QRM: AN ENABLER ON THE ROAD TO AGILITY ?

Quick Response Manufacturing - Part 2

Lindsay Harding, MIOM, Strategem

Quick Response Manufacturing is a technique developed in the USA at the University of Wisconsin-Madison, where as a result of successfully working with numerous organisations, the Centre for Quick Response Manufacturing has been established at the campus. Business literature introduced the Time-Based Competition strategy (TBC) in the late 1980s [1], in North America particularly, its specific application to the manufacturing arena has become known as Quick Response Manufacturing [3].

In the first part of this article the author gave readers an appreciation of the QRM philosophy and suggested that some of the tools and techniques developed in the USA for this approach could be adopted in the UK when embarking on a road towards Agility. This second part of the article introduces the special material control system developed, and the simulation software component of the approach. As a summary of the concepts, the following principles developed by Professor Rajan Suri (‡‡) are included to compliment the contents of the first part of the article.

The ten principles can be considered as the fundamental cornerstones of QRM in that they challenge much of the ‘conventional wisdom’ associated both with volume manufacture and, to an extent, both ‘Lean Manufacturing’ or ‘JIT’[2].

Traditional belief number 1:	QRM principle number 1:
Everyone will have to work faster, harder and longer hours, in order to get jobs done in less time.	Find whole new ways of completing a job, with the focus on lead time minimisation.

Quick Overview: Our organisations are not designed to manage time. Organisational structures, accounting systems and reward systems are based on managing scale and cost. A legacy of scale/cost-based management systems, and the greatest enemy of QRM efforts, is the functional organisation with specialised departments. Another legacy is the response time spiral, an increasing spiral of lead times that results from scale/cost-based management systems. Taking time out of the system requires completely rethinking how you organise production, materials supply, and office-based work.

Traditional belief number 2:	QRM principle number 2:
To get jobs out fast we must keep our machines and people busy all the time.	Plan to operate at 80% or even 70% capacity on critical resources.

Quick Overview: The first reaction of most managers is, "We can't afford to do that. We will be wasting our resources and our costs will go up!" QRM will eliminate the complex series of dysfunctional interactions, long lead times, growing queues, jobs spending a lot of time waiting for resources that result from the present 100% utilisation policy. QRM will show how idle capacity actually serves as a strategic investment that will pay for itself many times over in increased sales, higher quality, and lower costs.

Traditional belief number 3:	QRM principle number 3:
To reduce our lead times, we have to improve our efficiencies.	Measure the reduction of lead times and make this the main performance measure.

Quick Overview: This traditional belief goes hand-in-hand with the desire for maximising resource utilisation. The problem is not the concept of efficiency, but that most measures of efficiency work counter to lead time reduction. The QRM principle may seem a rather bold step, but in one company, lead times for a line of spare parts dropped from 36 days to 6 days using reduction of lead time as the main performance measure. To accomplish this, though, it is important for everyone in a manufacturing firm, especially senior managers, to understand the dynamics of factory operations, they need to study the interactions between capacity utilisation, efficiency measures and batch sizing policies, and their effects on lead time. It should be noted that batch sizes appropriate for QRM bear little relation to the values calculated by the economic order quantity (EOQ) formula, which fails to consider many costs of large batches and ignores the value of responsiveness. Nor can good batch sizes for QRM be predicted by an MRP system, since it assumes fixed queue times regardless of workload.

Traditional belief number 4:	QRM principle number 4:
We must place great importance on ‘on-time’ delivery performance by each of our departments and suppliers.	Stick to measuring and rewarding reduction of lead times.

Quick Overview: Almost every book on modern manufacturing discusses on-time delivery and says that it is a cornerstone of JIT. What has been observed though, is that while on-time performance is desirable as an outcome, emphasising it as a performance measure is dysfunctional. Human nature being what it is, instead of trying to reduce lead times, internal departments and external suppliers alike tend to pad their quoted lead times so that their on-time deliveries look good. As a result, the Response Time Spiral takes over the organisation. With QRM, organisational changes, along with appropriate performance measures, promote shorter lead times. These shorter lead times, in turn, kill the response time spiral, and delivery problems disappear-resulting in on-time performance.

Traditional belief number 5:	QRM principle number 5:
Installing a material requirements planning (MRP) system will help in reducing lead times. .	Use MRP to plan and co-ordinate materials. Restructure the manufacturing organisation into simpler product-oriented cells. Complement this with a new material control method that combines the best of push and pull strategies.

Quick Overview: MRP systems serve an important function of assisting with materials supply, but don't expect them to solve lead time problems because the underlying model in MRP is flawed. In the re-designed organisation, MRP is used for a higher level of planning and providing authorisation but not for micro-managing work centres. Teams should run their own cells, and they should be provided with simple tools to manage their capacity and continually improve their responsiveness. A novel material control strategy called POLCA, which is overviewed below, combines the better aspects of push and pull methods to limit congestion while at the same time providing a high degree of flexibility, enabling even custom-engineered products to be made.

Traditional belief number 6:	QRM principle number 6:
Since long lead times need to be ordered in large quantities, we should negotiate quantity discounts with suppliers.	Motivate suppliers to implement QRM, resulting in small batches at lower cost, better quality and short lead times.

Quick Overview: The more you purchase items in large batches, the longer the suppliers take to make them, motivating you to put in orders for even larger batches. This creates another dysfunctional spiral that is made worse by traditional purchasing policies and incentives.

Traditional belief number 7:	QRM principle number 7:
We should encourage customers to buy our products in large quantities by offering price breaks and quantity discounts.	Educate customers about your QRM programme, and negotiate a schedule of moving to smaller batch sizes at reasonable prices.

Quick Overview: This is the reverse of number 6. Now you are the supplier. The customer's behaviour of ordering larger batches will degrade your delivery performance. With QRM you form strategic partnerships with your customers and show how QRM will allow them to receive smaller batches at lower cost.

Traditional belief number 8:	QRM principle number 8:
We can implement QRM by forming teams in each department.	Cut through functional boundaries by forming a quick response office cell, which is a ‘closed-loop,’ collocated, multi-functional, cross-trained team responsible for a family of products. Empower it to make necessary decisions.

Quick Overview: Some of the team implementations that follow traditional belief 8 are the result of the quality (TQM) movement. True, a team with all its members in one functional department may result in local quality improvements. For the purpose of QRM, however, such a team will do little to cut overall lead time for office operations. Instead, the team for QRM must be the office cell with characteristics specified above. (Cells and other QRM changes are not restricted to the shop floor.) Such quick response office cells are the only way to get significant reduction of lead times for jobs such as estimating and quoting, order processing, and engineering. Closed-loop means that all the required steps can be done within the team, which means you will have to cut across functional boundaries and change reporting structures.

Traditional belief number 9:	QRM principle number 9:
The reason for implementing QRM is so that we can charge our customers for rush jobs	The reason for embarking on the QRM journey is that it leads to a truly more lean and mean company with a more secure future.

Quick Overview: Although customers may pay more for speedy delivery, and this may be a good short-term result of better response, it should not be the main reason for engaging in QRM. Searching for ways of squeezing time out uncovers quality problems and wasted efforts. Fixing these results in higher quality, lower WIP, less waste, lower operating costs, and greater sales. JIT and related methods have put a lot of emphasis on elimination of waste, but those approaches ignore certain types of waste caused by long lead times. With its broader definition of waste, QRM can create an even leaner enterprise that will remain a formidable competitor for years to come.

Traditional belief number 10:	QRM principle number 10:
Implementing QRM will require large investments in technology.	The biggest obstacle to QRM is not technology, but ‘mind-set.’ Combat this through training. Next, engage in low-cost or or no-cost lead time reductions. Leave high cost technological solutions for a later stage.

Quick Overview: New technologies, such as rapid prototyping and CAD/CAM, offer great opportunities for time reduction. These are important, but several steps must precede them, like education. In particular, you must realign the mind-set of all employees, from the shop floor to the boardroom, from desk workers to senior managers, to QRM principles. To bring about the mind-set change, organisations will need to thoroughly rethink existing performance measures. Performance measurement is intimately tied in with the cost accounting system, which is an obstacle to implementing an effective QRM programme. QRM goes beyond activity-based costing (ABC) to address this issue.

POLCA: THE MATERIAL CONTROL SYSTEM FOR QRM

The QRM philosophy advocates that cells are organised to eliminate the effects of the ‘Response Time Spiral’ [2] by focusing on reducing lead times. Here, there are drawbacks with both push and pull systems alike, and a traditional Kanban approach is not appropriate, as a consequence a hybrid system has been developed. This new material control system developed by the QRM centre in the USA is called 'Paired-cell Overlapping Loops of Cards with Authorisation' (POLCA), which could basically be described as a mechanism to link MRP with Kanban shop floor pull.

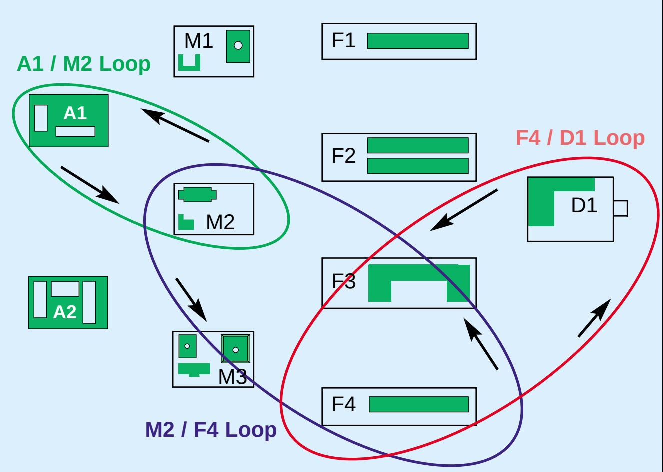
QRM strategy is particularly suitable to companies producing custom designed products in small batches, high variety options, and even one-offs, where inventory coverage to manage the variability is not feasible, hence this new method of material control was developed. The way in which this is envisaged to occur is that companies create cells focusing on subsets of the production process for like families of parts. Customer orders can then be processed through differing pairings of cells aligned to the order content. To allow such a replenishment system to be initiated, the following criteria should be positioned:

- A high-level material planning system (HL/MRP)
- Cellular organisations
- Flat BOMs.

Basically, authorisations are created via HL/MRP. POLCA cards are assigned to pairs of cells, as opposed to a Kanban card operating within a cell. These POLCA cards stay with the job during its journey through both cells in the pair before they loop back to the first cell in the pair. Since most cells will belong to more than one pair, card loops will overlap at each cell. Unlike Kanban, the cell is not controlled by the POLCA card. The POLCA card is not a replenishment signal, it is a capacity signal, a returning POLCA card indicates availability of capacity at the downstream cell. They are not product specific, they do not indicate a part number, rather the names of two cells (refer to Figure 3).

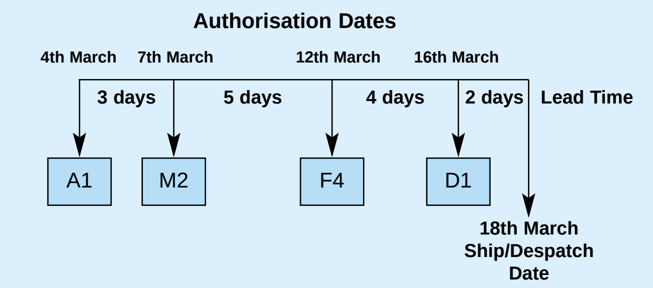
Suri [2] describes how when a ‘company receives a customer order, the HL/MRP system uses the planned lead times for each cell to determine times when each cell in the product routing may begin processing that job’. This part of the HL/MRP logic is no different from standard MRP, except that the HL/MRP is based on simpler BOMs and uses lead times for entire cells instead of individual work centres within a cell (ie. HL/MRP setting the lead time for the cell as a whole, not just the steps in the cell). When operating a QRM mentality, the system will output when production 'may' start as opposed to when it 'should' start, hence the authorisation aspect of the model.

Figure 1



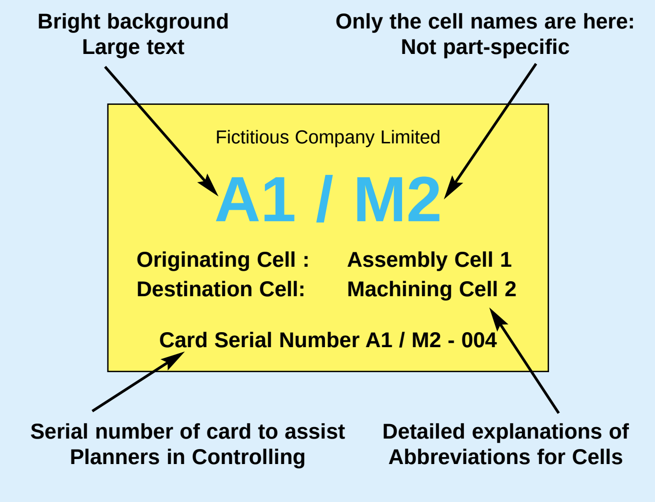
In the fictitious company illustrated (Figure 1 above), the order's routing takes it from Assembly Cell A1, onto Machining Cell M2, into the Finishing Cell F4, before finally being shipped from Despatch area D1. This particular order requirement will therefore be manufactured via the POLCA card loops A1/M2, M2/F4 and F4/D1, (this sequence of cells that a job must visit is quoted on the routing sheet). If raw material is available and the A1/M2 card is present at Cell A1 production will commence, by pairing cells together. The POLCA cards ensure that follow cells only commence processing of jobs for which the destination cell has capacity (normal MRP would continue to push at this point). The cards will contain material handing instructions etc. The card will be placed in a visible point (output buffer) on the job as the operation is completed, at A1 for example, if the team in cell M2 have a spare A1/M2 card they can commence the job. Note excess cards are held off the shop floor (typically by the planners). At the M2 cell shown, there will be two POLCA cards in the cell. The one which has moved with the job from A1, and the M2/F4 card which moves the job onto its next point in its routing, hence the term 'overlapping'. Note a completed M2/F4 card sitting at the output buffer of F4 will move back to the input buffer at M2, and so on. An example of the authorisation dates from HL/MRP for our fictitious company may be as follows in Figure 2 (similar to start dates obtained by back-scheduling in MRP but there are some key differences):

Figure 2



An example card would likely contain the following information:

Figure 3



The number of POLCA cards required for each loop can be set using Little's Law:

No. of A/B cards = [LT(A) + LT(B)] X NUM(A,B)/D

Where, LT(A) and LT(B) equates to the estimated lead time (in days) of cells A and B over the planning horizon, D is the number of working days in the planning horizon, and NUM(A,B) is the total number of jobs which go from cell A to cell B during the planning horizon [2].

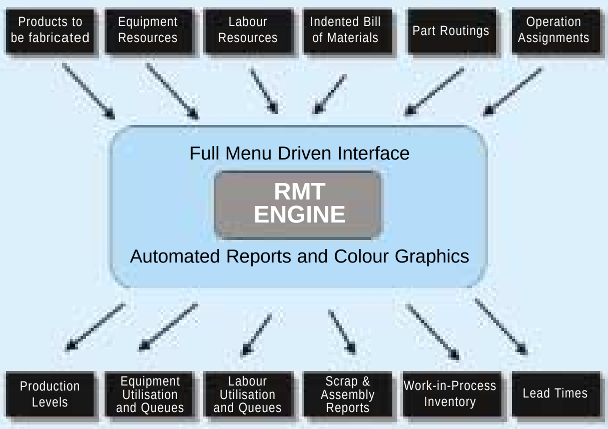
THE MPX SIMULATION SOFTWARE

The rapid modelling technology (RMT) software package used to analyse queues in work cells on the shop floor and even in administrative functions is called MPX. RMT is based on the established technique of queuing theory, a branch of engineering applied mathematics that describes factory floor dynamics particularly well [4]. MPX is a valuable tool which assists in achieving the objectives of the QRM philosophy (though it can be used separately). The software, which was developed by

'Network Dynamics' (of Burlington, Massachusetts, USA) is used to conduct 'what if' scenarios which impact a variety of manufacturing parameters, including parts routing, labour, equipment, equipment failure/repair, set-up, run time and lot size. The software model assists in evaluating the effects of alternative management decisions, and is of particular importance when analysing the operation of cells. It helps in obtaining an insight into the factors that influence the lead time performance of cells. MPX has been widely used in the USA and it can be used to establish whether the business will benefit from manufacturing in cells, and if so, what would be the ideal cell configuration.

Figure 4 summarises the inputs and outputs of MPX [5]. The tool requires basic information on the products to be made and the resources available - information that is easily turned into tables, guided by pull-down menus and windows (see Figure 5).

Figure 4



Lead time depends on both processing time and queue time (this is the time parts waiting to get their turn at machines, when the machines are busy). While processing time may be known based on the machining parameters, queue time depends on many 'dynamic' factors such as, which other parts are already in queue to use the machine, whether the machine has broken down, whether an operator is available, and so on. In order to predict whether a decision will improve lead times, it is necessary to be able to predict these queue times, which means any lead time reduction tool must model these dynamics and interactions. The RMT technology in MPX, models these complex dynamics of the manufacturing facility in terms of mathematical equations. Until a few years ago, these equations couldn't be solved. However, with the progress that has been made in queuing theory in recent years, very good estimates can now be obtained for system performance with amazingly little computer time, often just seconds on a personal computer [4].

'What-if' simulations are entered, changing set-up times, batch sizes, MTTF durations, etc from which resultant lead times etc are calculated. These are displayed graphically for the user, comparing with original settings and following proposed process changes to determine areas of greatest impact to direct effort within the business.

Finally, in summary, it must be re-emphasised that QRM is not an alternative to 'Lean' rather a complimentary set of techniques, which would be more effective when applied to appropriate areas (low volume items with high variety of options in particular). The suggestion of planning at a 70-75% utilisation level would be viewed with scepticism from a traditional financially biased approach, however these techniques can illustrate that driving up utilisation can go

Figure 5

The MPX Working Environment

beyond a critical point and actually begin increasing lead times. QRM itself is not a product, it is not a phrase copyrighted by the University of Wisconsin, it is however a mindset philosophy incorporating a collection of tools and techniques which when applied successfully can reap great reward for the organisation.

ACKNOWLEDGEMENT

I would like to acknowledge the input and support of both Frank Rath of the Wisconsin QRM Centre and Andy Shipley of Strategem in producing this article.

‡‡ (Professor Rajan Suri - Director of the Centre for Quick Response Manufacturing and Professor of Industrial Engineering at the University of Wisconsin-Madison, USA).

REFERENCES

[1] Stalk,G. Jr. "Time -The Next Source of Competitive Advantage", *Harvard Business Review,* pp 41-51, July-August, 1988.

[2] Suri, Rajan. *"Quick Response Manufacturing Companywide Approach to Reducing Lead Time",* Productivity Press, 1998.

[3] Tubino, Francisco and Suri, Rajan. "What Kind of 'Numbers' can a Company Expect After Implementing Quick Response Manufacturing?" From *Quick Response Manufacturing 2000 Conference Proceedings*, pp 943-972, Manufacturing Engineers Press, Dearborn, MI, 2000.

[4] *"MPX Users Manual",* Network Dynamics Inc., MA, 01803, USA

[5] Suri, Rajan. "Using Queueing Models to Support Quick Response Manufacturing", *Proceedings of the International IIE Conference*, Minneapolis, MN, May, 1996.

About the author

Lindsay Harding, MIOM, MSc CPIM CIRM MILT MCMI, is a Management Consultant working with Strategem Limited where he is involved in rolling out best practice techniques and business improvement initiatives to client companies. He gained his MSc in Logistics & Supply Chain Management at Cranfield University.