International product transfer and production ramp-up: a case study from the data storage industry

Christian Terwiesch,[1] Roger E. Bohn,[2] and Kuong S. Chea[3]

[1] The Wharton School, Department of Operations and Information Management, 1319 Steinberg Hall – Dietrich Hall, Philadelphia, PA 19104.6366, USA
terwiesch@wharton.upenn.edu
[2] Graduate School of International Relations and Pacific Studies, University of California San Diego, 9500 Gilman Dr 0520, La Jolla, CA 92093–0520, USA
rbohn@ucsd.edu
[3] Sun Microsystems, 3571 Lemona Ave., San Diego, CA 92105, USA
kchea@us.ibm.com

Many high-tech industries are shifting their focus from minimizing time-to-market to minimizing time-to-volume. This puts the tail end of product development, the production ramp-up, in a critical position. This article presents a case study of product transfer and production ramp-up in the hard disk drive industry. We provide a detailed description of the ramp-up period. By documenting detailed time-series of several operational measures, we also shed light on the various forces that allow an organization to increase its production volume. Finally, the setting allows us to study product transfer from development in the USA to an Asian production facility. We find that the physical distance is successfully overcome by several mechanisms.

1. Introduction

To achieve a fast pay-back of investments in new product designs and production facilities, companies must reduce their development time (time-to-market) as well as the time it takes them to achieve acceptable manufacturing volume, cost, and quality (time-to-volume). Whereas a number of studies have investigated time-to-market, the topic of time-to-volume has received relatively little attention. Yet the timing of revenues depends critically on time-to-volume, while development expenses are concentrated around the time just before product launch. This gives time-to-volume high leverage in determining total project net present value.

The fundamental difference between time-to-market and time-to-volume is that the former ends with the beginning of commercial production whereas the latter explicitly includes the period of production ramp-up. Production ramp-up is the period during which a manufacturing process makes the transition from zero to full-scale production at targeted levels of cost and quality.[1] Ramp-ups are needed for each new product, but also for new lines and new factories; these are often called 'start-ups'.

In the auto industry, time to full-scale production can be up to six months, while time to normal quality can range between one month and a year (Clark and Fujimoto, 1991). This accounts for about 10–20% of a car's lifecycle. In high-tech short lifecycle industries such as semiconductors or hard-disk drives, the portion of the product lifecycle that the product spends in ramp-up can be even larger. Since in these industries prices often fall rapidly, achieving high volume early has an especially high financial payoff.

This article describes a case study of production ramp-up, in the hard disk drive industry. We present longitudinal data on production volume, cost, and quality over the course of a ramp-up that enable us to shed light on the evolution and economics of the production ramp-up period. We provide a detailed description of the ramp-up process, from which we extract several organizational patterns that contribute to a successful ramp-up. Because product transfer from development to manufacturing sets the starting point and therefore preconditions the success of ramp-up, we also describe the transfer process.

Finally, we discuss issues related to product transfers and production ramp-ups that take place across physical and national barriers. In the hard disk drive industry, most development is done in the USA, and most manufacturing is done in same-company factories located in SE Asia. This violates the precept that close coordination between design and manufacturing requires physical closeness. Yet it can be done quite successfully, as we will show.

This article is organized as follows. We first review the relevant literature and motivate our research questions. We then give a background of the hard disk drive industry, followed by a discussion of our research methodology. Our findings are presented in three steps. First, we provide the case description of pilot production, transfer and ramp-up. Second, we address our research questions by linking them to the case. Finally, we derive managerial implications and discuss opportunities for future research.

2. Background

We define the ramp-up of a new product as the period when the normal production process makes the transition from zero to full-volume production, at or near the targeted levels of cost and quality. Thus, ramp-up directly succeeds process engineering and pilot production. This is in line with Wheelwright and Clark (1992), who define:

> In ramp-up the firm starts commercial production at a relatively low level of volume; as the organization develops confidence in its (and its suppliers) abilities to execute production consistently and marketing's abilities to sell the product, the volume increases. At the conclusion of the ramp-up phase, the production system has achieved its target levels of volume, cost, and quality (Wheelwright and Clark, 1992, p. 8).

Two conflicting factors are characteristic of this period: low production capacity, and high demand. High demand arises because the product is still 'relatively fresh' and might be even the first of its type. Thus, customers are ready to pay a premium price. Yet production output is low due to low production rates and low yields. The production process is still poorly understood and, inevitably, much of what is made does not work properly the first time. Machines break down, setups are slow, suppliers are late or have quality problems, special operations are needed to correct product and process oversights, and other factors impede output. Over time, with learning about the production process and equipment, yields and capacity utilization go up (although in many industries they never reach 100%). Due to the conflicts between low capacity and high demand, the company finds itself pressured from two sides, an effect referred to as the nutcracker (McIvor *et al.*, 1997).

The importance of production ramp-up

A recent example of the importance of ramp-up can be found in AMD's efforts to compete with Intel in the microprocessor market. AMD had several generations of product that were slow to ramp, leading to limited market acceptance and financial difficulties for AMD. But recently, Intel had problems ramping up the yield of its 0.18 micron version of the Pentium. In contrast, the effective ramp-up of AMD's K7 processor allowed AMD to compete in the high end segment of the PC market (*Electronic Buyers' News*, April, 1999).

The importance of fast production ramp-up is the result of several forces. At the beginning of a new product ramp-up, almost all of the investments for research, product development, and manufacturing equipment have already been made. Delays in reaching production at the targeted volume and quality will push the eagerly awaited revenues even further into the future. In this way, shorter ramp-up reduces the time to payback and improves traditional accounting measures such as return on investment (ROI) as well as return on assets (ROA).

Successful ramp-up is also important because it often substantially influences the market's acceptance of a new product. Early, high volume, high quality production will hasten market penetration, potentially raising subsequent market share and deterring competitors. Further, many high tech products such as integrated circuits and disk drives are sold primarily to system integrators (OEMs), for example PC assemblers. These intermediates have strict processes of vendor qualification, and allocate future demand based on their subjective evaluation of the product and process in the development and ramp-up phase. After this vendor qualification and selection, a large portion of future demand is already determined and changes to the qualified production process will have to get permission from the customer. Since ramp-up of critical components can be a gating item on ramp-up of a whole system, OEMs are risk-sensitive, and prefer vendors who can ramp rapidly.

Thus faster ramp-up increases lifetime sales volume of the product. The impact on revenue can be even larger, due to declining prices over time for many high-tech products. Prices fall as multiple competitors enter and ramp-up their factories. Therefore, early sales carry the highest prices. In some industries, price declines of 20–50% per year are normal. Therefore, a difference of a few months in ramp-up time can have a substantial effect on revenues and profits (e.g. Ulrich *et al.*, 1993).

Two recent trends have increased the importance of fast production ramp-up. First, some industries are experiencing an accelerated rate of product introductions. This makes the introduction of new products, and their associated production ramp-up, more the rule than the exception. Second, the growing complexity and tightening tolerances of 'high-tech' industries' products tends to increase the required up-front investments in research, product development, manufacturing technologies, and production equipment/tooling.

Previous research

Although the topic of time-to-market has received considerable attention, from the academic as well as the practitioner side, little is written about how to reduce time-to-volume. This is especially surprising, as the importance of production ramp-up has repeatedly been mentioned in research on product development e.g. Clark and Fujimoto (1991), Wheelwright and Clark (1992) and Pisano (1997). In their study of the global automotive industry, Clark and Fujimoto (1991) observed significant regional differences in ramp-up performance between Japanese companies versus US and European companies. Their data shows that the time to full-scale production varied from one to six months, while time to normal quality can range between one month and a year. For both measures, on average Japanese companies ramped-up faster than their American counter-parts.

Pisano (1997) identifies and contrasts two learning strategies during the production ramp-up. In his study of the pharmaceutical industry, the author finds that in more stable environments, such as chemical pharmaceuticals, the role of the plant in development is less critical (although still important) and ramp-up is easier to achieve. As the existing knowledge about chemical pharmaceuticals is deep enough to allow the development of analytical models, developers can anticipate many of the production issues prior to ramp-up, a process that Pisano calls 'learning-before-doing'. This is in contrast with pharmaceuticals resulting from biotechnology, as this field has not reached a sufficient maturity to allow the development of analytic models and the only way to find out about the producibility of a drug is by trial-and-error, or learning by doing.

Langowitz (1985) emphasizes the link between product development and the start-up period. She observes that the success of the start-up is a function of fit between requirements of the new product and capabilities in the plant. In order to make a ramp-up successful, a firm can design the product to match the existing competence of the factory, or prepare the factory in advance by increasing its capabilities.

Most later observers of ramp-up conclude, like Langowitz, that it cannot be viewed in isolation from earlier phases of development. The most important activities of ramp-up consist of discovering and removing 'bugs', problems, and missed opportunities that were introduced earlier in development. Thus well-executed product and process design phases will lead to easier ramp-ups. The way design and manufacturing are coordinated (e.g. concurrent engineering) will affect the number of problems and missed opportunities for improvement (Clawson, 1985). More specifically, the way the transition is made from development into manufacturing, usually in a physically separate facility, can have a large effect. Adler (1995) provides a taxonomy of design-manufacturing coordination mechanisms, which we will discuss extensively in the conclusion.

In addition to Langowitz, there are several Harvard teaching cases (e.g. Bohn and Jaikumar, 1986; Langowitz and Wheelwright, 1986; Wasserman and Clark, 1986; Freeze *et al.*, 1984; Pisano, 1992) that deal with ramp-up. However, to our knowledge, there exists no previous academic work providing a detailed description of the ramp-up period.

Research questions

The managerial importance of product transfer and production ramp-up, together with the scarcity of previous academic work on this topic, motivate our research efforts. Our first research objective is to better understand the steps required to bring a production process to full-scale production at targeted levels of cost and quality. On the one hand, product development research has traditionally focused on pre-manufacturing activities (see Krishnan and Ulrich (2001) for a recent overview). On the other hand, operations management research in general, and its sub-stream of process improvement research in particular, tends to look at mature processes, ignoring the dynamics of production ramp-ups.

Our first research question aims at the gap between these two literature streams by describing the transition process as well as organizational interactions between the development process (ending with pilot production) on one side and the manufacturing facility on the other hand. It is important to understand that this goes well beyond the traditional 'design for manufacturing' literature (e.g. Dean and Susman, 1989), which typically looks at the interface between product and process designs.

Question 1: What are the key activities carried out during the transfer from development to manufacturing, and during manufacturing ramp-up?

Wheelwright and Clark (1992) define ramp-up as a phase of increasing production volume ending with the achievement of the target volume. But what variables contribute to this increase in production volume? How do they interact? And ultimately, what forces drive their changes? Our goal is to identify these variables, to describe their trajectories as time series, and to explain what drives them. Much more is involved than simply training operators and adding more machines.

Question 2: (a) Which variables contribute to the desired increases in production volume? (b) How do these variables change over time? (c) What technical and organizational processes are behind the increases?

The third objective of the present article is to explore issues related to product transfers and production ramp-ups that have to overcome long distances. Such distances can be a result of outsourcing production to a remote supplier or, more generally, any off-shore production within or across the boundaries of the firm. Prior research has suggested that close co-location between development and manufacturing is important for effective rapid transfer from one to the other (see for example, Hatch and Mowery, 1998). However, hard disk drives, along with a small number of other industries such as ASICs, have shown that design and manufacturing need not be sited together. Essentially the entire HDD industry develops in one country (USA or Japan) and manufactures in another (Gourevitch *et al.*, 1997; McKendrick *et al.*, 2001). How is this accomplished, despite very narrow time windows for ramp-up, and fierce competition among firms?

Question 3: What specific issues arise in transferring and ramping up a product across long distances and international boundaries?

As we will see below, the analysis of these questions are rather complementary: studying the interactions between development/pilot production and volume production (question 1) is actually *easier* in the case of geographic separation, as organizational interactions become far more traceable.

3. Data storage technology and industry

To answer our research questions, we have conducted a detailed case study in the disk drive industry. As our research objective was to gain a detailed understanding of the ramp-up process, our case study uses a single site, exploratory case methodology, based on longitudinal on-site data collection. The following sections briefly present the technical and industrial background for our research, followed by a description of the research site. We then provide further details concerning our case methodology.

Technical background of a disk drive

A hard disk drive (HDD) is a data storage device that reads, writes and stores digital data using magnetic signals.[2] The main subsystems are the head disk assembly (HDA), and the printed circuit board assembly (PCBA). The HDA is an electromechanical system that includes heads, media (disks), a head positioning mechanism and a spindle motor. The PCBA is an electronic system with custom integrated circuits handling both analog and digital signals. The HDA includes a case that shields the mechanical portions from dust and other contamination over the life of the product.

The technology trajectory of the industry has been to provide rapidly increasing storage volumes, measured in megabytes, at slowly declining cost per drive. An important summary measure of technological difficulty and performance in the HDD industry is areal density, the number of stored bits per square inch on the recording surface of the disk. An increase in areal density provides for greater storage capacity from the drive without increasing the number of heads and disks, allowing companies to deliver higher capacity devices at lower cost per megabyte. During the 1990s, areal density has increased at about 60% a year. This has been accomplished through a stream of innovations including smaller head sizes, new materials for heads (magnetoresistive heads), more precise head positioning (servo mechanisms), lower flying heights, improved signal processing, better media, and other new technologies. Each of these changes requires a partially or completely new design to embody it, so that the high rate of overall change requires a high rate of new product introduction.

The manufacturing of HDDs is a difficult process. The low flying heights make the HDD vulnerable to contamination by small particles, requiring a Class 100 clean room environment for the HDA. A second challenge in the assembly of an HDD is given by the high degree of miniaturization of the components (especially the heads) and the extremely small tolerances in putting the parts together without damaging them. These and other issues lead to considerable 'yield fallout', i.e. products which fail tests at some stage of the process. Units that fail can usually be reworked, at some cost in labor and discarded materials.

Manufacturing of HDDs starts with the manufacture of components, of which the most difficult are heads, media, and integrated circuits, each done in its own factories. These components are brought to the HDD assembly facility, where the actuator mechanism, heads, disks, spindle motor and other components are assembled to form the HDA. Although assembly can be automated, it typically is a largely manual operation, partly because automated systems are harder to modify for new models. After the HDA is assembled, an operation known as servo writing puts a basic logical format on the disks. This is followed by several functional tests, which are highly automated. Finally, the PCBA is assembled to the HDA and the completed unit is again tested strenuously prior to packaging and shipment.

Industry background

The information storage industry in general, and the HDD market in particular, are characterized by intense competition and rapid technological progress, resulting in rapid price erosion, short product life cycles, and continuous introduction of new, more cost-effective products. The rapid increase in areal density pushes companies to improve products and processes, with short lifecycles and fast obsolescence. Typical product lives are 18 months or less, with a replacement product being introduced (by the same firm) every nine months or less. The short product lifecycles make the HDD industry attractive for longitudinal research, as researchers can – in a limited amount of time – observe several generations of products (and for our purposes, of ramp-ups).

Rapid technological progress, short lifecycles, and intense competition also force companies to introduce products to the market, before the products or their manufacturing process are fully understood and debugged. This puts a special burden on production in general and production ramp-up in particular. The industry can be considered a 'high tech commodity' industry, in that the technology evolves rapidly (high tech) but purchase decisions are heavily based on price, with little product differentiation among vendors and extreme competition-induced price pressure, leading to the very low margins typical of a commodity. As the personal computer industry moves toward much lower unit prices, this puts further pressure on PC component suppliers such as HDD makers.

Research site

As the site for our research, we used Hdrive Corporation.[3] Hdrive develops, manufactures, and markets HDDs for desktop and mobile computer systems. At the time of the case, Hdrive had approximately 10,000 employees worldwide and was roughly number 5 in worldwide unit volume. At the time of the case, it was in financial distress, in large part because of delays in getting previous product generations to market on time at high volume and low cost – in other words, problems with development and ramp-up.

Pilot production of the company's products, as well as cost reduction, quality and product improvement engineering on current products, were conducted at development facilities in the USA. All HDD assembly and test discussed in this article were performed at a company owned and operated facility in Singapore.

The data on which the present study builds results from a longitudinal case study at HDrive. Data was collected during one author's internship with the company. The third author spent three months with the development center in the USA, and six months with the manufacturing facility in Singapore. He followed one new model as it transitioned from development to manufacturing, attending relevant meetings and teleconferences, and collecting progress reports, bug reports and other documentation. The third author conducted additional site visits and interviews, both in Singapore and the USA. Additional data on the company and the overall industry were collected as part of a larger research effort on the disk drive industry. Further details about the methodology of data collection and other descriptions of the company and research sites are given in Chea (1997).

4. A description of pilot production, transition, and ramp-up

Our study focuses on pilot production, transition and ramp-up of a new product called Alpha5. Alpha5 contained a number of product and process innovations. The biggest process innovation was the use of tabletop clean areas called mini-Es, a miniaturized substitute for expensive cleanrooms. Among other advantages, the mini-Es allow modular capacity expansion, as each line has only about five operators and makes hundreds of drives per day (compared with total volume targets of thousands per day in full production). Other process innovations included computerized dispositioning of drives that failed test, redesigned tooling to fit in the restricted space of the mini-Es, and the first use of an outside vendor for head stack subassembly. Product innovations included a new head type and new media, both from outside vendors.

This section sketches the main events of pilot production (in the USA) and ramp-up (in Singapore). All dates are given relative to the official transition date which is referred to as Day 0. Pilot production for the Alpha5 began in the US on day −27.

Pilot production

The purpose of pilot production was to bring together and validate all the components of the production system, including materials, processes, tooling, vendors, and personnel. Specific numerical targets for pilot production included measures of first pass yield,[4] rework yield, process induced failures, test time, and tact time. Most of the targets were beyond performance levels achieved in pilot production by any previous HDrive product. Pilot production in the USA needed to meet these goals before Singapore officially accepted the transfer of Alpha5.

To respond to the need for coordination and to overcome its historical problems with time to market, HDrive had recently created New Product Introduction (NPI) groups in the USA and Singapore. NPI-USA was in charge of pilot production. Approximately 300 employees were in the NPI-USA structure, organized functionally for pilot line operations, test equipment, reliability, and clean room systems. NPI-USA had a counterpart in Singapore, with the mission of ramping-up and stabilizing the product in the Singapore production facility. NPI Singapore was substantially smaller in head count, serving primarily as a coordinating arm of the factory for new product introductions.

The NPI Singapore program manager for Alpha5 came to the USA approximately 6 months before transition. Among her tasks were to confirm test schedules, to coordinate Singapore's transition personnel in the USA, to relay communication needs and request from NPI Singapore, and to prepare for the transition.

The transfer of other personnel from Singapore to Longmont commenced before day −50 with everyone sent home shortly before the transition date (Figure 1). Engineers came from Singapore representing a variety of disciplines, including process engineering, failure analysis, material sciences, and tooling. Some of the personnel came directly from NPI Singapore while others came from their respective functional departments. They continued to report to their managers and directors in Singapore.

In addition to process engineering and fine-tuning of the production process, another purpose of pilot production was to familiarize the Singapore team with the Alpha5 assembly process. A team of 24 operators came to the pilot line for training approximately one month prior to Day 0.

The Longmont pilot line operated very irregularly because of its experimental nature. Problem solving took time, and so did the repair and recalibration of the production equipment. Another delaying factor was the supply of components, resulting from the suppliers' own ramp-up problems. These included vendors for head stack assemblies, PCBAs, baseplates, motors, disks, etc. Sometimes components were not available at all, and sometimes their quality was unacceptable. Delays in pilot production ranged from hours to days.

Prior to transition, and in parallel with pilot production in the USA, test and production equipment were installed and calibrated in Singapore. Initially only two lines were set up. Other lines were tooled and brought up after transition. Many of the problems with these two lines were seemingly trivial, such as power cords too short, and parts bins too bulky. However trivial, these were real problems to be dealt with, and are similar to many issues encountered during pilot production and ramp-up.

Two days prior to the transition date, one Singapore production line was tested. After six hours, only 2 disk drives had been built, a very frustrating experience for all involved. The tools kept breaking down. On top of this, it was not apparent if the drives actually worked, as testing took a day. The next day, the line built approximately six disk drives.

The transition

The transition officially occurred on the scheduled date, 'Day 0'. NPI-USA and NPI Singapore met by video-conference to review the check-off report for transition. The meeting started at 08 : 00 in Singapore and 17 : 00 in the USA.

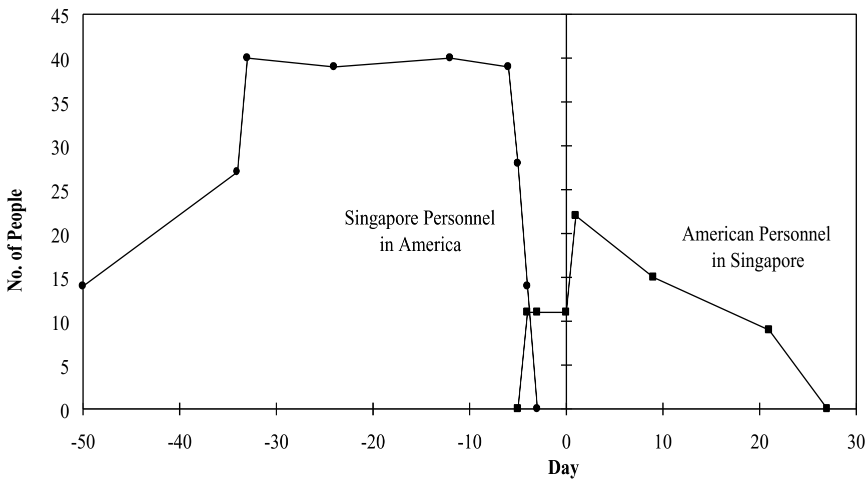

Figure 1. Movement of personnel.

The check-off report was 40 pages long, with categories including materials, processes, and tooling. It contained specific numerical targets to be met by the pilot line before transition. Every individual issue on the list had to be evaluated as 'Go', 'Exception', or 'No Go'. For Alpha5, the distribution turned out to be 15 Go's, 20 Exceptions, and 0 No Go. All goals and objectives not met were discussed during this meeting. Some of the exceptions were:

- New build first pass yield was only 87% of the pilot line target.
- Rework yield was only 83% of its target
- Head stack assembly component from the planned vendor was short by 40,000.
- The defective parts per million was double the target.
- The tact time was 150% of its target.

In the end, despite all the exceptions, the VP of NPI formally accepted Alpha5's transition from NPI-USA, with the understanding that the fixes to the exception issues were to be done as promised.

After Day 0, the USA continued to build a few drives for testing and process improvement.

Organizational structure and inter-site communication

NPI Singapore's mission was to ramp-up and stabilize the product before handing it over to volume production. Project managers could be classified as lightweight managers (Wheelwright and Clark, 1992). They had considerable expertise in coordinating the activities of the functional departments. However, the key resources, including engineers, were still under the control of the functional departments outside of NPI. Despite the limited resources, NPI Singapore had considerable influence on ramp-up success, as the manager controlling the production line reported within NPI-Singapore. He decided what, how, and when the drives were built.

There were ample methods of communication between locations. Engineers and managers communicated via email, telephone, fax, and video conferencing. Raw data, drawings, and text were shared in real time via Hdrive's corporate network. The biggest problem in communication was the time difference. At 08 : 00 when Singaporeans were starting their workday, workers in development were leaving work at 19 : 00 the previous day. It was difficult to hold real time meetings by phone or video conferencing this way. Electronic mail was used, but for the same reason it had a response time lag.

As during pilot production, moving people was a vital method of communication. Eleven USA engineers arrived just prior to Day 0 to assist in the ramp-up (Figure 1). These engineers included failure analysis, tooling, and information specialists. The trained Singaporean operators returned from the USA at the same time.

Reproducing the pilot line

In the various engineering meetings in Singapore, the phrase 'What was development using?' or 'We need to hear from development' was common. The strategy for the ramp-up was to solve as many problems as possible during pilot production and then reproduce the pilot production process in Singapore. However, reproducing something as complex as a production process is a difficult task.

One example of the multitude of problems that had to be dealt with to reproduce the pilot production process was the computer tracking system (FIS – Factory Information System). In the USA, there was only one pilot line for Alpha5. In Singapore, there were multiple lines. As a result, the software created in the USA did not allow tracking of drives switched from one line to another. However, such line switching occurred for special cases of rework, causing the FIS to fail in Singapore.

When the transition occurred, some assembly procedures had been modified in the USA. However, the documentation was not changed to reflect the new procedures. Thus, operators in Singapore were not alerted to the changes. Fortunately, news of the change came via the engineers visiting Singapore.

The attempt to reproduce a production line across 14 time zones went both ways. Not only did Singapore attempt to replicate the line in the USA, but phrases like 'What is the plant using?' or 'We need to copy what they are doing' were common in the USA. Thus, during the initial period of ramp-up, engineers tried to keep the two production facilities as identical as possible. This allowed engineers at the development facility to replicate problems encountered in Singapore and to increase the capacity for experimentation.

Implementing the new process in Singapore

The ramp-up goals for NPI Singapore included the following. Once achieved, the process would be finished with introduction and turned over to normal factory line management.

- Achieve β new build first pass yield and ν rework process yield.[5]
- Less than a specified number of hours of testing.
- Less than ω process induced failures.
- Tact time[6] of ε
- Downtime of less than $x\%$.

These numerical targets were higher than the corresponding goals of pilot production. The targets were also higher than for previous products at comparable points.

During the first week post transition the lines ran three shifts, and by the end of the week two more lines were brought up. However, most performance indicators were far below their targets:

- The first pass yield was at 0.76 β which was substantially below the pilot line's performance post transfer. Such a yield drop after transfer is normal in HDD production, but was still a disappointment.
- Frequent tool breakdowns resulted in line down time as high as 70% of total operating time.
- Tact time was erratic, ranging from 2.9 ε to 12.9 ε across lines and days due to inexperienced operators and tool problems.
- The defective parts per million (measured by an audit of ready to ship drives, after final testing) was still far over the target level.

Problems and engineering change orders

Most problems during pilot and ramp-up were unforeseen, and their solutions required multiple changes to the product and process. We documented about 120 'issues' that needed resolution in pilot production over a six-week period prior to transfer, and 55 in Singapore manufacturing during the four weeks after transfer. The pilot line (pre transfer) issues were overwhelmingly tooling, such as 'measurement tool not online yet' and testing, such as 'code difference between the two sites for testing at STW station'. Twenty of the 55 production line (post-transfer) issues were materials issues such as 'vendor not sending components meeting specification'. Process, tooling, and yield issues also had problems during ramp-up.

HDrive had a defined process for engineers to go through when requesting a change. Such engineering change orders (ECOs) provide a good method of measuring a development project in its final phases. Not all issues required engineering changes to resolve them, but there were 53 ECOs from week −4 to week 10. Figure 2 shows the timing of ECOs. The number of ECOs shows more changes soon after transition, consistent with Terwiesch and Loch (1999). Week 10's spurt of ECOs was largely due to qualifying new vendors. Ten out of 53 ECOs filed concerned documentation modification or addition. Twenty-five ECOs concerned material and component specifications.

History of production ramp-up for Alpha5

We tracked first-pass yields as the best single measure of line performance. The first two weeks of ramp-up production had yields as low as 0.58 β, which increased scrap cost and reduced good output. The biggest yield fallout occurred at two stations, Servo Track Writer (STW) and Self Test. As a result, a policy was implemented in which any time the yield was lower than 1.03 β^7 on these two test stations, the line was shut down. The testing engineers also verified whether the failures were due to the disk drives themselves or to the testing equipment by re-testing all failed drives on different equipment.

As mentioned, yields dropped during the first week of production in Singapore compared with yields on the pilot line. Once the shutdown system was instituted, issues were resolved faster, engineers understood the problems more, and the yields started to increase. By day 15, the overall yield reached 0.88 β, an unprecedented achievement for HDrive. Selected data on yields are shown in Figure 3 and Figure 7. The yield in the US pilot line prior to transition is also shown on the left. The upward arrow shows that higher is better on this figure.

Yield improvements resulted from better understanding of problems, and consequent ECOs and vendor improvements. For example, the head stack components received from the vendor had unacceptable variations. The obvious immediate solution would be to not accept the components from the vendor. However, that would require shutting down the production lines since the components were in short supply. The real solution, however, was to loosen the passing criteria on the testers, once it was understood that the criteria had been overly stringent. To

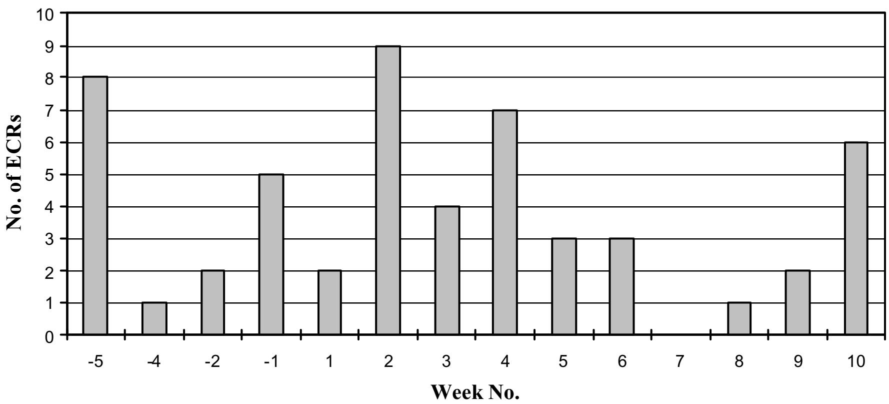

Figure 2. Engineering change orders.

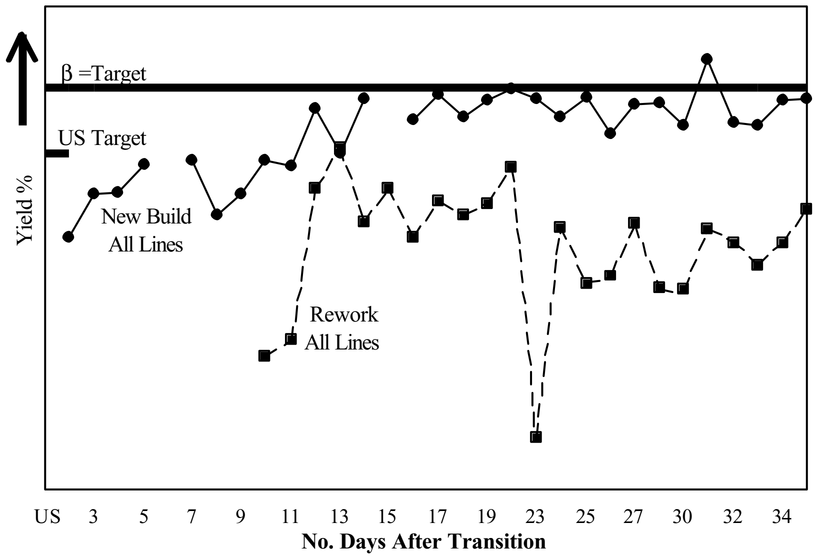

Figure 3. Yield history.

improve the assembly process further, an experimental line was set up to investigate more automation of the assembly process. This line was assigned to a process improvement team and was not a part of normal production.

Figure 4 shows the overall down time percentage, which equals line shutdown time divided by total (attempted) operating hours. Although Figure 4 shows little down time between days 3 and 7, during those days the line was not building at a steady pace (i.e. it had long and erratic tact times, see Figure 5). Only small amounts were built because of limiting factors such as materials availability and quality issues. The lines were almost completely shut down on day 7 as a management response to low yields, to provide more time for problem solving. When the transition to high volume began, the amount of down time increased considerably. However, once the major defects in the machines were sorted out and engineers had gotten into the routine of fixing machines immediately when the line was down, downtime dropped and stabilized around the targeted level.

The assembly tact time, i.e. the production interval between drives, started far above (worse than) its target. This is illustrated in Figure 5. For the first three weeks of production, there were substantial variations among lines and days. Spikes were due to problems caused by material issues, choice of line pacing, one-time occurrences of defects, etc. Eventually the times declined almost to their target levels, but this is one performance indicator that never reached its target. Management decided that slightly higher labor costs were not important enough to justify further engineering effort, especially compared with yield and other issues.

Figure 6 shows the overall volume ramp-up from day 6 on. The solid line is the plan set daily, and the broken line is the actual build. The daily plan reflects known disturbances such as parts shortages. By day 23, the steady-state volume target was achieved. The fluctuations in days 41 through 50 were caused by vendor problems; the heads were experiencing contamination issues. The vendors thus had a direct influence on yields and volume at HDrive.

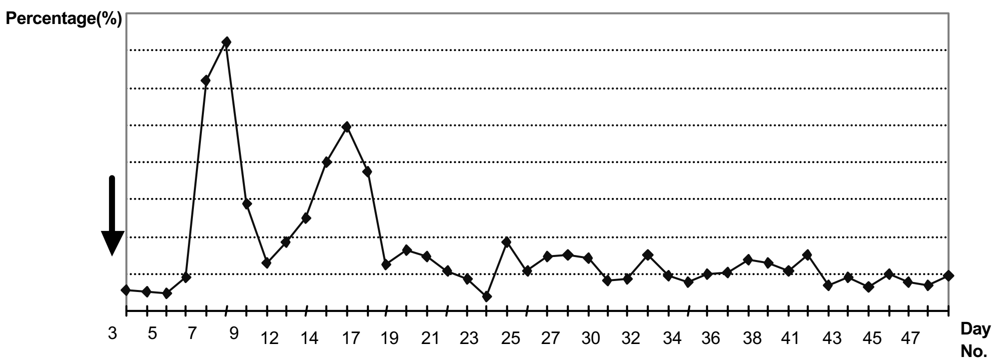

Figure 4. Downtime history.

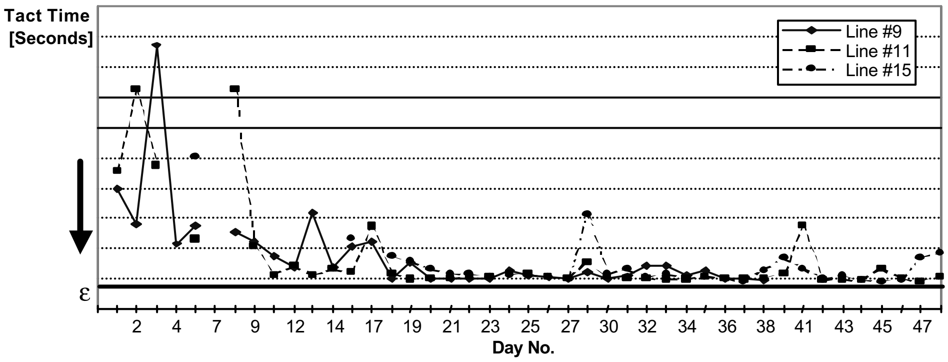

Figure 5. Tact time history.

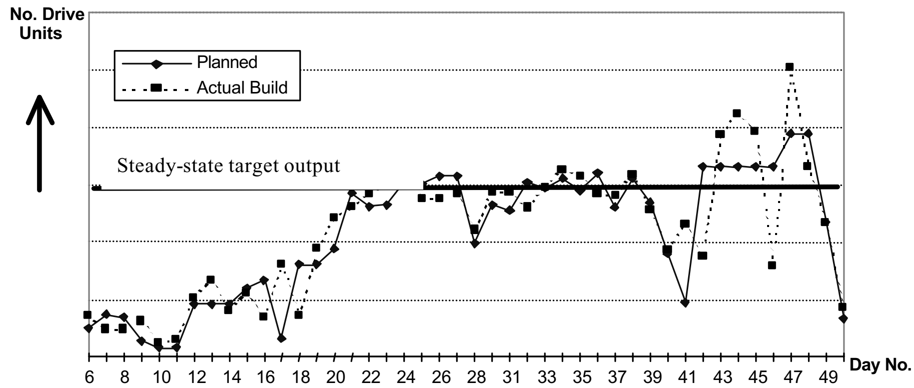

Figure 6: Volume Ramp

Figure 6. Volume ramp.

Other activities during ramp-up

Increasing volume and yields were not the only tasks during ramp-up. Other tasks included:

- Process maturity testing. The first several hundred drives produced in Singapore were run continuously under extreme conditions for four weeks;
- Establishing the capability to meet all unique customer requirements such as special software or firmware;
- Qualifying multiple vendors for components, and determining a 'vendor matrix' of which components from different vendors could be used together.

5. The activities of transfer and ramp-up (Question 1)

We now return to the research questions posed earlier. We have already covered the mechanics of what happened. But a chronological recitation of events does not explain why the ramp-up of Alpha5 went much better than previous HDrive ramps. What follows are our interpretations and analysis.

Fundamentally, the big 0–1 event of 'design it in the USA, transfer on Day 0, and manufacture it in Singapore' was broken down into many, well defined, discrete, measurable, and controllable steps, while preserving flexibility and resources for learning, to respond to the unexpected. This follows the 'beat Murphy's Law' concept (Chew *et al*., 1991). A single large innovation or change is inherently hard to predict and subject to unpleasant surprises. By breaking it down into many smaller changes, more issues can be anticipated up front. By providing flexibility and learning mechanisms at each stage, surprises could be dealt with as they occurred.

Furthermore the activities of transfer were planned and executed in a high degree of detail, but without sacrificing flexibility to deal with the unexpected. The 40-page transition check-off report set clear targets for every part of the product and process. Movement of people, tooling, and software was thought through in detail. The project manager from the factory went to the USA six months before transition to provide detailed coordination. Movement of people and equipment was choreographed in advance. Yet as it turned out, the check off for 'Day 0' became elastic. By postponing achievement of more than half of the target criteria, the transition became a *time interval*, rather than a *time point*.

Despite the seeming sharpness of the 'Day 0' timing, the transfer was actually done gradually with respect to people, tools, and goals. For example people moved gradually, with four distinct steps:

- Developers in the USA, manufacturing people in Singapore (up to day −50)
- Both types of engineers in the USA (from −50), with operators also in the USA (from −38)
- All manufacturing people, plus a declining number of development engineers, in Singapore (from −1 to 28)
- All manufacturing people in Singapore with prime responsibility; developers called in electronically (from day 28 onward).

Similarly, the locus of production shifted gradually from the USA to Singapore. A few units were built in Singapore before Day 0, while the pilot line continued in the USA after transition. Additional production lines beyond the first two were built, staffed, and started up in Singapore gradually, as the previous lines were debugged. The third and fourth lines were started several weeks earlier than first planned, due to the rapid rise in performance of the first two lines.

Because of the novelty of the product design, of the process (first use of mini-Es), and of the project management methods (use of the NPI department), task duration was inherently highly uncertain. In such environments, firm deadlines typically result in either schedule slippage or incomplete debugging, causing more severe problems later. Instead of using fixed deadlines, the transfer from the USA (development) to Singapore (manufacturing) was marked by *gradualism*, a high degree of planning of *what* was to take place and the sequence, and flexibility in *when* activities started and finished. For example US engineers stayed in Singapore until all problems involving their expertise were solved, not just up to a fixed return date determined in advance.

Overlap of product generations

For Hdrive, Alpha5 provided majors innovation in product design, process design, and transfer management. Many of the newly introduced process techniques and product features were to be reused in coming products. In this sense, Alpha5 was the start of a new *product platform*. Within the platform, subsequent products had a common architecture and reused a substantial portion of the HDA, custom electronics and firmware. In fact, within HDrive, Alpha5 was seen as a big success, although it was only produced for a few months and probably never directly paid off its fixed development costs. Management had planned for this and marketed the next product, Beta8, as the key high-volume product in the company's turnaround. Management felt that the first effort to use so many new innovations was too risky to depend on for major product volumes.

Wheelwright and Clark (1992) as well as Robertson and Ulrich (1998) propose several benefits of product platforms, including increased development efficiency by sharing components between products and the possibility of creating whole product streams rather than individual products. The power of the platform approach in managing transfer and ramp-up is visible from the products succeeding Alpha5 and based on its platform, namely Beta8 and Theta9. While most products prior to Alpha5 failed in meeting the yield target (Pi3 was the product before Alpha5), Beta8 and Theta9 showed the best yield levels HDrive had ever experienced. A comparison between the yield curves is given in Figure 7. Beta8, the product directly following Alpha5, showed 5–10% higher yield during ramp-up than Alpha5. Theta9 started another 5% higher and approached the target yield within four weeks.

Figure 7 also illustrates that much of the improvement from generation to generation within the family occurred during product/process design and transfer, not only during ramp-up. The starting (week 1) yields of successive products are each higher than the product

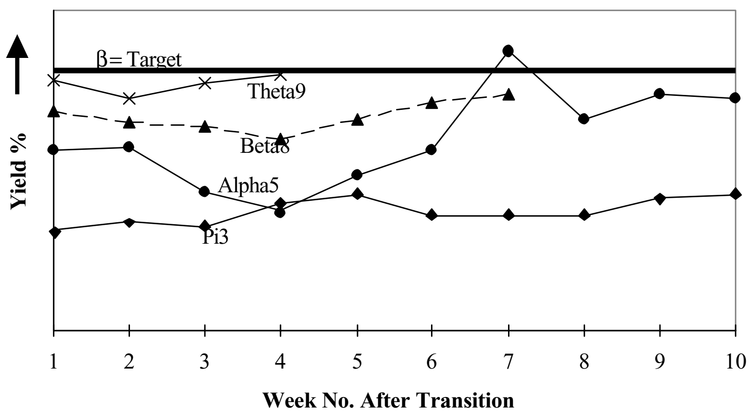

Figure 7. Yield history of successive products.[8]

before. We do not have enough information to know the precise cause of these improvements, but there are at least three phenomena at work. First, the NPI team and others at Hdrive are gaining experience with doing coordination and transfers in the new way. Second, because the successive products are similar to their predecessors, many of the process problem solutions in one generation carry over to the next generation. There are fewer sources of yield fallout remaining. Third, there is feedback from manufacturing to designers, in order to get more manufacturable designs. Many sources of yield fallout in one generation can be traced to tolerance, product design, and process capability issues that are then improved for the next generation. Because of the rapid succession of product generations in disk drives, many companies in the industry do not feed back such information to designers for the *current* product, but save it for the next generation. That is, once the product is transfered, only a very serious design problem would be fixed immediately.

The use of product platforms extends our discussion of gradualism from people and tools to the inter-product level. Product platforms avoid discrete large innovations in favor of a continuous stream of smaller innovations. Furthermore, the total engineering effort needed over several generations is considerably reduced, as is the amount of new tooling.

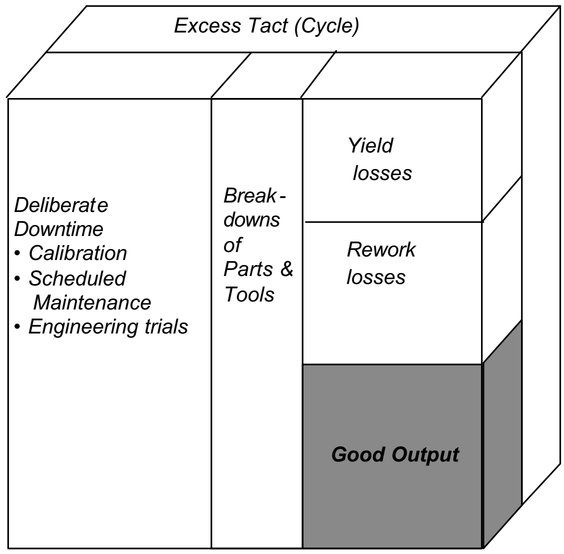

Figure 8. The determinants of effective capacity.

6. Volume increase: why and how? (Question 2)

Figure 6 illustrates how the total number of good drives produced a day rose over the course of the ramp-up. Clark and Fujimoto call this trajectory the ramp-up curve. The ramp-up curve is a very aggregate view of plant performance, hiding the detailed causes of *why* during ramp-up, less than half of the (long-term) targeted volume was produced.

As we have discussed in the case description, the most important ramp-up indicator for the NPI team was the yield curve (Figure 3). Low yields caused a substantial amount of scrap, and reduce output. However, yields are not a sufficient description of the plant's performance. Consider the first two weeks in the ramp-up of Alpha5: although yields were only moderately below their target level, the excessive tact time (by a factor of 2 or more) kept output low.

To understand the detailed drivers of the plant's performance during ramp-up, we need to pull the various performance measures together. Figure 8 graphically relates the effects of yields (Figure 3), downtime (Figure 4), and tact time (Figure 5) on output, and introduces the concept of *effective capacity utilization.* Consider the horizontal axis first. Over the course of a week, each production line is available for a set number of hours. However, not all of this time is used for production. Machines need to be maintained and calibrated, especially during ramp-up when setpoints and calibration methods are being learned. Engineers run experiments to further improve their understanding of the production process. The line may be stopped for process reasons, as discussed earlier. Finally, it is common during ramp-up that machines break down and suppliers cannot supply enough of critical components, which causes involuntary down-time. This leaves only a fraction of the machine-hours available for production. The exact downtime changes over the ramp-up and is measured in Figure 4.[9] The horizontal axis can also be 'stretched' by adding new lines or shifts. Alpha5 started with two lines at 3 shifts, added two more lines in the second week, and at peak had five lines.

Next, consider the *z* axis (going into the page). Early in the ramp-up, it takes many times longer to perform some operations. The improvement of tact time over the ramp-up is shown in Figure 5. Improvements come from operator learning, industrial engineering, and speeding up machine-paced operations. (In practice, the measurement of tact time is intermixed with the effects of unscheduled machine downtime.)

Finally, the vertical axis shows yield effects. First, some operations have to be performed several times for the same disk drive because of rework. Such repeated visits to operators and testing equipment is wasted from the perspective of output, and uses capacity that could have been used for new builds. Second, some drives have to be scrapped, which further reduces output (Bohn and Terwiesch, 1999). The improvement in yields over time is shown in Figure 3.

This leaves only a small volume, which corresponds to the output of good drives per week. It is a shaded cube in Figure 8 and its proportion of its surrounding cube is the effective utilization of the available production capacity.[10]

Note that the three dimensions in Figure 8 can be partially substituted for each other. For example, equipment can be adjusted less often and run faster, at the expense of lower yields. In fact, comparing the trajectories of the individual performance drivers, it is notable that the yields start relatively close to their target, compared with downtime and line speed. This fits an emphasis on yields during the pilot line period, as yields have a very high degree of product/process interaction and thus benefit from improvement before the product design is completed. There is also an economic rationale for this (Terwiesch and Bohn, 1999).

Locations and methods of learning

Uncovering and solving problems, and learning from the solutions, are essential activities during both prototyping and ramp-up. The product, the process, and the materials from vendors must all be 'debugged'. The standard sequence for this learning on Alpha5 was:

- Failed drives in production signal a problem;
- Failure analysis of each bad drive to repeat the problem and determine the basic location on the drive (e.g. circuit card, mechanical, head/disk, or process failure);
- Send the drives with similar failures to specific engineering groups for root cause analysis;
- Develop a quick fix, e.g. add an incoming visual inspection to weed out bad components;
- Develop a long term solution;
- Run a controlled experiment on 50 to 300 drives to verify the solution.

In addition, some proposed changes come from other sources besides failed drives, e.g. proposed changes in testing software. Such changes must also be verified.

One of the interesting features of the Alpha5 is that there were at least six different locations where systematic learning was done:

- Pilot production in the USA;
- Singapore production of five drives a few days before transfer;
- Startup in Singapore on only two lines;
- Some continued pilot line production and experimentation after transfer;
- A special line used for trying new concepts in Singapore;
- Some learning from regular high-volume production after ramp-up.

Some of these locations were in use simultaneously.

The multiplicity of learning locations is a response to several related forces on the learning process. The most important issue is that solving problems during ramp-up is expensive because of reduced yields and the lost production time for experiments (see the earlier discussion of effective capacity), and solving them after ramp-up is even more expensive for these reasons and because key engineers may not easily be available. Therefore, it is desirable to remove as many bugs as possible as early as possible, such as during pilot production. This accounts for the aggressive learning at early stages of the Alpha5 project. Second, it is desirable to keep production activities somewhat independent from learning activities, to maximize good output while allowing learning to continue. This accounts for the special learning line in Singapore and continued pilot line operation after transfer. Experiments done on them could have been done on the main production line, but at a higher cost in terms of lost output.

The third factor is the *fidelity* of the learning process to the final production process (Bohn, 1987; Thomke, 1998). It does little good to learn about problems if the same problems will not come up in normal production. Conversely, it is desirable to have the earlier learning locations detect as many problems as possible to reduce the problems in later stages. The way to achieve both goals is to have the early locations, such as the US pilot line, be high fidelity copies of the ultimate production lines. By bringing in operators and production equipment from Singapore, the development facility was able to create a high fidelity representation of the Singapore factory. This allowed the engineers to solve many problems before transfer, and also to solve them in the USA, where more engineering knowledge resided.

Once the problems were solved in the pilot production environment, the solutions and knowledge formed the starting point for the production lines. The common phrase in Singapore was 'what did the US use?' However, the transferred knowledge was incomplete for local needs in Singapore because fidelity among the locations was not perfect. Higher volumes expose more issues, including low frequency failures that were not clearly discernable at pilot line volumes. Vendor processes often shift as vendors go to higher volumes as well. Thus, not all learning can be done in the pilot phase.

Other innovations for learning

The use of so many locations for learning on the Alpha5 was unusual even for HDrive. The NPI team also created or made use of several other innovations to facilitate rapid learning.

The use of mini-Es, although usually discussed in terms of avoiding the cost of clean rooms, was also an important innovation for faster learning. In essence the mini-E approach shrank the minimum efficient capacity unit. Conventional HDD factories, including those at Hdrive, have large clean rooms for HDA assembly with dozens of workers on one line for each product. Adding more capacity requires integrating more workers and more stations into the existing line. Trying to run such a line at very low output, such as for pilot production, is very inefficient. Either the line must be run only sporadically, or different products must be intermixed on the same line. Similarly, running an engineering trial must be run on the same line being used for normal production, with consequent disruption and chance for errors. In contrast, the Alpha5 was produced on multiple lines in parallel, with about seven workers and mini-Es per line. Adding more capacity can be done by 'cloning' the line. And it is easy and inexpensive to dedicate a single line to a special purpose, such as running an engineering trial. A further benefit was that it allowed the pilot line in the USA to be laid out and tooled almost identically to the actual production lines. As discussed, this led to higher fidelity which contributed to rapid learning.

Finally, we observed HDrive using an unusual process of *deliberately* shutting lines down when their yield was low, in order to foster problem solving by engineers. All production was shut down for part of day 7, because of low yields. This was later institutionalized to shut down any line with yield below 0.82 β Certain test stations were also singled out, and their lines were shut down any time their yields were below 0.82 β. The purpose of these shutdowns was to allow engineers to focus on solving problems when problems became acute. All departments were notified and expected to respond immediately. This was effective in getting problems solved fast.[11] It also served to limit engineering fire-fighting, because once a line stops it no longer generates new problems. Without such limits, fire-fighting can become an endemic problem in ramp-up (Bohn, 2000).

7. Global transfer (Question 3)

Production of HDDs strongly requires integration among product development, process development, and production. Most of the recent product development literature recommends 'co-locating' the activities under these conditions, i.e. to bring them physically close together. However, the HDD industry has not followed this approach, and indeed almost all firms in the industry follows a model with development and manufacturing on different continents (Gourevitch *et al*., 1997). How can this be done while allowing rapid transfer and ramp-up?

In our case study, we observed several mechanisms designed to overcome the physical separation between development and manufacturing: exchange of personnel, transfer of production tools including software, creation of a common IT platform, and heavy use of physical and electronic communications.

Figure 1 shows the personnel transfer that occurred. About 40 people moved from Singapore to the USA before transfer and stayed 1500 person-days. Travel and housing costs for this would be roughly US$280,000, of which half was paid by the Singapore government as an incentive to companies to upgrade skills of Singaporean citizens. Post transfer, 22 engineers went to Singapore from the USA and stayed 600 person days, at a cost of roughly US$135,000, with no reimbursement from a government. Thus out-of-pocket costs to HDrive for personnel movement were roughly US$300,000.

In perspective, this is a small amount. $300,000 is less than the value of one day of steady-state output of Alpha5. Since the average selling prices for a new HDD decline roughly 1% each week, and since previous ramp-ups at HDrive took months instead of weeks, these costs are easily justified if they were partly responsible for the acceleration of the ramp-up process. Another comparison is that the salaries of the 62 employees for the time they were abroad came to roughly US$400,000. These salaries would have been paid for the time of people working on the transfer, whether they were in their home countries or not. Thus, the out of pocket costs of physical separation are moderate compared with other costs of transfer and the benefits of doing transfer well.[12]

We have already discussed the heavy use of electronic communications between the two sites. The common IT system between the two sites greatly facilitated communication. When physical objects had to be sent (e.g. sample components, tools, and assemblies for failure analysis), a one pound FedEx package cost $30 and took two days. This is an industry in which all final products and almost all intermediate materials are shipped by air, because of high value/weight ratios and rapid depreciation of inventory.

Finally, the international transfer was facilitated by movement of physical tooling. Production equipment, including testers, tooling, and the mini-Es, was procured, built and tested in Longmont, then shipped to the factory. Eventually, delivery from vendors of the less complex and more standard equipment including the mini-Es was shifted to Singapore, but all initial ramp-up and debugging in the factory was done on equipment which came from the prototype facility and therefore was 'known good'. Thus physical distance was not allowed to interfere with establishing tight fidelity between the different facilities.

In short, the huge geographical separation between the development/prototyping and manufacturing facilities was not a major impediment or cost in the transfer process. Language was also not a barrier, as all Singapore engineers spoke English, and operators in Singapore who did not speak English had interpreters easily available. There *were* many transfer impediments that had to be overcome effectively, as we discussed earlier, but they come up in any transfer from the culture of development to the culture of manufacturing, and are not due to international issues or distance. The highly structured transfer process, and other mechanisms we have discussed, were effective in dealing with the cultural and distance issues.

The Alpha5 project in perspective

We have documented what HDrive did on the Alpha5 project, and have discussed how the project's good results can be explained by the various actions. How could the transfer have been handled differently?

Adler (1995) proposes a typology and rationale for coordination between design and manufacturing in any industry. He studied coordination in two very different processes: printed circuit boards (electronics) and aircraft hydraulic tubing (mechanical). He identified five possible generic coordination mechanisms: none, standards (e.g. design rules), schedules and plans, mutual adjustment, and joint teams. These are in order of increasing degree of interaction, and presumably also in order of increasing expense. He also identified three different stages, two of which are relevant here: product/process design stage, and manufacturing stage. In terms of this typology, the Alpha5 project was marked by schedules and plans during the design stage (sign-offs by manufacturing), and by joint teams plus a moderate amount of mutual adjustment (ECOs) during the manufacturing stage.

How well does Adler's framework explain what was done for Alpha5? He found that increasing design novelty required the use of the more interactive coordination mechanisms. He defines degree of novelty in terms of how many design-manufacturing interaction problems existed that the firm had not encountered before. Thus it would tend to be higher for radically new products. At least in retrospect, the degree of novelty for Alpha5 was low: almost every problem that came up had been seen before for an earlier product. Furthermore, he found that in situations where design-manufacturing interaction was harder to analyze, the coordination tended to take place later in the process. Analyzability is defined as the difficulty of search for solutions to problems. In his words, 'analyzability ... will be particularly low when a new product requires a new manufacturing process, when the design tools do not allow a representation of the entire product, and when these tools do not allow simulation of product performance'. The third of these criteria is met for all hard disk drives, while the first of them (a rather new process) was in effect for Alpha5, making a moderate level of analyzability.

Thus the framework seems to predict that coordination will take place partly in the design phase and partly in the manufacturing phase, because of the moderate analyzability of the Alpha5. Indeed it took place in both phases, but more in the manufacturing phase. The factory got heavily involved less than two months before Day 0. The framework also predicts the use of low-interaction methods of coordination, such as standards, schedules, and plans, because of the low degree of novelty. This prediction is borne out for the design phase (which used sign-offs), but it is not borne out for the manufacturing phase, where transition teams were used heavily though the framework predicts that simpler coordination methods should have sufficed. It should be pointed out, though, that the creation of the NPI team mechanism, and the large amount of effort devoted to a smooth transfer and ramp-up, was new to HDrive corporation, and was a reaction to problems with past products. Relative to past products, much about Alpha5 was novel for HDrive Corporation.

If we look at a cross-section of projects and hard disk drive companies, we can make more sense of the differences among transfers. As part of a different project, we have interviewed at a number of companies about transfers from the USA to SE Asia. We also directly observed the Beta8 transfer in HDrive. All companies used electronic communications between design and manufacturing prior to transfer and during ramp-up. But Alpha5 was the most elaborate transfer we heard about, in that it involved more people movement, more formal sign-offs, and more elaborate preparation for the transfer. The less difficulty managers anticipate with the transfer, the later and the fewer engineers they send from manufacturing back to the USA, and the fewer engineers they send to Asia after the transfer. Where the transfer is expected to be straightforward, companies will sometimes rely entirely on movement of tools, prototypes, and information, with virtually no people movement. At least in 1999 we did not hear of any instances where large numbers of supervisors or operators were sent between locations, as was done for Alpha5. Another very significant trend is that several companies are shifting some of their pilot line activity to their manufacturing plants, and out of the US design facilities. Thus the timing of inter-site transfer is moving earlier, even while the degree of coordination goes down.

In short, there are differences in the elaborateness of transfers, which seem to be determined by on anticipated degree of difficulty. We did not investigate how 'degree of difficulty' of a transfer is estimated, but anecdotally it corresponds primarily to novelty of the product and process, and secondarily to process capability for the most difficult manufacturing processes.

8. Summary and managerial implications

In this study we have documented product transfer and production ramp-up of a 'high tech commodity'. Although some of our findings are specific to our host organization or the disk-drive industry, we believe that there are several implications for other high-tech companies.

With increasing development expenses and growing complexity of production equipment, the disk drive industry is pressured to achieve ever shorter times to ramp-up their factories. Although a 25-day ramp-up period looks short at first sight, it provides a significant portion of the product lifecycle. Further, given an annual price erosion of 60%, this ramp-up period coincides with the potentially most profitable time in a new product's life.

Our study identifies several organizational patterns that seem to facilitate a product's transfer to volume production and to shorten the ramp-up period. Each of them contributes towards a gradual, almost continuous process from pilot into volume production. First, by running pilot production and ramp-up in parallel for an interval, the 'real' transition point (day zero in our study) became much more fluid than in the case of a fixed hand-over. Second, the organizational responsibilities complemented this fluid transition. Both NPI-US and NPI Singapore were dedicated to production ramp-up. Further, the transition between the two was smoothed by a substantial movement of personnel and tools. Finally, product platforms allow a company to leverage previous ramp-up experience into the ramp-up of new products on the platform. Both the initial yields as well as the rate of yield improvement benefited from this platform approach.

Another interesting implication results from our observation of product transfer across a long geographic distance. Research on communication and coordination in R&D environments has long emphasized the importance of colocation of various organizational functions, specifically development and manufacturing. However, there can be strong economic forces encouraging off-shore production. Our study documents that even for a complex and high-tech product it is possible to overcome geographic distances. In fact, the international transfer was able to proceed using elaborate coordination mechanisms, namely cross-functional and cross-location teams. This provides management with an additional degree of freedom when making location decisions for both development and manufacturing.

9. Conclusion and future research

The academic contribution of our research follow from its managerial implications. Production ramp-up and the related concept of time-to-volume have been treated as a 'small brother' of time-to-market. The relatively unexplored topic of production ramp-up offers several opportunities for future research. First, additional studies are needed to overcome the limitations of our single company research approach. Second, future research could analyze more of the market/industry dynamics occurring with the launch of a new product. How does the ramp-up affect the diffusion process of a new product? How should one price a product, which is still undergoing ramp-up? Finally, there seem to be numerous opportunities to formalize some of the drivers of effective capacity into mathematical models. We have started to pursue this idea in Terwiesch and Bohn (2001).

Acknowledgement

This research was supported by the Information Storage Industry Center, funded by the Alfred P. Sloan Foundation, by the Program in Advanced Manufacturing funded by the National Science Foundation, and by HDrive Corporation. Our special thanks for great assistance from the people of Hdrive in both Singapore and the US.

References

Adler, P.S. (1995) Interdepartmental interdependence and coordination: the case of the design/manufacturing interface. *Organization Science*, **6**, 147–167.

Bohn, R. (2000) Stop fighting fires. *Harvard Business Review*, **78**, 4, 82–91.

Bohn, R.E. (1987) Learning by experimentation in manufacturing. *Working Paper*, Harvard Business School.

Bohn, R. and Jaikumar, R. (1986) Priam Corporation – hard disk drive manufacturing. Harvard Business School, *HBS case* 685–096.

Bohn, R.E. and Terwiesch, C. (1999) The economics of yield-driven processes. *Journal of Operations Management*, **18**, December.

Chea, K.S. (1997) Pilot Production, Transition, and Ramp-up of a New Product across an International Boundary: Hard Disk Drives in Singapore. MS thesis, UC San Diego.

Chew, W.B., Leonard-Barton, D. and Bohn, R.E. (1991) Beating Murphy's Law. *Sloan Management Review*, **32**, 5–16.

Clark, K.B. and Fujimoto, T. (1991) *Product Development Performance: Strategy, Organization and Management in the World Auto Industry*. Harvard Business School Press.

Clawson, R.T. (1985) Controlling the manufacturing start-up. *Harvard Business Review*, **63**, 6–20.

Daft, R.L. and Lengel, R.H. (1978) Organizational information requirements, media richness and structural design. *Management Science*, **32**, 554–571.

Dean, J.W. and Susman, G. (1989) Organizing for manufacturable design. *Harvard Business Review*, **67**, 28–36.

Electronic Buyers' News (1999) Pentium III stalled at 0.25 micron until November. 21 June.

Freeze, K. and Clark, K.B. (1986) Sun Microsystems, Inc. (B). Harvard Business School, *HBS Case* 686–134.

Freeze, K.J., Wasserman, N. and Clark, K.B. (1984) Seeq Technology. Harvard Business School, *HBS Case* 9–685–081.

Gourevitch, P., Bohn, R.R. and McKendrick, D. (1997) Who is US? The nationality of production in the hard disk drive industry. *Working Paper*, Sloan Center on the Information Storage Industry at UC San Diego.

Hatch, N.W. and Mowery, D.C. (1998) Process innovation and learning by doing in semiconductor manufacturing. *Management Science*, **44**, 1461–1477.

Krishnan, V. and Ulrich, K. (2001) Product development decisions: a review of the literature. *Management Science*, **47**, 1–21.

Langowitz, N. and Wheelwright, S.C. (1986) Sun Microsystems, Inc. (A). Harvard Business School, *HBS Case* 686–133.

Langowitz, N.S. (1985) An exploration of production problems in the initial commercial manufacture of products. *Research Policy*, **17**, 43–54.

McIvor, R., Martin, J.D., Matsuo, H. and Ng, J. (1997) Profiting from process improvement in the new semiconductor manufacturing environment. *Technology & Operations Review*, **1**.

McKendrick, D.G., Doner, R.F. and Haggard, S. (2001) *From Silicon Valley to Singapore: Location and Competitive Advantage in the Hard Disk Drive Industry*. Stanford University Press.

Pisano, G. (1992) BMW – The 7 Series Project. Harvard Business School, *HBS Case* 9–692–083.

Pisano, G. (1997) *The Development Factory*. Harvard Business School Press.

Randall, T. (1999) Product variety, supply chain structure and firm performance: an empirical examination of the US bicycle industry. *Working Paper*, The Wharton School, University of Pennsylvania.

Robertson, D. and Ulrich, K. (1998) Planning for product platforms. *Sloan Management Review*, **39**, 19–31.

Terwiesch, C. and Loch, C.H. (1999) Managing the process of engineering change orders: the case of the climate control system in automobile development. *Journal of Product Innovation Management*, **16**, 160–172.

Terwiesch, C. and. Loch, C.H. (1999) Measuring the effectiveness of overlapping development activities, *Management Science*, **45**, 4.

Terwiesch, C. and Bohn, R.E. (2001) Learning and process improvement during production ramp-up, *International Journal of Production Economics*, **70**, 1–19.

Thomke, S.H. (1998) Managing experimentation in the design of new products. *Management Science*, **44**, 743–762.

Ulrich, K., Sartorius, D., Pearson, S. and Jakiela, M. (1993) Including the value of time in design-for-manufacturing decision making. *Management Science*, **39**, 429–447.

Wasserman, N.H. and Clark, K.B. (1986) Everest Computer. Harvard Business School, *HBS Case* 685–085.

Wheelwright, S.C. and Clark, K.B. (1992) *Revolutionizing Product Development*. New York: The Free Press.

Notes

1. Some analysts measure time-to-market as the time until the first unit is sold, which includes a portion of ramp-up. However their emphasis is usually on time-to-develop.
2. Good descriptions of HDD technology and the industry can be found in the SEC filings or web pages of the independent drive manufacturers, such as Quantum and Seagate.
3. This name and some identifying information have been disguised.
4. First pass yield is defined as the pass rate the first time components are assembled into a drive and tested.
5. We use Greek symbols to disguise some of the confidential data in our study.
6. The tact time corresponds to the operation time of the slowest operation, which is called the bottleneck. Other industries refer to 'cycle time' rather than tact time, but in electronics cycle time is used differently.
7. An individual station's yield has to be higher than the overall yield target β. Overall yield is calculated by multiplying all the stations' yields together.
8. The graph shows overall yield, and thus provides a more aggregated view than given by Figure 3.
9. Accounting and performance tracking systems may not count all of these causes as 'downtime', but their physical impact on output is approximately equivalent.
10. The graphical representation is motivated by the CUBES model, which is a common industrial engineering representation of high-tech manufacturing (e.g. McIvor *et al*., 1998).
11. These deliberate shutdowns are similar to control-chart driven stops. However we have not seen many examples of such stopping rules used during ramp-ups, where the process is usually 'out of control' in the technical sense at all times.
12. There are two caveats to these calculations. First, all numbers are estimates and have not been verified with the company. Second, an underlying issue here is the cost of producing in Singapore and developing in the USA compared with having both activities co-located. This calculation does not say anything about what manufacturing costs would have been in the USA, or what development costs and speed would have been if all development were in Singapore. It also does not establish the *marginal* benefits of another person or another person-month of personnel transfer.