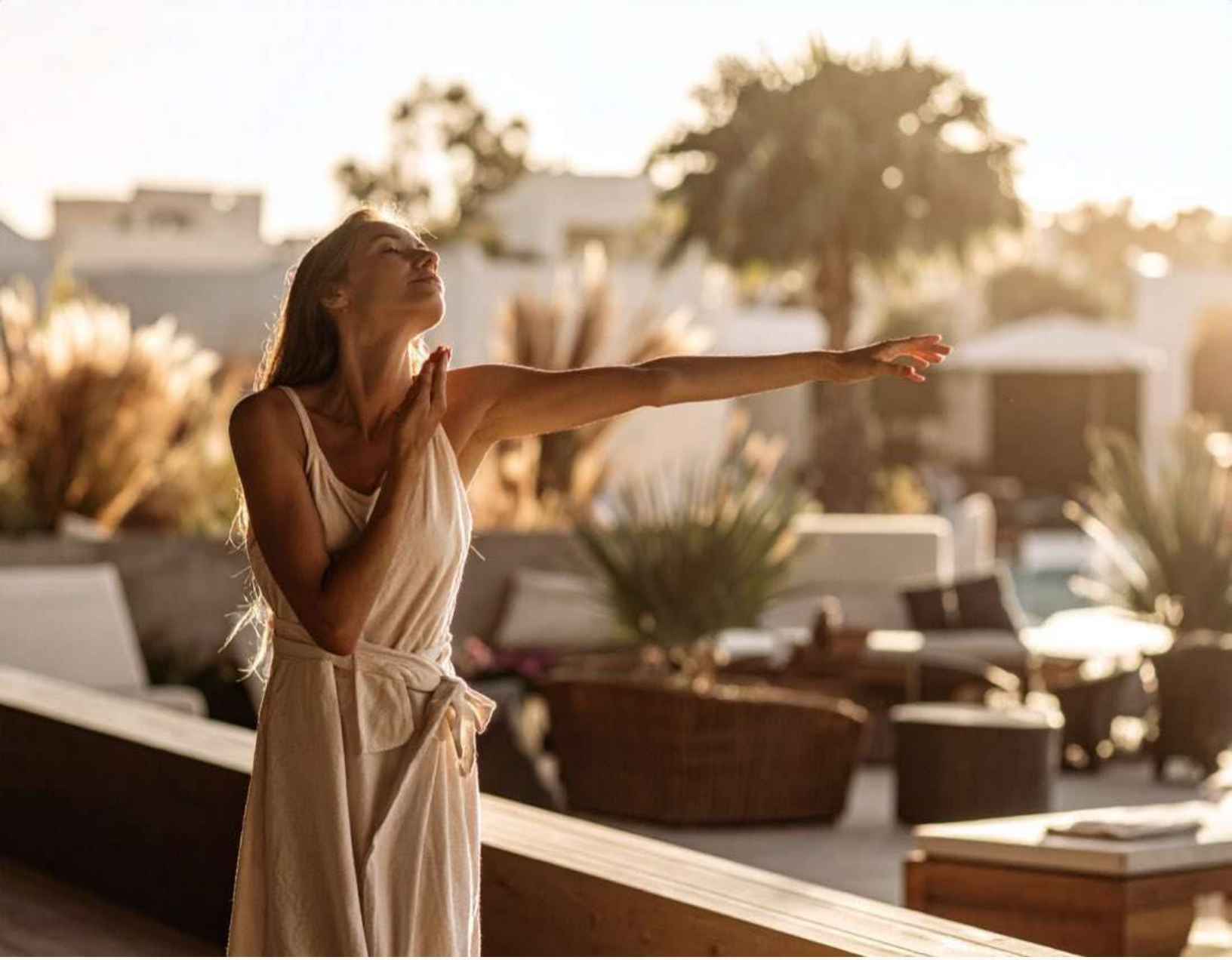


WHERE THE FLOW STOPS · CHAPTER TWO

One Order, One Desk, One Week

An order that holds fifty minutes of work waits six days, and the waiting is where the cost lives.

Ammanuel Santa Anna

Santa Anna, A. (2026). One Order, One Desk, One Week. In Where the Flow Stops (ch. 2). Luminous Prosperity Inc. https://doi.org/10.5281/zenodo.23989200
Luminous Library Licence 1.0 · business.hautelumiere.life/library/where-the-flow-stops

THE LICENCE

One Order, One Desk, One Week
An order that holds fifty minutes of work waits six days, and the waiting is where the cost lives.

Where the Flow Stops · Chapter Two. The Haute Lumière edition, 2026. © 2026 Ammanuel Santa Anna.

Released under the Luminous Library Licence 1.0 (hautelumiere.life/library-licence): free to read, share, quote, translate and teach, non-commercially, for ever. The Author keeps every other right, including every commercial right.

Published by Luminous Prosperity Inc. for Haute Lumière. Author ORCID 0009-0008-3244-0306.

DOI 10.5281/zenodo.23989200 · this record's page: business.hautelumiere.life/library/where-the-flow-stops/02-one-order-one-desk-one-week/

Please cite as: Santa Anna, A. (2026). One Order, One Desk, One Week. In Where the Flow Stops (ch. 2). Luminous Prosperity Inc. https://doi.org/10.5281/zenodo.23989200

Chapter Two of Where the Flow Stops. The whole book, with every chapter, is at business.hautelumiere.life/library/where-the-flow-stops (DOI https://doi.org/10.5281/zenodo.23989196).

A NOTE BEFORE YOU READ

This work is offered for education and leadership development. It is not legal, financial, accounting, investment, medical or psychological advice; take the advice of a qualified professional who knows your circumstances. Companies and events are described from the public record as the author understands it; characterisations are the author's analysis. Examples are illustrative, and no particular outcome is promised to any reader.

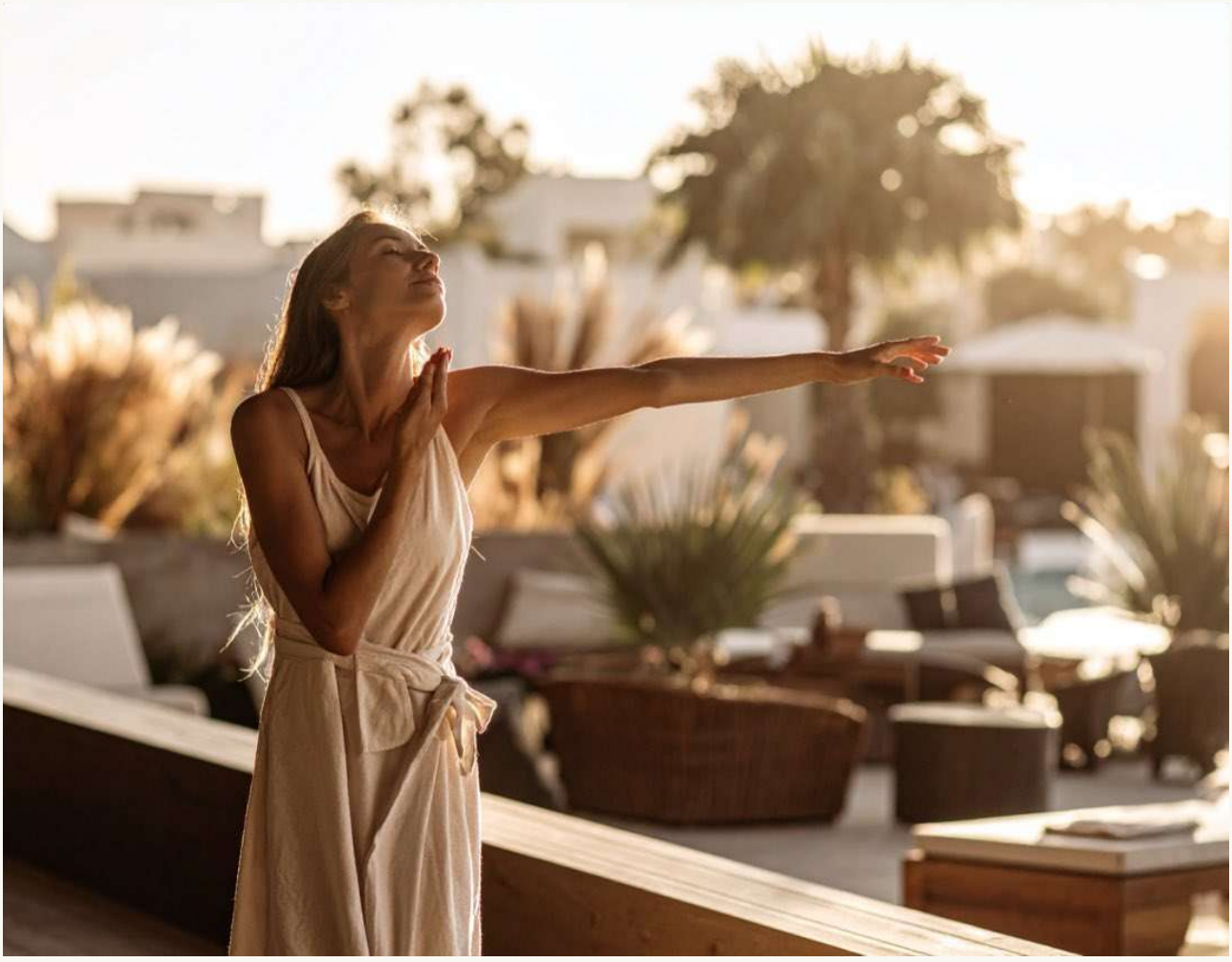
ABSTRACT

An order that holds fifty minutes of work waits six days, and the waiting is where the cost lives. It is ten to nine on a Monday, and the order has been in the building for six days. The customer is a hospital, and the part is a motor assembly for the dishwasher in its main kitchen, which has been washing trays by hand since the previous Monday. At nine o'clock the hospital's facilities manager will ring, as she has rung every morning since Wednesday, and the order coordinator who takes the call has the order sheet on the desk in front of her under the strip light. It is a printed page with a column of initials and dates down the right-hand side, one for every desk the order has visited. There are seven sets of initials. Her shoulders have been up near her ears since Friday afternoon.

CHAPTER TWO

One Order, One Desk, One Week

An order that holds fifty minutes of work waits six days, and the waiting is where the cost lives.



One Order, One Desk, One Week

HAUTE LUMIÈRE · WHERE THE FLOW STOPS · CHAPTER TWO

Ten to nine, the order sheet

It is ten to nine on a Monday, and the order has been in the building for six days. The customer is a hospital, and the part is a motor assembly for the dishwasher in its main kitchen, which has been washing trays by hand since the previous Monday. At nine o'clock the hospital's facilities manager will ring, as she has rung every morning since Wednesday, and the order coordinator who takes the call has the order sheet on the desk in front of her under the strip light. It is a printed page with a column of initials and dates down the right-hand side, one for every desk the order has visited. There are seven sets of initials. Her shoulders have been up near her ears since Friday afternoon.

She is a composite, drawn from the way an order is described when operations researchers follow one through a firm, and no firm is meant. What she is about to discover is not.

Before the phone rings she does something she has not done before. She adds up the time anybody actually spent on the order. Twelve minutes to enter it. Six to run the customer's credit. Fifteen for an engineer to confirm that the assembly fits that model of machine. Eight for stores to find it. Ten to pack it and book the courier. The total is a little under an hour, and the order has been in the building for six days. Almost everything that has happened to it since Tuesday afternoon is waiting.

Ammanuel Santa Anna's canon on mapping the flow of value through an organisation has a sentence for the column of initials on her sheet.[11] "Every handoff between departments is simultaneously a material exchange and a relational event." The order moved seven times. Each move was a piece of paper, and each was also a small negotiation between two desks about whose problem it now was.

This chapter makes one claim. In most organisations the largest share of the time between a request and its delivery is spent in queues between people, not in the work itself; the queues are invisible because every desk they sit beside is busy; and the time lost there is a gain the firm already holds, in its lead time, its working capital and its customers' patience, waiting to be collected.

Fifty minutes in six days

The ratio the coordinator has just calculated has a name in the operations literature. Flow efficiency is the share of an item's total time in a system during which somebody is actually working on it. Her order's flow efficiency is well under one per cent. That is not unusual for an order that crosses several departments, and it is not a sign that anybody is idle. Every person who touched the order was busy all week. The order was simply never at the front of anybody's pile.

Practitioners in lean operations and in software often quote typical flow-efficiency figures for whole industries, and they are usually low. This chapter does not repeat any of them, because the ones in circulation are rarely traced to a study that measured them. What can be said with confidence is the method: follow one

item, record when work starts and stops, divide. Anyone can do it with a printed sheet and a clock, and the figure it produces is the organisation's own, which is the only one that matters to it.

The idea that waiting is the largest waste in a working system is older than the term. Santa Anna's canon opens its account of value-stream mapping with the engineer who made it central, Taiichi Ohno at Toyota, and describes what Ohno saw on the floor in Nagoya: the waiting, the overproduction, the unnecessary movement, and "the inventory piling up between processes like sediment in a slow river." Elsewhere the canon puts the symptom more bluntly. When value stops flowing, "inventory piles up between process steps like plaque in arteries."

The house's Luminous 100, its reading of the frameworks every consultancy sells, adds a sentence to the textbook list of lean's seven wastes that changes how the list is read. In its account of lean thinking, waiting is not merely lost time: "waiting reveals a flow blockage where developmental energy is pooling." That is a more useful way to hold the coordinator's six days than as a failure of the people who held the order. It says the pile on each desk is information. It tells you where the organisation is constricted, the way a swelling tells a physician where to look.

The difficulty is that nobody can see the pile as a pile. The engineer who confirmed the fit on Thursday did not experience the order as having waited two days for him. He experienced a full inbox, which he worked through in order, competently. Every queue in a firm looks, from the desk beside it, like a reasonable amount of work.

The customer sees it differently, and the difference is the whole commercial point. The facilities manager at the hospital does not experience fifty minutes of work. She experiences six days of trays washed by hand, three porters moved from their own jobs to do it, and a daily phone call she would rather not make. What the firm sells her, in the end, is not the motor assembly, which a competitor could also supply. It is the elapsed time between her call and a working machine. That is the figure she will remember when the next order is placed, and it is the one figure in the transaction that the firm's own people, each working hard at a desk, are least able to see.

It is also, in a plain accounting sense, money. An order that sits inside a firm for six days is stock bought and not yet sold, a part reserved and not yet invoiced, a customer who has not yet been billed. Multiply one order by every order in the building and the waiting is a quantity of working capital that the firm has lent, interest free, to its own queues.



A proof in five pages

HAUTE LUMIÈRE · WHERE THE FLOW STOPS · CHAPTER TWO

A proof in five pages

In 1961 a young operations researcher named John Little published a five-page paper in the journal *Operations Research* with a modest title, *A Proof for the Queuing Formula: L = λW*.[12] The formula was already in use; what Little supplied was a proof that it held very generally. In any stable system, the average number of items inside it equals the average rate at which items arrive multiplied by the average time each item spends inside. The conditions in his abstract are technical, and they amount to this: over a long enough period, in a system that is neither filling up nor emptying out, the relationship always holds.

Fifty years later Little wrote a retrospective for the same journal, and his own description of the result is plainer than any of its later admirers': "Basically, it is a theorem in queuing theory."[13] It became famous, he noted, because of its practical importance. The reason is that it can be turned around. If you know how much work is sitting in a system and how fast the system finishes work, you know how long a new piece of work will take to get through, without timing a single item.

Apply it to the coordinator's building. Suppose the order desks between them hold about thirty open orders at any moment, and the firm ships about five a day. Little's Law says an order will spend, on average, about six days inside. Nobody has to be slow for this to be true. The six days are not caused by any person's effort; they are caused by the thirty. To halve the lead time, the firm does not need its people to work faster. It needs fewer orders sitting between them, at the same rate of finishing.

That turnaround is the law's practical gift, and it is also where its honest limits lie. It concerns long-run averages, not the next order. It says nothing about which order should go first, and a system that is growing or shrinking will break it for a while. It does not explain why the thirty orders are there. But within those limits it is not an estimate or a finding that a later study might overturn. It is arithmetic, and it is the most reliable sentence in operations management.

The arithmetic also explains a habit the coordinator's building has never questioned. The order desk enters new orders in a batch at four o'clock, because it is efficient to do them together, and engineering reviews its pile on Tuesday and Thursday mornings, for the same reason. Each batching rule is sensible at the desk that made it. Taken together they mean that an order arriving on Tuesday at two waits two hours to be entered and then, if it misses the Thursday review by a few minutes, five days to be confirmed. Reinertsen's book returns to this point repeatedly: large batches feel efficient to the people who process them and are expensive to the work that waits in them. Little's Law makes the cost visible. Every batch adds to the number of items sitting in the system, and so, at the same rate of finishing, to the time each one spends inside.

None of this is a reason to abolish batches, which often save real effort. It is a reason to know what each one costs in days, and to decide on purpose.

Why the busy desk is the slow desk

The question it leaves is why so much work sits in a firm at once. The answer most leaders reach for is that the people are too few, and sometimes they are. More often, the queues are produced by a habit that looks like good management: keeping everybody fully occupied.

In the same year as Little's proof, the British mathematician John Kingman published a short paper on a single-server queue in what he called heavy traffic, a system running close to its full capacity.[14] His result, now usually called Kingman's formula, is an approximation, and its message is one that most operations students learn with a small shock. As a server's utilisation rises towards one hundred per cent, the average waiting time does not rise gently. It rises steeply, and without limit, and it rises faster the more variable the arrivals and the work. A desk that is busy eighty per cent of the time is a desk with a queue. A desk that is busy ninety-five per cent of the time is a desk with a long one.

Donald Reinertsen, writing about product development in *The Principles of Product Development Flow* (2009), made this the centre of his argument.[15] In knowledge work, he observed, the queues are invisible because

the inventory is information: designs waiting for review, decisions waiting for a meeting, orders waiting for a question to be answered. Nobody walks past them on a factory floor. And because the conventional instinct is to load every person fully, organisations produce exactly the conditions under which, by Kingman's arithmetic, waiting explodes. Reinertsen's book is a practitioner's synthesis, built from queueing theory and his own consulting, not an empirical study, and it should be read that way. Its reasoning is sound where it rests on the mathematics.

There is empirical work on what full loading does to the people inside it. In 2009 Diwas KC and Christian Terwiesch analysed operational data from two very different hospital services, patient transport and cardiothoracic surgery.[16] As load rose, they found, workers sped up, and for a while that looked like a gain. Sustained over long periods, the extra load slowed the service down again, and in cardiac surgery it was associated with worse outcomes for patients. It is one study of two services in hospitals, and correlation in operational data is not proof of cause. But it fits the theory uncomfortably well: a system run at the edge of its capacity borrows speed from its people and repays it with interest.

The Luminous 100 describes the instrument that most simply answers this, the kanban board with its limit on work in progress, in a line any operations manager could pin above a desk: "The WIP limits are the magic — they prevent overload and make bottlenecks visible." Its living-systems reading goes further, and it is the reason the house cares about a technical-sounding limit. Every living organism, it says, "has a natural metabolic capacity, and exceeding it doesn't create more throughput but more toxicity". A team asked to hold forty things at once does not finish forty things sooner. It finishes each of them later, and more tired.

The leader's instinct runs the other way, and it is worth understanding why. An idle person is visible; a waiting order is not. A manager who walks past an engineer with nothing on her screen feels a cost, the salary being paid for nothing. The same manager walks past nineteen orders waiting in the same engineer's inbox and feels nothing, because a pile of work looks like diligence. So the organisation is tuned, desk by desk, to remove the visible cost of a person waiting for work, and in doing so it maximises the invisible cost of work waiting for a person. Kingman's arithmetic says which of the two is larger whenever a desk runs close to full. It is almost never the person.

The coordinator feels this in her own body before she can name it. When her desk is fullest she is busiest, and when she is busiest the phone rings most, because the orders she is holding are the ones customers are chasing. The chasing is itself work, and it joins the queue. A desk running near its limit spends a growing share of its day explaining why it is running near its limit. Anyone who has worked a week like that will recognise the feeling of running faster to stay in the same place. *The Margin*, another of the house's books, reads [the hours a team spends that nobody budgets](https://hautelumiere.life/editions/rsd-master-rsd-in-acedemia-final-manuscriptt/page/); a desk near its limit spends them on chasing.

The Margin, another of the house's books, reads [the list too long to finish, and what it does to the person holding it](https://hautelumiere.life/editions/rsd-master-rsd-in-acedemia-final-manuscriptt/page/). The

queue is that list at the scale of a firm, distributed across seven desks so that nobody holds all of it, and so nobody feels its weight.

A question that travelled by email

The coordinator's sheet has one entry that does not fit the pattern of honest queues, and it is the one that cost the most. The order reached engineering on Wednesday afternoon. The engineer's initials and the date beside them say Friday. Between the two is not a queue of other orders. It is a question.

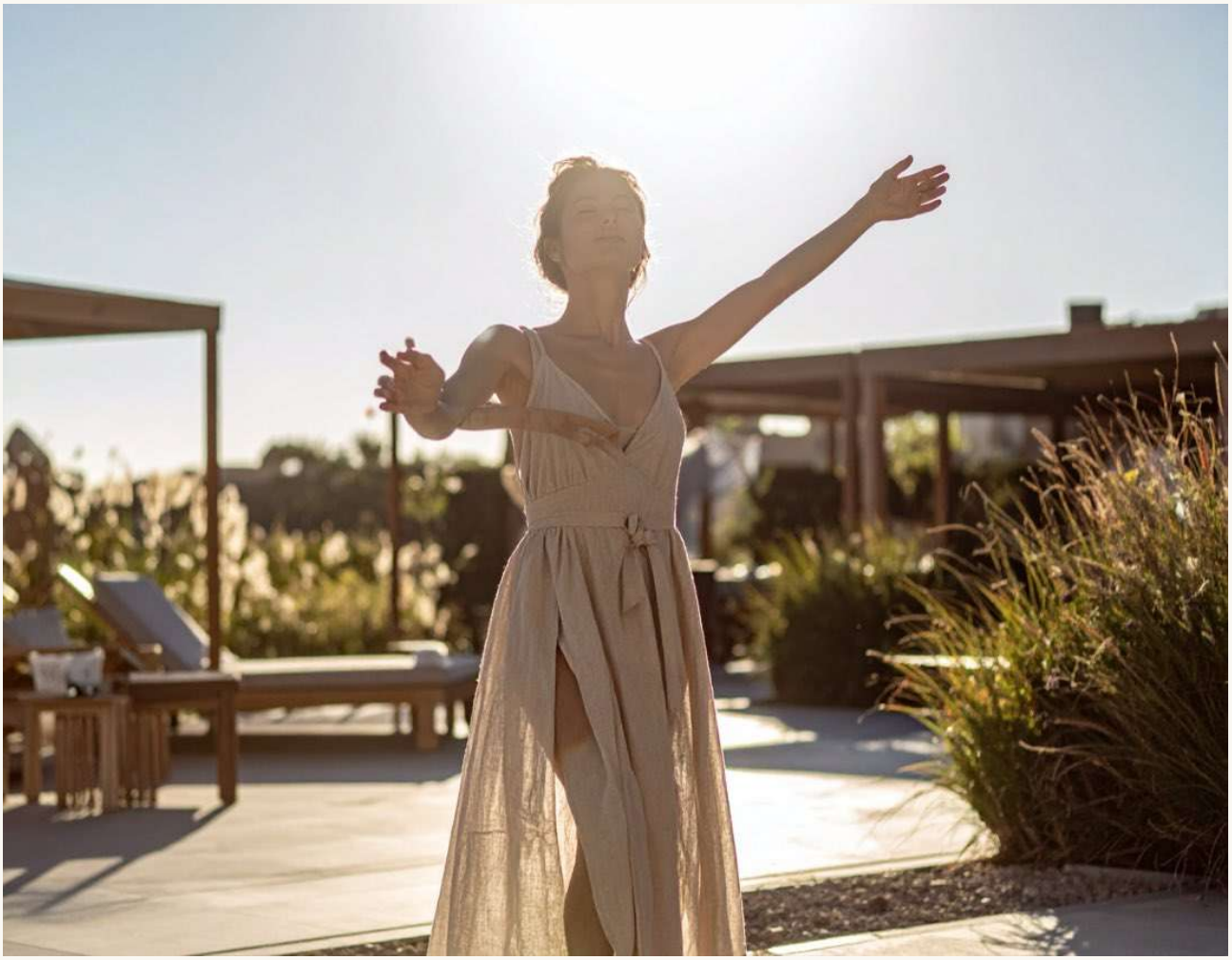
The order said *motor assembly, model 4*. There are two model-4 machines, an older and a newer, with different motors, and the hospital had not said which. The engineer sent an email to the order desk on Wednesday at five. The order desk read it on Thursday morning and rang the hospital, which had to send someone down to the kitchen to read the plate on the machine. The answer came back at two. The engineer, by then, was on site at another customer's, and read it on Friday morning. Fifteen minutes of engineering work took two days, because it needed one fact that the first desk could have asked for on the first call, if the first desk had known it mattered.

This is the relational half of the handoff that the canon insists on, and it is not a soft addition to the queueing mathematics. It is where much of the waiting comes from. A handoff in which the receiving desk's needs are known to the sending desk is fast. A handoff between two desks that communicate by leaving things for each other is slow in a specific, avoidable way: every missing fact becomes a round trip, and every round trip joins a queue at each end. The canon, describing a value stream that functioned but fumbled the same exchange every time, puts it ruefully: "The system works. It even wins races, sometimes." The race was being won. It was being won with a dropped baton at the same point on every lap.

The Luminous 100's entry on Eliyahu Goldratt's theory of constraints gives the other reason leaders misjudge their queues. The theory, it says, is "Elegant, counter-intuitive, and almost always reveals that the real bottleneck isn't where you think it is." Asked where the order was delayed, most people in the coordinator's building would say engineering, because engineering is always busy. The sheet says the delay was in the space between the order desk and engineering, which belongs to neither of them, and which neither of them is measured on.

The canon's point about relational value is easy to hear as sentiment and is better heard as mechanics. When two desks trust each other, a missing fact is settled by a call across the room in two minutes. When they do not, it is settled in writing, carefully, with a copy to a manager, and each careful message waits its turn in someone's inbox. Mistrust between departments does not show on any cost report. It shows as round trips, and every round trip shows, in Little's arithmetic, as days. A firm that wants its orders to move faster could do worse than ask which of its handoffs are conducted by phone and which by memorandum, and why.

The knowledge that would have closed that gap was in the building all along. The order desk's longest-serving clerk has known for years that model 4 comes in two versions. She checks it on every call she takes. The newer clerks do not, because nobody told them, and it is written nowhere. The person who already knows is the subject of Chapter 4, and she is usually closer to the queue than anyone thinks.



Walking the order

HAUTE LUMIÈRE · WHERE THE FLOW STOPS · CHAPTER TWO

Walking the order

What the coordinator does next is the practice this chapter is about, and it is less a technique than a way of looking.

She asks the operations manager to walk the order with her, backwards, from the courier booking to the first phone call, with the sheet in her hand. At each desk they ask three things and write the answers on the sheet. When did the order arrive here? When did someone start on it? What did it need from the previ-

ous desk that it did not have? The walk takes forty minutes. It is the first time either of them has stood at the engineering desk and looked at the pile.

The walk is uncomfortable in a particular way, and the operations manager notices it first. At every desk the person holding the order has a reason it waited, and every reason is true. Engineering was covering for a colleague on leave. Stores had a delivery come in on Wednesday. The order desk had a system outage on Thursday morning. Nobody is defensive, exactly, but everybody is explaining, and the manager realises that this is the first conversation she has ever had with these people about time rather than about effort. They have been asked many times whether they were busy. They have never been asked how long something sat.

That shift in the question is most of the practice. Asking how busy a desk is invites a person to defend their effort. Asking how long an order sat invites them to look at the system with you. The second question is answered more honestly, because it accuses nobody. It is the discovery Chapter 1 describes, when a leader moves the conversation from the lift to the table: people tell the truth readily when the truth is about the system.

Then they count. Not the hours people work, which everybody already knows are long, but the number of open items at each desk, a figure nobody in the building has written down before. Engineering is holding nineteen. Stores is holding six. The order desk itself, which would have said it was waiting on everybody else, is holding twenty-two, many of them orders waiting for a fact from a customer that nobody has yet asked for. Little's Law, applied desk by desk, tells them where the days go.

The changes they make are small and they are made in order. First, the order desk agrees to ask two questions on every call that engineering needs answered, the model and the serial plate, and the longest-serving clerk writes them down for everyone. Second, the managers agree that the measure of the week is not how busy each desk is but how long an order takes from call to courier, and that figure goes on the wall where all seven desks can see it. Third, engineering agrees a limit: when its pile passes a number the team chooses, the order desk stops sending it new work for an hour and helps it finish what it has. The house's Luminous 100 has the most useful sentence for anyone tempted to skip the order and buy a tool instead. In its account of lean thinking, "The philosophy is more important than the tools."

Each change has a predictable way of going wrong. The two questions become a form that nobody reads. The lead time on the wall becomes a target that people game by closing orders early. The limit on work in progress is broken in the first busy week by a manager who cannot bear to see an order waiting at the front door, which is where it was always waiting, only visibly. Chapter 7 carries the run sheet for this practice, and Chapter 6 places it in the order of a first ninety days.

There is a gain in all of this, and it is worth naming precisely without inventing its size. Every order that spends fewer days inside the building is cash collected sooner, stock held for less time, and a customer whose kitchen is working again. A firm that halves the work sitting between its desks halves, by Little's arithmetic, the time its orders spend inside, without anyone working harder. That gain is real, it belongs to

the firm's existing people and process, and it can be measured against what the firm already records. How a gain like that is baselined and attributed honestly is the subject of Chapter 5.

Santa Anna's *Ecstatic Economics* uses the same image for a whole economy that the canon uses for a value stream. In the gift economies it describes, to hold wealth back was a sign that the flow of life had become stuck in a person, "that you had become a blockage in the circulatory system of communal wealth." A queue is a smaller version of the same thing. It is value that has stopped moving, held by nobody in particular, on behalf of nobody at all. The book is Ecstatic Economics, and its argument that a firm is a body with a circulation is the one this chapter has been making at the scale of a single desk.

Walking an order in your own building

The order walk is simple enough to start on Monday and subtle enough to go wrong in several ways, and the book can carry the finer points the article left out.

Choose an ordinary item. The temptation is to walk the order that went most badly wrong, because it is vivid and everyone remembers it. It will mislead. An extreme case has extreme causes, and regression to the mean (Chapter 5 returns to it) guarantees that the next one will look better whatever is done. Walk a typical item instead, one that crossed at least three desks, finished recently enough that people remember it, and was not anybody's emergency. Then walk two more. Three ordinary orders tell a truer story than one famous one.

Take times from records where they exist, and from people where they do not. Most firms have more time stamps than they think: the time an email arrived, the time an entry was made in a system, the time a courier was booked. Use them first. Where there is no record, ask the person, and write down that the time is remembered rather than recorded. A walk that mixes the two without saying which is which will be argued over later by whoever dislikes the result.

Count what is waiting, honestly. The number of open items at each desk is the most useful figure the walk produces, and the easiest to undercount. Items waiting for a customer's answer are open. Items parked because they are awkward are open. Items a person is holding in her head because the system has no place for them are open, and they are usually the ones that wait longest. Little's Law works on the true count, not the visible one.

Do the arithmetic in the room. If engineering holds nineteen open orders and finishes about four a day, its share of the lead time is a little under five days. Writing that on a whiteboard, with the two inputs beside it, does more to change a team's mind than any amount of exhortation, because it locates the delay in the number of things being held rather than in anyone's effort. It is also the moment at which people see that they can shorten the wait without working harder.

Change the cheapest thing first. The order of changes matters. Missing information is usually the cheapest fault to remove and the most expensive to leave, because every missing fact creates a round trip through two queues. Fix the questions asked at the first desk before setting any limit on work in progress. Then make the lead time visible. Then set a limit, and set it close to the current count, a little below it, and tighten it as the team learns what it can hold. A limit set far below current reality is broken in the first week, and a broken limit teaches the team that limits are decoration.

Read the spread, not only the average. A customer does not experience an average lead time. She experiences her own order, and she remembers the slow ones. When the walk becomes a weekly figure, look at the slowest tenth of orders as well as the typical one. An improvement that shortens the average while leaving the slowest orders untouched has helped the firm's report more than its customers.

Expect the queue to move. Shortening one desk's queue often reveals the next constraint, and the Luminous 100's note on the theory of constraints is a fair warning that it will not be where anyone expected. That is not failure. It is the walk working. Walk again in a month.

Chapter 6 sets the walk in its place in the first ninety days, after the corridor question and before the measurement conversation, and Chapter 7 carries the walk sheet, the three questions asked at each desk, and the arithmetic laid out so that anyone can do it.

The next Monday

It is ten to nine on a Monday, and a new order is on the coordinator's desk under the strip light. The customer is a school, and the part is a door seal for a walk-in fridge. The sheet has a box at the top, filled in on the first call, with the model and the serial plate. There are four sets of initials down the right-hand side, not seven, because the credit check now happens while the order is entered and stores picks from engineering's confirmation without a second handoff.

The order arrived on Thursday afternoon. It shipped on Friday. The school's caretaker did not ring this morning, because there was nothing to ring about.

On the wall by the order desk the lead time for the month is written in marker, the way the engineers asked for it, in days rather than hours, and beside it the number of open orders at each desk. Nobody is praised for the figure and nobody is blamed. It is simply where all seven desks can see it.

The coordinator's shoulders are where shoulders usually are. She adds up the time on the sheet out of habit. About forty minutes of work, and a little over a day.

References

1. The author's canon in this chapter, and the works it draws on: Santa Anna, A. *Integrated Value Stream Mapping*, "The Relational Field", "Material Value" and the case of the fumbled baton. The Luminous Developmental Canon; Santa Anna, A. *The Luminous 100*, "Lean Thinking (Toyota Production System)", "Theory of Constraints (Goldratt)" and "Kanban". Luminous Prosperity Inc; Santa Anna, A. *Ecstatic Economics*, on the gift economies and circulation. The Luminous Developmental Canon.
2. Little, J. D. C. (1961). A proof for the queuing formula: L = λW. *Operations Research*, 9(3), 383–387. https://doi.org/10.1287/opre.9.3.383
3. Little, J. D. C. (2011). Little's Law as viewed on its 50th anniversary. *Operations Research*, 59(3), 536–549. https://doi.org/10.1287/opre.1110.0940
4. Kingman, J. F. C. (1961). The single server queue in heavy traffic. *Mathematical Proceedings of the Cambridge Philosophical Society*, 57(4), 902–904. https://doi.org/10.1017/S0305004100036094 (The steep rise of waiting with utilisation is stated here in words; the formula itself is not reproduced.)
5. Reinertsen, D. G. (2009). *The principles of product development flow: Second generation lean product development*. Celeritas Publishing. (Book; verified in Open Library; its argument is summarised, not quoted.)
6. KC, D. S., & Terwiesch, C. (2009). Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science*, 55(9), 1486–1498. https://doi.org/10.1287/mnsc.1090.1037

One Order, One Desk, One Week

Ammanuel Santa Anna · Luminous Prosperity Inc. · Haute Lumière

Set in Cormorant Garamond and Manrope.

Read the Business Library at business.hautelumiere.life/library.

Luminous Library Licence 1.0 · hautelumiere.life/library-licence

REQUEST A CONVERSATION

business.hautelumiere.life/book/?kind=gainshare

The Business Library

business.hautelumiere.life/library/

The packages

hautelumiere.life/pricing/