The Goal
Book by Eli Goldratt

This summary by Allen Cheng

Want to increase your personal output, or the output of your team? Do you feel like there’s a bottleneck constraining you? Have you tried brute force effort with little results?

The Goal: A Process of Ongoing Improvement, by Eliyahu Goldratt, is a classic management book, on Jeff Bezos’s shortlist of books recommended to his senior Amazon managers. This book introduced the Theory of Constraints, which identifies the constraint in a production system and restructures the organization around it. It upended traditional obsessions with cost efficiency to focus on what really matters.

Unique among management books, The Goal is written in the form of a novel, detailing a plant manager’s journey to save his factory from closing. The story itself is entertaining and teaches in a Socratic way, helping you identify and overcome your own constraints.

In this book summary of The Goal, you’ll learn
• Why optimizing activities that seem productive can be pointless
• How to identify your personal bottleneck
• How to increase capacity at your bottleneck to increase output
• How to apply these manufacturing principles to your everyday life
Preface

Goldratt’s *The Goal* concerns a manufacturing plant, and its lessons on throughput and inventory can be easily applied to literal analogues of this, such as supply chain, manufacturing, and automation problems.

It takes a bit more thinking to apply it to knowledge work or larger projects, but the principles are generalizable. To really get the most out of this *The Goal* book summary, think constantly about how you can apply it to your own situation - whether that’s closing sales clients, creating software products, managing a team, or writing books.

Take the advice of Ray Dalio in *Principles*: view yourself top-down as a machine, digesting inputs and creating outputs. From this perspective, you’ll be capable of studying and optimizing yourself.

### 1-Page Book Summary of *The Goal*

- **Productivity** is defined as bringing you closer to your goal. Every action that brings you closer to your goal is productive. Every action that does not bring you closer is not productive, even if it seems so.
  - Beware of defining subgoals that do not really drive toward the Goal, like production efficiency, team size, money raised, etc.
- **The Goal of every business is to make money.** Likewise, activities that do not bring you closer to making money are not productive.
  - Organizations can be measured by 3 metrics: Throughput, Inventory, and Operational Expense.
  - Throughput: the rate at which the system generates money through sales
  - Inventory: all the money system has invested in purchasing things it intends to sell
  - Operational expense: all the money the system spends to turn inventory into throughput
- Ideally, your activities improve all three at once.
  - Many companies focus primarily on decreasing operating expense, which can lead to unproductive behaviors that stifle throughput. Instead, switch your mindset to increasing throughput.

### The bottleneck of the system determines the throughput of the entire system.

- By definition, the throughput of the system cannot be greater than the capacity of the bottleneck. Where is the weakest link in the chain?
  - This means the value of an hour lost at the bottleneck is equal to the value of the entire system. Even if a bottleneck costs $5/hour to run, if the factory is producing $1000 of goods per hour, then an hour of the bottleneck idling is costing $1000/hour.
  - Constraints can be equipment, people, or policies.
- **Increase capacity at the bottleneck through a variety of interventions.**
  - Prevent idling by running the bottleneck all the time, ensuring inventory upstream of the bottleneck, and preventing back-ups at the bottleneck.
  - Bypass parts past the bottleneck if it’s not strictly necessary.
  - Improve quality of upstream work to prevent bottleneck working on poor-quality work.
  - Add new producers (eg machines, people) at the bottleneck, even if they’re less efficient.
  - Outsource bottleneck capacity to outside the organization.
  - (If you’re a bottleneck as a manager or a worker, consider implementing analogues of each of the above)
- The non-bottlenecks should be synchronized with the bottleneck, which means idling is acceptable.
  - If both the bottleneck and non-bottleneck go full steam ahead, the non-bottleneck will produce surplus inventory, which adds cost and causes traffic jams.
  - To coordinate this, use Drum-Buffer-Rope
    - Drum: pace non-bottlenecks production rates with bottleneck rates
    - Buffer: provide enough buffer inventory upstream of bottleneck to prevent idling
Rope: allow up to a threshold max surplus inventory, after which the non-bottleneck is idled

**The Goal and Metrics**

**Productivity is bringing a company closer to its goal.**
- Every action that brings a company closer to its goal is productive.
- Every action that does not bring a company closer to its goal is not productive.

**Defining the Goal is critical. The Goal of every business is to make money. Without money, the company is dead.**
- Do not deceive yourself into picking a subgoal - like decreasing cost per part, employing good people, manufacturing efficiency, product creation, quality, satisfaction, cutting-edge technology, market share. None of this matters unless it meets the Goal.
- Be wary of separate departments overoptimizing their subgoals without meeting the main goals. Like purchasing being more cost-effective and renting warehouses to store excess inventory.
- If you improve efficiency at one step without increasing overall output, you are not being more productive - you might even be causing an excess inventory and increasing cost per good sold.
  - Eg Imagine if you added a robot at one part of the assembly line. To justify the cost, you drive it at max capacity, which requires feeding it a lot of upstream supply, and also causes downstream accumulation. The cost-per-part at that step is high, but the system overall suffers.

**What is the minimum number of measurements you need to know if you’re making money?**
- Net profit (the more positive the better)
- ROI (relative, to tell what the base is)
- Cashflow (to make sure company stays alive)

**How do you convert these highest-level metrics to ones that are more actionable day-to-day?**
- **Throughput**: the rate at which the system generates money through sales
- **Inventory**: the money the system has invested in purchasing things to sell
  - IE the money that is currently stuck in the system
  - Any investment that you can sell is inventory.
    - R&D that can be sold, like a patent, is inventory, contrasted with R&D to improve throughput below.
- **Operational Expense**: the money the system spends in order to turn inventory into throughput
  - Includes labor, leases, R&D to improve throughput
  - It’s better not to consider where value is added - don’t get caught up over whether a dollar is investment or expense. If it’s not part of the packaged product being sold, it’s an operational expense.
- Ideally, you should try to improve all three at once.
  - Eg if you install robots, it should increase throughput by increasing sales, decrease inventory needed, and/or reduce labor costs.
- Be wary of a change that affects only one of these metrics - there may be second-order effects that backfire.
  - Eg reducing labor to reduce operational expense may decrease throughput.

**If you make a subpart more efficient, you do not raise your competitive advantage if it does not increase output or reduce cost.**
- Adding robots to improve step efficiency, without reducing workers or increasing output/sales, does not increase profits.
• “Economies of scale” can be deceptive when you focus on a narrow substep, like cost per part. The real metrics to care about are total products produced and cost per product; increasing efficiency at one step may aggravate the other steps.

• Fallacy: “we have to run our robots constantly, otherwise we’d lose our savings on cost per part and efficiencies go down.” This increases inventories and operational expense, because the parts pile up at the bottleneck.

• Says the mentor in The Goal: “A plant in which everyone is working all the time is very inefficient.” There must be a bottleneck along the chain; idling the rest of the chain prevents inventory from piling up.

Fallacy of Average Production Rates

In manufacturing, a balanced plant tries to match average capacity of every resource exactly with market demand. Any resource beyond the average rate is seen as extraneous, so it is either put to use or eliminated. This is the traditional mode of thinking at the protagonist’s company in The Goal.

However, two interconnected concepts make the balanced plant backfire, thus decreasing throughput, increasing inventory, and increasing carrying costs:

• **Dependent events** - one part of the chain depends on the upstream part.
  o The speed of the downstream part is constrained by the upstream part - if the upstream part isn’t delivering, the downstream part can’t do any work.

• **Statistical fluctuations** - many factors cannot be predicted precisely
  o Even at an average steady state rate, there are fluctuations in production. Someone may produce 2 widgets per minute on average, but at times he produces 2.5 and at times he produces 1.
  o Larger events - machines may break down; workers many get sick; inclement weather may arrive

Fluctuations happen regularly at each part in the chain. However, **each downstream part can only catch up to the extent that the upstream part permits it to**. Negative fluctuations are unconstrained; positive fluctuations are constrained. Over time, this causes a lower than expected average throughput.

• Let’s say there are three tools in a series - X, Y, and Z. They can each work at an average of 5 units/hour, with fluctuations around this average. Each hour, parts from one tool are moved to the next one.

• If X suddenly produces 3 units in 1 hour, then in the next hour, Y can only produce 3 units from X, **even if Y could actually work at 6 units/hour in that hour**.

• For the system to make up time for every negative fluctuation, both X and Y need to work at above-average rates in synchrony.
  o In a complex chain with many parts, this is difficult.

As explained in the next sections of this The Goal book summary, the solution is to balance **flow or throughput** with demand, not average capacity. The way to do this is a drum-buffer-rope system, where the bottleneck determines the throughput and inventory of the entire system.

Analogy of the Hiking Line

Imagine a troop of 10 boys hiking single-file on a narrow trail in the woods. The leader of the pack sets a comfortable pace that everyone on average should be able to meet.

Every boy is only able to catch up to the boy in front - he can’t pass the boy in front. Thus, the speed of each boy is constrained by the boy in front.
Analogy to manufacturing: the first boy is the most upstream step; the last boy measures throughput; the distance in between is inventory.

Here’s an example of a **negative fluctuation**
- Say the 3rd boy in line stops to tie his shoes, increasing the gap between him and the 2nd boy. The 4th boy runs into the 3rd boy and has to match his slower pace. Even if he could walk faster, the 4th boy is constrained.
- The 3rd boy gets back up and hurries to make up the distance to the 2nd boy. But the 4th boy is already tired and can’t jog, thus increasing the lag.
- Inevitably, even though all the boys are at the same average rate, the gap between first and last boy increases. **The fluctuations are accumulating because the ability to go faster than average is restricted by your upstream step.**
- This gap can be closed only if the last boy - and all the boys in front of him - all move much faster than the first boy.

Let’s say to relieve these pesky dependencies, you order all the boys in speed, with **fastest leading the pack**. All boys can now move unconstrained and operate at their individual peak efficiencies. However, the distance between the first boy and the last boy will incrementally grow, unbounded.
- This is what happens when you concentrate on single-step efficiency without focusing on throughput. All steps are working at full steam, but inventory progressively increases.

In contrast, let’s order the boys with slowest first and holding up the line for all other boys. On the surface, this seems very inefficient - the other boys are nowhere near their peak efficiency, and they’re all running into the next boy.

However, there are a few key advantages:
- The bottleneck is now quite obvious. We can now increase throughput by speeding up the slowest boy in front - for instance, by redistributing the slowest boy’s backpack load to the other boys.
- Inventory - the distance between the first and last boy - is not dramatically reduced. All boys are producing just enough to match the pace of the first slow boy.

**The Importance of the Bottleneck**

The critical concept in Goldratt’s **The Goal** is identifying the key constraint that holds back throughput. While this sounds like common sense, it can be hard to objectively identify your own bottlenecks when you’re in the thick of work.

**Building Blocks of Manufacturing Flows**

Relationships between a bottleneck (X) and non-bottleneck (Y) can be summarized in the following 4 diagrams:

**Y → X**
- The non-bottleneck feeds to the bottleneck
- Y has excess capacity. Letting both X and Y run continuously will build up inventory in front of bottleneck X.

**X → Y**
- The bottleneck feeds the non-bottleneck.
- Y is starved for inventory to process and marches to the beat of X’s drum. This is totally fine, even if Y is idling at times.
X → Assembly ← Y
- Parts passing through X and Y flow into a joint step of assembly.
- Working Y out of pace with X produces inventory in front of assembly.

X → Product A | Y → Product B
- Here X and Y are independently operating for separate marketing demands.
- The constraint for product A is bottleneck X, by definition.
- The constraint for product B is not Y - it is market demand. Y can produce more than the market demands to be “efficient,” but this will lead to finished product inventory.

In all these blocks, **Y never determines throughput for the system**. Throughput instead is determined by bottleneck X, or market demand.

### The Massive Cost of the Bottleneck

Any time that the bottleneck isn’t working is lost time forever that cannot be made up at any other part in the system.

**An hour lost at the bottleneck causes a loss in total throughput equal to the hourly capacity of that bottleneck.**
- If the total throughput is a thousand dollars per hour, then the bottleneck is processing at a thousand dollars per hour, **even if the literal operational costs or the parts going through it cost much less**.
- Alternatively, take the **entire** operating expense of the factory, divided by the hours worked by the bottleneck - that’s the **actual cost of the bottleneck**.

In other words, a loss in the bottleneck means a loss to the entire operation, and should be viewed with such gravity.

Other losses in effective throughput are also similarly costly. For example, feeding low-quality parts through the bottleneck will cause rejection later, leading to effectively lower throughput.

While time lost from the bottleneck can be made up for by hurrying non-bottlenecks, any extra effort here typically adds to operational expense (eg overtime pay). Ideally, the bottleneck is simply maintained at peak capacity.

### Identifying and Improving the Bottleneck

Now that we know the bottleneck is hugely significant, this **The Goal** summary will take you through how to identify and improve the bottleneck’s capacity.

### Five Focusing Steps for Improvement

In **The Goal**, Goldratt describes the 5-step process for continuous improvement:

1. IDENTIFY the system’s constraint.
2. Decide how to EXPLOIT the system’s constraint.
   - ie squeeze more capacity out of the constraint
3. SUBORDINATE everything else to the above decisions.
4. ELEVATE the system’s constraint.
5. If in the previous steps a constraint has been broken, go back to step 1, but do not allow inertia to cause a system constraint.
   - Different constraints can require very different optimizations.
   - Overcorrection can be counterproductive, eg obsessing about preventing the bottleneck from idling causes it to produce surplus goods above market demand.

We’ll dive more deeply into these.

**Identify the System’s Constraint**

The bottleneck is any resource whose capacity is equal to or less than the demand placed on it.

**Identify the bottleneck by seeing where you have the greatest upstream inventory piling up**, with low inventory at the next step. If the non-bottlenecks are producing at equal rates (eg 100 parts/hour), the slowest step will have the largest upstream inventory.

Alternatively, see which downstream steps are most in demand of upstream parts and are idling. If you decrease inventory sizes, you will see which work center, if stopped, halts the whole line.

- Analogy of rocks and water: the water level corresponds to inventory, while rocks are problems disturbing the flow. Lower the water level until a rock sticks out. Solve that problem, then lower the water level further.

Alternatively, in a more brute-force comprehensive way, define your market demand (by sales), then compare the productivity of each step of the chain to this demand.

**Increase Capacity at the Bottleneck**

The protagonist of *The Goal* book undergoes multiple iterations of increasing capacity as his bottleneck to increase overall throughput. Without detailing every struggle, in this book summary we’ll cover common causes of reduced capacity at bottlenecks, and fixes to increase capacity.

**This is a good point to consider your own work or life in this context, and to construct effective ways to relieve your personal bottlenecks.**

**Common contributors to bottlenecks**

- The machine can be made to run more efficiently, skipping unnecessary steps or decreasing setup/switching costs.
- The machine is running efficiently, but there simply aren’t enough machines
- The machine isn’t being run for the maximal number of hours
- The step requires a long batch period, and batches are not fully filled.
  - Eg running a dishwasher - each cycle takes 2 hours and more dishes can’t be added midway through
- Poor quality upstream parts waste bottleneck time
- Working on parts you don’t need urgently, diverting from the immediate backlog
- Occasionally, the bottleneck doesn’t have the upstream inventory to work at full capacity.
  - This may be because the non-bottlenecks are working on non-bottleneck parts or haven’t built up inventory.
- Machines run idle because people are redistributed to work on non-bottlenecks.
Fixes to bottlenecks

- Add on supplements to increase bottleneck capacity, even if they’re less efficient
- Guarantee round-the-clock production at the bottleneck.
  - Eg Eliminate lunch breaks and downtime.
- Redistribute capacity from non-bottlenecks to the bottleneck (eg more workers).
- Outsource the bottleneck outside the organization.
- Take some load on the bottleneck and redistribute to non-bottlenecks, if the same function can be performed.
- Allow parts to skip the bottleneck step if it’s not necessary.
- Do quality control upstream of the bottleneck to prevent time processing substandard parts.
- Focus the bottleneck only on parts that are needed today, not for safety’s sake tomorrow.
  - Order the work in terms of first in, first out - clear the backlog first to guarantee important work is being done.
- Make sure there is extra inventory ahead of the bottleneck so it can always be running at full capacity.
- Tag parts that will go to the bottleneck as higher priority, so they get processed first and you guarantee inventory in front of the bottleneck. The bottleneck gets an express lane.
- Announce the importance of bottleneck to the entire team, so they understand the priority of processing for the bottleneck.
- Collect accurate statistics on bottleneck operations to make better decisions.
- Permanently staff people at bottlenecks to decrease idle time.
  - Eg have people waiting by dishwasher to prepare loads and unload immediately.
  - Remember that the cost of a lost hour at this bottleneck is very expensive, and possibly well worth people idling on standby.
- Check upstream steps to see if adjustments can be made that decrease load on bottleneck.
  - In The Goal book, the team discovers that running an upstream milling step more efficiently leads to requiring heat treatment. If this is slowed down, then the parts can skip heat treatment. In this case, lowering efficiency at one step actually increases throughput.

Again, take some time to consider how to apply these to your own life. Consider the above fixes directly applied to a manager, whose time can often be the bottleneck of an organization’s throughput:

- Increase your time on tasks that you are uniquely good at (ideally these are higher leverage tasks). Train others to take over your lower level responsibilities.
- Allow workers to bypass you for permission for smaller decisions. Don’t require everything to pass by your desk.
- Prevent idle time by enforcing meetings starting on time with everyone present.
- Track your backlog so that more important items are worked on first, then process in first-in first-out order.

Market Demand and Throughput

Ideally, the flow through the bottleneck should match market demand. Producing more than this will increase inventory of finished product. Too much inventory will hinder flexibility through sunk cost fallacy - you’ll be reluctant to adapt to new market demands, because it would mean writing down your old inventory.

Instead, when you have surplus capacity, push to increase sales to make use of this capacity. Because you’re already paying for the fixed costs, you can lower prices to above material (marginal) cost to simulate more demand. This will decrease your overall cost per product.
Structuring Around the Bottleneck

Once you identify the bottleneck and improve its capacity, you may find other problems arising that decrease throughput. In Goldratt’s *The Goal*, the team goes through multiple iterations of solving problems, yielding the below principles in this book summary.

How to Keep Non-bottlenecks Synchronized with the Bottleneck

In *The Goal*, the team identifies a robot as the bottleneck. They devise a system whereby all parts destined for the bottleneck are always worked on at highest priority at non-bottleneck steps. This increases throughput temporarily, until they discover that at final assembly, suddenly there are shortages in non-bottleneck parts while there is massive inventory upstream of the bottleneck. How could this be?

They discover that they were running non-bottlenecks at full-speed, and having them crank out bottleneck parts far in excess of what the bottleneck could process. In turn, the non-bottlenecks has insufficient capacity to produce their non-bottleneck parts.

To avoid this, you must synchronize the non-bottlenecks with the bottleneck, to prevent massive deviations. Goldratt proposed the Drum-Buffer-Rope method, as follows:

**Drum** - the bottleneck dictates the pace of production of non-bottlenecks
- The slow boy scout beats a drum, and others take steps with the drum beat. If the boy scout beats more slowly, everyone steps more slowly.
- Similarly, a machine may regularly report its production rate, and the non-bottlenecks adjust their own rates up and down accordingly.
- Release starting materials to the non-bottlenecks strictly at the drumbeat rate.
  - The lead time can be calculated so the starting resources pass through upstream steps and arrive just in time at the bottleneck.
  - Similarly, all independent non-bottleneck routes can be timed so that all parts meet at assembly simultaneously.

**Buffer** - the bottleneck should have surplus inventory upstream so it doesn’t idle
- Non-bottleneck boys can scout ahead, clear brush so the bottleneck boy can keep walking at normal pace without having to stop.
- This buffer allows for “good enough” scheduling rather than needing to be perfectly accurate.
- Goldratt suggests choosing a time buffer equal to half the current lead time, then decreasing or increasing as deadlines or hit or missed.

**Rope** - when non-bottlenecks exceed a certain surplus level, they idle
- Tie a rope between the boy in the front and the bottleneck boy, and limit the maximum distance between the two.
- Similarly, prevent work-in-process inventory from exceeding a threshold level.
  - Henry Ford purposely limited space for inventory to detect bottlenecks.
  - In Toyota manufacturing, inventory is limited to containers containing a number of units, marked by a card. When this container is withdrawn for further processing, the card is returned to the upstream work center - only then can the center produce.
In Kanban software engineering, no work can be added to a pipeline until the existing work has been moved to the next pipeline.

In addition, production requires prioritization - complicated chains require more parts to be worked on in the correct order to avoid queue times. In The Goal, Goldratt suggests prioritizing batches by time elapsed since its release - the longer parts have been waiting, the higher the priority they get worked on.

The Four Elements of Process Time

Each piece of material spends time from when it enters a plant to when it leaves:

- Setup time: resource prepares itself to work on the part
- Queue time: the part waits for a resource while the resource is busy
- Process time: the part is being modified to become more valuable
- Wait time: the part waits for another part to be assembled together

For parts going through bottlenecks, queue time is dominant. For non-bottlenecks, wait time is dominant.

Cutting Batch Sizes

Traditionally, larger batch sizes are seen as more efficient per part. However, this decreases agility and increases inventory.

Imagine cutting batch sizes in half.

- You’ll halve the work-in-process inventory, which will ease cashflow.
- You reduce queue and wait times for parts.
- You reduce lead time from order to delivery.
- This will require more frequent deliveries from suppliers.

But wait - won’t this increase setup time at non-bottlenecks? As discussed in The Goal, an hour saved at a non-bottleneck is a mirage. Reducing an hour at a non-bottleneck doesn’t increase throughput, because the bottleneck determines throughput.

- For what it’s worth, this also decreases idle time, since non-bottlenecks are kept busier rather than waiting for the big batch upstream.

Appearance of New “Bottlenecks”

When you increase capacity at the bottleneck, you may start seeing shortages in non-bottleneck parts that hold up other parts of the chain.

While reflexively you might think this is a new genuine bottleneck, be wary - often production has much extra capacity that it takes a huge increase in throughput before this really happens.

For example, as in plot of the The Goal, a non-bottleneck may produce two parts - part Y that goes through a non-bottleneck chain, and part X that goes through a bottleneck. If you focus the non-bottleneck entirely on part X, then you create a scarcity of the non-bottleneck parts Y - which creates an artificial bottleneck.
• Instead, ideally you synchronize all parts that run through all chains so that right number reach the last step at the same time to meet market demand (with some buffer).

Similarly, taking on more orders may reduce spare capacity on non-bottlenecks, depleting inventory in front of the bottleneck and starving it of work.

**Virtuous and Vicious Cycles**

As the protagonist of *The Goal* book finds, productive practices can lead to virtuous cycles and increased competitiveness.

By adopting lean manufacturing principles like bottleneck alleviation, small batches, and Drum-Buffer-Rope systems:
- You improve throughput, lower inventory costs, and limit traffic jams.
- This decreases turnaround times
- This in turn reduces cost per product (because of fixed costs)
- Which allows lower prices and more sales.
- You can also respond faster to the market.

In contrast, Goldratt points out how poor practices can lead to aggravating policies that cause further problems. Focusing on cost-accounting and local efficiencies:
- Causes full activation of resources to prevent idling.
- This increases inventory and causes longer queues for parts.
- This causes a greater backlog and missed deadlines.
- This in turn prompts earlier release of starting material (a natural reaction), which increases work-in-progress inventory and aggravates the situation.
- Also, this prompts people to ask for more capacity and increase batch sizes (for the sake of efficiency), which further increases inventory, causes traffic jams requiring more management attention.
- All of these symptoms decrease sales.

**Miscellaneous Notes**

**Misleading Accounting Disguises Progress**

Measurements like financial accounting should induce the organization to do what’s good for itself. But as you improve throughput, traditional accounting measures may make your situation look worse than in reality.

Examples:
- Inventory is usually marked as assets, and reducing inventory will show up as a loss in assets.
- Drum-buffer-rope will cause idle time, thus decreasing local efficiencies.
- Consuming surplus inventory (beyond that determined by drum-buffer-rope) will decrease local efficiencies.
- Decreasing batch sizes increases setup time and overhead, thus increasing apparent cost-per-part.

However, these are usually temporary adjustments. The positive part of the tradeoff here is lower inventory costs and faster lead times, which improve topline sales.
Mistakes in Management

With process improvements and exposed extra capacity, a natural reaction of managers is to downsize the capacity. This punishes workers who just helped the company improve.

Instead, encourage sales to use the improved performance to gain more sales.

Distilling Facts into Principles

Collecting too much information without identifying the underlying intrinsic order leads to false patterns and bad decision making.

Before the periodic table, how to understand the elements was unclear. By color? State of matter? Instead, Mendeleev organized the elements first by atomic weight, then by reactivity (eg sodium and potassium behave similarly when thrown in water). This even allowed prediction of elements that didn’t yet exist.

Start from simpler If-Then principles. Then question, if this is true, what can I predict to be true? Then test these predictions and continue testing the hypothesis. [This section felt a bit shoehorned in by Goldratt as a pedagogical item.]

Misc Notes

- The manager’s job is to answer three questions: “what to change?” “what to change to?” “How to cause the change?”
- Management speak can disguise real problems and goals. Why do people use it?
  - It is shibboleth that conveys belonging to a group and signals expertise.
  - It disguises bad news (“our margins are down but we make up for it with improved operational efficiency which will help us reach our cost opportunity.”)
- Be careful about where you assign responsibility.
  - If you think it’s the union’s fault, people’s diligence, competition, foreign labor - question that and whether the strategy or your output could be better.
- Work smarter, not harder.
  - Continuous emergencies are a sign of poor planning.
- Even though this is all common sense, often it’s hard to view our own work objectively. We require triggers to realize something we knew intuitively.

Applications to Non-Manufacturing Work

While The Goal is literally concerned with manufacturing, its principles are generalizable to any work system in which multiple parts contribute to a single goal.

Consider asking yourself these questions:

- Are you refusing profitable work because you don’t have enough throughput? If so, find your bottleneck and alleviate it.
- How can you cut batch size to increase efficiency?
- Where is the weakest link in your chain?
- Are you currently the bottleneck for any throughput? How can you tell?
- What are 3 ways you can relieve yourself as the bottleneck?
Plot Summary of The Goal

Like Ayn Rand’s *The Fountainhead*, *The Goal* is explication of a philosophy with a fictional novel wrapper.

This has various effects beyond pure entertainment value. Portraying the protagonist’s struggle makes you empathize and absorb the teachings better. Unlikable characters stereotype detractors from the philosophy (like Eddie, who doesn’t ever question the traditional way of doing things). Overcoming the struggle paints a vivid picture of how the strategy can work.

However, at times the dialogue and epiphanies in *The Goal* book can feel forced, and the conversations don’t as naturally follow the logic of the situation.

Thus, I consider the plot to be a minor portion of the value of the book. Here’s the summary, for context:

- Alex Rogo is beleaguered plant manager in a small manufacturing town. They’re in a bad situation: large backlog of orders, all orders are late unless expedited.
- Policies are to blame. The division’s goal has been to increase cost efficiencies, so they focus on local efficiencies of production (like the fallacy, “we have to keep the robots running at all times or else the cost per part will go up and we’ll never make back the cost.”). Large inventories accumulate, and they add robots without increasing sales.
- Their entire division has 3 months to improve performance or it’ll be sold.
- He talks to his old physics professor Jonah throughout all this. Jonah feeds him insight piecemeal, and Alex uses it to solve each problem, only for another to appear.
- Why didn’t adding robots increase throughput? Alex is prompted to find the bottleneck of the plant.
- On a boy scout hike, Alex realizes the fallacy of the balanced plant and how to identify and resolve bottlenecks.
- At the plant, they identify two bottlenecks - the new robotic machine, and heat treatment.
- But after identifying bottlenecks, it’s not clear how to increase capacity. They realize the bottleneck machine is idling at times because of union rules; also, the heat treatment inefficiently runs small loads.
  - They reorder the work queue so that the bottleneck is working only on the oldest unshipped orders.
  - They clear the idle time with the union rep.
- They run into an issue where the bottleneck doesn’t have upstream parts available. They learn the non-bottlenecks are working on non-bottleneck parts. They put red tags on bottleneck parts so that those always get highest priority.
- They start filling their backlog and reduce their lead time. But it’s not enough to clear the backlog entirely.
  - They add an extra outdated machine that will increase capacity at the bottleneck, albeit inefficiently.
  - They start collecting better data about the performance at their bottlenecks.
  - They load and unload at heat treatment immediately, letting workers idle there instead of letting the bottleneck idle.
- New bottlenecks seem to emerge. It turns out they’re releasing too much upstream material and focusing on producing red bottleneck parts only, leading to a deficit of green non-bottleneck parts.
  - They utilize a drum-buffer-rope system to limit the release of materials upstream to synchronize with the bottleneck.
- They reduce batch sizes, which decreases their lead times. Sales increase as reputation grows.
- Alex’s plant gets a 1000-unit million-dollar order to be delivered within two weeks. They initially consider it impossible, but they cut batch sizes again and offer to ship the 1000 units in 4 weekly shipments, starting 2 weeks from then. They make this happen.
- Alex’s boss gets promoted to headquarters, and Alex becomes division manager. He gets tasked with finding out how to become a manager. He realizes he needs to learn to figure out problems for himself, rather than consulting Jonah.
Throughout all of this, there is a secondary storyline with Alex Rogo going through marital turmoil with his homemaker wife. The stress over losing the factory makes him miss dinner dates, they communicate poorly, and his wife leaves for her parents. Over time they reunite as Alex turns his factory around and gains clarity.

In this book summary, I’ve ignored the bulk of this storyline since it’s meant to add some emotional color and doesn’t affect the main teachings of The Goal.