

FIELD BESTIARY · DEPLOYMENT STRATEGY

Nobody Pays for Being Wrong

A bestiary of the go-live: why obeying the system and dodging it answer the same rule, and what a rule looks like when nobody has to carry it

---

Marcos Stipicich

Deployment Strategist · 13 years in supply chain and operations

Cover: Paolo Uccello, Saint George and the Dragon, about 1470

The argument in one page

When a system is wrong after go-live, the cost usually lands in a generic account, weeks later, with nobody's name on it.

That one fact feeds two habits that look like opposites. A supervisor says "the system already calculated it" and obeys against the evidence. A picker ships a dented unit to protect a pick rate. One passes for faith in technology, the other for resistance to change. My reading: both are cheap for the person doing them, because neither pays for being wrong, and neither does the team that deployed the system, which has moved on by then. And nobody learns which of them was right.

1. **The system is usually right, and sometimes it is not.** In health and behaviour, mechanical prediction beats judgment on average; in supply-chain forecasts, large human adjustments help and small ones hurt.
2. **One lever reduces both errors.** People who answer for the result check the aid more; forecasters who must record a reason change forecasts less, and better.
3. **Change the rule and the habit changes, with no sermon.** In US federal procurement, the year-end rush is absent where the budget can roll into next year.

How to read this bestiary

The first speech of Nietzsche's *Thus Spoke Zarathustra*, published in 1883, describes three metamorphoses of the spirit. The camel kneels and "wanteth to be well laden". The lion fights a great dragon whose name is "Thou shalt". The child is "a new beginning, a game, a self-rolling wheel".

I borrow the cast for six entries, with three liberties.

- **Positions, not stages.** Nietzsche describes one spirit changing over time. I describe positions in one building, held at the same time. One supervisor can be a camel in the steering committee at nine and a lion at month-end close at five.
- **Two extra beasts.** The scribe, who draws the others, and the smith, gone before the nail bends.
- **A child without the halo.** On this platform the child usually stands for an attitude: curiosity, humility, unlearning. If attitude moved behaviour, awareness campaigns would work. Nietzsche's child is not humble. It creates new values, and on a dock a new value is a new rule.

Medieval illustrators rarely saw the animals they drew; they copied descriptions and earlier drawings. Keep that in mind at the medieval plates. Many process maps are drawn the same way.

ENTRY I

The Dragon · *Draco debes*

“All values have already been created, and all created values—do I represent.”

On a dock the same claim is shorter: the system already calculated it.

Paolo Uccello, Saint George and the Dragon, about 1470. National Gallery, London.

Scales that move a whole government

Habitat. The steering-committee slide, the KPI wall, the go-live plan.

Scales. Every scale is a rule or a target, and every one glitters “Thou shalt”: so many lines an hour, spend the budget before the year closes, ship the replenishment the system assigns.

The dragon is not evil, and it is not stupid. It is a set of rules that decides, without anyone deciding it, who pays when something goes wrong. That is all its power.

Field notes. Jeffrey Liebman and Neale Mahoney studied 14.6 million US federal procurement contracts from 2004 to 2009. In the last week of the fiscal year, agencies spent 8.7% of their annual procurement, nearly five times the 1.9% of an ordinary week. A whole government does not grow anxious on the same day every September. A scale says “use it or lose it”, and everyone can read.

Taming. In Uccello’s painting the princess leads the dragon on a leash made from her belt. Every go-live has a version of this. On the slide, the dragon is on a leash. On the floor, it still has teeth.

Every decade a new dragon arrives, promising to be the last.



ENTRY II

The Camel · *Camelus obsequens*

“Then kneeleth it down like the camel, and wanteth to be well laden.”

It obeys the system against the evidence, and it is right more often than its critics admit.

Edward Lear, Camel studies, 1867.

Two weeks of iron by air, again

Call. “The new system already has it calculated. Don’t question it.”

I heard that call at the go-live of a Tier-1 enterprise WMS, in the parts distribution centre of a global agricultural-machinery manufacturer. The new rules began sending standard replenishment by air: cast iron, dense spare parts, freight well above the part’s margin. When I reported it, the instruction was to stop questioning the software and execute. It went on for about two weeks, until the carriers’ invoices reached senior management.

The first paper read this as a broken correction channel. This is the camel’s side of it, and it is my reading. Obeying cost the supervisor nothing that day: the freight landed in a company account. Questioning had a price: a delay, an argument with the project, the part of the one who does not trust a very expensive system.

Diet. Calm. Obeying is free today.

Madeleine Clare Elish named the other half of that arithmetic a moral crumple zone: when an automated system fails, blame tends to land on the human at the edge of the loop, which protects the design and the people who chose it.

What the camel gets right

Field notes. Linda Skitka, Kathleen Mosier and Mark Burdick gave people an automated aid in simulated flight tasks. Participants with the aid made commission errors: they followed its wrong directives even when other instruments said otherwise. That is the camel's signature, and Nietzsche saw why it feels like virtue: the spirit, as a camel, "once loved 'Thou shalt'" as its holiest thing.

A fluent answer makes kneeling easier. Parasuraman and Riley called over-reliance on automation misuse; in my reading, a language model that explains a wrong number in good prose invites more of it.

In the camel's defence. In William Grove and colleagues' meta-analysis of clinical and mechanical prediction, mechanical prediction was on average about 10% more accurate, and human judgment came out clearly ahead in only 6% to 16% of the studies. Health and behaviour are not a warehouse, but the camel has a point: a rule applied the same way every time often beats a judgment that changes with the day.

That is the trap. The camel is right often enough to be trusted, and nobody keeps score of the days it is wrong, because those days end in a generic account.

Natural predator. Accounts payable, about two weeks later.



ENTRY III

The Lion · *Leo umbraticus*

“To create itself freedom, and give a holy Nay even unto duty.”

On the page it fights the dragon. On the floor it lives on the dragon’s data.

A lion and a winged dragon, Franco-Flemish, after 1277. J. Paul Getty Museum, Ms. Ludwig XV 4, fol. 93v.

The holy Nay, at 10 to 15 minutes a unit

Call. “I’ll load it later.”

Habitat. The side spreadsheet. When Huber, Zimmermann, Rentrop and Felden studied shadow systems in organisations running a central ERP, 64% were directly connected to it, living off its data to make up for its rigidity. The lion roars at the dragon and eats from its bowl.

Field notes. On a picking line for machinery and automotive filters, dented units kept turning up in the pick faces. Reporting one meant stopping, calling the supervisor, filling in a non-conformity and walking the unit to quarantine: 10 to 15 minutes against the pick rate each picker was evaluated on. The dented units shipped.

Diet. Minutes.

Look closely and the lion’s no is another scale of the same dragon. The picker said no to the non-conformity procedure because a louder scale, the pick rate, said yes to speed. The lion is rarely free. It has chosen which scale to obey.

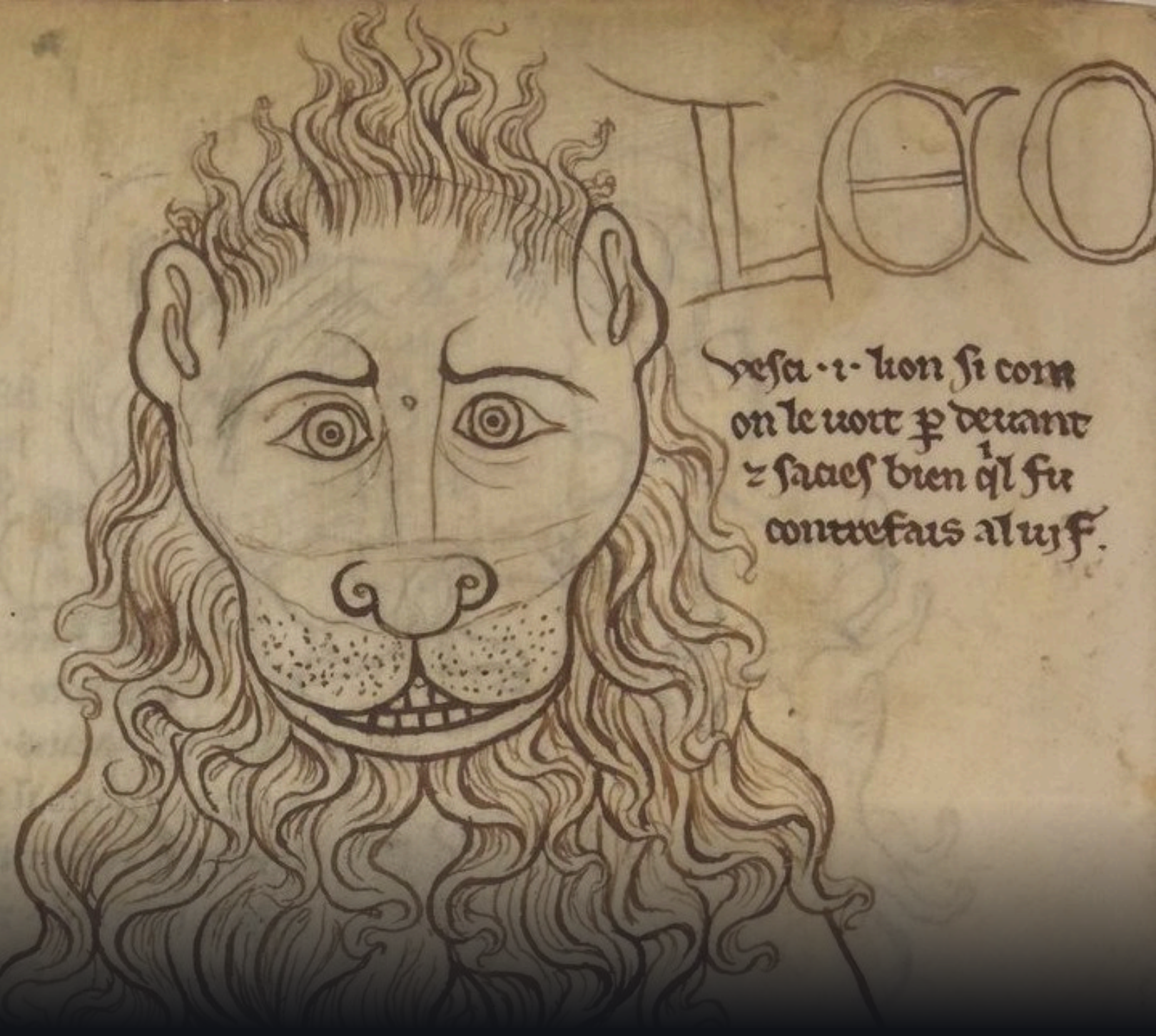
What the lion gets right

In the lion's defence. Robert Fildes, Paul Goodwin, Michael Lawrence and Konstantinos Nikolopoulos examined more than 60,000 forecasts and outcomes from four supply-chain companies. In three of the four, judgmental adjustments improved accuracy on average. Large adjustments helped most; small ones often made the forecast worse. Upward adjustments were much less likely to help than downward ones, and more often went the wrong way.

So the lion is sometimes the only one in the room who knows something the system does not: a customer about to cancel, a promotion nobody loaded. And sometimes it is a mood with a keyboard.

From the outside the two look the same. Both are overrides with no name and no reason, and both are corrected, or not, in a generic account weeks later. The organisation cannot tell the informed lion from the bored one, so it learns nothing from either.

Natural predator. None on record. The spreadsheet survives every go-live.



ENTRY IV

The Scribe · *Scriba tabularius*

“Know well that it was drawn from life.”

Villard wrote it next to this lion. Whether he saw a live one is still debated. The drawing looks like a template.

Villard de Honnecourt, portfolio, about 1230. Bibliothèque nationale de France, MS fr. 19093.

Drawing lions from exports

Habitat. The analytics team, the model. Anyone who draws the floor from the record.

Diet. Exports. The scribe never sees the shelf; it sees what the shelf was recorded as. The first paper showed how a record bends where it is written: in one retailer's audit, 65% of nearly 370,000 inventory records did not match the shelf.

Field notes. A scribe's model is a lion drawn from descriptions. It can be careful, elegant and wrong in exactly the places where the record is wrong, and the scribe will not find out, because the error is paid on the floor and booked in finance. Villard at least wrote down that he had seen the lion. Most models carry no such note, and nobody asks for one.

A subspecies. In Cistercian monasteries the conversi, the lay brothers, worked the fields so the choir monks could pray and study. For thirteen years I worked in operations, and on the docks and in the cold chambers I was often a conversus. These days I sit in the scriptorium and draw lions from exports. Every warning in this entry applies to me.

Natural predator. The physical count.



ENTRY V

The Smith · *Faber fugax*

“…if the only tool you have is a hammer…”

Everyone can finish Maslow’s sentence. The smith is already looking at the next job.

A smith at the anvil, Rabanus Maurus, De universo, 1022–1035. Abbey of Monte Cassino, Cod. Casin. 132.

Gone before the nail bends

Call. “Once it’s live, the system will drive the change.”

Habitat. The kickoff deck, the go-live plan, the slide where the project ends the week the floor begins. The implementation partner, the internal project team, the deployment strategist: anyone who arrives with a tool and leaves with a sign-off.

Diet. Go-live dates.

Field notes. In a 1997 management-journal article, Lynne Markus and Robert Benjamin named a belief they found in IT-enabled change: the magic bullet theory. Nobody is assigned to make the change happen, and everyone assumes the technology will do it by itself.

The smith drives the nail and the project closes. If the nail bends, it bends in month three, on the floor, in a generic account. By then the record shows a successful go-live, and the smith is hammering somewhere else. Like the camel and the lion, it does not pay for being wrong. It was thanked at the go-live party.

What the smith gets right

In the smith's defence. A smith without a hammer is a tourist, and the habit is not vanity. Abraham Kaplan called it the law of the instrument in 1964, picturing a small boy with a hammer; two years later Abraham Maslow added the nail. In a laboratory study by Merim Bilalić, Peter McLeod and Fernand Gobet, expert chess players who had found one solution said they were looking for a better one, while their eyes kept going back to the one they had.

A confession. I carry three hammers myself: floor experience, the incentive model and the guardrail. If you have read this far, you have watched me treat every beast as a question of who pays.

Natural predator. A named owner of the change: an executive on the client's side, named in writing before kickoff, who owns adoption on the floor after the deployer leaves and cannot be the deployer. And a screen where the operator's own move and doing nothing compete with the system's recommendations. If, over a 90-day pilot, neither ever wins, audit the deployer before the system. Both rules change who pays, the same lever as in the rest of this bestiary.



ENTRY VI

The Child · *Puer ludens*

“Innocence is the child, and forgetfulness, a new beginning, a game, a self-rolling wheel, a first movement, a holy Yea.”

The only portrait in this bestiary: a jeweller’s son, given a top instead of a hammer.

Jean-Siméon Chardin, Boy with a Spinning-Top, 1738. Musée du Louvre, Paris.

A rule that rolls by itself

The child is the rarest beast in a go-live, and the most misread. It is not the new hire who asks innocent questions, and it is not a manager who has discovered humility. Nietzsche's child creates new values. On a dock, a new value is a rule that nobody has to carry, like the camel, or dodge, like the lion, because the correct path is also the shortest one. A self-rolling wheel.

Habitat. Rare. Mostly in rules nobody notices, because nothing goes wrong.

Field notes. In the US federal government, Liebman and Mahoney found one budget that could be carried into the next year: one agency's IT funds. In that category they found no year-end spending rush and no drop in the quality of what was bought. Nobody gave a speech about responsible spending. The rule stopped paying for the rush.

Diet. Rules that change who pays.

One lever, both errors

The part I did not expect: one lever reduces both beasts' errors, and it is a rule, not a virtue. Make the person answer for the decision.

- **For the camel.** In a follow-up study, Skitka, Mosier and Burdick made participants accountable for their overall performance or for their accuracy. Automation bias fell: fewer wrong directives followed, and fewer events missed because the aid stayed silent.
- **For the lion.** Paul Goodwin tested cheap changes in a forecasting experiment. Two of them worked: making the statistical forecast the default, so that changing it took an explicit request, and asking for a recorded reason before any change. Both improved accuracy.

A default on the right side and a reason before the override are rules you can put on a screen without asking the board. A reason field that accepts "as discussed" is a checkbox, not a reason.

Before, not after

One warning keeps this from becoming a slogan. Jennifer Lerner and Philip Tetlock reviewed the research on accountability: its effects depend on the kind of accountability, the task, the person and the timing.

Timing is the part a screen controls. In an experiment by Tetlock, Linda Skitka and Richard Boettger, people answerable before committing, to an audience whose views they did not know, thought more self-critically. People answerable for a position already taken spent their effort defending it. People who knew what the audience wanted simply moved toward it.

So the screen asks for the reason before the override, not after it, and it does not tell the operator what the reviewer wants to hear. A reason written after the fact is a defence. A reason written to please the reviewer is an echo.

The camel’s load

Where does AI fit? My answer: under the camel’s load, so nobody has to fight the system like a lion and those who know the floor can build like the child. Three conditions keep that honest.

1. **The tuning fork is physical.** An AI fed the dragon’s rules and a number nobody counted is the perfect camel: obedient, tireless and fast. The blind count comes first.
2. **It takes the load, not the contact.** Skills fade when they are not used, as Lisanne Bainbridge warned in 1983. The count stays in human hands.
3. **Nobody loses their job for the improvement.** Otherwise, a faster cycle finances your own dismissal.

Profit and joy are signals that the rule rolls by itself, not levers. Decreeing joy is one more awareness campaign, and profit shows up as what stops burning: demurrage, obsolescence, emergency overtime, stockouts. If those do not fall against the baseline from before go-live, the design failed.

As I see it, the old regime uses software to watch the camel and punish the lion, and gets compliance on the screen and false data in the record. In the new one, people do not work better because they are happy. They can enjoy the work because the system no longer punishes the truth.

Field guide: which beast runs your deployment?

WHAT YOU SEE	WHICH BEAST, AND WHAT TO ASK
"The system already calculated it", said in a meeting	**Camel.** Who pays if it did not?
An override with no name and no reason	**Camel or lion: you cannot tell.** Where is the reason, and was it written before?
A side spreadsheet that reads the ERP every morning	**Lion.** What does it know that the system does not?
A spike at month-end or year-end	**A scale of the dragon.** Which rule pays for the rush?
A target raised right after it was met	**A scale of the dragon.** Who decided, and on what record?
An official number that disagrees with an independent one	**Scribe.** Which one was counted, and which one was typed?
A project plan whose last milestone is the go-live	**Smith.** Whose name owns the floor in month three?
A correction that arrives as an invoice	**Nobody.** How long did the floor know first?

A deployment with no child in this table is the normal case. It is not a verdict on anyone's character.

What this bestiary does not prove

- **The beasts are borrowed.** Nietzsche's metamorphoses are stages of one spirit, not positions in a building. The mapping is mine, and so are the scribe and the smith.
- **The mechanism is a reading; the lever is evidence.** That obeying and dodging answer the same rule is my reading of the first paper's arithmetic. What the studies support is narrower: accountability and a recorded reason cut errors both ways.
- **The settings differ.** Skitka's participants worked simulated flight tasks; Goodwin's and Tetlock's studies were experiments; Fildes's data came from four companies' forecasts; Liebman and Mahoney studied US federal procurement, and the IT exception is one agency in one category. None of them is a warehouse after go-live.
- **The new parts are softer.** The magic bullet theory comes from a practitioner journal, the hammer from an aphorism, Bilalić's experts from a chess laboratory. The named owner, the 90-day audit and AI under the camel's load are my proposals: hypotheses until a pilot tests them against the baseline.
- **The two scenes are mine, written from memory.** They illustrate; they do not measure. Employers are described, not named.

One question for the bestiary

Before the next go-live, ask two questions in the same meeting: who pays when the system is wrong, and who pays when the floor is wrong?

If both answers are “a generic account, weeks later”, the bestiary is complete. Everyone in the building is a camel or a lion, the scribe is drawing both, the smith has left, and the only one who ever talks back to the dragon is accounts payable.

About the author

Marcos Stipicich
Deployment Strategist · Supply Chain & Operations

I grew up on a sheep ranch, where my father taught me pragmatism early. I read across disciplines, though theory only comes alive for me on the floor, never behind an office desk. When I walk into an operation, I go in like a beginner—dropping the expert posture so people drop their guard and show how the shift actually runs. And when the crew is short a hand, I get on the forklift or the high-reach myself, pick the lines, and load the truck.

Over thirteen years across Chile, Argentina, and Australia, working and leading on the floor took me from -25 °C cold storage and a USD 15M three-site logistics budget to owning a USD 30M cross-border e-commerce P&L (freeing ~USD 5M of working capital tied up in inventory) and supervising floor operations through a 50,000-SKU Tier-1 WMS cutover at a global agricultural-machinery parts centre in Melbourne, where order processing time fell 20%.

Grounded early in C and C++, I architect the systems and direct AI coding agents to ship the Python, TypeScript, and Postgres—owning the math, schema contracts, MILP models, and rollback gates under the hood.

Adapted from my manuscript, The Deployment Chasm.

Available to operating partners and enterprise sponsors for pre-cutover forensic audits and deployment strategy, and for Forward Deployed roles.

linkedin.com/in/marcos-stipicich · github.com/esstipi-debug · marcos.stipicic@gmail.com

Sources

- Bainbridge, L. 1983. "Ironies of automation." *Automatica* 19(6): 775–779. doi.org/10.1016/0005-1098(83)90046-8
- Bilalić, M., P. McLeod, and F. Gobet. 2008. "Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect." *Cognition* 108(3): 652–661. doi.org/10.1016/j.cognition.2008.05.005
- DeHoratius, N., and A. Raman. 2008. "Inventory record inaccuracy: An empirical analysis." *Management Science* 54(4): 627–641. doi.org/10.1287/mnsc.1070.0789
- Elish, M. C. 2019. "Moral crumple zones: Cautionary tales in human-robot interaction." *Engaging Science, Technology, and Society* 5: 40–60. doi.org/10.17351/ests2019.260
- Fildes, R., P. Goodwin, M. Lawrence, and K. Nikolopoulos. 2009. "Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning." *International Journal of Forecasting* 25(1): 3–23. doi.org/10.1016/j.ijforecast.2008.11.010
- Goodwin, P. 2000. "Improving the voluntary integration of statistical forecasts and judgment." *International Journal of Forecasting* 16(1): 85–99. doi.org/10.1016/S0169-2070(99)00026-6
- Grove, W. M., D. H. Zald, B. S. Lebow, B. E. Snitz, and C. Nelson. 2000. "Clinical versus mechanical prediction: A meta-analysis." *Psychological Assessment* 12(1): 19–30. doi.org/10.1037/1040-3590.12.1.19
- Huber, M., S. Zimmermann, C. Rentrop, and C. Felden. 2016. "The relation of shadow systems and ERP systems: Insights from a multiple-case study." *Systems* 4(1): 11. doi.org/10.3390/systems4010011
- Lerner, J. S., and P. E. Tetlock. 1999. "Accounting for the effects of accountability." *Psychological Bulletin* 125(2): 255–275. doi.org/10.1037/0033-2909.125.2.255
- Liebman, J. B., and N. Mahoney. 2017. "Do expiring budgets lead to wasteful year-end spending? Evidence from federal procurement." *American Economic Review* 107(11): 3510–3549. doi.org/10.1257/aer.20131296
- Parasuraman, R., and V. Riley. 1997. "Humans and automation: Use, misuse, disuse, abuse." *Human Factors* 39(2): 230–253. doi.org/10.1518/001872097778543886
- Skitka, L. J., K. L. Mosier, and M. Burdick. 1999. "Does automation bias decision-making?" *International Journal of Human-Computer Studies* 51(5): 991–1006. doi.org/10.1006/ijhc.1999.0252
- Skitka, L. J., K. L. Mosier, and M. Burdick. 2000. "Accountability and automation bias." *International Journal of Human-Computer Studies* 52(4): 701–717. doi.org/10.1006/ijhc.1999.0349
- Tetlock, P. E., L. Skitka, and R. Boettger. 1989. "Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering." *Journal of Personality and Social Psychology* 57(4): 632–640. doi.org/10.1037/0022-3514.57.4.632

Books, essays and the plates

- Bugslag, J. 2001. “Contrefais al vif: Nature, ideas and the lion drawings of Villard de Honnecourt.” *Word & Image* 17(4): 360–378. doi.org/10.1080/02666286.2001.10435726
- Kaplan, A. 1964. *The Conduct of Inquiry: Methodology for Behavioral Science*. San Francisco: Chandler, p. 28.
- Markus, M. L., and R. I. Benjamin. 1997. “The magic bullet theory in IT-enabled transformation.” *Sloan Management Review* 38(2): 55–68.
- Maslow, A. H. 1966. *The Psychology of Science: A Reconnaissance*. New York: Harper & Row, pp. 15–16.
- Nietzsche, F. 1883. *Thus Spake Zarathustra*, Part I, “The Three Metamorphoses.” Translated by Thomas Common. Project Gutenberg eBook 1998.

Images. Public domain: Uccello, about 1470; Edward Lear, 1867; a Franco-Flemish illuminator, after 1277, courtesy of Getty’s Open Content Program; Villard de Honnecourt, about 1230; an illuminator at Monte Cassino, 1022–1035; Chardin, 1738.

Field cases. My own observations, written after the fact; their figures are estimates from memory. Employers are described, not named.