

THE GAIN WAS ALWAYS SHARED · CHAPTER NINE

How to Know It Worked

Chapter Nine of The Gain Was Always Shared

Ammanuel Santa Anna

Santa Anna, A. (2026). How to Know It Worked. In The Gain Was Always Shared (ch. 9). Luminous Prosperity Inc. https://doi.org/10.5281/zenodo.23989518
Luminous Library Licence 1.0 · business.hautelumiere.life/library/the-gain-was-always-shared

THE LICENCE

How to Know It Worked
Chapter Nine of The Gain Was Always Shared

The Gain Was Always Shared · Chapter Nine. 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.23989518 · this record's page: business.hautelumiere.life/library/the-gain-was-always-shared/09-how-to-know-it-worked/

Please cite as: Santa Anna, A. (2026). How to Know It Worked. In The Gain Was Always Shared (ch. 9). Luminous Prosperity Inc. https://doi.org/10.5281/zenodo.23989518

Chapter Nine of The Gain Was Always Shared. The whole book, with every chapter, is at business.hautelumiere.life/library/the-gain-was-always-shared (DOI https://doi.org/10.5281/zenodo.23989502).

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

Every leader who starts a participation system will, at some point, be asked whether it worked. The question usually arrives late: at a budget meeting eighteen months in, from a new chief executive, from a board member who was not there at the start. By then the honest answer is often unavailable, because nobody wrote down what the organisation looked like before, what it expected, or what else happened in the meantime. The leader is left to answer from impression, and impression is the weakest evidence there is. This chapter is about answering that question well, and about arranging matters at the start so that it can be answered at all. It is the measurement companion to the field manual in Chapter 8. It covers what to count, how to establish a before, how to read a change against the background of everything else that moves, how researchers have done this in the published studies, and how to read your own results with the same honesty this book has tried to apply to theirs.

CHAPTER NINE

How to Know It Worked



How to Know It Worked

HAUTE LUMIÈRE · THE GAIN WAS ALWAYS SHARED · CHAPTER NINE

The question nobody asks on time

Every leader who starts a participation system will, at some point, be asked whether it worked. The question usually arrives late: at a budget meeting eighteen months in, from a new chief executive, from a board member who was not there at the start. By then the honest answer is often unavailable, because nobody wrote down what the organisation looked like before, what it expected, or what else happened in the meantime. The leader is left to answer from impression, and impression is the weakest evidence there is.

This chapter is about answering that question well, and about arranging matters at the start so that it can be answered at all. It is the measurement companion to the field manual in Chapter 8. It covers what to count, how to establish a before, how to read a change against the background of everything else that moves, how researchers have done this in the published studies, and how to read your own results with the same honesty this book has tried to apply to theirs.

Ammanuel Santa Anna's *Ecstatic Economics* has a passage on national wellbeing budgets that serves as a fair warning for the whole subject. After describing what such budgets measure and why it matters, it turns to their limits:

> *The framework is necessary but not sufficient.*[72]

A measurement plan is the same. It will not make a participation system work. It will tell you, with some confidence, whether it did, and where it did not. That is worth a great deal, because a system whose effects cannot be shown will be cut by the first person who doubts it.

What to count

There are four families of measure, and a participation system should keep at least one from each.

The vital signs of the circulation. These are the measures Chapter 4 named for the up-current and the answer, and the secretary in Chapter 8 keeps them from the first week. Ideas received per month. The number of different people who contributed. The share answered within the agreed time. The share declined with a reason. The share implemented. The share of implemented ideas whose effect was reported back to the person who had them. And, from the work of Jeffrey Arthur and Lynda Aiman-Smith on suggestions over time, the share of ideas that question a process rather than fix a fault within it.[73]

These are the earliest signals. They move within weeks. If they do not move, nothing else will, and a leader who watches only the outcome measures will learn this months too late.

The work. The measure chosen in phase four of Chapter 8, and its neighbours: defects or errors per thousand units of output, rework, minutes of downtime, waiting time, complaints, output per hour where output can be honestly counted. Robert Kaufman's survey of 112 Improshare firms and Larry Hatcher and

Timothy Ross's single-plant study both located much of the gain in quality and in machines standing idle less often.[74][75] These are likely to move before output per hour does.

The people. Employment, voluntary turnover, absence, grievances, accidents. Michael Schuster's study of up to nine years of a Scanlon plant found employment and turnover stable while productivity rose, and he presented that stability as a finding in its own right.[76] Hatcher and Ross reported a significant fall in grievances at the moment their plant changed its pay system.[75] And Jeffrey Arthur and Gregory Jelf published a study in the *Journal of Labor Research* in 1999 whose title states its subject: the effects of gainsharing on grievance rates and absenteeism over time.[77] These measures matter for a reason beyond what they show about the system. A productivity gain achieved by cutting people, or by wearing them down, has moved value from one group to another rather than creating it. Counting the people is how an organisation knows which happened.

The relationship. Attitudes, trust, commitment, intention to leave, the sense that it is safe to say a process is wrong. These are measured by survey, and surveys are the weakest of the four families because people answer them strategically and because the questions shape the answers. They are still worth keeping, for one reason: the most striking finding in this literature is about them. Susan Hanlon, David Meyer and Robert Taylor found, three months after a group bonus was withdrawn, higher moral commitment, more prosocial behaviour and less intention to leave among the people who had been under the plan than among a control group.[78] Chapter 4 read that as the circulation outliving the money. It could only have been seen because somebody measured attitudes before, during and after.

Santa Anna's book, in the passage on wellbeing budgets, makes a claim about this family that a measurement plan can test rather than assume:

> *Trust, it turns out, is extraordinarily efficient. Suspicion is extraordinarily expensive.*[72]

The sentence is the author's judgement, not a measured result. A unit that keeps measures from all four families is in a position to see, over two or three years, whether it holds in its own case: whether the months in which the vital signs were strongest were also the months in which grievances fell and defects fell with them.

Why a before matters more than an after

The single most common error in judging any change inside an organisation is to compare one average after with one average before and call the difference the effect.

The difficulty is that almost everything moves. Defect rates drift with the age of equipment. Output varies by season. Grievances rise in the months after a reorganisation elsewhere. A new supplier, a new product, a new manager two floors up: each can move a number by more than a year of good ideas. Compare one quarter before with one quarter after, and you may be measuring any of these.

The researchers whose work this book relies on dealt with this in two ways, and a leader can borrow both.

The first way: watch the line, not two points

Schuster's study of the Scanlon plant, Hatcher and Ross's study of the plant that left piecework, and a 1988 study by John Wagner, Paul Rubin and Thomas Callahan whose title describes its method as an interrupted time series analysis of magnitude and trend, all used the same family of technique.[76][75][79]

In plain words, an interrupted time series works like this. Collect the measure at regular intervals, monthly or weekly, for a long period before the change. That period shows not only the level of the measure but its trend and its ordinary variation: whether it was already improving, how much it wobbles from month to month, whether it has a seasonal pattern. Then make the change and keep collecting. Now ask two questions. Did the level step up or down at the moment of the change, by more than its ordinary wobble? And did the trend after the change differ from the trend before it?

A step at the moment of the change is harder to explain away than a slow drift, because most background forces do not happen to begin on the day a system is introduced. This is why Schuster's word "abrupt", discussed in Chapter 5, carries weight. The level of output per hour stepped at the point the plan began. And a change in trend tells you something different: that the system did not merely produce a one-off gain but altered the rate at which things improve, which is what a working circulation of ideas should do.

The technique has limits a leader should know. It needs a long before: researchers typically want many data points on each side of the change, which is why Chapter 8 asks for at least three months of collection before the committees are formed and prefers a year. It cannot rule out a background event that happens to coincide with the change. And when several things change at once, as they did in Schuster's plant, where the committees and the bonus began together, it cannot tell you which of them produced the effect.

A unit can do a simple version of this without statistical software. Plot the measure, month by month, for as long a before as exists. Mark the date of each phase in Chapter 8: the answer rule, the posted measure, the committees, any design adopted after a vote. Mark every outside event the measurement group noted. Then look. A line that was flat and steps down after the answer rule, with no outside event at that point, is a different story from a line that was already falling at the same rate before anything began.

The second way: find a comparison

The stronger design is a comparison group: a similar unit, at the same time, that did not adopt the system.

Susan Hanlon and Robert Taylor's 1991 study of a self-designed gainsharing plan at a priority package delivery firm, described in its abstract as the first reported gainsharing study to use control groups and pre- and post-tests in a quasi-experimental field study, found that in the first six months positive job-related communication and the communication of ideas increased among the people under the plan.[80] The 1994 follow-up with David Meyer is the study, cited above, that kept measuring after the bonus was with-

drawn.[78] And M. M. Petty, Bart Singleton and David Connell published in the *Journal of Applied Psychology* in 1992 what their title calls an experimental evaluation of an organisational incentive plan in the electric utility industry.[81]

A comparison group answers the question an interrupted time series cannot: what would have happened anyway? If both units' defect rates fell by the same amount in the same months, something other than the system was at work. If only the unit with the system improved, the case is stronger.

Most organisations cannot randomise which units adopt a system, and should not try. But many have a natural comparison: a second site, a second shift, a department that will start a year later. Where one exists, collect the same measures in it, from the same date, with the same care. Starting the system in one unit first and another a year later is not only a sensible way to learn; it gives the organisation its own comparison for free.

A comparison is never perfect. The comparison unit knows it is being compared, which changes behaviour. Managers talk to each other, and good practices spread. Units differ in ways nobody listed. The honest reading of a comparison is that it narrows the range of explanations. It does not close it.



The survivors, again

HAUTE LUMIÈRE · THE GAIN WAS ALWAYS SHARED · CHAPTER NINE

The survivors, again

Chapter 3 named the fifth critique of gainsharing as a problem in the evidence itself: case studies are published more often when the plan worked, and surveys of current users cannot reach the firms that gave up. Kaufman's 112 Improshare firms were those still known to be using the plan in 1988 who chose to reply.[74] Bullock and Tubbs's thirty-three cases were the ones somebody published.[82]

The same bias operates inside an organisation, and it is worth naming because it is the one most leaders do not see.

Inside a unit, the ideas that get remembered are the ones that worked. The committee meetings that get described are the good ones. The supervisor who resisted and was moved is no longer in the room when the system is reviewed. The people who left in the first year, perhaps because the system asked something of them they did not want to give, perhaps because it did not ask enough, are not surveyed in the second. Every review of a system conducted only from the inside, by the people who remain, is a survey of survivors.

Three habits counter it.

Keep the declined ideas. The secretary's ledger in Chapter 8 records every idea and every answer, not only the adopted ones. A review that reads the declined ideas, and asks how many were declined for good reasons and how many because nobody wanted to deal with them, sees what a review of successes cannot.

Ask the leavers. Every person who leaves the unit in the first two years is asked, in a short conversation with someone outside their line, what they thought of the system and whether it had anything to do with their leaving. Most will say no. The ones who say yes are the evidence a survivor survey cannot reach.

Count the stops. Committees that stopped meeting. Supervisors who stopped answering. Months in which the measure was not posted. Every interruption is recorded with its date and reason. A system that has never been interrupted is either very well run or not being watched closely.

Medians, means and the size of a claim

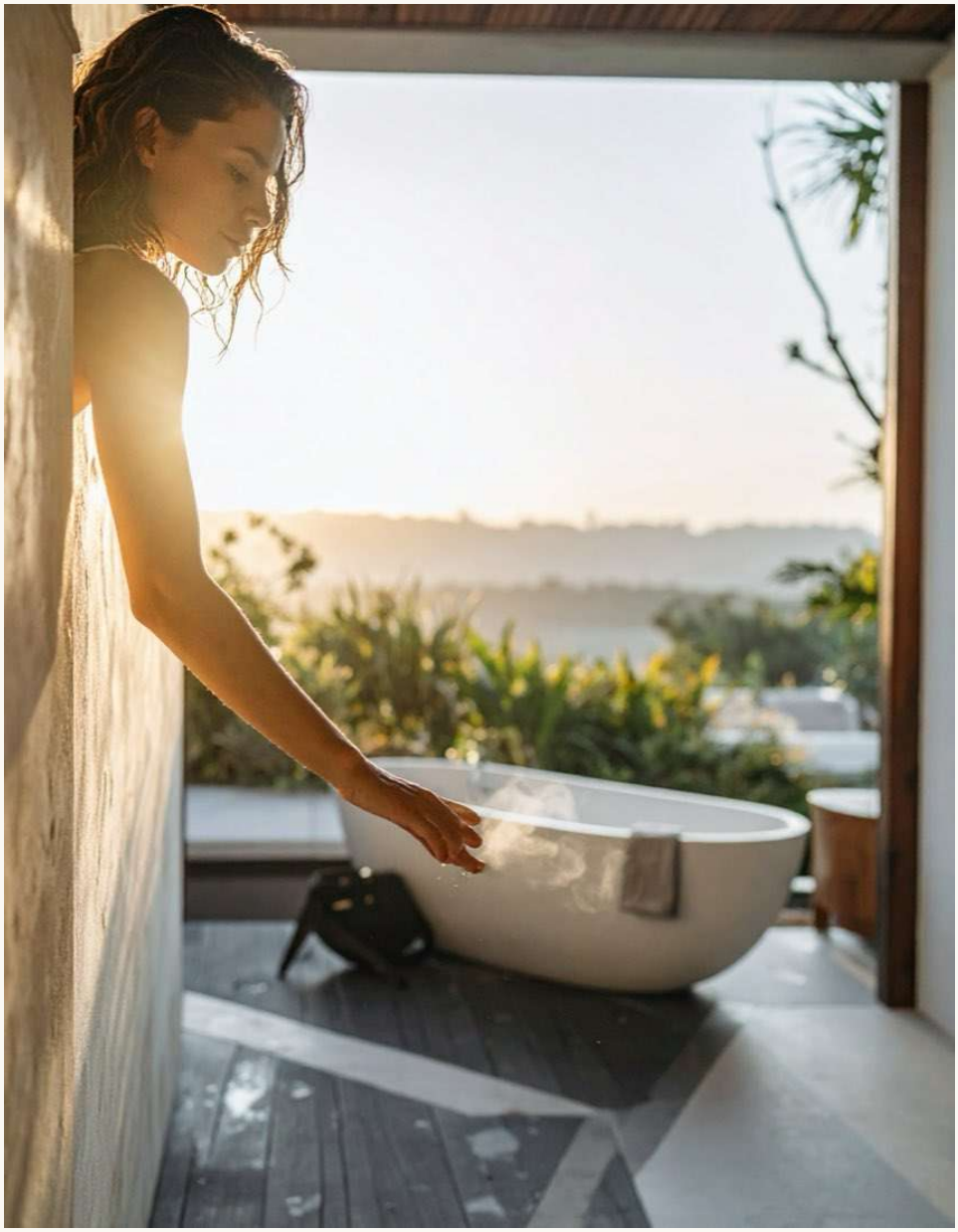
When Kaufman reported Improshare's results, he noted that mean productivity increases were larger than median increases and followed the same pattern.[74] Chapter 5 drew the practical lesson: if you use one figure from that study to think about your own firm, use the median, and treat even that as an upper estimate.

The distinction matters inside a unit too. Suppose a plant has six lines. Five improve a little. One improves a great deal, because its machine had a fault that a single idea fixed. The average improvement across six lines looks impressive. The typical line improved a little. Both statements are true. A leader who reports only the average will be believed until someone looks at the lines one by one, and then will not be believed about anything.

Report, at minimum:

- the median and the range, not only the average;
- the number of units, lines or teams the figure covers;
- the period, with its start and end dates;
- the outside events noted during that period;
- what the figure cannot show.

That last line is the habit of every note in this book, and a leader can adopt it without any training in statistics. A figure accompanied by a plain statement of its limits is more persuasive to a sceptical board than a larger figure without one, because it shows the person presenting it has already asked the board's question.



Writing the expectation down first

HAUTE LUMIÈRE · THE GAIN WAS ALWAYS SHARED · CHAPTER NINE

Writing the expectation down first

Chapter 8 asked the unit leader, secretary and finance person to write down, before the committees were formed, what they expected to change over the next year, by how much and by when, and to file that page where it could not be revised. It is worth saying why this single page matters so much.

Without it, every result can be read as a success. If defects fell by a tenth, that was the goal. If they fell by a fiftieth, the real goal was cultural. If they did not fall at all, the system has laid foundations whose effects will show later. Each of these readings may be true. None can be checked, because the expectation was formed after the result was known.

With it, the review has a fixed point. Expected ideas per month against actual. Expected median days to answer against actual. Expected change in the measure against actual. Where actual fell short, the review asks why. Where it exceeded expectation, the review asks whether something else was happening. Either way, the organisation learns something it could not have learned by looking only at the result.

This is the discipline of a pre-registered study, applied to a single unit. It costs one page and one hour. It is the cheapest protection available against the most common error in judging change, which is deciding what you meant after you see what happened.

Reading a mixed result

Most results will be mixed. The vital signs will rise and the measure will not move. Or the measure will improve and grievances will rise. Or everything will improve in one unit and nothing in another.

Chapter 4 offered a reading that helps here: a failing system is not a verdict, it is a location. Read a mixed result by asking where the circulation has blocked.

- Ideas are not arriving, or arrive from only a few people. The up-current is blocked. The usual causes are that the answer rule is not kept, or that it is not yet safe to say a process is wrong. Look at the on-time share and the declined-without-reason count first.
- Ideas arrive and are answered, but few are implemented. The committees cannot act. Look at the budget limit and at who holds the decisions above it.
- Ideas are implemented, but the measure does not move. Either the measure is not one the unit can affect, or the ideas being implemented are small fixes that do not touch the measure. Look at whether the measure passes the test in Chapter 2, and at the share of ideas that question a process.
- The measure moves, but people do not believe it. The books are not open enough. Look at whether the finance person explains the number every month, and whether anyone has asked to see the record and been refused.

- Everything moves, then levels off. The plateau Kaufman's firms reported. Look at training across boundaries and at the second-order ideas, as Chapter 8 describes.
- Outcomes improve, but people measures worsen. The gain may be coming from people rather than for them. Look at absence, grievances, accidents and turnover, line by line, before reporting any success.

Santa Anna's book offers a personal practice that maps onto this reading at the scale of one life. In the weekly Vitality Inventory, the reader reviews the economic events of the week and asks of each whether it generated or depleted vitality, and the book says the practice builds a map more reliable than a spreadsheet because it measures

> *the quality of aliveness that your economic life produces.*[83]

The claim about reliability is the author's. The structure of the practice is useful for a unit. A quarterly review that asks of each measure not only whether it rose but whether the people producing it are more or less able to keep producing it is asking the question the people measures exist to answer. Chapter 11 returns to the Inventory at the scale of a household.

Presenting the evidence to someone who was not there

The moment this chapter began with, the question from a new chief executive or a sceptical board, deserves a final word, because it is where measurement either protects a system or fails to. Chapter 10 gives the conversation in full; here is what to bring to it.

Bring the line, not the average: the monthly measure plotted from the earliest data, with each phase and each outside event marked. Bring the vital signs, month by month. Bring the people measures. Bring the sealed expectation and the comparison with what happened. Bring the comparison unit, if there is one. Bring the declined ideas and the leavers' answers. And bring, in a single paragraph, what none of this can show.

A system presented this way does not need to claim more than it achieved. It needs only to show that the organisation has been watching itself honestly. That is rare enough to be persuasive in its own right.

The Tool: The Honest Reading

A quarterly review, two hours, conducted by the screening committee with the sponsor. Fill every line from the secretary's ledger and the posted measure. Where a figure is an estimate, mark it as one.

1 · The before

- Months of data collected before the answer rule began: ___ · before the committees: ___

Was the measure already improving before anything began? yes / no / unclear · evidence:

Sealed expectation opened (date filed):

2 · The vital signs this quarter

- Ideas received ___ · from how many different people ___ · team size ___
- Answered on time ___ % · declined with reason ___ · declined without reason ___
- Implemented ___ % · effect reported to the person ___ of ___
- Ideas marked as questions about a process ___ of ___

3 · The work

- The chosen measure: median across units ___ · range ___ to ___ · number of units ___
- Defects or errors per thousand ___ · downtime minutes per shift ___
- Plotted monthly with phases and outside events marked? yes / no

Did the line step at any phase? which: · by more than its ordinary month-to-month variation? yes / no

Did the trend change after any phase?

4 · The people

- Headcount start ___ · end ___ · voluntary leavers ___
- Absence rate this quarter ___ · same quarter last year ___
- Grievances ___ · same quarter last year ___ · accidents ___

Roles removed by implemented ideas ___ · where each person went:

5 · The comparison

Comparison unit, if any: · same measures collected from the same date? yes / no

Change in the comparison unit over the same period:

6 · The survivors

- Declined ideas reviewed: ___ · declined for reasons the committee would now reverse: ___

Leavers asked about the system: ___ of ___ · said it played a part: ___ · what they said:

Interruptions this quarter (committee missed, measure not posted, answers late), with dates:

7 · Expected against actual

MEASURE	EXPECTED (SEALED)	ACTUAL	GAP	WHY
Ideas per month				
Median days to answer				
Share implemented				
The chosen measure				
Voluntary turnover				

8 · Where the circulation blocks

- First stage where flow falls off: ideas / answers / implementation / measure / belief / people

One change for next quarter: · owner:

9 · What this reading cannot show

In one paragraph, the alternative explanations this quarter's evidence cannot rule out:

References

1. Santa Anna, A. *Ecstatic Economics: Wealth Creation Rooted in Vitality, Not Extraction*, chapter 5, "Wellbeing Budgets: When Governments Measure What Matters". The Luminous Developmental Canon, Luminous Prosperity Inc.
2. Arthur, J. B., & Aiman-Smith, L. (2001). Gainsharing and organizational learning: An analysis of employee suggestions over time. *Academy of Management Journal*, 44(4), 737–754. https://doi.org/10.2307/3069413. One plant over four years; the classification of ideas follows the framework the article proposes, not a re-reading of its results.
3. Kaufman, R. T. (1992). The effects of Improshare on productivity. *ILR Review*, 45(2), 311–322. https://doi.org/10.1177/001979399204500208. Self-reported results from 112 firms still known to use the plan in 1988; abandoned plans are under-represented.
4. Hatcher, L., & Ross, T. L. (1991). From individual incentives to an organization-wide gainsharing plan: Effects on teamwork and product quality. *Journal of Organizational Behavior*, 12(3), 169–183. https://doi.org/10.1002/job.4030120302. One company; interrupted time-series analysis over more than four years. It cannot show the result would transfer to another plant.
5. Schuster, M. (1984). The Scanlon Plan: A longitudinal analysis. *The Journal of Applied Behavioral Science*, 20(1), 23–38. https://doi.org/10.1177/002188638402000107. One plant; the bonus and committees began together, so their separate effects cannot be distinguished.
6. Arthur, J. B., & Jelf, G. S. (1999). The effects of gainsharing on grievance rates and absenteeism over time. *Journal of Labor Research*, 20(1), 133–145. https://doi.org/10.1007/s12122-999-1029-3. Cited for its stated subject only; its data and findings were not read for this chapter, and nothing here reports them.
7. Hanlon, S. C., Meyer, D. G., & Taylor, R. R. (1994). Consequences of gainsharing: A field experiment revisited. *Group & Organization Management*, 19(1), 87–111. https://doi.org/10.1177/1059601194191005. One site with a control group, measured three months after the bonus ended; it cannot say how long the effects last.
8. Wagner, J. A., Rubin, P. A., & Callahan, T. J. (1988). Incentive payment and nonmanagerial productivity: An interrupted time series analysis of magnitude and trend. *Organizational Behavior and Human Decision Processes*, 42(1), 47–74. https://doi.org/10.1016/0749-5978(88)90019-2. Cited for its method, as stated in its title; its setting and findings were not read for this chapter.
9. Hanlon, S. C., & Taylor, R. R. (1991). An examination of changes in work group communication behaviors following installation of a gainsharing plan. *Group & Organization Studies*, 16(3), 238–267. https://doi.org/10.1177/105960119101600302. One firm in package delivery; communication was measured over the early months and does not show economic outcomes directly.
10. Petty, M. M., Singleton, B., & Connell, D. W. (1992). An experimental evaluation of an organizational incentive plan in the electric utility industry. *Journal of Applied Psychology*, 77(4), 427–436. https://doi.org/10.1037/0021-9010.77.4.427. Cited for its design as stated in its title; its findings were not read for this chapter.
11. Bullock, R. J., & Tubbs, M. E. (1990). A case meta-analysis of gainsharing plans as organization development interventions. *The Journal of Applied Behavioral Science*, 26(3), 383–404. https://doi.org/10.1177/0021886390263011. Thirty-three published cases; published cases over-represent successes.
12. Santa Anna, A. *Ecstatic Economics*, chapter 3, "Practice 5: The Vitality Inventory".

How to Know It Worked

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/