

FIELD NOTES FOR THE COO · DEPLOYMENT STRATEGY

Put the Toll on the Bridge

Every extra step on a shift screen is paid every time the crew meets it. Where that friction protects you, where it buys workarounds, and four questions for the next screen you approve

---

Marcos Stipicich
Deployment Strategist · 13 years in supply chain and operations

Cover: Cornelius Krieghoff, Bilking the Toll, 1860

The argument in one page

A step that guards an expensive mistake earns its place on a shift screen. The same step on every line teaches the crew to get around it.

New England's old toll roads showed where a gate works. On open ground, travellers took a short road around the gate: a shunpike. Where a marsh flanked the gate, they stayed on the road.

A shift screen has the same problem with smaller coins. A confirmation, a blind count or a wait before the AI's suggestion appears is a toll the crew pays on each line it sits on, with gloves on and a truck at the door. In my reading, placed where a mistake is expensive and hard to undo, it buys attention; placed everywhere, it buys a shunpike.

1. **Friction is paid per transaction.** At an illustrative 15-second wait, 500 transactions cost over two hours a shift.
2. **The friction that catches errors is the friction people like least.** In a lab study, the designs meant to make people think before the AI answered caught more of its errors, and the most effective got the least favourable ratings.
3. **An explanation is not a check.** In another, explanations raised acceptance of the AI's answer, right or wrong.

Where the gate stands

In 1919 Frederic Wood, a civil engineer, wrote the history of the toll roads of New England. One complaint runs through it: travellers often avoided the toll when a cheaper way past the gate existed, and they kept finding one.

Wood describes the shunpike as a short section of road that left the turnpike on one side of a gate and joined it again on the other. Legislatures passed laws against them. One act of 1806, for the Twelfth Massachusetts Turnpike, barred any landowner from keeping a road parallel to it within forty rods of a gate, about 200 metres.

Wood gives two gates that the ground protected. One stood on an island in the Lynn marshes, which Wood, citing a county history, calls a point of great advantage in preventing shunpiking. At Newton Upper Falls, a marsh and quicksand beside the gate obliged travellers to stay on the road. The same road through the Lynn marshes had a gate on Chelsea Bridge, and Wood does not say how it fared. My reading is that a bridge holds for the same reason, as long as nobody runs a ferry.

At Brunswick, in Maine, where the ground was open, a second gate did not help. Wood quotes a local history: travellers left the turnpike to cross the river by a ferry, so a second gate went up. They drove across the plains instead and avoided both.

Friction is paid per transaction

A screen step is a gate. On the floor the coin is time: seconds per transaction, paid by the person holding the scanner. When a step costs the crew more than it saves them, and there is a way past it, they find the way. Nobody has to be lazy for that to happen. Where there is a way around, the only fix I know that needs no policing is a right way that is quicker than it.

The cost is easy to underrate because it arrives in seconds. My manuscript uses an illustration: a 15-second wait or blind entry before each of 500 transactions in a shift is 125 minutes. Nobody approves two hours a shift for a wait. They approve one step. And the toll nobody designed, a screen that lags after each scan, is paid on every scan too.

My reading is that on the floor, where there is a way around it, friction spread over every line does not make people think harder. It makes them build a shunpike: a shared login so the screen stays open, a paper list keyed in at the end of the shift, a field filled with whatever closes it fastest. The record still looks complete. It is the toll that went unpaid.

The friction people like least

Zana Buçinca, Maja Barbara Malaya and Krzysztof Gajos tested three such steps with 199 people recruited online. Each saw a photo of a meal and had to replace its most carbohydrate-heavy ingredient with a lighter one. A simulated AI, right three times out of four, suggested an answer and explained it.

Some people saw the suggestion and its explanation at once. Others met one of three designs meant to make them think first: the suggestion appeared only if they asked for it; they gave their own answer first and could then change it, close to a blind count until the answer appears; or the screen made them wait 30 seconds before the suggestion appeared.

When the AI was wrong, people in the forcing designs made the right choice significantly more often, though still less often than a group with no AI at all. Over all the questions, right and wrong together, they did no better, and they rated the forcing designs as more complex. In the authors' own summary, the designs that most reduced overreliance (accepting the AI's answer when it is wrong) got the least favourable ratings.

The authors suggest the design themselves: use the most effective forcing only in the small share of cases where it is predicted to improve the final decision substantially. That is their proposal, not a result. The toll roads add a second rule: no way around the gate. A shift screen needs both.

An explanation is not a check

The friendly fix for an annoying step is an explanation: keep the click, and show the crew why the AI suggests it. Gagan Bansal and seven colleagues tested what explanations do. About a hundred people per condition judged the sentiment of beer reviews or book reviews, or answered logic questions from a law-school admission test, helped by an AI about as accurate as they were.

Explanations made people accept the AI's recommendation more often, right or wrong. They added no accuracy overall beyond showing the AI's confidence, though the AI's help itself raised it. Explaining the AI's first choice raised accuracy when the AI was right and, on the beer reviews and the logic questions, lowered it when the AI was wrong.

So an explanation can take the friction out of a step, and the protection with it.

My reading for the floor: at the step that matters, show something the crew can check against the floor, not the model's case for itself. A bin, a count and the lot on the label are checks. A paragraph of reasons is a pitch.

In a warehouse-system cutover I trained 23 operators, including on RF scanning, and stayed on the floor until the shortcuts died. My reading now: when the scan is part of the pick, it costs the picker nothing extra, and it checks the bin. A confirmation after the pick is a toll.

Four questions for the next screen you approve

Take the screen your crew uses most and list every step that is not the work itself. Then place each one.

WHAT YOU SEE ON THE SCREEN	WHAT IT IS, AND WHAT TO ASK
A confirmation on every line	A toll on open ground, unless it guards a rare, expensive mistake. Ask which mistake would get through if it disappeared. If the answer is a wrong lot or an expired date, those lines are a bridge: keep it there, and only there
A blind count (the screen hides the quantity it expects) on some lines only	A toll on the bridge. Ask who chose those lines and on what (value, error history, how hard the move is to undo), and when the list was last checked
An explanation beside every recommendation	Acceptance, not a check. Ask whether the crew can verify it on the floor, with a bin, a count or a label, or can only believe it
A field always filled with the same "x", 1 or code	A shunpike already built. Ask what filling it honestly costs per shift, then drop the field or move the check to where it pays

Start with the arithmetic. Count the steps per transaction that are not the work, time them on the floor at the busiest hour, multiply by the transactions in a shift, and ask which of those minutes ever caught a mistake.

What brings the crew back to the road. A step that is the work, like the scan that is the pick, or a gate the expensive move cannot skip. At Brunswick, a second gate on open ground only moved the shunpike.

What this does not prove

- **Neither study measured a shift.** Both are lab tasks done online by paid participants: a plate of food, a product review, a logic question. Nobody in them wore gloves or had a truck waiting, and in the lab the forcing ran on every question.
- **The forcing designs caught more of the AI's errors, not more right answers overall,** and they reduced blind acceptance without removing it.
- **The toll roads are an analogy.** A toll was money paid at a gate; a screen step is seconds and attention spent each time the crew meets it.
- **The shunpikes on the floor are my reading,** not a measured rate, and the shortcuts in the cutover are my experience, not a measurement. I have not run a usability study with outside users.
- **It would fail** if a step on every line caught enough mistakes to pay for its shift time, and the crew did not route around it. Then the gate on open ground would be worth its toll.

One question for the floor. Take the step your crew repeats most on its busiest screen. If it disappeared tomorrow, which mistake would get through, and how soon would you know?

About the author

Marcos Stipicich
Deployment Strategist · Supply Chain & Operations

I grew up on a sheep ranch in Patagonia, where my father's rule was simple: reading is worth nothing if you can't see it on a working day. I still work by it: I measure before I cut, and no tool or method beats what works on the floor.

Thirteen years in Chile, Argentina and Australia taught me that supply chain is not just logistics: it is the whole pipeline, which is why it ends up putting out fires that start elsewhere. A promotion planned in marketing becomes expedited orders in purchasing and a blocked aisle in the warehouse. I would rather find where the leak starts. On one gemba walk, mixed, short and damaged orders led back to a per-operator productivity KPI tied to a bonus; I retired the metric and measured customer satisfaction instead.

I have worked most of that pipeline: planning for more than 200 stores, a three-site logistics operation, a USD 30M cross-border e-commerce P&L and a 50,000-SKU warehouse-system cutover, where I stayed on the floor until the shortcuts died and order processing time fell 20%. Deployments live or die inside that pipeline, and that is where I work. I learned to code in C; today I build with AI coding agents, keeping the math and the rollback in my hands.

Available for pre-cutover forensic audits, deployment strategy, and Forward Deployed roles.

From my manuscript, *The Deployment Chasm*.

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

Sources

- Bansal, G., T. Wu, J. Zhou, R. Fok, B. Nushi, E. Kamar, M. T. Ribeiro, and D. S. Weld. 2021. "Does the whole exceed its parts? The effect of AI explanations on complementary team performance." *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*: 1–16. doi.org/10.1145/3411764.3445717
- Buçinca, Z., M. B. Malaya, and K. Z. Gajos. 2021. "To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making." *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW1): 1–21. doi.org/10.1145/3449287
- Wood, F. J. 1919. *The Turnpikes of New England and Evolution of the Same through England, Virginia, and Maryland*. Boston: Marshall Jones. Public domain. The passages used are on pages 32, 79, 86, 145, 162 and 163.

Image. Public domain: Cornelius Krieghoff, *Bilking the Toll*, signed in Quebec, 1860. Winnipeg Art Gallery (G-71-34), via Wikimedia Commons. The board on the toll house reads Tolls and Taux; the sleigh is through the gate while the keeper is still coming out of the woodpile. My reading: a toll that can be run gets run, and the fix is where the gate stands, not a faster keeper.

Field case. The cutover on page 6 is my own experience; the employer is not named.