FIELD MANUAL № 001

REALITY

Who writes it, and how to check it.

You live it. Memory edits it. You publish a version. A feed ranks it. A model summarises it. Someone else calls it the truth.

SAHIL TAYAL · A SAHILTAYAL.COM PRODUCTION · RESEARCH MEETS REALITY

Companion to sahiltayal.com/reality · Revised 18 November 2026 · Text CC BY 4.0

§00 · HOW TO USE THIS

A map, a catalogue and a kit.

The event happened one way. Every claim about it reaches you through a chain of authors. Each link has characteristic ways of failing. Those failures can be learned, spotted and partly reversed. This manual is **not** an argument that truth is relative, and it is not neutral: it is a self-published, interpretive essay by Sahil Tayal, with research claims linked to numbered sources (full list at the back).

Read it in this order if you're short of time: the Ten Questions (p. ~13), the Checking Kit, then the Stack.

THE FOUR BIG IDEAS

Reality is written in layers. Event, perception, memory, narrative, publication, amplification, synthesis, record.

Feeling true is manufactured. Thirteen mechanisms make claims feel credible independent of evidence.

Count origins, not mentions. Many sources that copy one source are one source.

Label the register. Fact, measurement, inference, interpretation, prediction, aspiration, value, performative, fiction.

§01 · THE REALITY STACK

Eight authors between an event and your belief.

01 EVENT — Nobody yet — only what happened

Something happens. No one owns that moment. Everything you will ever read about it is a copy made by someone who stood somewhere, noticed some of it, and wrote some of that down.

The referent: the territory, not the map[2]. Epistemically, layers 02–08 are access routes with characteristic error profiles; the event constrains them but does not speak for itself.

WHAT GETS LOST

Nothing is lost here, but nothing is accessible either. Every later layer is a route to the event, never the event itself.

THE CHECK

Ask: who had direct access to this — and who is only reporting someone else's access?

02 PERCEPTION — Your senses, attention and expectations

You don't record the world; you sample it. Attention is a torch, not a floodlight, and expectation decides where you point it.

Perception is selective and predictive; inattentional blindness shows salient events can go unencoded when attention is task-bound[5]. Observer vantage and goals are part of the data.

WHAT GETS LOST

Everything outside the beam of attention. People counting basketball passes regularly missed a gorilla walking through the scene[5].

THE CHECK

Ask: what was the observer looking for? Where else could someone have stood?

03 MEMORY — Reconstruction, every time you recall

Memory is not a video. Each time you remember, you rebuild the scene from fragments and fill gaps with what fits — and then that rebuild becomes the memory.

Recall is reconstructive[6]; post-event information is incorporated (misinformation effect)[7], and whole events can be suggested into memory[8].

WHAT GETS LOST

Detail, sequence and source. One verb in a question ('smashed' vs 'hit') changed speed estimates and false reports of broken glass[7].

THE CHECK

Prefer the earliest record over later recollection. Ask whether the question that produced the memory was leading.

04 NARRATIVE — The interpreter in your head

Your mind hates loose ends. It turns a sequence into a story with reasons — even when it doesn't know the reasons — and the story feels exactly like knowledge.

Explanation is generated post hoc[11]; introspective reports of causes are often inaccurate[10]; narrative is a mode of constructing reality, not merely reporting it[12].

WHAT GETS LOST

Randomness, ambiguity and the alternatives. Causes get inserted; outcomes start to look inevitable[14].

THE CHECK

Separate 'this happened, then that' from 'this caused that'. Write two other stories that fit the same facts.

05 PUBLICATION — Whoever selects, frames and posts

Publishing is choosing. What goes in, what stays out, which word, which photo, which order. The same facts can be framed to push opposite decisions.

Self-presentation is front-stage performance[15]; equivalent descriptions alter choice[16]. Omission is the least visible and most powerful edit.

WHAT GETS LOST

Everything left out, and every equivalent framing that was not chosen.

THE CHECK

Ask: what would the author have had to leave out to make it read this way? Reframe it and see if your reaction changes.

06 AMPLIFICATION — Rankers: search engines, feeds, editors, algorithms

Most of what you see was chosen for you by a ranking. The top result feels like the answer; the viral post feels like the consensus. Neither is a verdict.

Attention is the scarce resource[17]. Novel, emotional, moralised content diffuses further[18][19]; users trust rank order even when it is manipulated[20].

WHAT GETS LOST

Proportion. What spreads is what engages, and position is read as endorsement[20].

THE CHECK

Treat ranking as logistics, not judgement. Read laterally: leave the page and find out who is behind it[21].

07 SYNTHESIS — Summarisers: AI models, encyclopedias, explainers

A summary sounds neutral because it has no visible author. But it inherits every choice upstream — and an AI summary can also invent specifics that were never there.

Language models produce fluent text without grounding guarantees[34]; hallucination is a documented class of error[33]. Circular sourcing ('citogenesis') can make one claim look multiply confirmed[32].

WHAT GETS LOST

Attribution and uncertainty. Fluency is added; hedges and sources are stripped.

THE CHECK

Demand the citation, open it, and confirm it actually says that. A link is not a confirmation.

08 RECORD — Institutions: registries, courts, universities, archives

Some writing doesn't describe reality — it creates it. A degree, a contract, a verdict, a marriage certificate. Errors here are sticky because the paper does the work.

Institutional facts exist by collective acceptance ('X counts as Y in C')[38]; performatives enact what they state[39]; defined situations are real in their consequences[40].

WHAT GETS LOST

Reversibility. Once written into a record, a claim becomes real in its consequences, even if wrong[40].

THE CHECK

Ask: who can amend this record, and how? A record without a correction path is power without accountability.

§02 · SIX MUTATIONS

Nobody lies. The claim still changes.

Each hop between authors is a small, defensible edit. The distortion lives in the sequence. Safeguards must travel with the claim.

SOURCE LOSS A named, small source becomes a generic, bigger-sounding one.

CERTAINTY INFLATION 'Found tentative evidence' becomes 'finds', then 'proves'.

PARAPHRASE DRIFT Each retelling compresses and slightly upgrades the claim.

SCOPE CREEP '40 students, one college' disappears or becomes 'across populations'.

CAVEAT STRIPPING Hedges are cut first because they look like padding.

TIMELESSNESS The date falls off; a past finding becomes a permanent truth.

WORKED EXAMPLE (FICTIONAL SEED)

1 · ORIGINAL PAPER

A pilot study by two lecturers found tentative evidence that students who wrote for ten minutes before an exam scored slightly higher on average (40 students, one college, six weeks); the authors call it preliminary and unreplicated (March 2024).

2 · PRESS RELEASE

A new study found that writing before exams improves scores among students (2024).

3 · HEADLINE

Study finds writing before exams boosts scores.

4 · SOCIAL POST

Scientists prove journaling makes you smarter.

5 · AI ANSWER

Research shows that expressive writing significantly improves academic performance.

6 · CITATION IN A REPORT

Multiple sources confirm that expressive writing improves academic performance across populations.

§03 · THE AUTHORITY ENGINE

Thirteen ways a claim borrows credibility.

01 Repetition IN THE MIND

Hearing something again makes it feel more true — even when you already knew it was false.

Illusory truth effect: repeated statements receive higher truth ratings[23]; prior knowledge does not fully protect[24].

COUNTER-MOVE

Ask: have I seen evidence, or have I only seen it often?

02 Fluency IN THE MESSAGE

Easy-to-read claims feel truer. Clean fonts, short words and rhymes all borrow credibility they did not earn.

Processing fluency is misattributed to validity[25].

COUNTER-MOVE

Rewrite the claim clumsily in your own words. Does it still hold?

03 Source amnesia IN THE MIND

You remember the claim but forget where you got it — so a rumour and a textbook end up feeling equally solid.

Source-monitoring errors: content is retained while its origin is lost or misattributed[26].

COUNTER-MOVE

Before repeating anything, name where you learned it. If you can't, say so.

04 Coherence IN THE MIND

A story that hangs together feels complete, even if half the facts are missing.

WYSIATI: confidence tracks the coherence of the available story, not the quantity or quality of evidence[27].

COUNTER-MOVE

List what you would need to know that you don't. Missing facts don't announce themselves.

05 Confidence IN THE MESSAGE

We trust people who sound sure. Certainty is a performance, and it works.

Confidence heuristic: judges prefer more confident advisers[28]; confident tone substitutes for track record.

COUNTER-MOVE

Ask for the track record, not the tone. How often has this source been right before?

06 Precision IN THE MESSAGE

'73.4%' sounds more researched than 'about three-quarters', whether or not anyone measured anything.

Precise values act as stronger anchors and imply measurement[29].

COUNTER-MOVE

Ask for the denominator, the method and the date. Precision without a method is decoration.

07 Social proof IN THE NETWORK

If everyone seems to believe it, it seems believable. Agreement looks like evidence.

Conformity to unanimous groups[30]; social proof as a decision shortcut[31].

COUNTER-MOVE

Count independent reasons, not people. A thousand shares of one claim is still one claim.

08 Rank position IN THE MACHINE

The first search result feels like the answer, because it is first.

Users favour top-ranked results even when rank order is manipulated[20].

COUNTER-MOVE

Scan the whole results page before clicking (click restraint)[21].

09 Citation loops IN THE NETWORK

Site A says it, site B quotes A, site C quotes B, and now it is 'widely reported'. One origin, many echoes.

Circular reporting / citogenesis: later sources are cited to confirm the claim they copied[32].

COUNTER-MOVE

Trace every citation back to the first appearance. Count origins, not links.

10 Fluent synthesis IN THE MACHINE

AI answers sound calm, balanced and complete. That tone is produced the same way whether the content is right or wrong.

Generated fluency is decoupled from grounding[34]; hallucinated specifics are a known failure class[33].

COUNTER-MOVE

Treat an AI answer as a lead, not a source. Verify each specific (name, number, quote, date).

11 Well-knownness IN THE NETWORK

Being famous becomes evidence of deserving fame. Visibility gets mistaken for merit.

The pseudo-event and the celebrity known for well-knownness[36]; cumulative advantage[42].

COUNTER-MOVE

Separate 'often mentioned' from 'often right'.

12 Emotion & novelty IN THE NETWORK

Surprising, outrageous, moralised claims travel fastest — and speed feels like importance.

False news spread farther and faster than true news[18]; moral-emotional language increases diffusion[19].

COUNTER-MOVE

When a claim makes you feel strongly, slow down. Strong feeling is a reason to check, not to share.

13 Wanting it to be true IN THE MIND

We search harder for evidence we like and inspect evidence we dislike. The verdict comes first; the reasons come later.

Motivated reasoning biases search and evaluation toward desired conclusions[37]; but much susceptibility is plain inattention rather than bias[50].

COUNTER-MOVE

Ask: what would I believe if this evidence pointed the other way?

THE INDEPENDENCE RULE

Two sources are independent only if each had its own access to the facts — its own observation, document or measurement. A source that read another source adds reach, not evidence. Seven mentions that trace back to one self-description are one source[32].

§04 · NINE REGISTERS

Writing a story is power. Labelling it is honesty.

REGISTER	TEST	WRITE IT AS
FACT Something that happened or is the case, supported by evidence others can inspect.	Could someone else check this without trusting me?	State it plainly, with the evidence or where to find it.
MEASUREMENT A number produced by a method.	Who counted, how, out of how many, and when?	Give the method, the denominator and the date.
INFERENCE A conclusion reasoned from evidence.	What else could explain the same evidence?	Show the step from evidence to conclusion, and the alternatives.
INTERPRETATION A meaningful reading others may reasonably contest.	Could an informed person read it differently?	Mark it as your reading: 'I read this as…'.
PREDICTION A claim about the future.	How confident, by when, and how will we know?	Give a probability and a resolution date.
ASPIRATION A desired future, not a present accomplishment.	Does this exist yet?	Use future tense: 'I want to…', 'I am working toward…'.
VALUE A judgement about what matters or what is good.	Is this a disagreement about facts or about priorities?	Own it: 'I believe…', 'I value…'.
PERFORMATIVE Words that make something true by being said in the right context.	Does the speaker have standing to make this so?	Name the authority and the context ('By the power vested…').
FICTION An imagined world, openly presented as one.	Is the reader told it is invented?	Label it, every time it travels.

PRACTICE — COVER THE ANSWERS

1. “The Earth orbits the Sun once roughly every 365.25 days.”

FACT / MEASUREMENT — Inspectable and measured. Both labels are defensible; the precision comes from a method.

2. “38% of respondents said they read the news daily.”

MEASUREMENT — A number produced by a survey. Useless without who was asked, how many, and when.

3. “Sales fell after the price rise, so the price rise caused the fall.”

INFERENCE — A causal inference from sequence. Seasonality, competitors or supply could explain the same pattern.

4. “This film is really about grief, not war.”

INTERPRETATION — A reading. Defensible, contestable, not checkable like a fact.

5. “There is a 70% chance it rains tomorrow.”

PREDICTION — A probabilistic forecast. Judge it over many forecasts, not one day.

6. “I am building a school that will change education.”

ASPIRATION — A wanted future stated in present tense. Honest version: 'I want to build…'.

7. “Rest is as important as work.”

VALUE — A priority claim. Evidence can inform it but cannot settle it alone.

8. “I hereby resign from the committee.”

PERFORMATIVE — Saying it, with standing, does it[39].

9. “In the city of Tidemark, the sea returns every lost letter.”

FICTION — An invented world. Honest as long as the label travels with it.

10. “Experts agree this is the best diet.”

INTERPRETATION / VALUE — Dressed as a fact, but 'experts' is unnamed and 'best' is a judgement. The register is opinion wearing a lab coat.

11. “She was the first woman to graduate from the programme.”

FACT — Checkable against institutional records — the kind of claim that should be checked before it is repeated.

12. “This website gets millions of visitors.”

MEASUREMENT — A measurement claim with no method, period or source. Until those appear, treat it as unverified.

§05 · WORDS DOING THE PERSUADING

Patterns that deserve a question.

A lexical checklist, not a fact-check. Each pattern tells you where to look.

Unnamed authority — *studies show, research shows, research suggests, experts say, experts agree, scientists say, sources say, it is known, many people…*
Which studies, which experts?

Certainty boosters — *prove, proves, proven, definitely, clearly, undeniably, undeniable, always, never…*
What evidence justifies this much certainty?

Status words — *best, leading, world-class, renowned, top, first, only, revolutionary, groundbreaking…*
According to whom, measured how?

Causal verbs — *causes, caused, leads to, led to, because of, boosts, makes you, results in, due to…*
Causation shown, or sequence and correlation?

Future / aspiration — *will, aims to, plans to, wants to, going to, intends to, hopes to, soon…*
Prediction with a probability, or aspiration posing as fact?

Hedges (usually good) — *may, might, could, suggests, preliminary, estimated, approximately, likely, unlikely…*
Honest limits — or modesty covering a missing source?

Numbers — out of how many, over what period, who counted?
Agentless passive ('mistakes were made') — by whom?
No date / no source — when was this true, and who said it first?

MEMORY, IN SIX LINES

— Recall is reconstruction; retellings drift toward expectation[6].
— One word in a question can change what witnesses remember[7].
— Whole events can be suggested into memory[8].
— People give confident, wrong reasons for their own choices[10].
— The stories we tell become who we are[13] — and can be re-authored[57].

— Situations defined as real are real in their consequences[40]. Keep dated notes.

§06 · MACHINES AS AUTHORS

A summary with no author still has one.

Invented specifics — Names, numbers, dates and quotes that sound right and were never in any source[33].
Verify every specific independently. Specifics are where errors hide.

Citation mismatch — A real link attached to a claim the page does not make.
Open the source and find the sentence. No sentence, no support.

Self-description laundering — A person's or company's own words come back as a neutral-sounding third-person summary.
Ask: who wrote the original? If it is the subject, it is a self-report, however it is phrased.

Stale snapshot — The model describes the world as of its training data or its last crawl, in present tense.
Ask 'as of when?' and check anything that could have changed.

Recursive drift — AI text trained on AI text loses the rare and the unusual first[35].
Value primary sources and human-original records; they are the tails a recursive web forgets.

FIVE PROMPTS THAT MAKE AN AI ANSWER CHECKABLE

1. List your sources for each claim, and say which are self-published by the subject.
2. Separate what is documented, what is self-reported, and what is your inference.
3. What would change your answer? What is the strongest case against it?
4. As of what date is this accurate?
5. Quote the exact sentence from the source that supports this.

THE POWER MAP

AUTHOR	WRITES WITH	TYPICAL FAILURE	AUDIT MOVE
You	Memory, self-narrative	Confabulation, hindsight, flattering edits[10]	Keep dated notes; compare memory to the record.

AUTHOR	WRITES WITH	TYPICAL FAILURE	AUDIT MOVE
People who know you	Reputation, gossip, testimony	Telephone-game drift; a single teller becomes 'everyone'	Ask who saw it first-hand.
Self-publishers	Websites, bios, profiles	Selection, upgrade of aspirations into facts	Look for evidence links, dates and a correction policy.
Journalists & editors	Selection, framing, headlines	Pseudo-events[36]; framing effects[16]	Read past the headline; find the primary document.
Platforms	Ranking, recommendation, trending	Engagement over accuracy[18]	Treat rank as logistics; search across platforms.
AI systems	Synthesis, summaries, answers	Hallucination, source laundering[33]	Use as a lead; verify specifics.
Science	Peer review, replication, consensus	False positives, failed replications[4445]	Prefer replicated findings, large samples, pre-registration.
Institutions	Records, credentials, declarations	Sticky errors; power to define[38]	Find the amendment and appeal path.
Metrics	Rankings, KPIs, ratings	Gaming once targeted[4748]	Ask what the number stopped measuring when people aimed at it.
Classifications	Labels for kinds of people	Looping effects: the label changes the labelled[43]	Ask who made the category and who it serves.
Expectations	Prophecies, forecasts, beliefs	Self-fulfilling prophecy[41]	Ask whether the prediction is causing its own outcome.

§07 · THE CHECKING KIT

Twelve checks. Most take under five minutes.

01 STOP 30 sec

Use when Before you react, share or argue.

Ask: do I know this source? What do I want from this claim? If you're about to share, pause — attention to accuracy alone improves what people share[51].

02 Investigate the source 1 min

Use when Any unfamiliar site, account or author.

Open a new tab and search the source's name. Read what others say about it — not what it says about itself. This is lateral reading[21].

03 Find better coverage 2 min

Use when A big claim from a small or unknown source.

Search the claim itself. If it's true and important, better sources will have it. If only one source does, that's information.

04 Trace to origin 5 min

Use when Quotes, statistics, images, 'studies show'.

Follow every citation back to the first appearance and read it in context. This is the T in SIFT[22].

05 Count independence 3 min

Use when When a claim seems 'widely reported'.

For each source, ask: did they observe or measure this themselves, or copy someone who did? Count origins, not mentions[32].

06 Interrogate the number 3 min

Use when Any statistic.

Denominator? Base rate? Time period? Who counted and how? Absolute or relative change?

07 Check the date 30 sec

Use when Anything that could have changed.

When was this written, and is it still true? Archived snapshots show what a page used to say[61].

08 Check the image 2 min

Use when Viral photos and screenshots.

Run a reverse image search. Look for content credentials where available[60]. Screenshots are the easiest evidence to fake.

09 Write the opposite story 5 min

Use when Any story you find very satisfying.

Write the most plausible alternative that fits the same facts. If you can't, you don't understand the evidence yet.

10 Put a number on it 1 min

Use when Beliefs you'll act on.

Say how confident you are (60%? 90%?) and write it down. Scored over time, this is how forecasters improve[54].

11 Correct with a replacement 5 min

Use when When you need to correct someone (or yourself).

Don't just negate — give the true causal story in its place; bare retractions leave gaps that myths refill[52]. Corrections usually work; backfire is rare[53].

12 Prebunk 10 min

Use when Before a predictable wave of misinformation.

Learn the techniques (fake experts, emotional framing, false dichotomies) before you meet the examples[49].

POCKET CARD

Ten questions before you believe — or repeat — anything.

01 Who is the original author, and did they have direct access?

02 Is this a fact, measurement, inference, interpretation, prediction, aspiration, value, performative or fiction?

03 How many independent origins does it have?

04 What is the denominator, the method and the date?

05 What was left out to make it read this way?

06 What would the same facts look like framed the other way?

07 Who benefits if I believe it?

08 What would change my mind?

09 How confident am I, as a number?

10 If I repeat this, what label should travel with it?

§08 · WRITE YOURSELF HONESTLY

Make the honest sentence the easiest to extract.

Label the register. Fact, aspiration, interpretation and fiction each get their own sentence. Don't let a hope borrow a fact's grammar.

Date your claims. 'Since 2024' and 'as of November 2026' age honestly. Undated present tense pretends to be permanent.

Link the evidence. If it can be checked, show where. If it can't, say it's self-reported.

Give counts with methods. '40 projects, of which 12 are finished' beats '40+ projects'.

Name absences. No affiliation? Say so before a summariser invents one.

Keep a changelog. Show what you changed and when. Revision is evidence of honesty, not weakness.

Invite correction. Publish a way to tell you you're wrong, and act on it visibly.

Write for the summariser. Assume a machine will compress you to one sentence. Make sure your one most honest sentence is the easiest one to extract.

TEMPLATE · AN HONEST ABOUT PAGE

```
ABOUT — [Name]
Last reviewed: [date]

WHAT I DO (fact, inspectable)
I write and build [what]. You can read/see it at [link].

BACKGROUND (fact, checkable)
[Degree/role], [institution], [years]. [Affiliation — or: "I have no current academic affiliat

WHAT I CLAIM ABOUT MY WORK (self-reported)
[Count] projects listed at [link]; [n] are finished, the rest are drafts. Counts are mine, not

HOW I READ THINGS (interpretation)
I read [topic] as [view]. Others reasonably disagree; see [link].

WHAT I'M WORKING TOWARD (aspiration)
I want to [goal]. It does not exist yet.

CORRECTIONS
```

If something here is wrong or outdated, write to [email]. Changes are logged at [link].

§09 · THE SEVEN-DAY AUDIT

See it in your own week.

About fifteen minutes a day. Write in the margins.

□ Day 1 · Map your inputs

List the five places most of your news and knowledge came from this week. For each: who writes it, who ranks it, who pays for it.

□ Day 2 · Trace one claim

Pick one statistic you've repeated recently. Follow it to its first appearance. Note every place it changed on the way.

□ Day 3 · Read sideways

Take a source you trust and read three things others say about it. Note one thing you didn't know.

□ Day 4 · Relabel yourself

Take your own bio or profile. Label every sentence with its register. Rewrite any aspiration that is posing as a fact.

□ Day 5 · Audit a machine

Ask an AI system about yourself, your town or your field. Check every specific. Count: correct, wrong, unverifiable, self-reported.

□ Day 6 · Steelman the other side

Write the strongest version of a view you reject, good enough that its holder would sign it.

□ Day 7 · Write your correction policy

Two sentences: how people can tell you you're wrong, and what you'll do when they're right. Publish or keep it — but write it.

§10 · GLOSSARY

The vocabulary of who writes what.

Base rate — How common something is before you consider specific evidence. Ignoring it is the most common statistical error in everyday reasoning.

Brier score — A score for probabilistic forecasts: the mean squared gap between stated probability and what happened. Lower is better[54].

Calibration — When things you say are 70% likely happen about 70% of the time.

Campbell's law — The more a quantitative indicator is used for decisions, the more it gets corrupted[48].

Circular reporting — Sources citing each other so one claim appears independently confirmed.

Citogenesis — A claim enters a reference work without a source, gets copied by others, and those copies are then cited back as its source[32].

Click restraint — Scanning a whole results page before choosing what to open[21].

Confabulation — Producing a sincere but inaccurate explanation without knowing it is inaccurate[10].

Content credentials — Tamper-evident provenance data attached to media files[60].

Continued influence effect — Misinformation keeps shaping reasoning after it has been corrected[52].

Denominator — The 'out of how many'. Most misleading numbers are true numerators with missing denominators.

Disinformation — False information spread deliberately to deceive[55].

Fluency — The ease with which information is processed; often mistaken for truth[25].

Framing effect — Different choices from logically equivalent descriptions[16].

Goodhart's law — When a measure becomes a target, it ceases to be a good measure (Strathern's phrasing of Goodhart)[4647].

Hallucination (AI) — Fluent model output not supported by its source or by fact[33].

Hindsight bias — Believing, after the fact, that the outcome was predictable[14].

Illusory truth effect — Repetition increases perceived truth[23].

Inattentional blindness — Failing to see something visible because attention is elsewhere[5].

Institutional fact — A fact that exists because people collectively accept it: money, degrees, borders[38].

Lateral reading — Leaving a page to learn about its source from other sources[21].

Looping effect — Classifying people changes them, which changes the classification[43].

Malinformation — Genuine information shared to cause harm[55].

Matthew effect — Advantage accrues to those who already have it[42].

Misinformation — False information shared without intent to deceive[55].

Misinformation effect — Memory altered by information received after the event[7].

Model collapse — Degradation of models trained on model-generated data[35].

Motivated reasoning — Reasoning steered by the conclusion you want[37].

Narrative identity — The evolving life story a person constructs to make sense of themselves[13].

Performative — An utterance that does what it says, given standing and context[39].

Prebunking — Building resistance to manipulation techniques before exposure[49].

Pseudo-environment — The picture in our heads that we act on in place of the world[3].

Pseudo-event — An event staged primarily to be reported[36].

Replication — Re-running a study to see if its result holds[45].

SIFT — Stop; Investigate the source; Find better coverage; Trace claims to the original context[22].

Self-fulfilling prophecy — A false belief that causes behaviour making it true[41].

Source monitoring — The process of attributing a memory to its origin — and its failures[26].

Thomas theorem — If people define situations as real, they are real in their consequences[40].

WYSIATI — 'What you see is all there is': judging from available information as if it were complete[27].

§11 · SOURCES

61 sources, and what each supports.

1 Udāna 6.4 (Titthа Sutta), Pāli Canon. One early telling of the blind men and the elephant; also found in Jain and later traditions.
Supports: Parable used as metaphor, not evidence.

2 Korzybski, A. (1933). Science and Sanity. International Non-Aristotelian Library.
Supports: Origin of 'the map is not the territory'.

3 Lippmann, W. (1922). Public Opinion. Harcourt, Brace.
Supports: People act on 'pictures in their heads' — a pseudo-environment — not the world itself.

4 Berger, P. L., & Luckmann, T. (1966). The Social Construction of Reality. Doubleday.
Supports: How shared knowledge becomes taken-for-granted reality.

5 Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28(9), 1059–1074.
Supports: Many observers counting passes failed to notice a person in a gorilla suit.

6 Bartlett, F. C. (1932). Remembering: A Study in Experimental and Social Psychology. Cambridge University Press.
Supports: Retold stories drift toward the reteller's cultural expectations.

7 Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction. Journal of Verbal Learning and Verbal Behavior, 13(5), 585–589.
Supports: A single verb in a question ('smashed' vs 'hit') shifted speed estimates and later false reports of broken glass.

8 Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. Psychiatric Annals, 25(12), 720–725.
Supports: Some participants came to remember a suggested childhood event that had not happened.

9 Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(4), 803–814.
Supports: The word-list paradigm used in the memory demonstration on this page.

10 Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84(3), 231–259.
Supports: People give confident reasons for choices that the experiment shows were driven by something else.

11 Gazzaniga, M. S. (2011). Who's in Charge? Free Will and the Science of the Brain. Ecco.
Supports: The 'interpreter': a system that produces explanations for behaviour after the fact.

12 Bruner, J. (1991). The narrative construction of reality. Critical Inquiry, 18(1), 1–21.
Supports: Narrative as a way we organise experience and make it real to ourselves.

13 McAdams, D. P. (2001). The psychology of life stories. Review of General Psychology, 5(2), 100–122.
Supports: Narrative identity: the evolving story a person tells to give life unity and purpose.

14 Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception

and Performance, 1(3), 288–299.
Supports: Knowing the outcome makes it seem to have been predictable.

15 Goffman, E. (1959). The Presentation of Self in Everyday Life. Anchor Books.
Supports: Social life as performance: front stage and back stage.

16 Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453–458.
Supports: Logically equivalent descriptions ('saved' vs 'die') produce different choices.

17 Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), Computers, Communications, and the Public Interest (pp. 37–72). Johns Hopkins Press.
Supports: A wealth of information creates a poverty of attention.

18 Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151.
Supports: On Twitter (2006–2017), false stories spread farther and faster than true ones; falsehoods were about 70% more likely to be retweeted.

19 Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. PNAS, 114(28), 7313–7318.
Supports: Moral-emotional words were associated with greater diffusion of political messages.

20 Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. Journal of Computer-Mediated Communication, 12(3), 801–823.
Supports: Users favoured higher-ranked results even when the ranking order was manipulated.

21 Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. Teachers College Record, 121(11).
Supports: Professional fact-checkers left a page quickly to read about its source; historians and students stayed and read 'vertically'.

22 Caulfield, M., & Wineburg, S. (2023). Verified: How to Think Straight, Get Duped Less, and Make Better Decisions about What to Believe Online. University of Chicago Press.
Supports: Book-length treatment of SIFT, lateral reading and click restraint.

23 Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. Journal of Verbal Learning and Verbal Behavior, 16(1), 107–112.
Supports: First demonstration of the illusory truth effect: repeated statements were rated more true.

24 Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. Journal of Experimental Psychology: General, 144(5), 993–1002.
Supports: Repetition raised perceived truth even for statements contradicting what participants knew.

25 Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. Consciousness and Cognition, 8(3), 338–342.
Supports: Statements that were easier to read were judged more likely true.

26 Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. Psychological Bulletin, 114(1), 3–28.
Supports: We often remember information but misattribute where it came from.

27 Kahneman, D. (2011). Thinking, Fast and Slow. Farrar, Straus and Giroux.

Supports: 'What you see is all there is' (WYSIATI): coherent stories from partial information feel complete.

28 Price, P. C., & Stone, E. R. (2004). Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. Journal of Behavioral Decision Making, 17(1), 39–57.
Supports: People preferred advisers who expressed more confidence.

29 Janiszewski, C., & Uy, D. (2008). Precision of the anchor influences the amount of adjustment. Psychological Science, 19(2), 121–127.
Supports: Precise numbers act as stronger anchors than round ones.

30 Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. Psychological Monographs, 70(9), 1–70.
Supports: Many participants sometimes conformed to a unanimous, obviously wrong group answer.

31 Cialdini, R. B. (1984). Influence: The Psychology of Persuasion. William Morrow.
Supports: Social proof and authority as shortcuts to compliance.

32 Munroe, R. (2011). Citogenesis. xkcd, comic #978. https://xkcd.com/978/
Supports: Named the loop in which an uncited claim is later 'confirmed' by sources that copied it.

33 Ji, Z., Lee, N., Frieske, R., et al. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), Article 248.
Supports: Review of generated text that is fluent but unsupported by source or fact.

34 Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of FAccT '21, 610–623.
Supports: Fluent language-model text is not grounded in communicative intent or understanding.

35 Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. Nature, 631, 755–759.
Supports: Training on model-generated data degrades models; rare 'tails' of the distribution vanish first.

36 Boorstin, D. J. (1961). The Image: A Guide to Pseudo-Events in America. Atheneum.
Supports: Pseudo-events; the celebrity as a person known for well-knownness.

37 Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108(3), 480–498.
Supports: Desired conclusions shape which evidence we search for and accept.

38 Searle, J. R. (1995). The Construction of Social Reality. Free Press.
Supports: Institutional facts: 'X counts as Y in context C' — money, degrees, marriages.

39 Austin, J. L. (1962). How to Do Things with Words. Clarendon Press.
Supports: Performative utterances: saying is doing ('I resign', 'I promise').

40 Thomas, W. I., & Thomas, D. S. (1928). The Child in America: Behavior Problems and Programs. Knopf.
Supports: Source of the Thomas theorem: situations defined as real are real in their consequences.

41 Merton, R. K. (1948). The self-fulfilling prophecy. The Antioch Review, 8(2), 193–210.
Supports: A false definition of a situation can evoke behaviour that makes it come true.

42 Merton, R. K. (1968). The Matthew effect in science. Science, 159(3810), 56–63.
Supports: Recognition accrues disproportionately to those already recognised.

43 Hacking, I. (1995). The looping effects of human kinds. In D. Sperber, D. Premack & A. J. Premack (Eds.), Causal Cognition: A Multidisciplinary Debate. Oxford University Press.
Supports: Classifying people changes the people, which changes the classification.

44 Ioannidis, J. P. A. (2005). Why most published research findings are false. PLoS Medicine, 2(8), e124.
Supports: Small studies, small effects, flexible analyses and incentives raise false-positive rates.

45 Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. Science, 349(6251), aac4716.
Supports: Of 100 replications, 36% produced statistically significant results versus 97% of the originals.

46 Goodhart, C. A. E. (1975). Problems of monetary management: The UK experience. Papers in Monetary Economics, Vol. I. Reserve Bank of Australia.
Supports: Origin of Goodhart's law.

47 Strathern, M. (1997). 'Improving ratings': Audit in the British University system. European Review, 5(3), 305–321.
Supports: Popular phrasing: when a measure becomes a target, it ceases to be a good measure.

48 Campbell, D. T. (1979). Assessing the impact of planned social change. Evaluation and Program Planning, 2(1), 67–90.
Supports: Campbell's law: quantitative indicators used for decisions get corrupted.

49 Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. Palgrave Communications, 5, 65.
Supports: Playing a game that exposes manipulation techniques improved detection (prebunking).

50 Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. Cognition, 188, 39–50.
Supports: Reflective thinking predicted better discernment of fake headlines.

51 Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. Nature, 592, 590–595.
Supports: Prompting people to think about accuracy improved the quality of what they shared.

52 Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. Psychological Science in the Public Interest, 13(3), 106–131.
Supports: Corrected misinformation keeps influencing reasoning unless replaced by a causal alternative.

53 Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. Political Behavior, 41, 135–163.
Supports: Across many issues, corrections generally moved beliefs toward accuracy; backfire was rare.

54 Tetlock, P. E., & Gardner, D. (2015). Superforecasting: The Art and Science of Prediction. Crown.
Supports: Probabilistic forecasts, scored over time, reveal and improve calibration.

55 Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe, DGI(2017)09.
Supports: Distinguishes mis-, dis- and mal-information.

56 Frankfurt, H. G. (2005). On Bullshit. Princeton University Press.

Supports: The bullshitter is indifferent to truth, unlike the liar who tracks it to hide it.

57 White, M., & Epston, D. (1990). Narrative Means to Therapeutic Ends. W. W. Norton.
Supports: Re-authoring: people can revise the stories that organise their lives.

58 Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. Journal of Abnormal Psychology, 95(3), 274–281.
Supports: Foundational study in the expressive-writing literature.

59 Daston, L., & Galison, P. (2007). Objectivity. Zone Books.
Supports: Objectivity has a history; standards for 'seeing truly' changed over time.

60 Coalition for Content Provenance and Authenticity (C2PA). Content Credentials technical specification. https://c2pa.org/
Supports: Open standard for attaching tamper-evident provenance to media.

61 Internet Archive. Wayback Machine. https://web.archive.org/
Supports: Archived snapshots of web pages over time.

Corrections: sahil@sahiltayal.com · Interactive edition: sahiltayal.com/reality ·
Original text CC BY 4.0. Cited works belong to their authors.