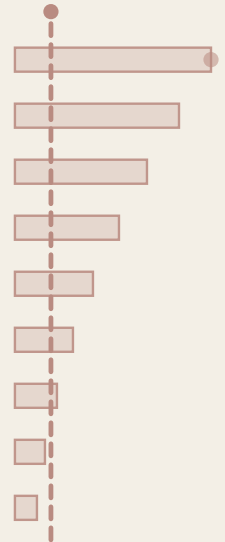


YUVAL NOAH HARARI · SAPIENS, CHAPTER 19

The Happiness Ledger

Taking apart the claim that all of it bought us nothing – axiom by axiom, in the simplest English possible.



“We are far more powerful than our ancestors, but we are not much happier.”

SAPIENS: A BRIEF HISTORY OF HUMANKIND, 2011

-
- 01 Thirteen hidden axioms, dug out and tested one by one
 - 02 The argument steelmanned – and then steelmanned again
 - 03 The measurement that does not exist before 1946
 - 04 Japan, the lottery study, the despair data, and Bhutan
-

How to read this document

Everything here is written in short, plain sentences. No jargon is used without being explained on the spot. If you can read a newspaper, you can read this.

This is a study of **one sentence**. Not a book review. Not a summary of *Sapiens*. One sentence, pulled apart until nothing is left hidden inside it.

The sentence is: “*We are far more powerful than our ancestors, but we are not much happier.*” It is the closing move of the book — the line that turns four hundred pages of achievement into a question. Vaccines, literacy, electricity, the end of famine in most of the world: all of it, and the ledger still does not balance.

It sounds simple. It is not. Buried inside it are about thirteen separate assumptions. Each one has to hold for the sentence to work. Most readers swallow all thirteen at once without noticing. This document digs each one out, holds it up to the light, and asks: is this true?

Then it does something harder. It builds the **strongest possible version** of Harari’s argument — stronger than the version he actually wrote. That is called a *steelman* (the opposite of a strawman, where you attack a weak fake version of someone’s view). Then it takes that strong version and makes it stronger again. Only after the argument is at its absolute best does it get attacked.

THE FOUR BOXES YOU WILL KEEP SEEING

IN ONE LINE — the point, stripped to a single sentence.

SAY IT SIMPLE — the same idea with an everyday example.

STEELMAN — the best case for the argument.

WHERE IT CRACKS — the honest weaknesses, including ones fans of the book usually skip.

A warning about honesty. This document does not decide for you. By the end you will find that the sentence contains one genuine and unanswerable challenge, one empirical claim that the data have since contradicted, and one comparison that cannot be made at all because the measurement did not exist. All three parts are labelled.

One more thing, and it is the most important thing here. The sentence compares us to **our ancestors**. Nobody has ever measured the happiness of an ancestor. The first survey that asked anybody how happy they were ran in 1946. Everything before that — every century of the story the book has just told — is outside the data entirely. Harari knows this and says so. The sentence does not. Keep that in view and the whole argument becomes readable.

Contents

A SINGLE-SENTENCE STUDY IN TEN PARTS

00	The sentence itself, and what it actually says	The exact words
01	The thirteen axioms hidden inside the claim	The load-bearing walls
02	Why Harari said it — motive, lineage, and the long game	Easterlin to Kahneman to Harari
03	Steelman, level one: the strongest honest case	The treadmill
04	Steelman, level two: the argument that survives every attack	The Unkept Ledger
05	Where it cracks — the archive, the flat line, the set point	1946 and after
06	What every field of study says about it	Eighteen verdicts
07	The argument tested against six real bodies of data	Where the ledger was opened
08	The ledger you are keeping right now	A usable tool
09	The scorecard, and the sentence rewritten	Thirteen verdicts
—	Appendix: glossary, timeline and sources	Everything, explained

• • •

IF YOU READ NOTHING ELSE

Harari asks the right question and answers it with data that do not exist. There is no happiness measurement before 1946. Within the period we can measure, the flat line he describes has largely failed to hold up: reported happiness rose in 45 of 52 countries between 1981 and 2007, richer countries report higher wellbeing than poorer ones, and the money-and-happiness curve does not flatten where it was said to. What survives — and it is formidable — is not a finding about happiness. It is an indictment of every institution that decides what gets counted.

PART 00

The sentence itself, and what it actually says

Before you can judge a claim, you have to know exactly what was claimed. Most fights about this sentence are fights about a comparison nobody can actually run.

Where the sentence sits

This is the last argument in the book. Chapter 19, “And They Lived Happily Ever After”, comes after everything — after the foragers, the wheat, the empires, the money, the science, the industry. The whole structure has been built so that this question can be asked at the end: *was any of it worth it?*

Harari’s answer is a shrug, and the shrug is deliberate. He does not say we are miserable. He says we are *not much happier*, and — more importantly — that historians have almost never asked, so nobody knows. That second point is the real one, and it is much stronger than the first.

The exact words, taken slowly

THE SENTENCE

“We are far more powerful than our ancestors, but we are not much happier.”

Four moving parts, and each one is doing work:

- **We.** Which we? Humanity’s average? The rich world? The median person alive today? The answer changes the verdict completely, and the sentence does not say.
- **Far more powerful.** This half is not in dispute. Life expectancy, child survival, food supply, literacy, energy per person — all transformed, all measured, all documented.
- **Our ancestors.** Which ancestors? Foragers of 30,000 years ago? Medieval peasants? Your grandparents? Nobody asked any of them, ever.
- **Not much happier.** How much is “much”? The sentence never says, which means no result could contradict it.

SAY IT SIMPLE

Imagine claiming that a hospital’s patients are no happier than they were in 1650. You would need two things: a way of measuring patient happiness, and records from 1650. We have the first. The second does not exist, and never did.

What Harari actually argues in the chapter

The sentence is the headline. The chapter underneath it makes five separate moves, and they are of very different quality. Setting them out now saves a great deal of confusion later.

THE MOVE	WHAT IT SAYS	STATUS
The neglect charge	Historians have written about power, wealth and empires and almost never about whether people felt better.	Correct, important, and the best thing in the chapter.
The expectations argument	Happiness depends on the gap between what you have and what you expect, and expectations rise with conditions.	Real effect, well evidenced, but much smaller than “cancels everything”.
The biochemical argument	Happiness is a chemical state with a personal set point that events barely move.	The 2011 popular consensus. It has aged badly.
The community argument	Modernity gained wealth and lost family and community, and that trade may have been bad.	Plausible, partly supported, and the part most worth taking seriously.
The Buddhist argument	Chasing pleasant feelings is itself the source of suffering, so the ledger is the wrong instrument.	A philosophical position, offered as one, and honestly flagged.

Notice what happens when these are separated. The neglect charge does not need the biochemistry. The Buddhist argument actually *contradicts* the biochemistry — one says happiness is a chemical you cannot change, the other says happiness is a relationship to your own mind that you can train. Harari presents both. He does not reconcile them.

Three different claims are wearing one coat

1. **An empirical claim.** Measured wellbeing has not risen much. This is checkable, has been checked, and the results are in Part 07.
2. **A methodological claim.** Nobody has been keeping this ledger, so the entire story of progress has been told using a scoreboard that omits the thing it claims to be for. This is correct and devastating.
3. **A philosophical claim.** Happiness is the right yardstick for judging history. This is asserted, not argued, and it is the shakiest of the three.

IN ONE LINE

Harari's sentence is a superb question, an unmeasurable comparison, and a contested value judgement, held together by the word “but”.

What is not being claimed

He is not saying the past was better. Harari is explicit that pre-modern life was short, violent and disease-ridden, and Chapter 19 does not romanticise it. Anyone answering with infant mortality stat-

istics has not touched the argument — though, as Part 05 shows, those statistics matter enormously for a reason the sentence obscures.

He is not claiming to have proved it. The chapter’s honest core is agnostic: we do not know, and we should be embarrassed that we do not. The confident version — *progress bought nothing* — is what readers carry away, and it is stronger than anything Harari actually defends.

1946

the first year in which anybody was surveyed about their own happiness

45 / 52

countries where reported happiness rose, 1981–2007

22

lottery winners in the study that founded the whole idea

• • •

PART 01

The thirteen axioms hidden inside the claim

An axiom is a thing you have to assume before your argument can even start. It is the floor you stand on. Harari's one sentence stands on thirteen floors. Several of them were solid in 2011 and are not solid now.

Axiom 1 Happiness is one thing

IN ONE LINE

There is a single quantity called happiness that a person has more or less of.

Researchers now routinely separate at least three quantities, and they do not move together.

- **Evaluative wellbeing** — how you rate your life when you stop and think about it. Measured by asking people to place their life on a ladder from 0 to 10.
- **Experienced wellbeing** — how you actually feel, moment to moment, sampled during the day. Often quite different from the rating.
- **Eudaimonic wellbeing** — whether your life feels worthwhile. Meaning, not mood.

The gap is not academic. Raising children reliably lowers experienced mood and reliably raises meaning. So does caring for a dying parent. So does most work worth doing. A sentence with one word for all three cannot say anything about a civilisation.

Why he needs it: a single quantity is what makes the comparison possible at all. Three quantities and the sentence has to be three sentences with three answers.

What would break it: it is already broken in the literature. Harari's own chapter half-concedes it by distinguishing pleasant feelings from meaning — and then goes back to using one word.

Axiom 2 Happiness can be measured at all

IN ONE LINE

An inner state can be turned into a number that means something.

This one is in better shape than sceptics assume. Self-reported wellbeing correlates with things nobody chooses: how often you smile, what your friends say about you, brain activity, cortisol, and how likely you are to still be alive in ten years. It predicts. A number that predicts is measuring something.

STEELMAN

Harari is entitled to this axiom. The standard objection — “you can’t put a number on feelings” — has been answered for thirty years. The measurements are noisy, culturally loaded and crude, and they are still real measurements, in the same way that a thermometer under the tongue is a real measurement of something a body is doing.

What would break it: very little. This is the strongest floor in the sentence.

Axiom 3 Asking people is the way to measure it

IN ONE LINE

What a person says about their happiness is good evidence for their happiness.

Mostly, with a known set of distortions. People rate their whole life by how the last few minutes went. A questionnaire filled in on a sunny day scores higher. What you were asked immediately before changes the answer. The wording of the scale changes the answer.

None of this is fatal — the errors are largely noise and wash out in large samples. But it matters for a claim about centuries, because the distortions are not constant across cultures or eras, and there is no way to correct for a bias you cannot observe.

What would break it: nothing on the timescale of a survey. Quite a lot on the timescale of a civilisation.

Axiom 4 A 7 means the same thing in Denmark and Nigeria

IN ONE LINE

Numbers reported by different people in different languages can be added up and compared.

This is the quiet load-bearer under every international happiness ranking. Cultures differ in how willing people are to claim extremes; some languages have no clean equivalent of “happy”; some have several with different weights. East Asian respondents tend to avoid the ends of scales; Latin American respondents tend to use them. Those tendencies are stable and large enough to move a country several places in a league table.

Why he needs it: without it there is no cross-country evidence at all, and the chapter has almost nothing left to work with.

What would break it: it is a known limitation, not a refutation. Rankings should be read as rough bands, not as positions. Almost nobody reads them that way.

Axiom 5 There is evidence about how our ancestors felt

IN ONE LINE

We can say something about the happiness of people who lived before us.

This is the axiom that fails hardest, and everything in the sentence depends on it.

The first surveys asking people how happy they were date from 1946. Systematic international data begins in the 1970s and 1980s. For most of the world, comparable annual data begins in 2005–2006. Before 1946 there is nothing — not a thin record, not a biased record: no record.

So the sentence compares a measured quantity with an unmeasured one and reports the difference. In any other field this would end the discussion.

THE ONE HONEST ATTEMPT

In 2019 a team led by Thomas Hills published in *Nature Human Behaviour* the only serious effort to extend the record backwards: sentiment analysis of millions of digitised books and newspapers in four countries, producing a National Valence Index from about 1820. It is ingenious and it is a proxy — it measures the emotional temperature of published language, not of people. It reaches 1820. The book's comparison runs to 30,000 BC.

What would break it: nothing can fix it. This is a permanent hole, and Harari says so in the chapter. The sentence does not.

Axiom 6 The average is the right unit

IN ONE LINE

If the average has not moved, nothing important has happened.

Averages hide distributions, and distributions are where the moral content lives. Two societies can report identical mean wellbeing while one has a contented majority and a small population in severe distress, and the other has neither. The 2023 resolution of the money-and-happiness debate turned on exactly this: income keeps improving wellbeing for most people, while for roughly the least happy fifteen per cent it stops helping above a threshold. One curve, two populations, opposite conclusions.

What would break it: any case where the mean is flat and the tails have moved. There are many, and the deaths-of-despair data are the starkest.

Axiom 7 Power and happiness are the right two things to compare

IN ONE LINE

The gains of history are best summarised as “power”, and the test of those gains is “happiness”.

“Power” is a strange summary of what changed. The largest measurable differences between then and now are not power at all: a child born today is roughly ten times less likely to die before the age of five than one born in 1800, and far more than that in the richest countries; most people alive have never seen a famine; most can read. Those are not increases in power. They are decreases in suffering, and the sentence has no column for them.

What would break it: the asymmetry in Axiom 12, which is the deepest problem in the whole sentence.

Axiom 8 Happiness is the correct yardstick for judging history

IN ONE LINE

The question “was it worth it?” means “did it make people feel better?”

This is a substantial ethical commitment — roughly, a form of utilitarianism — and it is not argued for. Rival answers are available and serious: history might be judged by freedom, by capability (what people are actually able to do and be), by justice, by knowledge, by the sheer number of lives that got to happen at all.

SAY IT SIMPLE

A person freed from prison who reports the same mood as before has still been freed. If your instrument records no change there, the problem is the instrument, not the freedom.

What would break it: nothing decisively — it is a values question. But it should be stated as one, and in the sentence it is invisible.

Axiom 9 Each person has a biological happiness set point

IN ONE LINE

Your happiness is a chemical thermostat, set at birth, that events cannot move for long.

This was the received view when *Sapiens* was written, and it came from a real finding: twin studies in the 1990s put the heritable share of wellbeing around 44 to 52 per cent, with the stable component higher still. One of those papers contained the line that everything else has been quoting ever since — that trying to be happier might be as futile as trying to be taller.

Two things have happened since. The heritability estimates held up, roughly, at somewhat lower values. The *immovability* claim did not. Half the variance being genetic leaves the other half somewhere, and that other half is where every life event lives.

What would break it: long-running panel studies in which the same people are followed for decades. They exist, and they are in Axiom 10.

Axiom 10 Adaptation always returns you to the set point

IN ONE LINE

Whatever happens to you, good or bad, you drift back to where you were.

This is the hedonic treadmill, and it is the engine of the whole sentence: if adaptation is complete, then no improvement in conditions can ever register, and the flat line is guaranteed in advance.

It is not complete. Large national panel studies following the same people for many years — in Germany, Britain and Switzerland — find that adaptation is real, quick and near-total for some events, and clearly incomplete for others. Unemployment, disability, chronic pain and widowhood are associated with lasting shifts. People do not fully return.

WHERE IT CRACKS

This is the single most consequential correction to the chapter, and it cuts both ways. If bad conditions leave permanent marks, then removing bad conditions produces permanent gains — which means the vast historical reduction in disability, hunger, chronic pain and premature bereavement should register on the ledger. The treadmill was the reason to expect nothing. The treadmill has slipped.

What would break it: it is already qualified. Adaptation is a real phenomenon with limits, not a law.

Axiom 11 Expectations rise exactly enough to cancel the gains

IN ONE LINE

Getting more makes you want more, in the same proportion, forever.

The first half is well supported: people judge their situation against what they expected and against what those around them have. The second half — that the cancellation is complete — is the part that would need to be true for the sentence to work, and it is not what the data show. Richer countries do report higher wellbeing than poorer ones, consistently, and the relationship does not run out at the top.

SAY IT SIMPLE

Expectations are a tax on improvement, not a confiscation. The gain is smaller than you hoped. It is not zero, and treating the tax as a hundred per cent is what produces the flat line.

What would break it: a clear income–wellbeing gradient across countries. It exists.

Axiom 12 Suffering avoided and pleasure gained belong in the same column

IN ONE LINE

A child who does not die of measles and a person who buys a better phone are both entries in one ledger, to be added up.

This is the deepest problem in the sentence, and almost nobody raises it.

Most of what changed since 1800 is not the addition of pleasures. It is the subtraction of catastrophes: the child who lives, the tooth extracted with anaesthetic, the mother who survives the birth, the crop failure that does not become a famine. A ledger that asks “are you happier?” is asking about a level. Catastrophe-avoidance does not show up as a level. It shows up as an absence — of the event that would have destroyed the level.

WHERE IT CRACKS

Ask a hundred parents in 1800 and a hundred today how happy they are and you may well get similar numbers. More than a third of the first hundred have buried a child. The instrument does not see that, because the people it destroyed are not in the sample and the ones who survived have adapted. A measure that cannot detect the largest change in human experience is not a measure of that change.

What would break it: nothing rescues it. This is where the sentence is not merely unproven but structurally unable to answer its own question.

Axiom 13 The claim is one that evidence could contradict

IN ONE LINE

“Not much happier” is a statement that could turn out to be false.

How much is “much”? If reported wellbeing rose by half a point on a ten-point scale across a century, is that much? The sentence does not say, and so no result can be inconsistent with it. This is the difference between a hypothesis and a mood.

What would break it: a stated threshold. Harari never gives one, and neither do most people who quote him.

• • •

WHERE WE ARE SO FAR

Of the thirteen, two are solid (2, and the first half of 11). Two are workable with care (3, 4). Three have been contradicted or heavily qualified by evidence published after the book (9, 10, and the second half of 11). Three are values or definitions presented as findings (1, 7, 8). One is unfalsifiable as written (13). One is a hole that cannot be filled (5). And one — the twelfth — describes a question the instrument cannot answer.

That is a heavy scorecard. And yet the chapter is the most important in the book, for a reason the next three parts set out: the part of it that survives is not a claim about happiness at all.

PART 02

Why Harari said it — motive, lineage, and the long game

Nobody writes a sentence like that by accident. It is aimed at something. Understanding what it is aimed at explains most of its shape — including its exaggerations.

The idea has a paper trail

Unlike the sentence about imagined orders, this one sits on a specific and recent research literature, and the chapter is a fair summary of what that literature looked like around 2011. The lineage:

WHO AND WHEN	WHAT THEY SAID	WHAT IT ADDED
Jeremy Bentham 1789	The right action is the one producing the greatest happiness; in principle it can be calculated.	The original proposal that society keep this ledger at all.
Brickman & Campbell 1971	People adapt to improvements, so gains in conditions produce no lasting gain in feeling.	Named the treadmill. The engine of Harari's chapter.
Richard Easterlin 1974	Richer people within a country are happier, but as a whole country gets richer its average happiness does not rise.	The paradox the chapter rests on.
Brickman, Coates & Janoff-Bulman 1978	Lottery winners were not happier than neighbours; people paralysed in accidents were only modestly less happy.	The single most quoted study in the field, and the smallest.
Lykken & Tellegen 1996	Around half the variation in wellbeing is genetic; the stable part far more.	The set point. Also the line about trying to be taller.
Daniel Kahneman 1999 onwards	The self that experiences a life and the self that remembers it give different answers.	Broke “happiness” into pieces that behave differently.
Richard Layard 2005	Governments should target wellbeing directly instead of income.	Turned the research into a policy movement.

WHO AND WHEN	WHAT THEY SAID	WHAT IT ADDED
Stiglitz, Sen & Fitoussi 2009	GDP is a bad measure of how a society is doing; measure wellbeing too.	Made the neglect charge official.

So Harari is reporting, not inventing — and reporting accurately, as of 2011. The problem is what happened next. Between 2008 and 2023 almost every load-bearing element of that summary was challenged, qualified or overturned, and the sentence has gone on travelling as though nothing occurred.

THE DATES THAT MATTER

Stevenson and Wolfers published their reassessment of the Easterlin paradox in **2008**. Inglehart and colleagues published rising happiness in 45 of 52 countries in **2008**. Both were in print before *Sapiens* appeared in Hebrew in 2011, and neither shapes the chapter. This is not hindsight. It was available.

What the sentence is aimed at

Target one: the progress story. The book has spent four hundred pages accumulating achievement. The chapter exists to refuse the conclusion the reader is about to draw. That is a legitimate structural move and it is executed superbly.

Target two: the historical profession. This is the sharpest and least noticed target. Harari's complaint is that his own discipline has written millions of pages about kings, crops, wars and trade routes and almost nothing about whether any of it made anyone feel better — and that this is not an oversight but a symptom. Historians study what left records, and feelings left none.

Target three: the reader. The chapter turns outward at the end. The question stops being about the Sumerians and becomes about you: you are richer than your grandparents, and are you better off? Very few readers answer that with a confident yes, and the chapter banks on it.

STEELMAN OF THE MOTIVE

There is nothing cheap here. A historian noticing that his profession has never measured the one variable everything was supposedly for is doing his job. The neglect charge would survive even if every empirical claim in the chapter were wrong — and, as Part 05 shows, several are.

The Buddhist strand, stated openly

Harari has practised Vipassana meditation since 2000, in the tradition of S. N. Goenka, sits for two hours a day, and takes a long silent retreat most years. He says so in the book's acknowledgements and in interviews. He is not hiding it, and this study does not treat it as a hidden agenda.

It matters analytically for one reason. The Buddhist argument in the chapter — that chasing pleasant feelings is itself the mechanism of suffering — is not a supplement to the scientific case. It is a rival to it. The biochemical argument says the ledger cannot be improved because your chemistry is fixed. The Buddhist argument says the ledger is the wrong instrument and a person can learn to stop keeping it. Both cannot be the reason the sentence is true.

SAY IT SIMPLE

One argument says the thermostat is welded. The other says stop staring at the thermostat. If the first is right, the second is impossible. The chapter offers both without noticing.

The long game

The set point is not a detail; it is the hinge to the next book. If happiness is a biochemical state with a fixed range, then the obvious twenty-first-century project is to re-engineer the chemistry — and that is precisely what *Homo Deus* (2015) proposes as one of humanity's next great undertakings. Chapter 19 of *Sapiens* is the load-bearing wall for the sequel.

Which makes the weakening of the set-point evidence more consequential than it looks. If wellbeing responds durably to unemployment, disability, chronic pain and bereavement — as the panel studies indicate — then the shortest route to raising it is not neurochemistry. It is jobs, medicine, pain control and public health. That is a far less exciting conclusion and a much better supported one.

The timing, which is not an accident

Sapiens arrived in the middle of a wave. Bhutan had been talking about Gross National Happiness since the 1970s and wrote it into its constitution in 2008. France commissioned Stiglitz and Sen in 2008. The United Nations passed a resolution on happiness in 2011. Britain's statistical office began asking wellbeing questions in 2011. The first World Happiness Report appeared in 2012.

The book landed exactly as governments started keeping the ledger Harari says nobody keeps — and the chapter does not mention any of it. That absence is the strongest evidence that the chapter is a philosophical argument wearing an empirical coat.

IN ONE LINE

Harari did not invent the doubt about progress. He found the phrasing that a hundred million readers would carry — and froze a fast-moving research field at one particular year.

PART 03

Steelman, level one: the strongest honest case

Here the argument is built at its best, using the best evidence available, with no cheap shots. A steelman means: if you are going to disagree, disagree with this version – not with a weaker one you invented.

The case in one paragraph

People judge their lives against a reference point, and the reference point moves. It moves with what you have recently had, with what you expected, and with what the people around you have. Because it moves, improvements in conditions are converted into new baselines rather than into lasting satisfaction. This is not a theory about weak-willed moderns; it is a well-documented feature of how perception works in every sensory and evaluative system we have. And it means that a civilisation can multiply its wealth many times over and arrive at a population reporting roughly what it reported before.

The mechanism is not mysterious

Adaptation is not a moral failing. It is what nervous systems do. Step out of a dark room into sunlight and you are dazzled, then you are not; the light has not changed. Every sense recalibrates to its recent input, because the useful information is in the *change*, not the level. There is no obvious reason evaluation of one's own life should work differently, and considerable evidence that it does not.

STEELMAN

If wellbeing is a change-detector rather than a level-detector, then permanent improvement in conditions is *expected* to produce temporary improvement in feeling. Harari is not making a cynical claim about human nature. He is applying a general principle about perception to the historical record, and the prediction falls straight out of it.

The comparison problem is real and underrated

The second leg is relative position. A large body of work finds that people's satisfaction depends not only on their own income but on the incomes of those around them – colleagues, neighbours, people like them. A raise that everyone receives is worth much less than the same raise received alone.

This has a brutal implication for the historical question. Growth that lifts everybody cannot deliver the satisfaction that individual advancement delivers, because the comparison group rises with you. A society can therefore become far richer and produce very little additional reported wellbeing without anyone behaving irrationally.

The community leg — Harari's best empirical card

The strongest and least contested part of the chapter is not about money at all. Across the wellbeing literature, the most robust correlates of a good life are not income: they are close relationships, marriage or an equivalent bond, community connection, health, and having work. Modernity has delivered enormously on income and health, and has plausibly eroded the first three.

The recent data are not comfortable. In the United States, roughly one adult in four now reports having eaten every meal of the previous day alone — an increase of over half since 2003. Americans under thirty rank around 62nd in the world for life satisfaction while Americans over sixty rank around 10th, in the same country, in the same year. The United States fell out of the top twenty of the world rankings for the first time in 2024 and slipped further in 2025.

STEELMAN

This is the version of Harari's argument that no critic has answered. The richest society in human history is producing a cohort of young adults with wellbeing scores comparable to far poorer countries, and it is doing so while its wealth continues to rise. Whatever else is true, the mapping from material improvement to human flourishing is broken somewhere, and it is broken in the place the chapter said it would be: not health, not money — the connections.

And the hardest evidence of all is not a survey

Between 1999 and the late 2010s, mortality among middle-aged Americans without a college degree rose — from suicide, drug overdose and alcohol-related liver disease. Anne Case and Angus Deaton named these deaths of despair. In 1999, death rates for white non-Hispanic Americans aged 50–54 with only a high-school education were about 30 per cent *below* those of Black Americans in the same age group; by 2015 they were about 30 per cent *above*.

This is not a mood questionnaire that might be culturally biased. These are death certificates, in the richest large country on earth, moving the wrong way for two decades while GDP per person rose.

LEVEL ONE, SUMMARISED

A reference point that moves, a comparison group that rises with you, a set of social goods that modernity has not delivered, and a body count that runs the wrong way inside the richest society in history. That is a serious case, and none of it depends on romanticising the past.

PART 04

Steelman, level two: the argument that survives every attack

Now the hard part. We hand every good objection to Harari's side and let them rebuild. What comes out is a stronger argument than the one in the book — and a more disturbing one, because it no longer depends on whether anybody is happier.

Everything is conceded

Give the critics all of it. Concede that reported happiness rose in 45 of 52 countries between 1981 and 2007. Concede that richer countries score higher than poorer ones, reliably, with no ceiling in sight. Concede that the money–wellbeing curve keeps climbing for most people well past any supposed threshold. Concede that adaptation is incomplete, that the set point moves, that the founding lottery study had twenty-two people in it, and that there is no data at all before 1946.

Concede, in short, that the empirical claim in the sentence is probably false.

Now watch what is left, because the surviving argument is the one worth being frightened of.

STEELMAN, LEVEL TWO — THE UNKEPT LEDGER

The interesting claim was never that we are no happier. It is this: **every institution that built the modern world optimises for something it can count, none of them counts this, and if the results have been good it was an accident.**

Not “progress failed”. Something narrower and worse: nobody was ever steering by the variable the whole project claims to be for — and a system that is not steering can only arrive somewhere good by luck.

Why this version is stronger

Because it is now indifferent to the answer. If it turns out we are much happier than our ancestors, the argument is untouched: it was not aimed at, not measured, and not defended, so it can be lost as easily as it was gained. A good outcome that nobody was tracking is not an achievement. It is a coincidence with no guarantee of continuing.

It also stops being a claim about happiness and becomes a claim about institutions. Which makes it checkable. It has three legs.

Leg one: what gets counted gets pursued

Gross domestic product was constructed in the 1930s and 40s as a wartime planning tool. Within about a decade it had stopped being a measurement and become an objective: governments fell over it, careers were made on it, and every ministry in the world learned to report in its terms. Nobody

decided that a nation's purpose was to raise GDP. It happened because GDP was the number that existed.

That is the general mechanism, and it does not require anybody to be cynical. Institutions are steering systems, steering systems need a signal, and the signal has to be countable, timely and hard to fake. Wellbeing is none of those things. So it is not the signal, so nothing steers by it, so any effect on it is a side effect.

SAY IT SIMPLE

A ship with an excellent speedometer and no compass will go very fast. Everyone aboard will be able to tell you exactly how fast. Nobody will be able to tell you whether it is going anywhere worth arriving at.

Leg two: the average is a hiding place

Even where wellbeing is now measured, the measurement that gets reported is a mean, and a mean is precisely the statistic that can rise while a substantial minority collapses. The 2023 resolution of the money-and-happiness argument found this inside the data: income keeps raising wellbeing for most people, and for the least happy portion of the population it stops helping above a threshold. Same dataset, two populations, and the headline number reports neither.

The deaths-of-despair data show the same shape at the extreme. National averages of income, health and life satisfaction can all look tolerable while mortality rises for a large group inside the same country. The instrument is not lying. It is answering a question about the middle.

Leg three: nobody was asked, and nobody could have been

The third leg connects this study to the first in the series. The transition into farming, into cities, into industry, into the attention economy — none of these was ever put to a vote, and none could have been, because the costs appear generations after the benefits and the exit closes before the bill arrives.

So the ledger is unkept not because anybody suppressed it but because there is no moment at which it could have been consulted. Every large trade humanity has made was made by people who could not see the terms, on behalf of descendants who were not born, and evaluated afterwards by historians who measured the harvest and not the mood.

STEELMAN

This is the passage that makes the chapter worth its errors. Harari's complaint is not that historians reached the wrong verdict. It is that there was never a trial. Four hundred pages of history in which the central variable is never once measured is not a bias in the historiography. It is the historiography.

The prediction this version makes

A rebuilt argument is worth having only if it can be checked. This one makes a specific, uncomfortable, testable claim:

Any society that begins seriously measuring wellbeing will discover its wellbeing and its trajectory pulling apart – and when they conflict, it will choose the trajectory, and then adjust the measurement.

That is checkable, and Part 07 checks it against the one state that adopted the ledger as constitutional policy. The result is not encouraging, and it is exactly what the prediction says.

LEVEL TWO, SUMMARISED

Drop “we are not much happier”. Keep this: the thing every civilisation claims to be for has never been on any civilisation’s dashboard; where it is now measured it is measured as an average, which is the one form that hides the people who matter most; and no generation has ever been in a position to consent to the trade it was making. That claim survives every objection in the next part – including the objections that demolish the sentence itself.

PART 05

Where it cracks — the archive, the flat line, the set point

Now the attack. Some of these are problems with the evidence. Some are holes in the reasoning. And some are things almost never mentioned in popular discussion of this chapter, because they are inconvenient for the critics as well as the fans.

5.1 The archive begins in 1946

This is the central crack, and it is not a technicality.

The sentence compares us with our ancestors. The oldest data of the kind it needs comes from Gallup surveys around 1946. International comparison becomes possible in the 1970s and 1980s. Genuinely global, annual, comparable data begins in 2005–2006 with the Gallup World Poll.

So the claim covers, at absolute maximum, eighty years, in a handful of rich countries, out of a story the book has told across seventy thousand. Everything before 1946 is not sparsely evidenced. It is unevidenced.

WHERE IT CRACKS

Harari states this limitation clearly inside the chapter — and then writes the sentence anyway, and the sentence is what travelled. A claim about all of history, resting on a measurement invented after the Second World War, is not a finding. It is an intuition with a footnote.

5.2 The flat line is not flat

The Easterlin paradox — countries get richer, average happiness does not rise — is the empirical spine of the chapter. It is genuinely contested, and the contest was already underway when the book was written.

- **Inglehart and colleagues, 2008.** Using World Values Survey data across 1981–2007, they report that happiness rose in **45 of the 52 countries** with substantial time-series data — and that the strongest predictor was how much free choice a society permitted.
- **Stevenson and Wolfers, 2008.** Using a broader set of countries and datasets, they find a clear positive relationship between wellbeing and income per person *across* countries, a similar relationship *within* countries, growth associated with rising wellbeing over time, and **no satiation point** above which extra income stops mattering.
- **Japan, the flagship case.** Real income per person multiplied several times over between 1958 and 1987 with no visible movement in reported satisfaction. It is the single most cited demonstration of the paradox — and the underlying Japanese survey changed its question wording and response categories several times over that period. When the series is repaired, the flat line becomes considerably less flat.

The paradox is not dead; Easterlin and colleagues have defended it repeatedly, and the disagreement turns on real technical questions about time horizons and which datasets to trust. But it is a live dispute between serious economists, not a settled result – and the chapter presents it as settled.

5.3 The founding study had twenty-two people in it

The lottery-winner finding is the most quoted result in the history of happiness research. Here is what it was.

Brickman, Coates and Janoff-Bulman, 1978: **22 major lottery winners**, 22 controls, and **29 people paralysed in accidents**. On a six-point scale, winners averaged 4.00, controls 3.82, and accident victims 2.96. The winners were not significantly happier than their neighbours. The paralysed group was clearly less happy – by more than a point – though the study is usually remembered as showing that they were fine.

WHERE IT CRACKS

Seventy-three people, no regression adjustment for large demographic differences between groups, and a known bias nobody controlled for: the typical lottery winner has spent a great deal on tickets first, which selects for people whose finances were already strained. Modern work using large German panel data finds that substantial lottery wins *do* raise life satisfaction durably. The most famous result in the field would not pass review today.

5.4 The set point moves

Set-point theory said that life events cause temporary deviations and people return to a genetically anchored baseline. Long-running panel studies – the same individuals surveyed annually for decades in Germany, Britain and Switzerland – have tested this properly, which the 1978 cross-section could not.

The result is a split verdict. Adaptation is fast and near-complete for some events, including marriage. It is clearly **incomplete** for others: unemployment, onset of disability, chronic illness and widowhood are associated with lasting changes in the baseline. People do not fully come back.

WHY THIS MATTERS MORE THAN IT LOOKS

The treadmill was the reason to predict a flat historical line. If the treadmill does not run in the domains where history made the largest difference – disability, disease, chronic pain, bereavement, unemployment – then the mechanism guaranteeing the flat line is gone, and the flat line has to be established by evidence rather than assumed. It has not been.

5.5 The biochemistry has aged badly

The chapter's account of happiness as a matter of serotonin, dopamine and oxytocin, with a fixed personal range, reflects the popular science of its moment. That account has taken a serious knock.

In 2022 an umbrella review led by Joanna Moncrieff in *Molecular Psychiatry* examined the main lines of evidence for the serotonin theory of depression and concluded there was no consistent support for it. The review was itself heavily criticised — for summarising rather than reanalysing, and for selective inclusion — and the argument continues. But the popular version Harari relies on, in which mood is a straightforward readout of neurotransmitter levels, is no longer a claim any careful researcher would make.

Two honest qualifications. First, the heritability finding is a separate matter and has broadly held: genes account for a substantial share of the variation in wellbeing. Second, “substantially heritable” does not mean “immovable” — height is highly heritable and rose by more than ten centimetres in many countries during the twentieth century because nutrition changed.

5.6 The threshold that was not there

A closely related pillar has also moved. In 2010 Kahneman and Deaton reported that emotional wellbeing improved with income only up to around \$75,000 a year. The figure entered general circulation immediately.

In 2021 Matthew Killingsworth, using real-time experience sampling, found wellbeing continuing to rise well past that point. Rather than trading papers, the two sides ran an **adversarial collaboration** with a neutral third researcher, published in 2023. The resolution: for most people, wellbeing keeps rising with income well beyond \$100,000, and for the happiest group it accelerates. The flattening is real only for roughly the least happy fifteen per cent, whose wellbeing stops improving above about \$100,000.

SAY IT SIMPLE

Money keeps helping most people. It stops helping those whose unhappiness has a cause money does not touch. Both halves are true, and the popular version of the chapter contains neither.

5.7 The instrument cannot see the largest change

This is the objection that survives even if every number above went the other way, and it is barely ever raised.

The dominant change between the pre-modern world and now is not more pleasure. It is fewer catastrophes: children who live, mothers who survive childbirth, teeth extracted under anaesthetic, famines that do not arrive, infections that do not kill. In 1800 more than a third of children died before their fifth birthday, and in some places closer to half; today it is under one in twenty-five globally, and a small fraction of that in rich countries.

A satisfaction survey measures a level. Catastrophe-avoidance does not appear as a level. It appears as the absence of the event that would have destroyed the level — and, brutally, as the presence in the sample of people who would otherwise not be there to answer.

WHERE IT CRACKS

If a society halves its child mortality and its average reported happiness does not move, the honest conclusion is not that the change did not matter. It is that a mean of self-ratings is the wrong instrument for detecting it. Harari's ledger has no column for a death that did not happen, and that is where most of the modern gain sits.

5.8 Two of the chapter's arguments cancel each other

The biochemical argument says wellbeing is a fixed chemical range, so nothing can improve it. The Buddhist argument says wellbeing is a trainable relationship to one's own experience, which is why chasing feelings fails. The second is a claim that a person can change something fundamental about their inner life. The first is a claim that they cannot.

Both are offered as reasons the ledger will not balance. They cannot both be the reason. The chapter never chooses, and the reader is left with the impression of converging evidence where there is in fact a contradiction.

5.9 The numbers may not be rankable at all

A technical objection with large consequences. Happiness answers are ordinal — “very happy” is more than “pretty happy”, but nobody knows by how much. Comparing two groups' averages requires assuming something about the spacing, and in 2019 Timothy Bond and Kevin Lang showed in the *Journal of Political Economy* that the conditions for making such comparisons without those assumptions are extremely strong and generally not met. Applying plausible alternative transformations, they were able to reverse standard findings in nine prominent areas of happiness research.

This does not mean the data are worthless. It means that “group A is happier than group B” is a much more fragile statement than the league tables suggest — and the sentence is precisely such a statement, made across centuries.

5.10 “Historians never asked” is overstated

The neglect charge is the best thing in the chapter, and it needs one correction. There is an established field called the history of emotions, launched by Lucien Febvre in 1941 and built out from the 1980s by Peter and Carol Stearns, William Reddy and Barbara Rosenwein. Historians have asked.

What they concluded is more interesting than the charge. That field largely holds that Harari's question cannot be answered in Harari's form — because the emotional categories themselves are not stable across time. What a medieval European meant by joy, what the word covered, when it was permitted and what it was contrasted with, does not map cleanly onto a modern ten-point ladder. On this view the problem is not missing data. It is that the quantity being compared is not the same quantity in both centuries.

5.11 The hidden side: what the framing licenses

This is the part usually skipped, because it is uncomfortable for admirers and critics alike.

“Progress did not make anyone happier” is available to anybody. It is as useful to someone arguing against spending on clean water as to someone arguing against consumerism. If material improvement is a treadmill, the case for the next improvement weakens — and the people for whom the next improvement is not a treadmill but a surviving child are exactly the people least likely to be reading the book.

Harari is not responsible for every use of his words, and his own answer is a good one: the point of the chapter is not to stop the work but to notice that nobody has been checking the results. That is right. It is also a much narrower claim than the sentence makes, which is why the rebuilt version in Part 04 is the one worth carrying.

IN ONE LINE

The crack is not that Harari asked a bad question. He asked the best question in the book. The crack is that he answered it — with a measurement that starts in 1946, a mechanism that has since been qualified, a founding study of seventy-three people, and a ledger with no column for the catastrophe that did not occur.

PART 06

What every field of study says about it

One sentence, eighteen departments. Each one has looked at this question with its own tools and reached its own verdict. Read together, they are more informative than any single field — and they do not agree.

6.1 Happiness economics

The field that owns the question, and the one most divided. Easterlin's paradox survives in the long-run time-series form and is contested almost everywhere else. Stevenson and Wolfers find a clear income gradient with no satiation; Easterlin and colleagues reply that short-run fluctuations are being mistaken for long-run trends. **Verdict: an open dispute, presented in the book as a result.**

6.2 Development economics

Angus Deaton — who has argued both that the escape from poverty and early death is the central fact of modern history and that rich societies are producing rising despair — is the best single witness. Both halves are his. **Verdict: the gains are enormous and the distribution of them is deteriorating.**

6.3 Psychology of subjective wellbeing

Self-reports are valid enough to be useful: they predict health, longevity and behaviour. They are also noisy, context-dependent and not designed for century-scale comparison. **Verdict: a good instrument used far outside its range.**

6.4 Behavioural economics

Kahneman's central contribution here is the split between the self that experiences a life and the self that remembers and rates it. They disagree systematically. Any sentence using one word for both is describing two different quantities at once. **Verdict: the question needs to be asked twice.**

6.5 Behaviour genetics

Twin studies put the heritable share of wellbeing at roughly 40 to 50 per cent, higher for its stable component. This supports a set point and does not support immovability — heritable traits can shift dramatically when environments do. **Verdict: half right, and the half everyone quotes is the wrong half.**

6.6 Psychiatry and neuroscience

The simple neurotransmitter account of mood is no longer defensible in the form the chapter uses. The 2022 serotonin umbrella review is contested on its methods, but nobody now argues that happiness is a readout of a chemical level. **Verdict: the mechanism has not survived the decade.**

6.7 Epidemiology and public health

Provides the hardest evidence in the whole debate, because death certificates do not have response bias. Rising mortality from suicide, overdose and alcohol among middle-aged Americans without degrees, from 1999 onwards, is a fact no survey critique can dissolve. **Verdict: strongly supports Harari, in one country, for one group.**

6.8 Demography

Child mortality fell from more than a third of children before age five in 1800 to under one in twenty-five globally today. This is the largest change in human experience ever measured, and a satisfaction average is structurally incapable of registering it. **Verdict: the ledger is missing its largest entry.**

6.9 Medicine and pain research

Chronic pain is one of the clearest cases where adaptation fails — people do not return to baseline. Anaesthesia, analgesia and dentistry therefore represent durable rather than temporary gains. **Verdict: contradicts the treadmill exactly where history delivered most.**

6.10 History of emotions

Founded by Febvre in 1941 and active since the 1980s. Its position is not that people were equally happy but that the categories do not transfer: what a word for joy covered, and when it was permitted, differs so much between eras that a single scale is not available. **Verdict: the comparison is not merely unmeasured but ill-formed.**

6.11 Anthropology

Contemporary foraging peoples are not living museums of the Palaeolithic; they occupy marginal land under modern pressure. Whatever they report cannot stand in for our ancestors, and the practice of treating them as a control group is exactly the error the discipline spent the twentieth century correcting. **Verdict: there is no proxy population.**

6.12 Sociology

The strongest support for Harari's community argument. Measured declines in social connection, associational membership and shared meals track the wellbeing deficits better than income does. **Verdict: right about what was traded away.**

6.13 Moral philosophy

Judging history by aggregate feeling is one ethical framework among several. The capability approach associated with Amartya Sen and Martha Nussbaum asks instead what people are actually able to do and be — under which the modern record looks very different, because a satisfied person with no options scores badly. **Verdict: the yardstick is a choice, and it is never defended.**

6.14 Statistics and measurement

Bond and Lang, 2019: because happiness answers are ordinal, comparing average happiness between two groups requires assumptions that are rarely met, and plausible alternative assumptions reverse many standard findings. **Verdict: comparisons of the exact kind the sentence makes are the least robust in the field.**

6.15 Political science and public policy

Governments have been keeping this ledger since the 2010s: national wellbeing statistics, the World Happiness Report from 2012, Bhutan's constitutional commitment from 2008. The neglect charge was true when Bentham made it and was already becoming false as the chapter was written. **Verdict: overtaken by events, and Part 07 shows what those events look like.**

6.16 Contemplative science

Meditation research supports modest benefits for some outcomes and has a well-documented problem with weak methods, small samples and publication bias — set out most bluntly in the 2018 review “Mind the Hype”. **Verdict: the Buddhist leg is promising and not yet load-bearing.**

6.17 Media and technology research

The claim that modern life is corroding young people's wellbeing is where the current fight is hottest. Some analyses find substantial effects; large-scale reanalyses find effect sizes small enough to be comparable to trivial variables. Meanwhile the ranking data show under-thirties in rich countries reporting sharply lower life satisfaction than their elders. **Verdict: the pattern is real; the cause is unresolved.**

6.18 Animal welfare and ecology

Harari's own extension, and the one place he applies the ledger consistently. If the accounting includes the more than eighty billion land animals slaughtered annually — the great majority of them chickens, most raised in industrial systems — the modern balance sheet gets worse by an enormous margin, whatever happened to human mood. **Verdict: the strongest use of his own instrument, and the one readers skip.**

. . . .

READING THE EIGHTEEN TOGETHER

The fields that measure feelings say the instrument is real but not built for this distance. The fields that measure bodies — demography, medicine, epidemiology — say the largest changes in human experience are invisible to a satisfaction average, and that where they can be seen they point in both directions at once: vastly less catastrophe, and a rising toll of despair inside the wealthiest societies.

And the two fields best placed to adjudicate the sentence as written — the history of emotions and the statistics of ordinal data — both return the same answer, which is not yes or no. It is that the comparison cannot be made in this form.

PART 07

The argument tested against six real bodies of data

A claim about how the world works has to survive the world. These are the six datasets that actually bear on the sentence. Two of them are the evidence it was built on. Three of them have gone the other way. The last one is the test of the rebuilt version.

7.1 Japan, 1958–1987: the flagship case

This is the exhibit everyone cites. Across roughly three decades of the fastest sustained growth any large economy had achieved, Japanese real income per person multiplied several times over. Reported life satisfaction sat at around six out of ten at the start and around six out of ten at the end.

If any single graph made the Easterlin paradox famous, it is this one.

WHERE IT COMPLICATES

The underlying survey — the Japanese government’s long-running national life survey — did not ask the same question throughout. Wording and response categories changed more than once, producing a series with breaks in it. Analysts who have attempted to repair the breaks report a considerably less flat picture. The most famous flat line in social science is partly an artefact of a questionnaire that was quietly revised.

7.2 Illinois, 1978: the seventy-three people

The other founding exhibit. Twenty-two people who had won major lottery prizes, twenty-two neighbours as controls, and twenty-nine people who had been paralysed in accidents.

The headline result — winners no happier than neighbours, victims not much less happy — became the single most repeated claim about human wellbeing of the last fifty years. What is repeated is not quite what was found: the paralysed group scored a full point lower than controls on a six-point scale, which is a large difference. And the sample is seventy-three people, unadjusted for the substantial demographic differences between the groups.

WHERE IT FAILS

There is also a bias built into the design that nobody controlled for: lottery winners are drawn overwhelmingly from heavy ticket buyers, who differ systematically from their neighbours before the win. Later work with large German panel datasets — thousands of people, tracked before and after — finds that sizeable wins do produce durable increases in life satisfaction. The founding study has been overturned by the method it lacked.

7.3 The World Values Survey, 1981–2007: the line that moved

The largest body of directly relevant evidence available when *Sapiens* was written. Representative national surveys across more than fifty countries, repeated over twenty-six years.

The finding, published by Ronald Inglehart and colleagues in 2008: **happiness rose in 45 of the 52 countries** for which substantial time-series data existed. The strongest single predictor of the rise was not income but the extent to which a society allowed people free choice over their own lives — which is a striking result, and one that fits the capability framework rather better than the hedonic one.

WHERE IT CONTRADICTS

This is the direct empirical answer to the sentence, it covers the period in which the sentence can be tested, and it was published three years before the book. Forty-five out of fifty-two is not a flat line. Whatever is true about the deep past, in the measurable era the number went up nearly everywhere.

7.4 The adversarial collaboration, 2010–2023

The most instructive episode in the field, and a model of how an argument should end.

In 2010 Kahneman and Deaton reported that emotional wellbeing stopped improving above about \$75,000 a year. In 2021 Killingsworth, sampling people's feelings in real time on their phones, found it kept improving well beyond that. Instead of a decade of duelling papers, the two sides ran a joint reanalysis with Barbara Mellers as neutral arbiter and published the result together in 2023.

The answer was that both were partly right and both had missed the shape of the distribution. For most people, wellbeing rises with income well past \$100,000 — and accelerates among the happiest. For roughly the least happy fifteen per cent, it flattens above about \$100,000: their unhappiness has a source money does not reach.

WHAT IT SHOWS

Two things. First, the popular claim that money stops mattering is false for most people. Second, and more important for Harari: an average was concealing two opposite realities, which is exactly the objection to running civilisational verdicts off a mean. Level two of the steelman is confirmed by the very study that refutes level one.

7.5 The United States, 1999–2025: the ledger going backwards

The strongest evidence for Harari, and it does not come from a survey.

From 1999, mortality among middle-aged Americans without a four-year degree began rising, driven by suicide, drug overdose and alcohol-related liver disease. In 1999 the death rate for white non-Hispanic Americans aged 50–54 with only a high-school education was about 30 per cent below the

rate for Black Americans of the same age; by 2015 it was about 30 per cent above. Case and Deaton called these deaths of despair, and the pattern appeared in every state.

The survey data point the same way for the young. In 2024 the United States dropped out of the top twenty of the World Happiness Report for the first time, and fell again in 2025. Americans under thirty rank around 62nd in the world for life satisfaction; Americans over sixty rank around 10th. About one adult in four reports having eaten every meal of the previous day alone, up by over half since 2003.

WHAT IT SHOWS

In the richest large society in history, over two and a half decades of rising output, a large group got sicker and started dying earlier, and the youngest cohort reports wellbeing comparable to countries a fraction as wealthy. No critique of survey methodology touches a death certificate. Harari's claim is wrong as a statement about humanity and disturbingly right as a statement about this.

7.6 Bhutan, 1972–2008: the state that kept the ledger

The test of the rebuilt argument in Part 04, which predicted that a society measuring wellbeing would eventually find it conflicting with its trajectory, choose the trajectory, and adjust the measurement.

Bhutan is the only country to have made this its constitutional purpose. Gross National Happiness was articulated from the 1970s, developed into a formal index across nine domains and thirty-three indicators, and written into the 2008 constitution, which directs the state to promote the conditions that enable its pursuit. It is measured properly, on a multi-year cycle, with real methodology.

It is also the period in which Bhutan pursued a “One Nation, One People” policy from 1989, under which roughly one hundred thousand Lhotshampa — the Nepali-speaking minority — were pressured and expelled to Nepal during the 1990s. Many spent close to two decades in refugee camps.

WHY THIS IS THE HARD CASE

A national happiness average can be raised by removing the people who bring it down. This is not a hypothetical objection to utilitarian measurement; it is what the arithmetic permits, in the one country that adopted the instrument as policy, during the years it was being built. The prediction in Part 04 said the trajectory would win and the measurement would be adjusted around it. That is what the record shows.

The fair addition is that Bhutan's framework is genuinely serious — it counts ecology, culture, time use and governance, which no GDP figure does, and it has been influential precisely because it is more honest than the alternative. Both things are true, and this is exactly why the level-two argument is about institutions and not about mood.

• • •

THE SIX, SCORED

Japan: the flagship line is partly an artefact. Illinois 1978: overturned by better data. The World Values Survey: forty-five of fifty-two went up, and it was in print before the book. The 2023 collaboration: refutes the sentence and confirms the rebuilt version. The United States: the strongest case Harari has, and it is made of death certificates rather than questionnaires. Bhutan: the ledger was kept, and the prediction held.

PART 08

The ledger you are keeping right now

If the mechanism is real, it did not stop being real when the chapter ended. This part is a practical tool: what the evidence actually supports about which conditions move wellbeing durably, which do not, and which the ledger cannot see at all.

The three columns

Harari has one column. The panel studies have three, and the difference between them is the most useful thing to come out of fifty years of this research.

COLUMN	WHAT IS IN IT	WHAT THE EVIDENCE SAYS
Adapts fast	A pay rise, a purchase, a promotion, a move, most single pieces of good news.	Real effect, then substantial return towards baseline. This is the treadmill, and here it runs.
Adapts slowly or not at all	Unemployment, onset of disability, chronic pain and illness, bereavement, sustained isolation.	Lasting shifts in the baseline itself. People do not fully come back. The treadmill does not run here.
Invisible to the ledger	The illness that did not happen, the child who did not die, the famine that did not arrive.	The largest historical changes of all, and a satisfaction score cannot register any of them.

SAY IT SIMPLE

The evidence for removing *bad conditions* is far stronger than the evidence for *adding good ones*. That is not a counsel of despair. It is a direction — and it happens to be the direction in which most of the last two hundred years actually went.

Two traps, both common after reading this chapter

The cynic’s trap. If it is all a treadmill, nothing is worth pursuing. This is the most frequent takeaway and it does not follow from anything in the research. Adaptation is partial, uneven and absent in exactly the domains where suffering is concentrated. “Some gains fade” is not “no gains exist”.

The blame trap. The opposite error, and the more damaging one. If happiness is a fixed genetic set point, then a person’s dissatisfaction becomes a fact about them rather than about their situation — which quietly converts a hard circumstance into a personal failing. The panel data say the reverse. Unemployment, chronic pain, illness and isolation move the baseline durably. Conditions matter, measurably, and someone struggling under bad ones is not misreading their own thermostat.

Reading the sentence in your own life without either trap

Three things the evidence supports well enough to be worth knowing, stated without any promises attached:

1. **The comparison group is doing more work than the amount.** Satisfaction tracks position relative to expectation and to the people around you. This is why a society can quadruple its income and report the same numbers — and why the same is true of a person.
2. **The durable entries are mostly relational and physical.** Close relationships, being in work, health, freedom from chronic pain, and not being alone show up across the literature far more reliably than income does above a modest threshold.
3. **An average is not a person.** Every result in this document is a population statistic. For roughly the least happy sixth of people, the aggregate patterns do not describe the situation at all — which is the finding of the 2023 collaboration, and the reason population averages should never be used to tell an individual what their life is like.

Four questions, for any ledger you are inside

1. What is the number this system actually optimises — my job, my country, my week? Say it out loud.
2. What is the thing that number was supposed to be a proxy *for*, and when did anyone last check the two were still connected?
3. If the number is an average, who is inside it that it is hiding?
4. Which column is the change I am considering in — fast-adapting, slow-adapting, or invisible?

The second question is the real one. Every measure begins as a proxy for something that matters and ends as the thing that matters. Harari's chapter is not really about happiness. It is about the interval between those two states, and how long a civilisation can run inside it without noticing.

THE VERSION WORTH KEEPING

You cannot escape being measured by something. You can find out what it is, check whether it still tracks the thing it was standing in for, and notice who the average is hiding. That is small, unglamorous, and available today — which is more than can be said for the answer to Harari's question.

PART 09

The scorecard, and the sentence rewritten

Thirteen axioms went in. Here is what came out, with each one marked honestly, including the ones that did not survive.

#	AXIOM	VERDICT	WHY
1	Happiness is one thing	Fails	Evaluative, experienced and eudaimonic wellbeing move differently and sometimes oppositely. One word cannot carry three quantities.
2	Happiness can be measured	Holds	Self-reports predict health, behaviour and longevity. Crude, noisy, and a real measurement.
3	Asking people works	Holds, with a caveat	Known distortions wash out in large samples. They do not wash out across centuries, where they cannot be observed.
4	Scores are comparable across cultures	Workable, barely	Response styles differ enough to move a country several places. Rankings are bands, not positions.
5	There is evidence about our ancestors	Fails	The record begins in 1946. The one serious backward extension reaches 1820, by proxy, in four countries.
6	The average is the right unit	Fails	The 2023 collaboration found one mean hiding two opposite realities. Deaths of despair show the same at the extreme.
7	Power and happiness are the right pair	Wrong framing	The biggest measurable change is not power gained but catastrophe removed, and the sentence has no column for it.
8	Happiness is the correct yardstick	Values, not facts	One ethical framework among several. Freedom, capability and justice give different verdicts, and are never mentioned.
9	There is a biological set point	Half holds	Heritability around 40–50 per cent is well supported. Immovability is not, and heritable does not mean fixed.

#	AXIOM	VERDICT	WHY
10	Adaptation always returns you to it	Fails	Panel studies find adaptation incomplete for unemployment, disability, chronic pain and widowhood — the domains history changed most.
11	Expectations cancel the gains exactly	Overstated	Reference points are real; complete cancellation is not. The income gradient across countries has no visible ceiling.
12	Suffering avoided belongs in the same column as pleasure gained	Structural failure	A satisfaction level cannot record a death that did not occur. The instrument is blind to most of the modern gain.
13	The claim could be contradicted	Unfalsifiable	“Not much” is never quantified, so no result is inconsistent with it. That is a mood, not a hypothesis.

. . .

The sentence, rewritten honestly

Harari wrote: “We are far more powerful than our ancestors, but we are not much happier.” After all of the above, here is the version that survives every objection in this document:

We are far more powerful than our ancestors, and whether we are happier is not known — the measurement is eighty years old, the categories may not survive the journey, and the largest change of all is a subtraction that no satisfaction score can record. In the period we can measure, the number has mostly gone up, and inside the richest societies a large minority is going the other way fast enough to show in the mortality figures. What can be said without qualification is narrower and worse: in seventy thousand years, nothing was ever steered by the answer.

That version is less quotable. It is also true, and it does the same work: it refuses the comfortable conclusion without replacing it with a bleaker one that the evidence does not support.

What Harari got right, stated plainly

- The question is the right question, and it had been almost entirely absent from the way history is written and taught.
- Reference points move, and a great deal of material improvement is converted into new expectations rather than lasting satisfaction.

- The social goods that predict wellbeing best are the ones modernity has delivered least, and possibly eroded.
- Aggregate progress can coexist with a group inside it that is unambiguously worse off, and usually does.
- Any honest accounting has to decide whether non-human suffering is in the ledger — and if it is, the modern balance is much worse than anyone admits.

Five true things, in a book read by a hundred million people. That is a real public service, and the errors in the sentence do not cancel it.

The last thing worth saying

People argue about this chapter as though the question were whether the past was better. It was not better. On every measure that leaves a physical trace — how long people lived, how many of their children lived, what they ate, what they could read, what was done to them without their consent — the past was worse, and often unimaginably so.

The question the chapter actually poses is about **instruments**. Every large human system runs on a number. The number always begins as a stand-in for something that matters and cannot itself be counted. And the number always, eventually, becomes the thing that matters, because it is the thing that can be reported, compared, targeted and rewarded. Harvest yields. Gross domestic product. Engagement. Test scores. Quarterly growth.

Harari's chapter is the moment the book turns around and asks what happened to the original quantity — and finds that nobody wrote it down, because nobody could. That is not a claim about happiness. It is a claim about the permanent gap between what a civilisation measures and what it is for, and it is the most durable thing in the book precisely because no amount of better data will close it.

THE POINT OF THE WHOLE THING

A score nobody keeps is a score nobody can be answerable for. Harari's chapter is worth its errors because it names the missing column. Keep that, and drop the verdict — because “we are not much happier” is an answer, and the honest position is that the question was never asked in time to have one.

THE SHELF, COMPLETED

Three sentences, one book, one shape. **The Wheat Bargain**: a trap you cannot see, because its costs arrive after its exit closes. **The Imagined Order**: a structure you cannot name, because naming it is what it is built to prevent. **The Happiness Ledger**: a score nobody keeps, because the things that can be counted are never the things that were wanted. Harari is wrong in the details of all three and right that all three are there.

APPENDIX

Glossary, timeline and sources

Every term used in this document, explained in one line. Then the dates, then where all of it comes from.

Glossary

TERM	PLAIN MEANING
Axiom	Something you must assume before an argument can start. The floor under the argument.
Steelman	The strongest possible version of an argument you may disagree with. The opposite of a strawman.
Subjective wellbeing	The umbrella term for how good a person reports their own life or mood to be.
Evaluative wellbeing	Your considered rating of your life as a whole, usually on a 0–10 ladder.
Experienced wellbeing	How you actually feel moment to moment, sampled during the day rather than recalled.
Eudaimonic wellbeing	Whether your life feels worthwhile and meaningful. Not the same as feeling good.
Cantril ladder	The standard question: imagine a ladder where 10 is the best possible life for you and 0 the worst. Where do you stand?
Hedonic adaptation	The tendency to return towards a baseline level of feeling after a change in circumstances.
Hedonic treadmill	The image for the above: you keep improving your conditions and end up in the same place.
Set point	The baseline a person is said to return to. Partly heritable; less fixed than once thought.
Easterlin paradox	Richer individuals are happier within a country, but a country getting richer does not raise its average happiness over the long run.
Reference point	The standard you compare your situation to — what you expected, what you had, what others have.
Relative income effect	The finding that your satisfaction depends on your income compared with your comparison group, not only on the amount.

TERM	PLAIN MEANING
Panel study	A survey that follows the same individuals for years, so change within a person can be measured.
Cross-section	A snapshot of different people at one moment. Cannot show change within a person.
Adversarial collaboration	Researchers who disagree running a joint analysis, usually with a neutral arbiter, and publishing the result together.
Umbrella review	A review of existing reviews, used to summarise a whole literature at once.
Ordinal data	Answers with an order but no known spacing – “very happy” beats “fairly happy”, but nobody knows by how much.
Deaths of despair	Case and Deaton’s term for deaths from suicide, drug overdose and alcohol-related liver disease.
Capability approach	Sen and Nussbaum’s alternative yardstick: judge a society by what people are actually able to do and be.
Utilitarianism	The ethical view that the right outcome is the one producing the most wellbeing overall.
Gross National Happiness	Bhutan’s official index across nine domains and thirty-three indicators, constitutionalised in 2008.
Lhotshampa	Bhutan’s Nepali-speaking minority, roughly a hundred thousand of whom were expelled during the 1990s.
Vipassana	The meditation practice, in the S. N. Goenka tradition, that Harari has practised since 2000.
Emotionology	The Stearns’ term for a society’s rules about which emotions may be felt and shown, and how.
Survivorship bias	Drawing conclusions from those who came through, while those who did not are absent from the data.
Goodhart’s problem	When a measure becomes a target, it stops being a good measure.

Timeline

DATE	EVENT
1789	Bentham proposes that the greatest happiness be the standard, and that it be calculated.
1941	Lucien Febvre calls for a history of emotions, founding the field.

DATE	EVENT
1946	The earliest surveys asking people directly how happy they are. Everything before this is outside the record.
1971	Brickman and Campbell describe hedonic adaptation and name the treadmill.
1972	Bhutan begins articulating Gross National Happiness.
1974	Easterlin publishes the paradox, drawing on surveys from 1946 to 1970 across nineteen countries.
1978	Brickman, Coates and Janoff-Bulman: 22 lottery winners, 22 controls, 29 paralysed accident victims.
1981	The World Values Survey begins its first wave.
1985	Peter and Carol Stearns introduce “emotionology”.
1996	Lykken and Tellegen: 44–52 per cent of wellbeing variance heritable; the line about trying to be taller.
2001–2006	Reddy’s <i>The Navigation of Feeling</i> and Rosenwein’s <i>Emotional Communities</i> establish the modern history of emotions.
2005	Layard’s <i>Happiness</i> argues governments should target wellbeing directly. The Gallup World Poll begins.
2008	Inglehart et al.: happiness rose in 45 of 52 countries, 1981–2007. Stevenson and Wolfers reassess the Easterlin paradox and find no satiation. Bhutan constitutionalises Gross National Happiness.
2009	The Stiglitz–Sen–Fitoussi commission reports that GDP is an inadequate measure of national performance.
2010	Kahneman and Deaton report emotional wellbeing plateauing around \$75,000.
2011	<i>Sapiens</i> published in Hebrew. The UN passes a resolution on happiness. Britain’s ONS begins national wellbeing measurement.
2012	The first World Happiness Report.
2014	<i>Sapiens</i> published in English.
2015	Case and Deaton publish the first deaths-of-despair findings. <i>Homo Deus</i> published in Hebrew.
2019	Bond and Lang show that many happiness comparisons reverse under plausible rescaling. Hills et al. extend a wellbeing proxy back to 1820 using digitised books.
2021	Killingsworth finds wellbeing rising well past \$75,000.

DATE	EVENT
2022	The Moncrieff umbrella review on serotonin and depression, and the disputes that followed.
2023	Killingsworth, Kahneman and Mellers publish the adversarial collaboration resolving the income question.
2024–2025	The United States falls out of the top twenty of the World Happiness Report, then falls further, driven by the under-thirties.

Sources

THE BOOK UNDER STUDY

- Yuval Noah Harari, *Sapiens: A Brief History of Humankind* (Hebrew 2011; English, Harvill Secker, 2014), Chapter 19, “And They Lived Happily Ever After”.
- Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (Hebrew 2015; English 2016), where the biochemical account becomes the basis for a project rather than a diagnosis.

THE FOUNDING LITERATURE

- Richard A. Easterlin, “Does Economic Growth Improve the Human Lot? Some Empirical Evidence” (1974).
- Philip Brickman and Donald Campbell, “Hedonic Relativism and Planning the Good Society” (1971).
- Philip Brickman, Dan Coates and Ronnie Janoff-Bulman, “Lottery Winners and Accident Victims: Is Happiness Relative?”, *Journal of Personality and Social Psychology* 36 (1978), pp. 917–927.
- David Lykken and Auke Tellegen, “Happiness Is a Stochastic Phenomenon”, *Psychological Science* 7 (1996), pp. 186–189.

THE CORRECTIONS

- Ronald Inglehart, Roberto Foa, Christopher Peterson and Christian Welzel, “Development, Freedom, and Rising Happiness: A Global Perspective (1981–2007)”, *Perspectives on Psychological Science* 3(4) (2008), pp. 264–285.
- Betsey Stevenson and Justin Wolfers, “Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox”, *Brookings Papers on Economic Activity* (Spring 2008).
- Richard E. Lucas, “Adaptation and the Set-Point Model of Subjective Well-Being”, *Current Directions in Psychological Science* 16(2) (2007), and the associated German, British and Swiss panel analyses.
- Daniel Kahneman and Angus Deaton, “High Income Improves Evaluation of Life but Not Emotional Well-Being”, *PNAS* (2010); Matthew Killingsworth, “Experienced Well-Being Rises with Income, Even above \$75,000 per Year”, *PNAS* (2021); Matthew Killingsworth, Daniel Kahneman and Barbara Mellers, “Income and Emotional Well-Being: A Conflict Resolved”, *PNAS* 120(10) (2023).
- Timothy N. Bond and Kevin Lang, “The Sad Truth about Happiness Scales”, *Journal of Political Economy* 127(4) (2019), pp. 1629–1640.
- Joanna Moncrieff et al., “The Serotonin Theory of Depression: A Systematic Umbrella Review of the Evidence”, *Molecular Psychiatry* (2022), together with the published critiques of its method.

THE RECORD, AND THE ATTEMPTS TO EXTEND IT

- Thomas T. Hills, Eugenio Proto, Daniel Sgroi and Chanuki Illushka Seresinhe, “Historical Analysis of National Subjective Wellbeing Using Millions of Digitized Books”, *Nature Human Behaviour* 3 (2019), pp. 1271–1275.
- The Gallup World Poll (2005–) and the *World Happiness Report* (2012–), including the 2024 and 2025 editions on the United States and the under-thirties.

- Lucien Febvre (1941); Peter and Carol Stearns on emotionology (1985); William M. Reddy, *The Navigation of Feeling* (2001); Barbara Rosenwein, *Emotional Communities in the Early Middle Ages* (2006).

THE HARD CASES

- Anne Case and Angus Deaton, “Mortality and Morbidity in the 21st Century”, *Brookings Papers on Economic Activity* (2017), and *Deaths of Despair and the Future of Capitalism* (Princeton, 2020).
- Angus Deaton, *The Great Escape: Health, Wealth, and the Origins of Inequality* (Princeton, 2013).
- Bhutan’s Centre for Bhutan Studies on the GNH Index and its nine domains; Article 9 of the Constitution of Bhutan (2008); and the documented expulsion of roughly 100,000 Lhotshampa during the 1990s.
- Amartya Sen, *Development as Freedom* (1999); Martha Nussbaum, *Creating Capabilities* (2011).

The Living Archive — Single-Sentence Study No. 03, completing the series *Famous books, taken apart: Sapiens*.

Written and compiled by Lovepreet Singh · misterlove.in/books/sapiens/ · August 2026.

Method: one sentence; every axiom under it named and graded; the argument steelmanned twice before it is attacked; each field of study allowed its own verdict; the whole thing tested against data that actually exist. Where the author is right, the scorecard says so.

Quotations from *Sapiens* are short and used for criticism and review. All errors are the compiler’s.