lo demostraron en entornos estables, pero ¿cuál es el impacto de mejorar el flujo en entornos relativamente inestables?

El primer aspecto de la inestabilidad es la inestabilidad debida a la vida corta del producto. Cuando la vida del producto es corta, la sobreproducción puede volverse obsolescencia. Más aún, dado que el tiempo de vida es corto, los tiempos largos de producción llevan a incumplir con la demanda del mercado. Por ejemplo, supongamos que el tiempo de vida de un producto es cercano a los 6 meses, y el tiempo de producción de ese producto es de dos meses. El tiempo largo de producción resulta en ventas perdidas, no debido a que no haya demanda, sino que, para un período importante de tiempo, la producción no puede satisfacer la demanda.

El segundo aspecto de la inestabilidad es la inestabilidad en la demanda en el tiempo de cada producto. La práctica común en los ambientes que tienen un elevado número de SKUs que están sujetos a una demanda esporádica, es reducir las complicaciones al tratar de satisfacer esta demanda a partir de inventario. La desventaja de esta práctica son los altos inventarios de producto terminado que rotan muy lentamente, junto con un alto nivel de agotados. Un sistema de producción que es capaz de organizar el piso de la planta hasta el punto en que se alcanza un flujo mucho mejor tiene un impacto drástico sobre estos ambientes.

Los ambientes que sufren del tercer aspecto de inestabilidad – inestabilidad en la carga total – son aquellos que pueden ganar más de tener un mucho mejor flujo. Las sobrecargas temporales sobre los distintos recursos llevan a estas empresas generalmente a tener un pobre desempeño en el cumplimiento de las fechas de entrega prometidas (<90%) y, como resultado, están propensas a agregar más capacidad. La experiencia indica que cuando estas empresas tienen éxito en mejorar drásticamente el flujo, no solamente el cumplimiento en las fechas de entrega prometidas alcanza los altos noventa, sino que se revela capacidad, a veces tanto como un 50%.¹²

Ohno había demostrado que los conceptos introducidos por Ford no se limitan a la producción en masa de un único tipo de producto. Aún cuando los obstáculos para aplicar estos conceptos en un ambiente menos restrictivo parecían insuperables, el genio y la tenacidad de Ohno nos demostraron no sólo que podía lograrse, sino cómo lograrlo.

¹² Mabin, Victoria J. and Balderstone, Steven J., *The World of the Theory of Constraints*, CRC Press LLC, 2000. Una revisión de la literatura internacional sobre TOC analizó los resultados promedio alcanzados: 70% de reducción en el tiempo de entrega, 44% de mejora en cumplimiento de fechas de entrega prometidas y un 76% de incremento en ingresos / tróput / utilidades.

Ahora entendemos que:

- TPS está limitado a entornos relativamente estables.
- La mayoría de entornos sufren de inestabilidad, y
- Los entornos relativamente inestables tienen mucho más que ganar de una mejora del flujo que los entornos estables.

Ahora que nos hemos dado cuenta de lo anterior, ¿No deberíamos seguir los pasos de Taichi Ohno? ¿No deberíamos volver a los conceptos de la cadena de suministros y derivar una aplicación efectiva que sea adecuada para los entornos relativamente inestables?

Una aplicación basada en el tiempo de los conceptos de la cadena de suministros

La base más intuitiva para el mecanismo de restringir la sobre producción no es el espacio o el inventario, sino el tiempo – si uno quiere impedir que se produzca antes de tiempo uno no debe liberar el material antes de tiempo. Utilizar el tiempo como la base no solo es más intuitivo y, por lo tanto, más fácilmente aceptado por el piso de la planta, sino que tiene la ventaja de que lo hace adecuado para los entorno inestables – es mucho menos sensible a las perturbaciones al flujo.

La robustez del mecanismo basado en el tiempo surge del hecho que limita directamente la cantidad de carga total de trabajo en el sistema, en vez de hacerlo limitando la cantidad de trabajo entre cada dos centros de trabajo. En las líneas de ensamble o en los sistemas basados en Kanban, el inventario asignado entre los centros de trabajo se limita al mínimo posible (usualmente corresponde a mucho menos de una hora de trabajo). Por tanto, cuando un centro de trabajo está detenido por más de un tiempo breve, los centros de trabajo siguientes se quedan casi inmediatamente sin trabajo y los centros precedentes quedan “bloqueados” para trabajar. Cuando para cualquiera de los centros de trabajo el tiempo acumulado consumido por no trabajar y por bloqueos es mayor que la capacidad excedente de ese centro de trabajo, se reduce el tróput de la empresa. La sensibilidad de las líneas de ensamble y de los sistemas basados en Kanban surge del hecho de que una

perturbación que ocurra en un centro de trabajo consume capacidad también de los centros de trabajo que quedan antes y después – un fenómeno que (casi) no existe en los sistemas basados en el tiempo, dado que el trabajo, una vez liberado en el piso de la planta, no se restringe artificialmente.

La dificultad en utilizar un sistema basado en el tiempo es que, para cada pedido, deberíamos restringir la liberación del material correspondiente para que sea un tiempo apropiado antes de la fecha de entrega prometida del pedido. Pero, ¿cómo hace uno para calcular el tiempo apropiado?

Cuando los computadores aparecieron en el escenario industrial (a comienzos de los sesenta) parecía que teníamos, al fin, la herramienta adecuada para manejar la inmensa cantidad de detalles y de cálculos necesarios para determinar los tiempos apropiados para cada material y cada orden. En unos diez años se desarrollaron muchos programas de computador para hacer precisamente esto, en numerosas empresas alrededor del mundo. Desafortunadamente los resultados esperados de un mejor flujo y de un menor trabajo-en-proceso no se materializaron.

El problema es que el tiempo que toma convertir a el material en un producto terminado, listo para ser entregado al cliente, depende más del tiempo que tiene que esperar en cola (esperando por un recurso que está ocupado procesando otro pedido o esperando en frente de ensamblaje a que llegue una pieza), y no tanto del tiempo de toque en procesar la orden. Es un hecho común, que en casi cualquier operación industrial (a excepción de las líneas de proceso y de las empresas que utilizan el sistema Kanban) el tiempo que toma un lote de partes en procesarse es solo cerca de un 10% del tiempo total de producción. Como resultado, la decisión de cuándo liberar el material determina en dónde y qué tan grandes van a ser las colas, lo que a su vez determina qué tanto tiempo va a tomar completar el pedido, lo que determina cuándo liberar el material. Estamos enfrentando un problema del huevo y la gallina. En los setenta se sugirió manejar el problema a través de reiterar el procedimiento (MRP de ciclo cerrado) – para correr el sistema de computo, para revisar las sobrecargas planeadas resultantes en los distintos recursos (el tamaño de las colas), para ajustar las fechas de entrega prometidas y eliminar las sobrecargas, y para repetir este proceso hasta que todas las sobrecargas significativas fueran eliminadas. Esta sugerencia no duró ya que la experiencia demostró que el proceso no converge: que no importaba

cuántas iteraciones se hagan, las sobrecargas finalmente se mueven de un tipo de recurso a otro.

Como resultado, ya en los setentas, el uso de estos sistemas de computo no era para guiar la liberación precisa y oportuna del material al piso de la planta, sino que se limitaba a ofrecer una mejor información sobre las cantidades (y la oportunidad) para pedir el material a los proveedores. El nombre oficial de estos sistemas se acuñó reflejando su uso principal: Planeación del Requerimiento de Materiales (MRP = *Material Requirements Planning*)¹³.

El hecho de que un esfuerzo descomunal como este no condujera a un mecanismo práctico basado en el tiempo para guiar a las operaciones sobre cuándo no producir, no debe considerarse como una prueba de que un mecanismo como éste no se pueda desarrollar para entornos menos estables – entornos que deben cumplir con las fechas de entrega prometidas con un flujo no uniforme de pedidos de los clientes. Ni tampoco debe desalentarnos de intentar utilizar el tiempo como la base del mecanismo práctico. Pero debe ser una advertencia de no tratar de desarrollar el mecanismo a partiendo del manejo de una inmensa cantidad de detalles y de cálculos. Lo que se requiere es de un enfoque de una visual de ángulo más amplio.

Volviendo a lo básico, siguiendo los conceptos de la cadena de suministros, el objetivo es mejorar el flujo – reducir el tiempo de respuesta. El considerar al tiempo (en vez del espacio o el inventario) como la base para el mecanismo que guía a la operación sobre cuándo no producir, obliga a que debamos insistir en liberar el material respectivo en un tiempo adecuadamente corto, justo-a-tiempo, antes de la fecha de entrega prometida de esa orden. Pero, ¿qué queremos decir con “justo-a-tiempo”? Aun cuando el término “justo a tiempo” es un concepto clave en LEAN, su uso es figurativo y no cuantitativo. En LEAN, cuando se habla de producción justo a tiempo no se refiere precisamente a que la parte que se acaba de procesar se necesita que esté en la plataforma de despacho lista para en el siguiente segundo,... o minuto... u hora. Es muy probable, incluso bajo el mejor sistema de Kanban, que esta parte no se procese inmediatamente por el centro de trabajo siguiente (como puede deducirse del hecho de que contenedores completos generalmente están esperando entre los centros de trabajo). Entonces, ¿qué intervalo de tiempo debemos considerar como justo a tiempo? Más explícitamente: si queremos restringir la sobre producción

¹³ Orlicky, Joseph, *Material Requirements Planning*, McGraw-Hill Book Company, 1975.

restringiendo la liberación de material, ¿qué tanto tiempo antes de la fecha de entrega de un pedido debemos liberar el material para ese pedido?

Una forma de llegar a una respuesta razonable es examinando el impacto que tiene la elección de este intervalo de tiempo en la magnitud de atención gerencial requerida para cumplir con todas las fechas de entrega prometidas. Supongamos que liberamos el material antes de la fecha de entrega prometida con solo el tiempo suficiente que se necesita para procesar el pedido. La elección de este período de tiempo va a requerir de gran atención gerencial para dar seguimiento de cerca a las operaciones, debido a que cualquier atraso en cualquier operación, o incluso un atraso en mover las partes entre las operaciones, va a resultar en un atraso sobre la fecha de entrega prometida. Más aún, se va a requerir de una programación de producción precisa para asegurar que no vayan a existir colas, puesto que cualquier cola causa un atraso en las partes que están esperando en la fila. Definitivamente esta no es una elección práctica; aún con atención gerencial infinita no va a ser suficiente para cumplir con todas las fechas de entrega prometidas. Debemos escoger un intervalo de tiempo mayor; un intervalo de tiempo que contenga seguridad para ajustar por los atrasos. La necesidad de incluir seguridad es la razón para referirse a este intervalo de tiempo para la liberación del material antes de la fecha de entrega prometida como el “amortiguador de tiempo (time buffer)”.

La selección de amortiguadores de tiempo más largos hace que sea mayor el tiempo de respuesta e incrementa el trabajo en proceso, pero en vista que los tiempos de amortiguador más largos significan más tiempo de seguridad, las expectativas son que, con mucho menos tiempo gerencial, un porcentaje más alto de órdenes van a completarse en o antes de sus fechas de entrega prometidas. Esto es correcto para amortiguadores de tiempo relativamente cortos, pero cuando los amortiguadores de tiempo son considerables, comienza a levantar su cabeza de otro horrible fenómeno. Lo que debemos tener en cuenta es que entre más largo sea el tiempo de amortiguador elegido, más temprano es liberado el material, lo que significa que más órdenes están presentes de forma simultánea en el piso de la planta. Cuando hay muchas órdenes en el piso de la planta, las congestiones y atascos de tráfico comienzan a ocurrir. Mientras más congestiones de tráfico, más se requiere de atención gerencial para establecer las prioridades. La magnitud de la atención gerencial requerida como una función de la longitud del tiempo de amortiguador seleccionado, se representa de forma esquemática en la Figura 1.

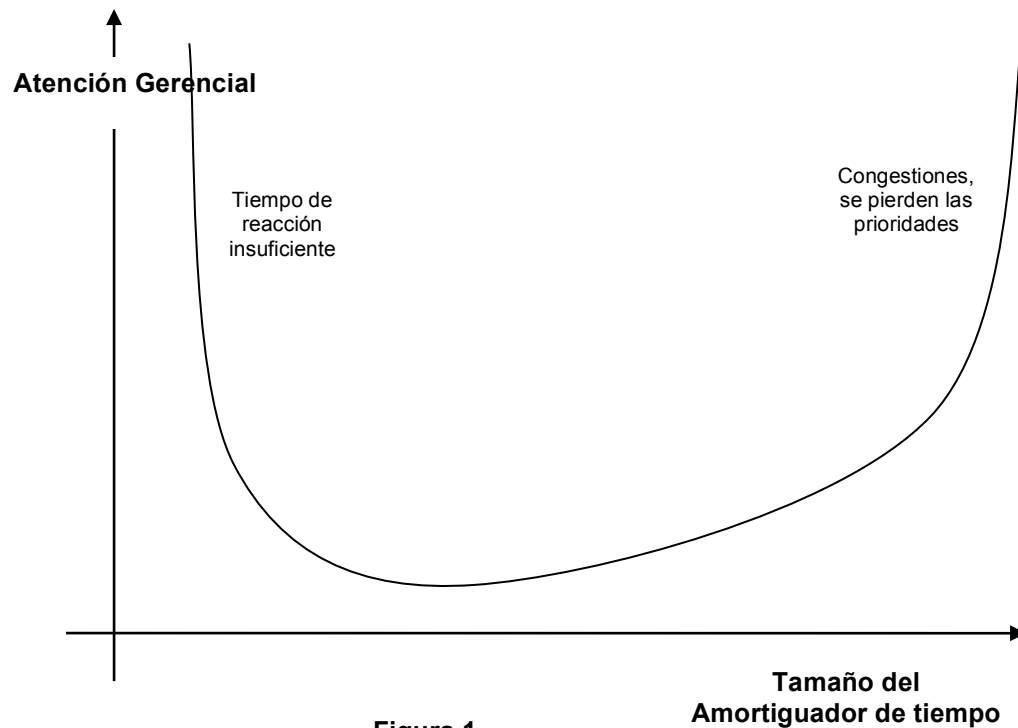


Figura 1

Las operaciones que implementaron los sistemas de Ford y de Ohno disfrutaron un tiempo de respuesta promedio que es solo unas pocas veces mayor al tiempo de toque actual, y la gerencia tiene que invertir casi nada o muy poca atención para orientar al personal del piso de la planta sobre qué debe trabajar ahora. Ellos definitivamente residen del lado izquierdo de esta gráfica, en la parte baja de la meseta. Pero, ¿en qué lugar de la gráfica se encuentran la gran mayoría de las operaciones, aquellas que están utilizando prácticas más convencionales?

Como hemos mencionado, en plantas convencionales los lotes de partes consumen solo cerca de un 10% de su tiempo en ser procesadas. Alrededor del 90% del tiempo los lotes están, o bien esperando en cola por un recurso, o esperando por otro tipo de parte para ser ensambladas. Lo que aprendimos de Ford, y más aún de Ohno, es que no debemos aceptar como dado el tamaño de los lotes; que los tamaños económicos de lote no son económicos, y que en cambio debemos y podemos alcanzar el flujo de-a-una-sola pieza. Armados con esta convicción es fácil comprender que cuando un lote de partes está siendo procesado (excepto en procesos como

curado o mezclado) se está realmente trabajando en una unidad solamente, mientras que las demás unidades del lote están esperando. Esto significa que en empresas convencionales que usan tamaños de lote de más de diez unidades en un lote (lo cual es el caso en la mayoría de los ambientes de producción), el tiempo de toque es realmente menos de 1% del tiempo de respuesta. Existe otro fenómeno que caracteriza a estas empresas: cualquiera sea el sistema de prioridades, si es que existe un sistema formal de prioridades del todo, el sistema real de prioridades es: “caliente”, “al rojo vivo” y “deje todo y hágalo ya”. Estas empresas parecen estar en la parte alta del lado derecho, sobre la pendiente inclinada, de la gráfica de la atención gerencial vs. el amortiguador de tiempo (Figura 1).

Estar en el extremo derecho sobre la pendiente significa estar en una situación de perder-perder: los tiempos de respuesta son muy largos (en relación con el tiempo de toque), los inventarios son elevados y en muchos casos la empresa sufre de un pobre desempeño en el cumplimiento de las fechas de entrega prometidas (<90%), a pesar de la alta atención gerencial. Teniendo en mente que si la gerencia hubiese escogido un tiempo de amortiguador más corto (moverse a la amplia región plana de la gráfica, la meseta), la situación sería considerablemente mejor, ¿cómo es que la gran mayoría de empresas gerenciadas de forma convencional están en esta situación de perder-perder?

La respuesta fue dada por Ford y Ohno. A través de su trabajo, decisivamente, demostraron que contrario a la creencia común, luchar constantemente por activar todos los recursos todo el tiempo no es una receta para tener unas operaciones efectivas. Por el contrario, lo exactamente opuesto es cierto: para lograr tener operaciones efectivas se deben abolir las eficiencias locales. Pero las empresas convencionales tratan de alcanzar una activación total de los recursos. Siempre y cuando los recursos aguas arriba no sean cuellos de botella (y este es el caso en la gran mayoría de entornos), se quedarán, de cuando en cuando, sin tener qué trabajar. Para impedirlo, se libera material; material que se requiere para órdenes más lejanas (incluso para órdenes pronosticadas). La consecuencia inevitable es colas más largas. Las colas más largas ocasionan que algunas órdenes no se completen a tiempo lo que a su vez se interpreta como: debemos liberar el material aún más temprano. Y también se interpreta como: no tenemos capacidad suficiente. No es difícil entonces imaginar cómo estas fuerzas empujan a las compañías a subir por la pendiente.

Un buen punto de partida para mejorar el flujo será escoger el amortiguador de tiempo que es igual a la mitad del tiempo de producción o respuesta actual; una elección como ésta asegura que la empresa se ubique en alguna parte de la meseta de la gráfica. No tiene sentido perder el tiempo tratando de encontrar el punto óptimo, los beneficios inmediatos son muy significativos para posponerlos, y los esfuerzos siguientes para balancear el flujo van a modificar la gráfica misma.

Restringir la liberación del material para que sea sólo el tiempo de amortiguador (la mitad del tiempo de respuesta actual) antes de la fecha de entrega prometida de estas órdenes va a mejorar considerablemente el cumplimiento de las fecha de entrega prometidas, va a reducir el tiempo de entrega a la mitad de lo que es actualmente, y por lo tanto a medida que el inventario en exceso se evacua, reduce el inventario en proceso a menos de la mitad de su nivel actual.

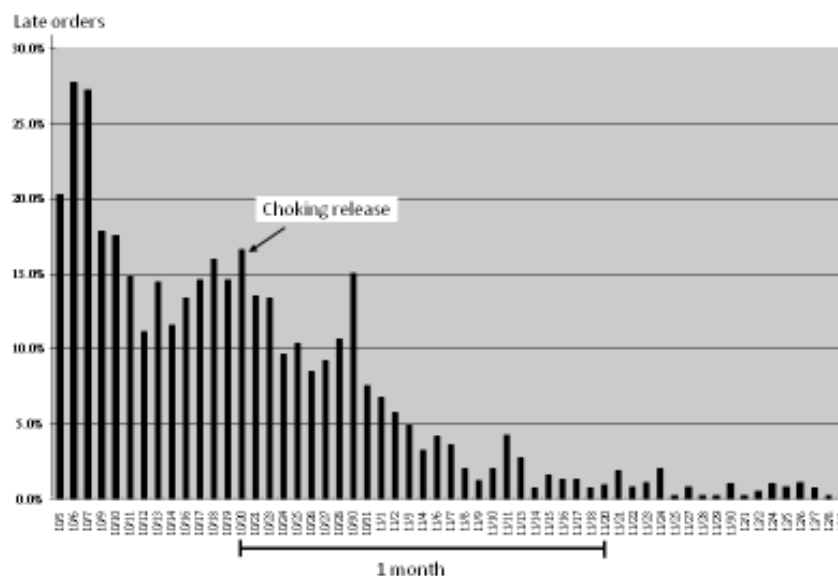
Pero uno no debe esperar que sólo este cambio lleve el cumplimiento en las fechas de entrega prometidas a estar encima del 90%. Simplemente porque aún hay muchas órdenes en el piso de producción, hay colas frente a los recursos y dejar al azar la secuencia en que el trabajo se procesa va a resultar en que muchas órdenes se terminen tarde. Se necesita un sistema de prioridades.

La necesidad de un sistema de prioridades no debe abrir las compuertas para algoritmos sofisticados para fijar las prioridades. Simplemente, el número de órdenes que ingresan cambia constantemente, el contenido del trabajo difiere de una orden a la siguiente, la longitud de las colas cambia constantemente y no olvidemos que aún existen interrupciones, perturbaciones; concretamente, este es un entorno con alta variabilidad. La lección que Shewhart trajo a la manufactura desde la física, y que Deming hizo conocer mundialmente, es que tratar de ser más preciso que el ruido (en nuestro caso, tratar de utilizar algoritmos sofisticados que consideren todo parámetro posible en un entorno de alta variabilidad) no mejora las cosas sino que las hace peores – los resultados con toda certeza no serán de una mejora sino de un deterioro en el cumplimiento de las fechas de entrega prometidas.

Un sistema de prioridades directo emerge cuando reconocemos que el amortiguador de tiempo, siendo la mitad del tiempo de respuesta actual, es aún mucho más largo que el tiempo de toque, y debido a que reduce dramáticamente las congestiones y atascos de tráfico, sin ninguna interferencia, muchas órdenes van a ser terminadas en solo un tercio del amortiguador de tiempo, y la mayoría van a ser

terminadas en los primeros dos tercios del amortiguador de tiempo. Basados en este entendimiento, las prioridades se asignan a través de la “gerencia de los amortiguadores”. Para cada lote se hace el seguimiento del tiempo que ha transcurrido desde su liberación al piso de la planta. Si ha transcurrido menos de un tercio del amortiguador de tiempo, el color de la prioridad es verde, si ha transcurrido más de un tercio, pero menos de dos tercios, el color de la prioridad es amarillo, si ha transcurrido más de dos tercios, el color es rojo, si la fecha de entrega prometida se ha pasado, el color es negro. Las órdenes en negro tienen mayor prioridad que las rojas, etc. Si hay dos lotes que tienen el mismo color, tratar de decidir cuál se debe trabajar primero es un excelente ejemplo de tratar de ser más precisos que el ruido.

Implementar un sistema como este en el piso de la planta es relativamente fácil. Para el primer paso, no se requiere hacer ningún cambio físico, solamente ahogar la liberación del material para que sea la mitad del tiempo de respuesta histórico antes de la fecha de entrega prometida correspondiente, y guiar al piso de la planta para que siga el sistema de prioridades de códigos de colores. En comparación con los esfuerzos, el impacto es impresionante. Para tener una percepción de primera mano sobre el impacto (y la velocidad) del primer paso solamente, la Figura 2 indica el porcentaje actual de órdenes atrasadas en una planta de 2,000 empleados que produce miles de tipos diferentes de utensilios de cocina.



Por supuesto que las eficiencias locales deben ser abolidas, o de lo contrario va a continuar la presión por liberar material más temprano. La experiencia muestra que la velocidad a la que todos en el piso de la planta se dan cuenta del impacto positivo, hace que el cambio sea casi libre de resistencia.

Pero en la mayoría de los ambientes todavía hay órdenes que incumplen sus fechas de entrega, y aún hay un potencial enorme de mejora para capitalizar. El cuarto concepto debe también llevarse a la práctica – debe implementarse un proceso de enfoque para balancear el flujo.

El primer paso para balancear el flujo es relativamente fácil. Ahogar la liberación de material revela capacidad disponible abundante, que antes estaba enmascarada. Pero es probable que algunos centros de trabajo tengan menor capacidad que el resto. Estos centros de trabajo se hacen evidentes debido a que tienen una cola de inventario frente a ellos. El abolir las eficiencias locales ayuda a identificar acciones simples que se requieren para incrementar su capacidad – acciones simples como asegurar que un centro de trabajo con capacidad restrictiva no se quede sin trabajo durante el almuerzo o los cambios de turno, desviar trabajo a centros de trabajo menos eficientes que tienen amplia capacidad excedente, etc.¹⁴

Debido a que las acciones anteriores agregan capacidad efectiva a los centros de trabajo que generan las colas, estas se hacen más cortas y menos órdenes alcanzan a estar en rojo. Esto significa que el amortiguador de tiempo ahora se hace innecesariamente largo. Una regla efectiva para ajustar el amortiguador de tiempo, sin correr el riesgo de deteriorar el alto nivel de desempeño en el cumplimiento de las fechas de entrega prometidas, es reducir el amortiguador de tiempo cuando el número de órdenes en rojo es inferior al 5% del número total de órdenes liberadas, e incrementarlo cuando la proporción de órdenes en rojo es más del 10%.

Una compañía que aplica lo anterior se va a encontrar a sí misma, en unos pocos meses, con un cumplimiento muy alto en las fechas de entrega prometidas, unos tiempos de respuesta considerablemente más cortos y una amplia capacidad disponible. Aquí es donde el verdadero reto comienza. En el pasado, algunas veces (demasiadas veces) la reacción de la alta gerencia ante la capacidad disponible revelada, expuesta es “reajustar” la capacidad y obtener ahorros en costos. Este es un

¹⁴ Eliyahu M- Goldratt y Jeff Cox, *La Meta, Un Proceso de Mejora Continua*, North River Press, 1984

grave error. La “capacidad excedente” son empleados – empleados que lo que han hecho es ayudar a la empresa a mejorar y como una consecuencia directa son “premiados” perdiendo su trabajo o el de sus amigos. En todos los casos en que este tipo de acciones se tomaron, la reacción inevitable que genera, produce retroceso y deteriora rápidamente el desempeño de la planta a una situación peor que la del punto de partida. Ojalá que este comportamiento gerencial sea cosa del pasado.

La manera más razonable de manejar la capacidad revelada excedente es capitalizar en ella; estimular a la fuerza de ventas para que tome ventaja de la mejora en el desempeño para generar más ventas. El incremento en las ventas puede causar fácilmente la aparición de un cuello de botella real. Ignorar la capacidad limitada del cuello de botella cuando se establezcan compromisos de fecha de entrega para los pedidos nuevos deteriorará el desempeño en el cumplimiento, y las ventas a clientes inconformes caerán en picada. Es esencial fortalecer la conexión entre ventas y operaciones – ese es el verdadero reto. Debe establecerse un sistema que asegure que cada compromiso de fecha de entrega prometida se hace sólo de acuerdo con la capacidad del cuello de botella que aún esté sin asignar.

El cuello de botella se convierte en el “golpe de tambor” para las órdenes, el “amortiguador de tiempo” traduce las fechas de entrega prometidas en fechas de liberación y la acción de ahogar la liberación del material se convierte en la “cuerda” que ata la orden con la liberación del trabajo. Esta es la razón por la que esta aplicación de la Teoría de Restricciones, fundamentada en el tiempo, se conoce como el sistema Tambor – Amortiguador – Cuerda, o TAC. (DBR= *Drum-Buffer-Rope*).

Actualmente hay un experimento de amplia base para pulir un proceso para mejorar aún más las operaciones, basado en el registro y análisis de las razones para las órdenes en rojo.

El ejemplo de Hitachi

La empresa *Hitachi Tool Engineering Ltd.*, de 24 mil millones de yens, diseña y manufactura cerca de 20,000 diferentes herramientas de corte. La demanda para la mayoría de los productos es esporádica, y los estándares de su industria condicionan a la empresa a lanzar nuevas familias de herramientas cada seis meses. Cuando se lanzan

las nuevas familias de productos, las familias antiguas se hacen obsoletas. No es extraño entonces que los esfuerzos por implementar LEAN fueran poco exitosos.¹⁵

Hitachi Tool Engineering Ltd. comenzó a implementar DBR en una de sus cuatro plantas en Japón en el año 2000. El salto en el desempeño en el cumplimiento de las fechas de entrega prometidas (de 40% a 85%), asociado con la reducción del IPP y del tiempo de respuesta a la mitad, junto con la habilidad de despachar un 20% más de productos, con la misma fuerza laboral, les motivó a expandir la implementación. Para el 2003, ya habían implementado DBR en las cuatro plantas.¹⁶

La drástica reducción en el tiempo de producción y la capacidad de respuesta mejorada le permitió reducir el inventario en la cadena de suministros – de los distribuidores – de 8 a 2,4 meses. La reducción del inventario mejoró dramáticamente el retorno sobre la inversión de los distribuidores, liberó su efectivo y fortaleció sus relaciones con Hitachi. No es de extrañar que los distribuidores expandieran el rango de herramientas de Hitachi que estaban ofreciendo, llevando a un incremento del 20% en las ventas (en un mercado estable).

El verdadero impacto se revela cuando evaluamos el desempeño de esta empresa en sus indicadores financieros, considerando que durante el período del 2002 al 2007 el precio de las materias primas (metales) se incrementó en aprox. Un 60%, mientras que el precio de venta de las herramientas de corte se mantuvo igual. Bajo estas condiciones, las utilidades de la empresa debieran haberse esfumado. Por el contrario, las utilidades netas antes de impuestos de *Hitachi Tool Engineering Ltd.* se incrementaron de 0,98 mil millones de Yens en el año fiscal que finalizar en Marzo 2002 a 5,3 mil millones de yens en el año que finaliza en Marzo 2007 – un incremento de cinco veces la utilidad neta en cinco años. La tasa de utilidades de *Hitachi Tool Engineering Ltd.* se incrementó de cerca de 10% en el 2002 a un 22% en el 2007, la tasa más alta jamás reportada en este tipo de industria.¹⁷

¹⁵ Umble, M, Umble E-. and Murakami, S., "Implementing theory of constraints in a traditional Japanese manufacturing environment: the case of Hitachi Tool Engineering", *International Journal of Production Research*, Vol 44, No. 10, 15, May 2006, pp. 1863-1880.

¹⁶ Ibid.

¹⁷ A GUIDE TO MAKING EVER FLOURISHING COMPANY – PRODUCTION, DISTRIBUTION, MARKETING AND SALES. Chuhei Publishing, 2008, Satoru Murakami, Jun Takahashi, Shotarou Kobatashi p 196~p207.

Las fronteras de DBR

Como se resaltó anteriormente, una aplicación hace supuestos (a veces supuestos ocultos) acerca del ambiente, y no debemos esperar que la aplicación funcione en ambientes para los cuales estos supuestos no son válidos. El supuesto que hace DBR es evidente; asume que el tiempo de toque es muy pequeño (<10%) comparado con el tiempo de producción-respuesta actual. Este supuesto es válido para muchos, sino para la mayoría, de los entornos típicos de producción. Pero definitivamente no es válido para un rango muy amplio de ambientes que se denominan tradicionalmente como ambiente de proyectos.

En los ambientes de proyectos el tiempo de toque es relativamente largo, y el afán de los clientes para que los proyectos se terminen pronto fuerza a las operaciones a prometer tiempos de entrega que son sólo el doble (o en pocos casos, el triple) de largos que el tiempo de toque. No sorprende entonces que el desempeño en proyectos sea malo, hasta el punto que nadie espera que el proyecto se termine a tiempo, dentro del presupuesto y con todo el contenido; siempre hay algo que se espera falle. Pero este hecho no debe distraernos de la conclusión de que dado que el supuesto de DBR no es válido, DBR no es apropiado para el entorno de proyectos. Una aplicación diferente, una aplicación que maneja directamente el tiempo de toque relativamente largo es lo que se requiere.¹⁸

¹⁸ Goldratt, M. Eliyahu, Cadena Crítica, North River Press, 1996.