

# Three Valleys: Spread, Value and Staying Current in AI Adoption

*The single valley of death, reworked for AI inside an organization and set against the audit record on who pays to cross it and the research on implementation, with two simulations on the honey bee voting model: who pays to carry a build to the rest of the organization, and who is still on the best way of working after it has changed six times.*

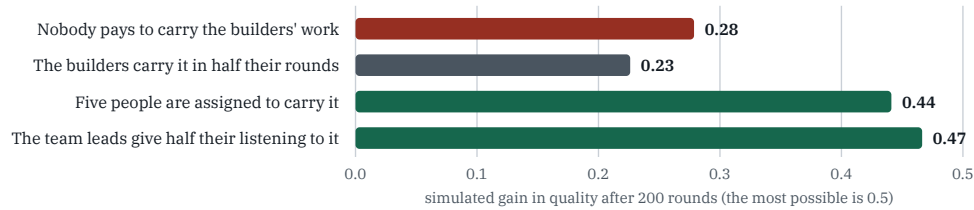
**Berg Atkinson**, Wolfberg LLC · [berg@wolfberg.ai](mailto:berg@wolfberg.ai)

Preprint. Reviewed by a panel of AI reviewers (Claude Opus 5.5) and revised 7 October 2026; not externally peer reviewed. First published 3 October 2026.

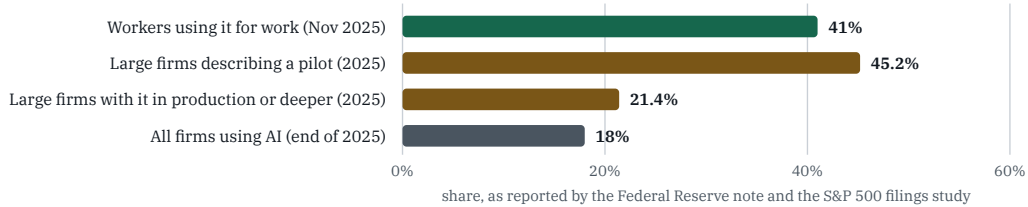
**About the author.** Berg Atkinson: 27 years in defense, intelligence, commercial and international technology, including one of five Chief Engineers on the Air Force ABMS Digital Infrastructure Consortium and advisor to Leidos sector CTOs; since April 2026, sole operator of Wolfberg LLC with an AI team. The positions in this paper rest on that experience; “From one practice” means Wolfberg. [linkedin.com/in/bergatkinson](https://linkedin.com/in/bergatkinson)

---

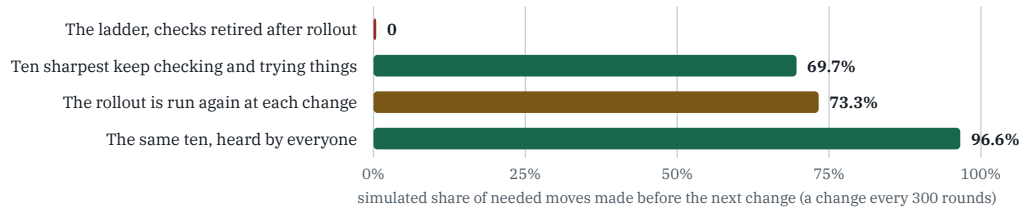
### 1 Spread: the gain by round 200 from what ten people build, by who pays to carry it



### 2 Adoption as others count it: the share depends on the unit



### 3 Staying current: after the best way changes, how often the organization moves to the new one



**At a glance.** One reading for each valley. Top: how far a simulated 200-person organization had gained by round 200 from what ten of its members built, by who paid to carry their work to everyone else; Bars are the gain among members doing their own work; per head, with the five carriers producing nothing, the third is 0.418. Middle: adoption as others have counted it; the share depends on the unit (Section 4.2), and the S&P 500 pilot shares bear on spread (Section 4.1). Bottom: how often the simulated organization moved to the new best way of working after the best way changed; the ladder is the staged rollout in Who Gets It First (Atkinson, 2026b), whose checks stop when the rollout ends (Section 5). **Top and bottom simulated, 400 runs per bar, not a study of people; the middle panel is reported by the sources named in Sections 4.1 and 4.2.**

## IN PLAIN TERMS

The usual picture of why a new technology stalls on its way into use has one valley of death: a stretch between research and a product where the money and the people run short. In a companion note the author proposed that for an AI tool inside an organization there are three valleys, and that the last one repeats. The first is spread: something that works stays in the team that built it. The second is value: the tool is in use across the organization and nothing anyone measures has changed. The third is staying current: it works, people stop checking it, and then the model or the work changes and the organization does not follow. This note sets that proposal against what was already known, and tests the first and third valleys in simulation.

Each valley had already been described, under another name. For the first, the United States Government Accountability Office compared four leading companies with the Department of Defense in 2006 and found that one difference was who manages and pays for the handover. In the companies the laboratory that develops a technology also manages and funds its move into a product. In the department the receiving program does, and the developer has no funds at stake. The office

recommended setting money aside for the developers to manage that move, and of the seven recommendation items its tracking page lists for that report, six are marked implemented and that one is not. For the second valley, researchers showed in 1999 that a technology can be widely bought and sparsely deployed, and later that measured productivity can understate what a broadly useful technology is doing in its early years. For the third, a 1994 study found that people adapt a new technology for a short time after it arrives and then stop, and research on automation shows that people who supervise a machine lose the skill they no longer use. One part is new with AI: the vendor retires the model. On one vendor's public record, 11 models lasted between 12.0 and 25.4 months from the date in their name, with a median of 14.0, and the median notice of retirement was 62 days. Each valley also has close relatives in research on how practices move inside firms and how organizations learn when the target keeps moving (Section 4.4). What we did not find is anyone who puts the three together for AI inside one organization, prices the carry by who pays it, and ties the loop to a vendor's retirement clock.

We ran two simulations on the model we used in the earlier notes, the weighted voter model of how honey bees choose a nest, with an org chart added: 200 people in 25 teams, team leads whose word counts more, an executive, and people who differ in how well they read results. In the first, ten sharp members build better ways of working, and the question is who pays to get what they build in front of everyone else. With nobody paying, the organization had gained 0.279 after 200 rounds, on a scale where 0.5 is the most it could gain. When the builders paid, by presenting in half of their rounds instead of building, it gained 0.226, which is less than with no carry at all, because the builds they gave up are the thing being carried (at the model's default audience; with a much larger one, a quarter of their rounds beat no carry). Five members assigned to carry the work, producing nothing of their own, brought it to 0.441, or 0.418 per head with those five counted as zero output. The team leads giving half of their own listening to the builders brought it to 0.466, in a model that charges nothing for that listening. When the carriers were cut at round 100, the organization kept what had already been carried and went back to gaining at about the unfunded pace, and that held when we let members slide back to the old way.

In the second simulation the best way of working changed six times after the rollout was over. Every plan in which nobody tries anything new after the rollout failed to move every time, 0 of 23,752 needed moves, whether its checks were retired, kept or brought back at each change. In the model that is certain by construction, because a check only tests what somebody proposes. Ten sharp members who kept checking and trying things moved the organization in 54%, 70% and 82% of the needed moves when the changes came 150, 300 and 600 rounds apart, and in 87%, 97% and 99% when everyone could hear them. Running the rollout again at each change moved it in 72% to 73% at every spacing, in a median of 13 rounds, for 3.2 times the checking when changes came 300 rounds apart, and what it missed were the small changes. Everyone trying something new now and then without checking found the new best way and then lost it again in 22.5% to 50.9% of the needed moves.

Three cautions go with all of it. These are simulations of a simplified model, not a study of people or of any organization. The model has nothing to say about the second valley, because quality in it is given. And the strongest evidence against the three-valley picture is real: people have taken up generative AI at work faster than they took up the personal computer, measured gains on single tasks are large, and

the research on general-purpose technologies reads the second valley as a delay and not a death. Section 9 sets that evidence out.

## ABSTRACT

The standard account of stalled innovation has one valley of death, between research and product development (Markham et al., 2010; Auerswald & Branscomb, 2003). For AI inside an organization the author has proposed three: spread (a working build stays in one team), value (the tool is in wide use and nothing measured changes) and staying current (it works, the checking stops, and the organization cannot follow when the model or the work changes), the last repeating (Atkinson, 2026b). We set the proposal against the audit record and the implementation literature and test the first and third valleys in an agent-based model. Each valley has an existing name. The audit record locates the first in who funds and manages the handover (GAO, 2006), and inside one firm it is close to the internal stickiness of best-practice transfer (Szulanski, 1996). The second is the assimilation gap (Fichman & Kemerer, 1999) and the productivity J-curve (Brynjolfsson, Rock & Syverson, 2021). The third is the window that closes after introduction (Tyre & Orlikowski, 1994), the loss of skill in people who supervise automation (Endsley & Kiris, 1995; Budzyń et al., 2025) and a vendor's retirement clock, a median of 14.0 months from a model's name date to its retirement on one vendor's record. The model is the weighted voter model, a noisy social-learning model first used for honey bee nest-site selection (Valentini, Hamann & Dorigo, 2014), with an organization of 200 members added, as in our earlier notes (Atkinson, 2026a, 2026b). In the first simulation ten members build and the carry of their work to the other 190 must be paid for. Uncarried work still spreads in this model, so the results are differences in speed. By round 200 the organization had gained 0.279 of a possible 0.5 with nobody paying, and 0.418 per head with five dedicated carriers counted as producing nothing. Builders who paid with their own rounds gained nothing extra, and lost at half their rounds or more, at the default reach of a presentation; a much larger reach reversed that. Team leads giving half their hearing to the builders were the most effective carry slot for slot, in a model that charges nothing for a lead's attention. In the second simulation the best of five ways of working changes six times. Plans in which nobody proposes anything new cannot follow a change, by construction. Ten members who kept checking and exploring followed 69.7% of needed moves at a change every 300 rounds, and 96.6% when heard by everyone; a rollout run again at each change followed about 73%, quickly, at 3.2 times the ten's checking, and missed most of the smallest changes. Of 25 predictions written before the runs, 15 held, 8 held in part and 2 failed. The second valley is not

simulated. We set out the evidence against the picture, which is strongest for individuals and single tasks.

**Keywords:** AI adoption, valley of death, technology transition, assimilation gap, implementation, organizational learning, model deprecation, human oversight, agent-based model

---

CONTENTS

|   |                                                |    |                     |
|---|------------------------------------------------|----|---------------------|
| 1 | Introduction                                   | 7  | Results: the loop   |
| 2 | The standard picture, and who pays to cross it | 8  | What it means       |
| 3 | Three valleys for AI inside an organization    | 9  | What contradicts it |
| 4 | What was already known                         | 10 | Limits              |
| 5 | Methods                                        |    | References          |
| 6 | Results: the carry                             |    |                     |

# 1 Introduction

Accounts of why a new technology stalls on its way into use usually draw one valley of death, a stretch between research and a product where projects run out of money and people. This note argues that for an AI tool adopted inside an organization the picture has three valleys, shows that each already has a literature, and tests two of them in simulation.

The proposal is the author's, from Who Gets It First, a note in this series published the same day (Atkinson, 2026b, Section 7.4). Our swarm note (Atkinson, 2026a) is the other earlier note in the series. A working build is cheap now, so the first valley is no longer between research and a product. It is spread: something that works stays a demonstration in one team. The second is value: the tool is in use across the organization and nothing measured has changed. The third is staying current: it works, the checking stops, and when the model or the work changes the organization cannot follow. The standard picture ends when the product ships. This one does not end, because each new model and each change in the work starts it again. The author added an observation from practice about the first valley. Research and development is funded, the side of the organization that would take a result to everyone else is usually not funded to do it, and so efforts lose their footing at the start of the valley.

There is reason to look. Most people work at a firm that reports using these tools, and fewer than half use them for their work. In a Federal Reserve note of April 2026, about 78% of the labour force worked at firms that senior leaders reported had adopted AI, and about 41% of workers reported using generative AI for work as of November 2025 (Allen, 2026). Counted one firm at a time the share is far lower, about 18% at the end of 2025, because most firms are small and the surveys differ. Among S&P 500 firms, a study that scored what each firm's annual report says about AI put 45.2% in a pilot phase in 2025, 10.0% using it in production and 11.4% with it deeply integrated in their business processes (Yu et al., 2026). The United Kingdom's Department for Business and Trade found that the people in its

Copilot pilot were satisfied and reported saving time, and it did not find robust evidence that the time saved had led to improved productivity, which it notes was not a key aim of the evaluation (Department for Business and Trade, 2025). Our position is that seats and usage are not the measure (Section 8.3).

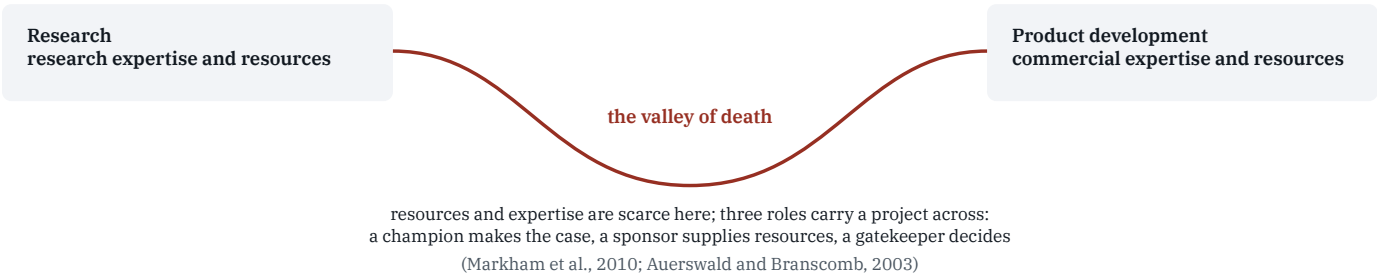
We ask four questions. Does each valley correspond to something already established, and what did the earlier work find? Who pays to carry a build from the people who made it to everyone else, and what does each way of paying buy? When the best way of working keeps changing, which arrangements keep an organization on it, and what do they cost in checking? And what contradicts the picture? Section 2 gives the standard picture and the audit record on who pays to cross it. Section 3 states the three valleys. Section 4 says what was already known, valley by valley, and how others have cut the same ground. Section 5 gives the methods of the two simulations, Sections 6 and 7 their results, Section 8 what they imply, Section 9 the evidence against, and Section 10 the limits.

## **2 The standard picture, and who pays to cross it**

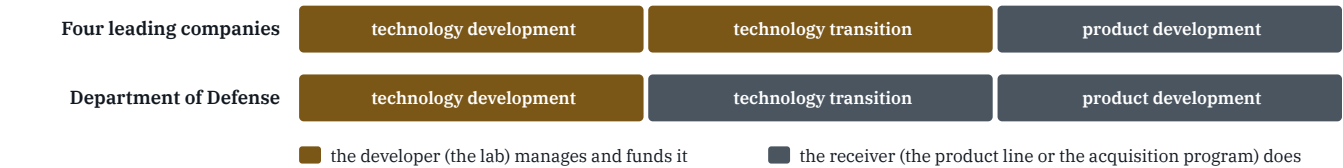
The valley of death is a metaphor for the front end of product innovation, the stretch between research and product development where resources and expertise are relatively scarce. Markham, Ward, Aiman-Smith and Kingon (2010; abstract) surveyed 272 members of the Product Development and Management Association with long experience at that front end. They found that significant development takes place before a project enters a firm's formal process, and that three roles work together to carry a project across: a champion who conceptualizes the idea and prepares the business case, a sponsor who provides resources to demonstrate that it is viable, and a gatekeeper who sets criteria and makes the acceptance decision. Auerswald and Branscomb (2003) describe the same stretch as a financing gap that comes from asymmetries of information and motivation and from institutional gaps. In their study for the National Institute of Standards and Technology a workshop participant defined the valley inside a company as the point where the bill adds up and management asks what it will get for it, while the project is still too early to say. The authors themselves preferred the image of a Darwinian sea, in which business and technical ideas contend and survival goes to the creative, the agile and the persistent (Branscomb & Auerswald, 2002).

The United States government's auditors have written the most direct account of what the valley is made of. In 2015 the Government Accountability Office described a chasm between the Department of Defense's science and technology community and its acquisition community, which department insiders call the valley of death, and gave its cause in one sentence: the acquisition community often requires a higher level of technology maturity than the science and technology community is willing to fund and develop (GAO, 2015). Nine years earlier the same office had compared the department with four leading companies, Boeing, IBM, Motorola and 3M (GAO, 2006). The companies planned at the corporate level before developing a technology, reviewed it at gates, wrote transition agreements that fixed what the lab had to deliver, and kept what the report calls relationship managers on both sides of the handover. All four protected the funding for early projects at the corporate level. And in the companies the lab managed and funded a technology until the product line accepted it, so the product line could act as a customer and did not have to accept, manage or fund the technology's risk.

The standard picture: one valley, between research and product development



Who manages and funds the crossing (after Figure 11 of GAO-06-883)



In the companies the lab pays for the transition and is held to it. In DOD the transition is funded and managed by the receiving program, and the developer "has no funds at stake". GAO recommended setting money aside on the developer's side. DOD did not agree.

**Figure 1.** The standard picture (after Markham et al., 2010, and Auerswald & Branscomb, 2003), and who manages and funds each stage in four leading companies and in the Department of Defense (after Figure 11 of GAO, 2006). **A drawing of two sources; no data of ours.**

The department did the opposite (Figure 1). There the transition was funded and managed by the acquisition programs, the receivers, and the report's sentence about the developers is that the science and technology community "has no funds at stake" once a program has pulled the technology in (GAO, 2006, p. 29). The department planned to spend about \$13 billion on science and technology in fiscal year 2006. Two programs the report names as focused on transition, the Foreign Comparative Testing Program and the Technology Transition Initiative, received about \$37 million and \$29 million that year. The auditors recommended, among other things, that the Secretary of Defense set aside a portion of advanced component development and prototype funds for the science and technology community to manage the transition of technologies. The department did not agree with the set-aside, nor with a recommendation to adopt process-oriented metrics, and generally agreed with the rest. On the set-aside it held that setting aside a small portion of science and technology funds through programs such as those two was appropriate, and the auditors answered that they continued to believe its approach to funding transition was flawed. The office's tracking page for the report lists seven recommendation items. Six are closed as implemented. The set-aside is closed as not implemented, with the comment that none of the services took steps to implement it.

Money was not the whole finding. A later audit counted 20 technology transition programs in the department, which together obligated about \$7.9 billion in fiscal years 2010 through 2012 and reported a historical transition rate of over 70 percent for their projects. For the most part those programs did not track projects beyond transition, so what the transitioned technologies went on to do was not known (GAO, 2013). And the study for NIST found that the companies most active in early-stage research had more formal mechanisms in place to sustain its funding: control of the money split

between business units and a central corporate organization, some discretionary fund for new ideas, and dollar or percentage spending targets (Branscomb & Auerswald, 2002).

Read together, these sources say the valley is not a shortage of technology. It is a question of who holds the money and the responsibility at the handover, and of whether anyone looks after the handover is made. Our position: the author's observation from practice, that the build is funded and the carry is not, is the auditors' finding of 2006 in another setting; in both, the builders are not funded to hand their work over. Against it: the auditors' finding was also about who has money at stake and is held to deliver, which an observation of unfunded carry does not show.

### 3 Three valleys for AI inside an organization

Three things are different when the technology is an AI tool and the adopter is one organization. A working build is cheap. A person with a current model can produce something that works on their own team's task in days, so the scarce thing is no longer the demonstration; the hard part comes after it. Part of that is the old gap from a demonstration to something fit for wider use. One preprint ran six document workflows across four model families and three tool setups and found that 57 of 72 configurations cleared a demonstration bar and 32 cleared a production bar that asked for sustained accuracy, repeatability, attribution and a usable confidence signal (Ahuja, 2026, abstract), so 25 of the 57 that worked in demonstration fell short in production. Here that gap is part of the first valley, because a build not fit for wider use cannot spread. The receiver is the same organization. There is no customer or contract on the far side of the valley, only other teams with their own work, so the handover is a matter of whose time and attention is spent on it. And the tool keeps changing after it is adopted, on a schedule the organization does not set (Section 4.3).

With those three differences the single valley comes apart into three (Figure 2). **Spread.** A build works where it was made and stays there. What has to be paid for is the carry: somebody's time to take the work to everyone else, and somebody's attention to receive it. **Value.** The tool is in use across the organization and nothing measured has changed. What has to be paid for is the change in the work around the tool, and the measuring. **Staying current.** The tool works, the checking stops, and then the model or the work changes. What has to be paid for is somebody who keeps checking and trying things, or a rollout that is run again each time. The third valley leads back to the first, because each model retirement or change in the work makes yesterday's best way one of several candidates again.

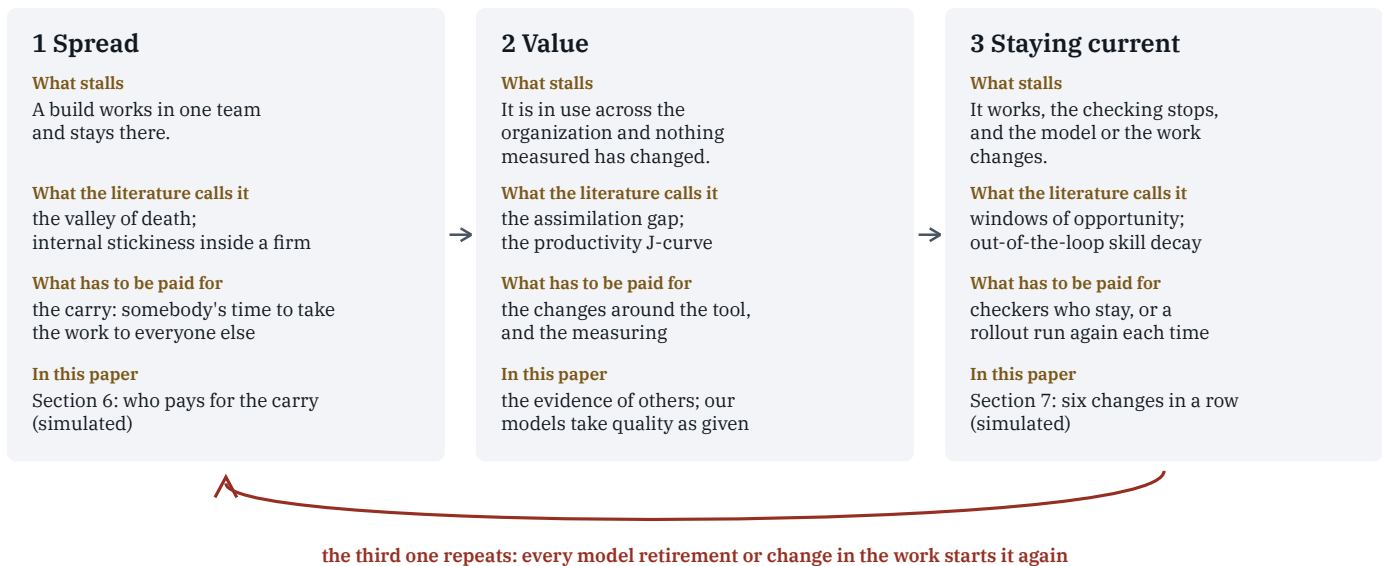


**Figure 2.** Three valleys for AI adoption inside an organization: what stalls in each, what earlier work calls it, what has to be paid for, and where this paper bears on it. **A proposal; Sections 6 and 7 test the first and third in simulation.**

The three are not equally final. An effort that dies in the first valley is over: the build is abandoned or stays a local habit. The second may be a death or a delay, and Section 9 takes up the argument that it is a delay. The third is not an end but a recurrence, which is why a plan that treats adoption as a project with a finish date has no place to put it.

## 4 What was already known

**Table 1.** What earlier work had already established, by valley. One line each; the fuller account follows.

| Valley | What was known                                                                                                                                                         | Source                       | Where it bears here      |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|--------------------------|
| Spread | Where the developer funds and manages the handover it is held to it; where the receiver pays, the developer has nothing at stake                                       | GAO (2006)                   | Who pays (6.1 to 6.3)    |
| Spread | A champion, a sponsor and a gatekeeper carry a project across                                                                                                          | Markham et al. (2010)        | The roles (8.2)          |
| Spread | Firms that sustain early-stage work have formal mechanisms to fund it                                                                                                  | Branscomb & Auerswald (2002) | The funding window (6.5) |
| Spread | Transfer of a best practice inside one firm is sticky; the main barriers are the recipient's lack of absorptive capacity, causal ambiguity and an arduous relationship | Szulanski (1996)             | The receiver pays (6.3)  |

| Valley          | What was known                                                                                                                                                                   | Source                                                                    | Where it bears here      |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|--------------------------|
| Spread          | Sixteen strong hospitals implemented one technology with strikingly different success; the receiving team's learning explained it                                                | Edmondson, Bohmer & Pisano (2001; abstract)                               | The receiver pays (6.3)  |
| Spread          | A system can work in pilot conditions and stall when the project's support ends                                                                                                  | Legara, Jose & Martinez (2026)                                            | The funding window (6.5) |
| Value           | A technology can be widely acquired and sparsely deployed, so counts of purchases mislead                                                                                        | Fichman & Kemerer (1999)                                                  | What to measure (8.3)    |
| Value           | What explains adoption does not explain infusion; use ranges from resistance to commitment                                                                                       | Cooper & Zmud (1990); Klein & Sorra (1996)                                | What to measure (8.3)    |
| Value           | Structures adopted for legitimacy are decoupled from the work, and confidence replaces inspection                                                                                | Meyer & Rowan (1977)                                                      | What to measure (8.3)    |
| Value           | Complementary investment is unmeasured, so early productivity reads low; adoption task by task misses that tasks interact                                                        | Brynjolfsson, Rock & Syverson (2021); Agrawal, Gans & Goldfarb (2024)     | A death or a delay (9)   |
| Staying current | Adaptation happens in a brief window after introduction and then stops until something triggers it                                                                               | Tyre & Orlikowski (1994)                                                  | Nobody looks again (7.1) |
| Staying current | Innovations give early results and then fail to sustain them; organizations encode history into routines                                                                         | Repenning (2002); Repenning & Sterman (2001); Levitt & March (1988)       | The loop (7)             |
| Staying current | Exploration and exploitation have been simulated when the environment keeps changing; change does not always call for more exploration                                           | March (1991); Posen & Levinthal (2012)                                    | The loop (7.2 to 7.4)    |
| Staying current | People who monitor automation lose skill and awareness, experts included                                                                                                         | Endsley & Kiris (1995); Parasuraman & Manzey (2010); Budzyń et al. (2025) | Limits (10)              |
| Staying current | The same model service changes its behavior within months                                                                                                                        | Chen, Zaharia & Zou (2024)                                                | The loop (7)             |
| The model       | A member takes up a way heard from others with probability proportional to how highly its holder rates it (the weighted voter rule, first a model of honey bee nest-site choice) | Valentini, Hamann & Dorigo (2014); Soma et al. (2024); Atkinson (2026a)   | Methods (5)              |

4.1 Spread

The audit record in Section 2 is the core of what was known about the first valley. Three more results bear on it. Edmondson, Bohmer and Pisano (2001; abstract) studied 16 hospitals implementing the same new technology for cardiac surgery. All had top-tier departments, and they differed strikingly in how far they got; the successful ones went through a different team learning process, with enrollment, preparation, trials and reflection. The technology was the same in every hospital, so what differed was the receiving side. Legara, Jose and Martinez (2026) describe two AI systems in a large public education system that reached technically viable stages and could not advance to broader rollout for institutional reasons, and their readiness framework has a row for exactly the author's observation: a system that works in pilot conditions and cannot be maintained, retrained or expanded once the initial project support ends.

And the pilot share is large. Yu, Fleming, Hampton, Combemale and Thompson (2026) had a language model score the AI-related paragraphs of each S&P 500 firm's annual report on a five-level rubric, 4,463 scores for 510 firms from 2016 to 2025. In 2025, 45.2% of firms were scored as pilot-testing AI in products or processes, up from 21.9% in 2022, against 10.0% using it in production and 11.4% with it deeply embedded in business processes (Figure 3). The measure is what filings say, not verified use, and a pilot in a filing is not the same thing as a build stuck in one team. What it does say is that most large firms that report anything about AI report being at the stage before spread.

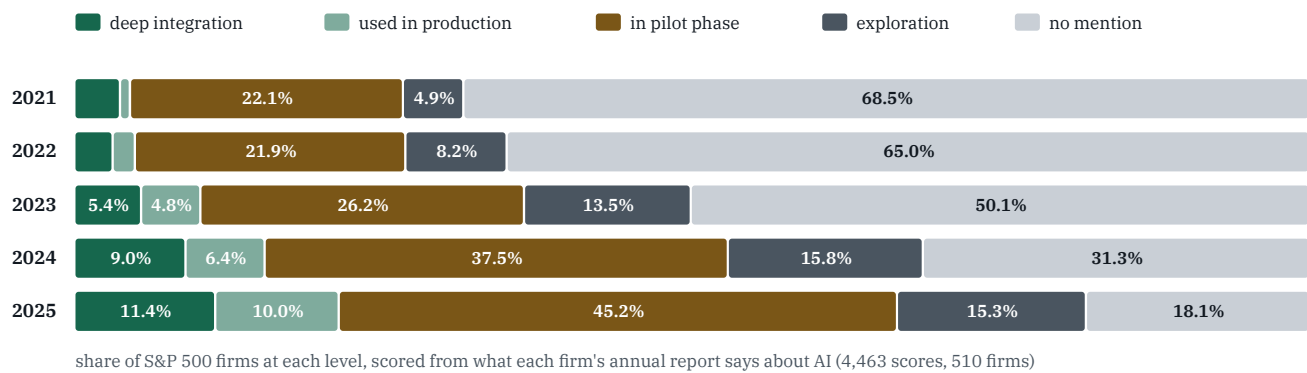


**Figure 3.** What S&P 500 firms' annual reports say about AI, 2021 to 2025: the share of firms at each level of a five-level rubric. **Reported by Yu et al. (2026), Table 12; scored by a language model from the filings; not our data.**

4.2 Value

The second valley has two established names. Fichman and Kemerer (1999; abstract) called the first the assimilation gap. A new information technology may be widely acquired and then only sparsely deployed, so a curve of cumulative purchases and a curve of cumulative deployments separate, and counting purchases gives what they called an illusory picture of diffusion. Of the three software technologies they measured, the gap was very pronounced for one and moderate for two. Earlier, Cooper and Zmud (1990) had found in a sample of manufacturing firms that the interaction of managerial tasks with the technology affected whether a production and inventory control system was

adopted and did not seem to affect how far it was infused into the work, and concluded that rational decision models may explain adoption while political and learning models may explain infusion better. Klein and Sorra (1996) defined implementation effectiveness as the consistency and quality of the targeted members' use of an innovation, with outcomes that range from resistance and avoidance through compliance to commitment. Meyer and Rowan (1977) described what happens when a structure is adopted because it is expected: it is decoupled from the ongoing work, and a logic of confidence and good faith takes the place of inspection and evaluation. An organization can hold seats, usage figures and a governance committee and be in that state.

The second name is the productivity J-curve. Brynjolfsson, Rock and Syverson (2021) show that a general-purpose technology requires complementary investments that are often intangible and poorly measured, so measured productivity growth is understated in the technology's early years and overstated later, when the benefits of those investments are harvested. Adjusting for intangibles related to computer hardware and software, they find total factor productivity 15.9% higher than the official measure at the end of 2017 (abstract). Agrawal, Gans and Goldfarb (2024) give the organizational version: analyses of AI adoption look at single tasks, organizations are made of interacting tasks, and adoption may require a system-wide change.

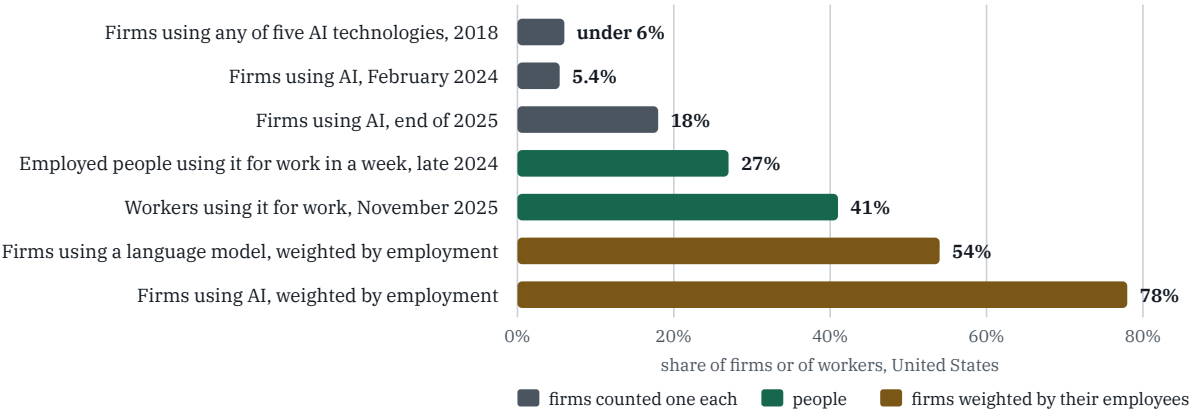


**Figure 4.** How much AI adoption there is depends on what is counted: seven counts for the United States, by what each is a count of. **Reported by McElheran et al. (2024), Bonney et al. (2024), Bick, Blandin & Deming (2026), the first three as reported in their abstracts, and Allen (2026); different surveys with different questions and units; not our data.**

How much adoption the counts show depends on the unit (Figure 4). Counted one firm at a time, the share of firms reporting AI use is far below the share of people reporting it. Fewer than 6% of firms used any of five AI technologies in 2018, in a survey of 850,000 firms (McElheran et al., 2024, abstract). On the Census Bureau's Business Trends and Outlook Survey the share of businesses using AI rose from 3.7% to 5.4% between September 2023 and February 2024 (Bonney et al., 2024, abstract), and it stood at about 18% at the end of 2025, after the question was revised (Allen, 2026). People moved faster: 27% of employed respondents used generative AI for work in the previous week in late 2024, and 10% on every workday (Bick, Blandin & Deming, 2026, abstract), and about 41% of the workforce used it for work by November 2025 (Allen, 2026). Weighted by employment the firm numbers are far higher, about

78% for AI and 54% for a large language model in a survey of senior leaders (Allen, 2026), because larger firms adopt more (McElheran et al., 2024) and because the surveys differ in sample and unit. Like for like, then, most workers are at firms that report using AI and fewer than half use it for their work, which is the gap between acquisition and use that Fichman and Kemerer named. The Federal Reserve note that assembles these figures reports no measure of productivity or outcome.

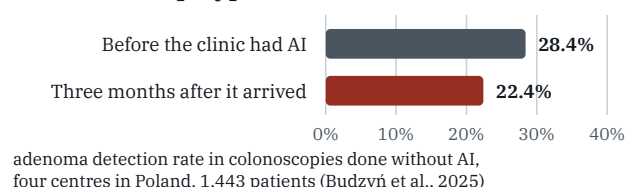
Where outcomes have been measured at the level of the firm, the result so far is the gap this valley names. Across the S&P 500, firms with AI in production or deeply integrated show higher net profit margins, with a J-curve across firms from no adoption to deep adoption (a pattern across levels of adoption at one time, not the productivity J-curve over time described above), and no significant association between adoption and productivity measured as revenue per employee; the authors make no causal claim (Yu et al., 2026). The United Kingdom's evaluation found satisfaction and reported time savings, and no robust evidence of improved productivity, which the report says was not a key aim, so limited data were collected (Department for Business and Trade, 2025). And the defense transition programs of Section 2 mostly did not look past the handover at all (GAO, 2013).

### 4.3 Staying current

The third valley has the oldest literature. Tyre and Orlikowski (1994; abstract) studied the introduction of process technologies in three manufacturing and service organizations and found that adaptation is not gradual. There is a relatively brief window after initial implementation in which users explore and modify the technology. After that, experience makes the use routine, the technology and its context congeal, often with unresolved problems embedded in the practice, and further change comes in episodes triggered by discrepant events or by users' new discoveries. Repenning (2002) built a formal model of how the members of an organization come to commit to an innovation and showed in simulation how seemingly rational decision rules produce innovations that generate early results and fail to produce sustained benefit. Repenning and Sterman (2001) found that the failure of most improvement efforts has little to do with the specific technique and comes from how the effort interacts with the organization around it. Levitt and March (1988) describe organizational learning as the encoding of inferences from history into routines that guide behavior. A routine fitted to one model's behavior is such an encoding.

The person who checks changes too. Endsley and Kiris (1995) showed that operators of an automated system are handicapped in taking over when it fails, through a loss of skill and of situation awareness that they traced to the shift from active to passive processing, and that keeping the operator at a higher level of control moderates the loss. Automation complacency and automation bias appear in experts as well as novices and are not prevented by training or instructions (Parasuraman & Manzey, 2010). There is now a measurement from medicine. At four endoscopy centres in Poland that introduced AI for polyp detection, the adenoma detection rate of colonoscopies performed without AI fell from 28.4% (226 of 795) in the three months before the introduction to 22.4% (145 of 648) in the three months after, a drop of 6.0 points with a 95% interval of 1.6 to 10.5 (Budzyń et al., 2025, abstract; Figure 5, left). The study is observational, and its authors' conclusion is that continuous exposure to AI might reduce the detection rate of standard colonoscopy.

### The checker: polyps found without AI



### The tool: one task, three months apart

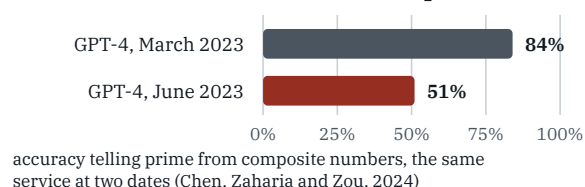


**Figure 5.** The checker's skill and the tool's behavior both move. Left: detection without AI before and after a clinic began using AI. Right: one model service on one task, three months apart. **Reported in the abstracts of Budzyń et al. (2025), an observational study, and Chen, Zaharia & Zou (2024); not our data.**

And the tool changes. Chen, Zaharia and Zou (2024) evaluated the March 2023 and June 2023 versions of two widely used model services on the same tasks and found that the behavior of the same service can change substantially in a short time: one model identified prime and composite numbers with 84% accuracy in March and 51% in June (abstract; Figure 5, right). The two were dated versions behind one service name, and the authors trace part of the change to the model following chain-of-thought prompting less well. That is drift behind a named service; pinning a dated version stops it, and then meets the retirement clock below. There is also a scheduled change, which is new to this technology. Vendors retire models. On Anthropic's public deprecations page, read on 3 October 2026, the 11 dated models that have been retired or given a retirement date lasted between 12.0 and 25.4 months from the date in the model's name to its retirement, with a median of 14.0 months, and the 5 dated 2025 or later lasted 12 to 14 months. The notice between the announcement and the retirement ran from 60 to 189 days with a median of 62, and the page commits to at least 60 days for publicly released models (Figure 6). We read one vendor's page, and the date in a model's name is the nearest thing to a release date that the page carries. Tyre and Orlikowski's finding was that change resumes on a trigger. A vendor now delivers one on its own schedule. The lifetimes are an upper bound on how long one model can be relied on, since an organization adopts it after release; switching by choice, or an alias moving to a newer model, adds triggers. Each retirement comes with about two months' notice to every organization that built a way of working on the retiring model.

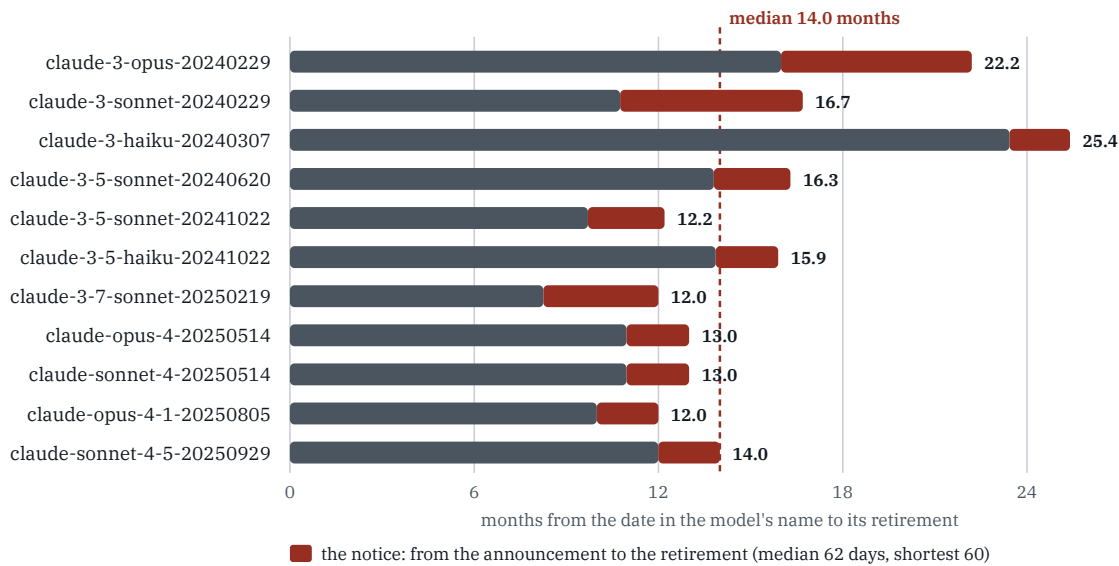


**Figure 6.** How long a model stays available: months from the date in each model's name to its retirement, with the notice period marked, for the 11 dated models one vendor has retired or scheduled for retirement. **Computed from Anthropic's model deprecations page as read on 3 October 2026; one vendor.**

#### 4.4 Other ways to cut the same ground

Others have divided this ground differently. MIT's Center for Information Systems Research describes four stages of enterprise AI maturity: experiment and prepare (28% of the enterprises in its survey), build pilots and capabilities (34%), develop AI ways of working (31%) and become AI future ready (7%). It reports financial performance below the industry average in the first two stages and well above it in the last two, from a 2022 survey of 721 companies that, as the authors say, mainly reflects analytical AI, with 16 executive interviews in 2024 (Weill, Woerner & Sebastian, 2024). Mollick (2025) starts from the observation that AI use which raises individual performance does not by itself raise organizational performance, and divides the work among leadership, which sets direction and incentives, a crowd of employees who discover uses, and a lab that turns what the crowd finds into products and builds the organization's own benchmarks, so that each new system can be compared with the last. The United States Department of Energy scores a technology's adoption readiness on 17 dimensions in four areas (value proposition, market acceptance, resource maturity and license to operate) on a scale of 1 to 9, beside the familiar technology readiness level (U.S. Department of Energy, n.d.). Legara and colleagues (2026) score the receiving institution on five dimensions, and Costa (2026) proposes a six-stage model of where value leaks between a capable model and production.

These cuts agree with the three valleys more than they differ. The maturity stages put the second valley between stages two and three. Mollick's lab, which tests again when a new system arrives, is a loop on model change, and his crowd and lab are a way of paying for the carry. The readiness frameworks score the receiver, which is where the first valley sits. Inside one firm the first valley already has a name, the internal stickiness of best-practice transfer, which Szulanski (1996) traced mainly to the recipient's lack of absorptive capacity, causal ambiguity and an arduous relationship, close to the receiver's side of Section 6.3. Lazer and Friedman (2007) simulated how network structure shapes the spread of

solutions, and March (1991) and Posen and Levinthal (2012) simulated exploration when the environment keeps changing. The carrier has precedents in change agents and opinion leaders (Rogers, 2003; Valente, 2012) and boundary spanners (T. J. Allen, 1977; Tushman, 1977), and Repenning (2002) is a precedent for simulating implementation. What the three-valley cut adds is a combination: who pays at each valley, a price on the carry, and a recurring third valley on a clock the vendor sets, where the stage models end at a final stage. Our search was not systematic: one automated research run that opened 18 of the 74 sources it found, plus the author's own reading; the closest work above was added after a review on 7 October 2026.

## 5 Methods

Both simulations use the organization model of our earlier notes, unchanged, and each adds one thing to it. It is a noisy social-learning model on an org chart: members read results with their own noise and take up ways of working they hear about from colleagues and their boss, as people in organizations often do, rather than testing every option themselves. The rule is the weighted voter model of honey bee nest-site selection (Valentini, Hamann & Dorigo, 2014; Soma et al., 2024), as reproduced in our swarm note (Atkinson, 2026a) and given an org chart in Who Gets It First (Atkinson, 2026b). They are simulations, not studies of people.

### WHO GETS IT FIRST IN BRIEF

That note asked how the order of a rollout, and its checking, decides which way of working an organization ends up on. A *check* of  $n$  tries means trying the candidate and one's own way  $n$  times each and switching only if the candidate reads at least as well. *The ladder* is its staged rollout, whose checks run five, two and one tries over three 100-round stages and then stop. *The close call* is its choice among five ways with true qualities 0.60 to 0.68. The audience *given away* gave its ten sharpest builders 10% of everyone's hearing at no cost to anyone. Two of its counts recur here: an average member's five-try check let a switch to a worse way through 37.4% of the time, and an order to adopt randomly chosen builders' work left 8.2% of runs below their start.

#### The carry: ten people build; who pays to get it heard?



Paid for by nobody: the work spreads through the org chart.  
By the builders: they present instead of building.  
By carriers: a few people stop their own work and relay it.  
By the receiver: team leads spend their listening on it.

#### The loop: the best way changes six times



After the rollout: nobody looks again; or some people keep checking and trying things; or the rollout is run again at each change. Counted: how often the organization moves to the new best way, how long it takes, and the checking spent.

**Figure 7.** The two simulations. Left: ten members build, and the carry to the other 190 has to be paid for by somebody. Right: after a rollout the best of five ways of working changes six times. **Diagrams of the models; no data.**

## 5.1 The organization

Two hundred members are arranged in 25 teams of eight. The first member of each team is its lead, and member zero, a lead, is also the executive. Every time a member uses a way of working they read its payoff with noise: the reading is uniform around the way's true quality with a half-width drawn for each member, once, between 0.1 and 0.5, so some members read results far more accurately than others. Each member keeps a running estimate of the way they are on, weighting each new reading by 0.3. In each round every member revises once. A revising member hears eight others and adopts one of the ways they hold, with probability proportional to how highly each of them rates their own way, which is the weighted voter rule. Each of those eight is a teammate 60% of the time, the member's boss 15%, the executive 5% and anyone in the organization 20%; a lead's vote counts three times; and 15% of the time a member simply adopts their boss's way. With 200 members hearing eight each, the organization has 1,600 hearing slots a round, and that number is the currency of the first simulation.

## 5.2 The carry

The first simulation is the building model of Who Gets It First with a price on the carry. Everyone starts on an incumbent way of working with true quality 0.50. The ten members who read results with the least noise build: in each round each of them makes a new way with probability 0.2, whose true quality is their current one plus a draw from a normal distribution with mean 0 and standard deviation 0.02, so a build is as likely to be worse as better. The builder tries the new way and the current one five times each and keeps the new one only if it read at least as well. Any other member who is about to switch to a way they heard tries it and their own once each, and keeps their own unless the candidate read at least as well. That is the configuration in that note with the ten sharpest doing all of the organization's building, about two builds a round. In that note, left to spread through the org chart, it raised the organization's mean quality by 0.279 by round 200, and when the ten were given 10% of everyone's hearing for nothing it raised it by 0.426. Here that hearing has to come out of somebody's capacity.

**Who pays.** The builders: in each round each builder carries with a set probability (an eighth, a quarter, a half or three quarters) and does not build in that round. Dedicated carriers: one, two, five or ten members who are neither builders nor leads stop their own work and relay the builders' current ways; their own way of working stands still and their output is counted as nothing. The receivers: each of the 25 team leads gives half, or all, of their own eight hearing slots to the builders, and the rest of the organization hears the leads as it always does. A person who carries reaches 32 hearing slots in a round, a room, unless a run says 8 or 128. A listener whose slot is filled by the carry hears one of the ten builders, with that builder's own running estimate of the way. **Variants.** The carriers funded only before round 100, only before round 50, or only from round 100; carriers who report their own single reading of the way they relay and not the builder's estimate; receivers who do not check before switching, or check five times; ten builders chosen at random with five carriers; and, for the funding runs, members who slide back: in each round every member who is not a builder and is not carrying goes back to the incumbent way with a probability of 0.5% or 2%, and can take up a better way again by hearing it.

**Readings.** The workers' gain is the mean true quality in use among members doing their own work, minus 0.50. Net gain per head is the sum of true quality in use over the members doing their own work,

divided by all 200 members, minus 0.50, so a carrier counts as zero output. We also read the best way built so far, the builds attempted, and the share of runs below their starting quality. A way's quality cannot pass 1.0, so every plan approaches a gain of 0.50 by round 400; as in that note we read the plans at round 200, where they are still apart, and give rounds 50 to 400 for the funding runs. In this model an uncarried build cannot die: with nobody paying the gain was 0.412 by round 300. Every carry result is a difference in how fast the work spreads. Read as time to a gain of 0.40, five carriers got there between rounds 150 and 200 (0.345, 0.441), and nobody paying between rounds 200 and 300 (0.279, 0.412). The share of all hearing the carry fills is the share of the 1,600 slots in which the listener hears a builder through the carry. For the builders and the leads it is the slots they give. For dedicated carriers it is slightly more than the slots they take, because a carrier who is heard in the ordinary way, by a teammate, relays the builders too; we counted that with the model's own hearing rule, and five carriers fill 11.8% of all hearing, not 10%. The model charges a builder who carries with the builds they do not make and a carrier with their output. It charges nothing for a lead's attention, and Section 6.3 says what that leaves out.

Before any run, the code had to return that note's records for those two baseline configurations exactly, and it did, value for value at every reading. We then ran 21 configurations. The 8 sliding-back configurations were added after those results were in, because the first draft of Section 6.5 rested on a model in which nothing is forgotten; their code had to return four of the earlier configurations exactly with sliding back switched off, and it did.

### 5.3 The loop

The second simulation is the rollout model of Who Gets It First with the change repeated. The organization chooses among five existing ways of working whose true qualities are 0.60, 0.62, 0.64, 0.66 and 0.68, the close call in that note, and nothing is built. Each plan rolls the tool out over the first 400 rounds as it did in that note. At round 400 the world changes, and it changes again every 150, 300 or 600 rounds, six times in all, so a run lasts 1,300, 2,200 or 4,000 rounds. At each change the five qualities are dealt again to the five ways at random, such that a different way is now the best. The old way stays available: a vendor's retirement forces a change of model, and what the re-deal stands for is the way of working built around it (prompts, workflow, who checks what), which carries across the swap unless somebody revisits it. The sequence of changes is drawn from its own random stream for each run, so every plan meets the same sequence of worlds. In that note there was one change, and the best and second-best ways traded places; we also ran that rule six times over, for four plans, at a change every 300 rounds.

**Plans.** What differs between plans is what is still running after the rollout. *Nobody looks again:* everyone gets the tool at once and nobody ever checks; the staged ladder of Who Gets It First, whose checks run five, two and one tries over three 100-round stages and then stop; the same ladder with its one-try check kept on for good; and the ladder with its checks brought back for 75 rounds after each change (five tries, then two, then one, 25 rounds each) while nobody tries anything new. *People who stay:* the ten members who read results with the least noise get the tool first, as in that note, and afterward keep checking every switch five times and trying a random way in 5% of their updates; the same ten filling

10% of everyone's hearing; ten members chosen at random doing the same; everyone trying a random way in 1% of their updates with no check; and everyone doing that under the ladder whose one-try check is kept on. *A rollout run again on a trigger:* at each change 20% of members pick a way again at random and everyone checks five, two and one times over the next 75 rounds; the same re-run starting 50 rounds late; and the ten who stay, with everyone checking once for 75 rounds after each change. The re-run is told when the world has changed. A vendor's retirement notice does that for a change of model; a change in the work announces nothing, which is what the late re-run stands for.

**Readings.** At each change a move is needed in a run if the way most members hold is not the new best way. The run has followed if most members hold the new best way when the next change arrives. Followed is the share of needed moves that were made, pooled over the six changes and the 400 runs, and we also give the share for each spacing of changes, by how far below the new best way the organization's way was at the change (0.02, 0.04, 0.06 or 0.08), and the median number of rounds until most members first held the new best way, over the moves that were made. A run reached and left if most members held the new best way at some point in the cycle and not at its end. After round 400 we count the share of all member-rounds spent on the way that was best at the time, the mean shortfall per member-round (how far below the best way the way in use was, in the ways' own units), and the check trials per 100 rounds, where one trial is one try of one way.

Before any run, the new code had to return what the code for Who Gets It First returns on twelve of that note's configurations, key for key, and the full run began by replaying four of that note's change configurations and comparing each with its record. All matched. We then ran 40 configurations.

## 5.4 Runs and predictions

Every configuration ran 400 times on seeds 0 to 399: 69 configurations and 27,600 runs in all, beside the two baseline configurations. Intervals on mean gains and on shares of member-rounds are 95% normal intervals over runs. Followed is given pooled over the six changes and the 400 runs, with its counts; because a run contributes up to six moves, the pooled counts overstate how precise these shares are. Predictions for both simulations were written into the scripts before any run, the smoke runs included: nine for the carry, three more for its sliding-back runs, twelve for the loop, and a thirteenth for the loop written after a 16-run smoke test and before any 400-run reading. Section 10 says which held.

## 6 Results: the carry

### 6.1 When the builders pay, the organization gains nothing extra

**When the builders paid for the carry out of their own rounds, the organization gained nothing extra at an eighth or a quarter of their time, and lost at a half or**

more. That holds at the model's default reach of 32 hearing slots; Section 6.4 gives the exception.

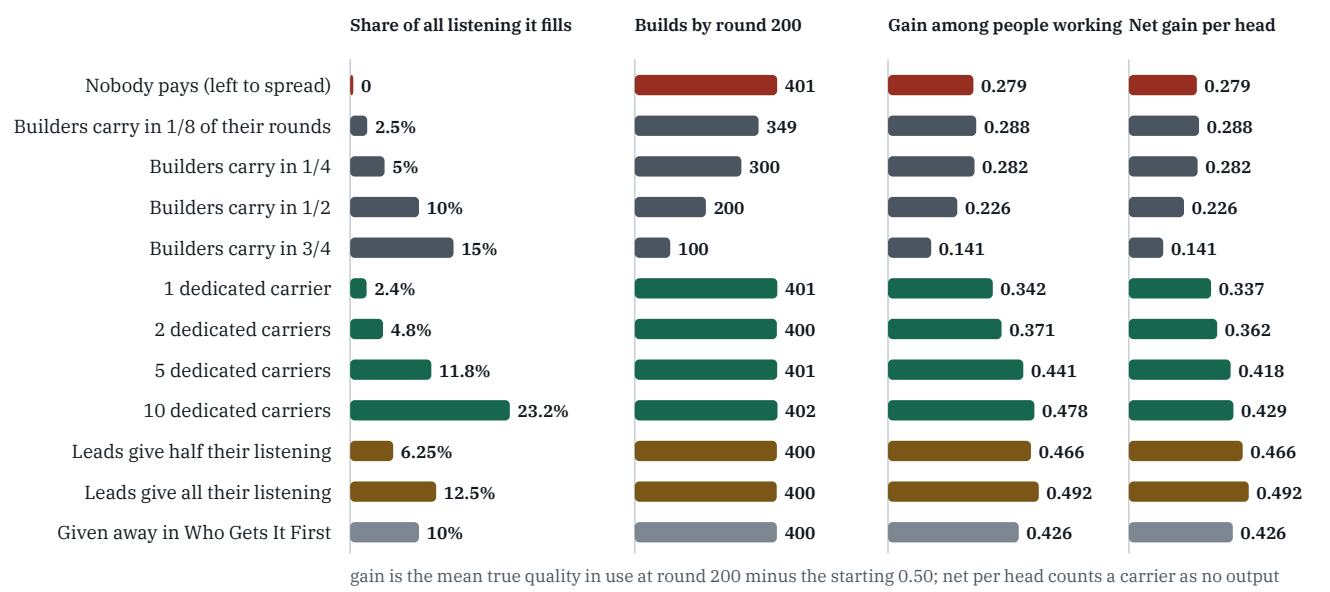


**Figure 8.** Who pays for the carry: the share of all hearing it fills, the builds the ten builders still make, and what the organization gains by round 200, with and without the carriers' lost output. **Simulated; 400 runs per row; not a study of people.**

Left to spread through the org chart, the ten builders' work raised the organization's quality by 0.279 by round 200 (95% interval 0.271 to 0.287). When each builder gave an eighth of their rounds to carrying it, the gain was 0.288, and with a quarter, 0.282; neither can be told apart from no carry at all. With half of their rounds it was 0.226 and with three quarters 0.141, both well below (Figure 8, Table 2). The carry they paid for did reach people. At half their rounds it filled 10% of all hearing, the same share that raised the gain to 0.426 when Who Gets It First (Atkinson, 2026b) gave it away. What it carried was worse. The builders made 401 builds by round 200 with no carry, and 349, 300, 200 and 100 as their carrying rose, and the best way anyone had built by round 200 fell from 0.342 above the start to 0.266 at a half and 0.171 at three quarters. The builds a builder gives up are the thing being carried, so the organization heard more of less.

**Table 2.** Who pays for the carry. The ten sharpest members build in every row. Gain is the mean true quality in use among members doing their own work, minus the starting 0.50; net per head counts a carrier as no output. 400 runs per row.

| Who pays                                       | Share of all hearing the carry fills | Builds by round 200 | Gain by round 100 | By 200 | 95% interval at 200 | Net per head at 200 | Best way built by 200 |
|------------------------------------------------|--------------------------------------|---------------------|-------------------|--------|---------------------|---------------------|-----------------------|
| Nobody (the work is left to spread)            | 0                                    | 401                 | 0.121             | 0.279  | 0.271 to 0.287      | 0.279               | 0.342                 |
| The builders, 1/8 of their rounds              | 2.5%                                 | 349                 | 0.129             | 0.288  | 0.279 to 0.296      | 0.288               | 0.342                 |
| The builders, 1/4                              | 5%                                   | 300                 | 0.127             | 0.282  | 0.274 to 0.290      | 0.282               | 0.331                 |
| The builders, 1/2                              | 10%                                  | 200                 | 0.104             | 0.226  | 0.219 to 0.234      | 0.226               | 0.266                 |
| The builders, 3/4                              | 15%                                  | 100                 | 0.066             | 0.141  | 0.135 to 0.147      | 0.141               | 0.171                 |
| 1 dedicated carrier                            | 2.4%                                 | 401                 | 0.153             | 0.342  | 0.333 to 0.350      | 0.337               | 0.398                 |
| 2 dedicated carriers                           | 4.8%                                 | 400                 | 0.179             | 0.371  | 0.363 to 0.378      | 0.362               | 0.418                 |
| 5 dedicated carriers                           | 11.8%                                | 401                 | 0.226             | 0.441  | 0.435 to 0.447      | 0.418               | 0.470                 |
| 10 dedicated carriers                          | 23.2%                                | 402                 | 0.267             | 0.478  | 0.474 to 0.481      | 0.429               | 0.491                 |
| The leads, half their hearing                  | 6.25%                                | 400                 | 0.246             | 0.466  | 0.462 to 0.471      | 0.466               | 0.485                 |
| The leads, all their hearing                   | 12.5%                                | 400                 | 0.301             | 0.492  | 0.490 to 0.494      | 0.492               | 0.497                 |
| Given away (the audience in Who Gets It First) | 10%                                  | 400                 | 0.211             | 0.426  | 0.420 to 0.433      | 0.426               | 0.459                 |

## 6.2 A few carriers pay for themselves

**Carriers who did nothing else cost the organization their own output and returned several times that.**

One dedicated carrier, half a percent of the organization, raised the gain at round 200 from 0.279 to 0.342, which is 43% of the way to the free carry's 0.426. Two carriers gave 0.371, five 0.441 and ten 0.478. A carrier produces nothing, so the fair reading is net gain per head with each carrier counted as zero output: 0.337, 0.362, 0.418 and 0.429, against 0.279 with nobody carrying. Five carriers, 2.5% of the organization, raised net gain per head by 0.139, from 0.279 to 0.418 of the 0.5 the model allows. The

returns diminish: going from five carriers to ten added 0.037 to the gain among those working and 0.011 to net output per head.

Five carriers did about as well as the carry that Who Gets It First gave away, 0.441 against 0.426, at a slightly larger share of hearing (11.8% against 10%). The carry also sped up the building itself. The builders are listeners too, so a carry that fills hearing slots puts each builder's best work in front of the other nine, and the best way built by round 200 rose from 0.342 with no carry to 0.470 with five carriers, on the same number of builds.

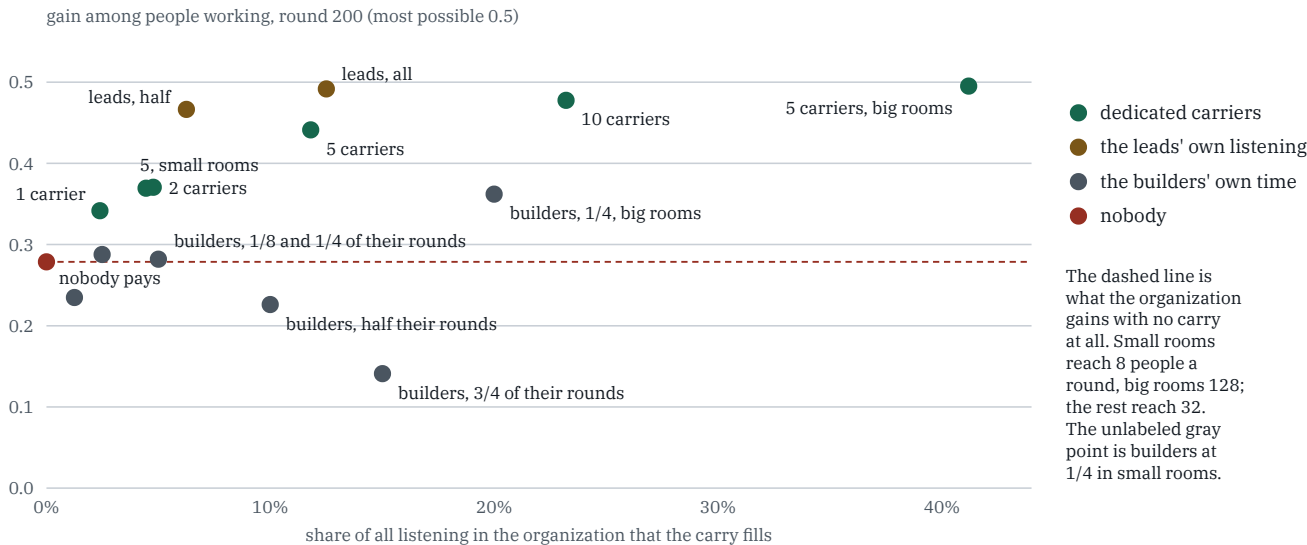


**Figure 9.** The same hearing, bought three ways: the organization's gain at round 200 against the share of all hearing the carry fills, by who pays for it. **Simulated; 400 runs per point.**

### 6.3 The receiver can pay with attention

**The team leads giving half of their own listening to the builders did as well as five carriers, in a model where a lead's attention is free.**

When each of the 25 team leads gave half of their own hearing to the builders, the organization gained 0.466 by round 200, and when they gave all of it, 0.492 (Figure 9). Slot for slot this was the most effective carry we ran: half of the leads' hearing is 6.25% of all hearing, and it did better than the 11.8% that five carriers fill. In the model the reason is the org chart. A lead is heard by the team in 15% of each member's hearing, a lead's vote counts three times, and members defer to their lead 15% of the time, so a lead who has taken up a builder's way passes it down with all of that behind it. In Who Gets It First the chart pulled the organization toward whatever its leads happened to hold. Here it carries the builders' work, once the leads are listening to the builders.

Two cautions belong with this result. The model charges nothing for a lead's attention. If half of a lead's hearing cost half of that lead's output, net gain per head would be about 0.406, and with all of it about 0.368; that is arithmetic on the counts and not a run, and on that pricing half of the leads' attention is slightly below five carriers (0.418) and all of it is further below. And the model makes the leads listen.

Whether a receiver with its own work to do will spend its attention on someone else's build is exactly what the audit record doubts (receiver-side management is the arrangement the auditors criticized in the department), and what Edmondson and colleagues found separated the hospitals that implemented a new technology from those that did not (Section 4.1).

## 6.4 Reach is the lever

### What a round of carrying buys depends on how many people it reaches.

Five carriers reaching 8, 32 and 128 hearing slots a round gave 0.369, 0.441 and 0.495 (Table 3). The builders' own time shows the same thing from the other side. A quarter of their rounds costs the same quarter of their builds at any reach. Reaching 8 slots it left the organization 0.044 below no carry at all (0.235), reaching 32 it was level (0.282), and reaching 128 it was 0.083 above (0.362). So a builder's time spent presenting is not always a loss. It is a loss in a small room. Among dedicated carriers the gain follows the share of hearing they fill: two carriers filling 4.8% and five carriers in small rooms filling 4.4% read the same, 0.371 and 0.369. The builders' own time falls below that line because it costs builds, and the leads' attention sits above it for the reason in Section 6.3 (Figure 9).

**Table 3.** Reach and the other settings, at round 200. A person who carries reaches 32 hearing slots a round unless the row says otherwise. 400 runs per row.

| Setting                                    | Share of all hearing the carry fills | Gain by round 100 | By 200 | Net per head at 200 | Runs below the start |
|--------------------------------------------|--------------------------------------|-------------------|--------|---------------------|----------------------|
| 5 carriers, reach 8                        | 4.4%                                 | 0.171             | 0.369  | 0.348               | 0                    |
| 5 carriers, reach 32                       | 11.8%                                | 0.226             | 0.441  | 0.418               | 0                    |
| 5 carriers, reach 128                      | 41.2%                                | 0.320             | 0.495  | 0.470               | 0                    |
| The builders, 1/4 of their rounds, reach 8 | 1.2%                                 | 0.103             | 0.235  | 0.235               | 0                    |
| The builders, 1/4, reach 32                | 5%                                   | 0.127             | 0.282  | 0.282               | 0                    |
| The builders, 1/4, reach 128               | 20%                                  | 0.173             | 0.362  | 0.362               | 0                    |
| 5 carriers, receivers do not check         | 11.8%                                | 0.229             | 0.448  | 0.424               | 0                    |
| 5 carriers, receivers try five times       | 11.8%                                | 0.243             | 0.462  | 0.438               | 0                    |
| 5 carriers who report their own reading    | 11.8%                                | 0.220             | 0.438  | 0.414               | 0                    |
| Ten random builders, 5 carriers            | 11.8%                                | 0.116             | 0.246  | 0.227               | 0                    |

6.5 A cut keeps what was carried and loses what comes next

Cutting the carriers did not undo what they had carried. It put the organization back on the unfunded pace.

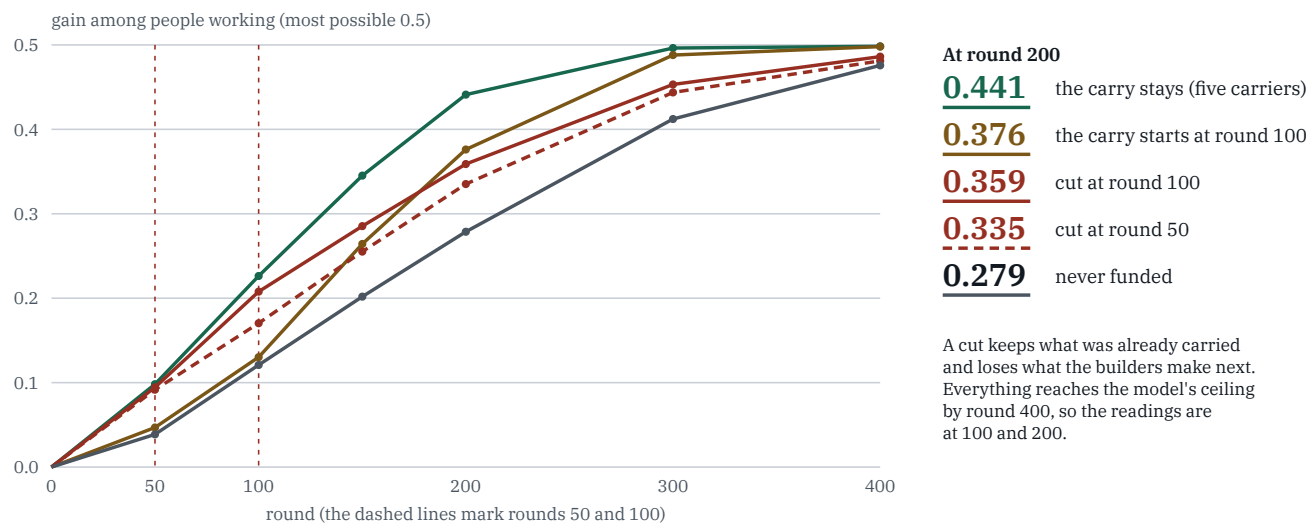


**Figure 10.** When the carry is funded: the organization's gain by round, with five carriers throughout, starting at round 100, cut at round 100, cut at round 50, and never. **Simulated; 400 runs per line; every line approaches the model's ceiling of 0.5 by round 400.**

With five carriers funded throughout, the organization had gained 0.226 by round 100 and 0.441 by round 200. With the carriers cut at round 100 it reached 0.359 at round 200, which is 0.082 below the funded organization and 0.080 above one that was never funded (Figure 10, Table 4). Over rounds 100 to 200 the cut organization gained 0.151, the never-funded one 0.158 and the funded one 0.215. The cut did not take back what had been carried; it returned the organization to the pace of one with no carry, from a higher base. A cut at round 50 left 0.335 at round 200, 0.057 above never funded. Funding that started at round 100 reached 0.376 by round 200, 0.065 below funding from the start, because what is carried is the builders' current work and not its history. By round 400 every line is near the ceiling, which is a property of the model.

**Table 4.** When the carry is funded. Five dedicated carriers; gain among members doing their own work, by round, with nobody sliding back and with every member who is not a builder going back to the incumbent way with a probability of 0.5% or 2% in each round. 400 runs per row.

| Funding             | Sliding back | Round 50 | 100   | 150   | 200   | 300   | 400   |
|---------------------|--------------|----------|-------|-------|-------|-------|-------|
| Never funded        | none         | 0.039    | 0.121 | 0.202 | 0.279 | 0.412 | 0.476 |
| Funded throughout   | none         | 0.098    | 0.226 | 0.345 | 0.441 | 0.496 | 0.498 |
| Starts at round 100 | none         | 0.047    | 0.130 | 0.264 | 0.376 | 0.488 | 0.498 |
| Cut at round 100    | none         | 0.094    | 0.208 | 0.285 | 0.359 | 0.453 | 0.486 |
| Cut at round 50     | none         | 0.092    | 0.170 | 0.255 | 0.335 | 0.444 | 0.481 |
| Never funded        | 0.5% a round | 0.041    | 0.122 | 0.202 | 0.281 | 0.410 | 0.472 |
| Funded throughout   | 0.5% a round | 0.091    | 0.213 | 0.336 | 0.430 | 0.492 | 0.496 |
| Cut at round 100    | 0.5% a round | 0.093    | 0.209 | 0.286 | 0.361 | 0.454 | 0.483 |
| Cut at round 50     | 0.5% a round | 0.088    | 0.165 | 0.248 | 0.325 | 0.438 | 0.479 |
| Never funded        | 2% a round   | 0.041    | 0.122 | 0.201 | 0.278 | 0.401 | 0.461 |
| Funded throughout   | 2% a round   | 0.089    | 0.213 | 0.334 | 0.429 | 0.485 | 0.488 |
| Cut at round 100    | 2% a round   | 0.087    | 0.203 | 0.282 | 0.354 | 0.450 | 0.476 |
| Cut at round 50     | 2% a round   | 0.087    | 0.164 | 0.245 | 0.322 | 0.431 | 0.472 |

That reading could have been an accident of a model in which nothing is forgotten, so we let members slide back (Table 4, lower rows). With every member who is not a builder going back to the incumbent way with a probability of 2% in each round, the organization cut at round 100 read 0.354 at round 200 and the one never funded 0.278, a lead of 0.076 where it was 0.080 with nobody sliding back; at 0.5% a round the lead was 0.080. The organization funded throughout read 0.429, 0.012 below its reading with nobody sliding back. Sliding back moved every line by about a hundredth or less, because a member who slides back hears the better way again from the people around them and takes it up again.

Differences that small should not be read as effects: rows of Table 4 that are the same treatment before an intervention already differ before it (0.098 and 0.094 at round 50 for funded throughout and cut at round 100; 0.039 and 0.047 for never funded and starts at round 100), and Table 4 gives no intervals.

The author's observation was that the funding stops at the start of the valley. In this model the loss from stopping it is the future gain, at the difference between the two paces, and what was already carried stays, with or without members sliding back. What the model does not have is anything else competing for the same hearing: the builders' work is the only new thing there is to hear about.

## 6.6 What mattered little

Three settings we expected to matter moved the result by a few hundredths. Carriers who report their own single reading of the way they relay, a handoff as a game of telephone, gave 0.438 against 0.441 for a clean relay, because the vote is proportional to the reported value and unbiased noise mostly washes out. A handoff that changes what the work is, and not only how accurately it is reported, is not in the model. Receivers who did not check before switching gave 0.448, receivers who checked once 0.441 and receivers who checked five times 0.462. And with ten builders chosen at random in place of the ten sharpest, five carriers raised the gain to 0.246, 2.1 times the 0.115 that Who Gets It First found when random builders' work was left to spread, with 0 of 400 runs below their start. In that note an order to adopt random builders' work left 8.2% of runs below their start. A carry is not an order: each member still tries what they hear against their own way before switching.

## 7 Results: the loop

### 7.1 A plan in which nobody tries anything new never follows

**After the best way changed, every plan in which nobody tries anything new stayed where it was: 0 of 23,752 needed moves, with its checks retired, kept or brought back. In the model this holds by construction: with nobody proposing a way they do not already hold, a check has nothing to test.**

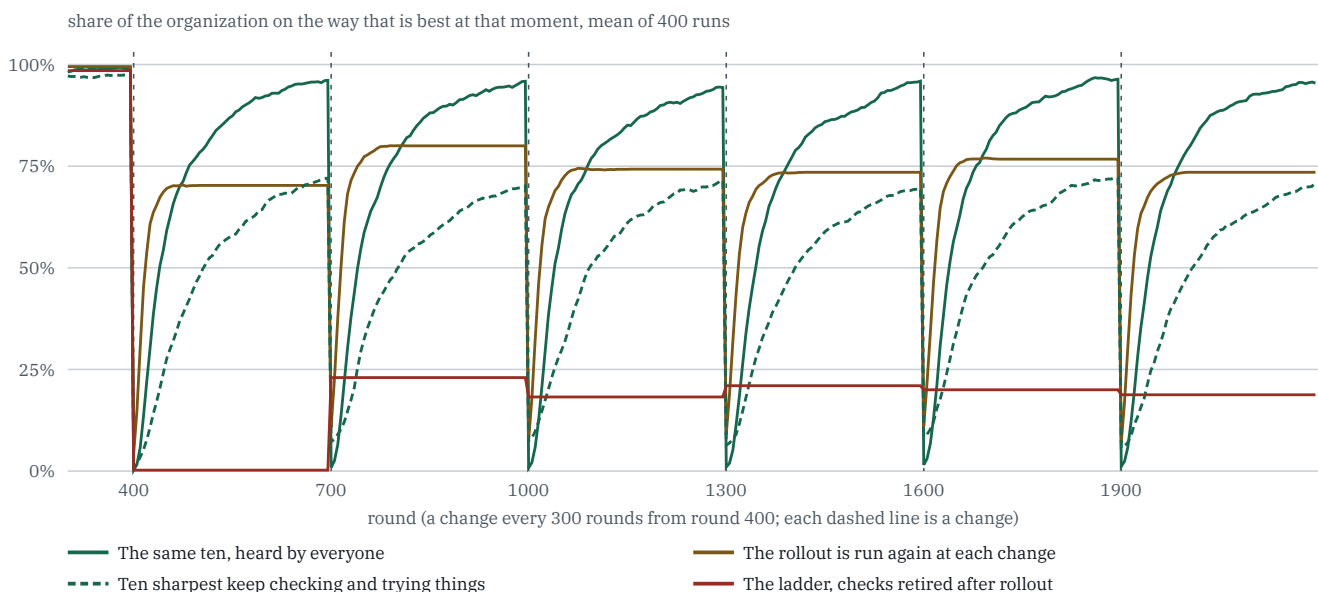


**Figure 11.** Who is still on the best way, change after change: the share of all 200 members on the way that is best at that moment, with a change every 300 rounds, for four plans. **Simulated; mean of 400 runs per line; not a study of people.**

The four plans in the first group finished their rollouts very differently. With everyone getting the tool at once and nobody checking, the organization was on a worse way at round 400 in 66.2% to 71.5% of runs, and under the three ladders in 1.0% to 2.8%. After the first change none of that mattered. In each of the twelve configurations, four plans at three spacings, a move was needed 1,921 to 1,999 times and was made in none (Figure 11, Table 5). The ladder that keeps its check on tests every proposed switch, and in a settled organization no switch is proposed. The ladder whose full checks came back for 75 rounds after each change checked nothing, for the same reason. After round 400 these organizations spent 16.7% to 20.0% of their member-rounds on the best way, which is what a way that never moves earns when the best of five is dealt again at random, and their mean shortfall was between 0.039 and 0.042, about two steps of the five-way scale.

Who Gets It First (Atkinson, 2026b) found this for one change, where the finished ladder followed in 0 of 400 runs. Six changes and two plans built to answer it, a standing check and checks that return, change nothing. It is Tyre and Orlikowski's window in a model: the organization adapted during the rollout, the practice set, and nothing in these plans reopens it. Levitt and March (1988) describe the organizational form of this as a competency trap.

**Table 5.** Followed: the share of needed moves made before the next change, pooled over six changes and 400 runs, with the counts. A move is needed when the organization's most common way is not the new best way at the change.

| Plan                                                                | A change every 150 rounds | Every 300              | Every 600              |
|---------------------------------------------------------------------|---------------------------|------------------------|------------------------|
| The same ten, heard by everyone                                     | 87.1% (2,031 of 2,333)    | 96.6% (2,303 of 2,384) | 99.2% (2,380 of 2,398) |
| Everyone tries something new 1% of the time, and checks             | 68.4% (1,543 of 2,256)    | 83.4% (1,927 of 2,311) | 92.5% (2,183 of 2,359) |
| The rollout is run again at each change                             | 72.4% (1,645 of 2,271)    | 73.3% (1,665 of 2,272) | 73.1% (1,644 of 2,249) |
| The rollout is run again, 50 rounds late                            | 71.0% (1,598 of 2,251)    | 72.0% (1,624 of 2,257) | 73.0% (1,652 of 2,264) |
| Ten sharpest, plus a 75-round check by everyone                     | 60.3% (1,332 of 2,209)    | 73.3% (1,652 of 2,253) | 83.1% (1,927 of 2,318) |
| Ten sharpest keep checking and trying things                        | 54.0% (1,175 of 2,175)    | 69.7% (1,570 of 2,254) | 81.7% (1,886 of 2,309) |
| Ten random people keep checking and trying things                   | 29.5% (613 of 2,076)      | 44.4% (940 of 2,117)   | 57.4% (1,239 of 2,158) |
| Everyone tries something new 1% of the time, unchecked              | 30.3% (613 of 2,023)      | 37.5% (760 of 2,024)   | 42.1% (871 of 2,067)   |
| The ladder, checks return at each change, nobody tries anything new | 0 (0 of 1,995)            | 0 (0 of 1,992)         | 0 (0 of 1,992)         |

| Plan                                                  | A change every 150 rounds | Every 300      | Every 600      |
|-------------------------------------------------------|---------------------------|----------------|----------------|
| The ladder, one check kept, nobody tries anything new | 0 (0 of 1,989)            | 0 (0 of 1,999) | 0 (0 of 1,995) |
| The ladder, checks retired after rollout              | 0 (0 of 1,993)            | 0 (0 of 1,995) | 0 (0 of 1,995) |
| Everyone at once, nobody ever checks                  | 0 (0 of 1,940)            | 0 (0 of 1,921) | 0 (0 of 1,946) |

## 7.2 People who stay follow, by how long they have and who hears them

Ten members who kept checking and trying things moved the organization more often the longer they had between changes, and nearly every time when everyone could hear them.

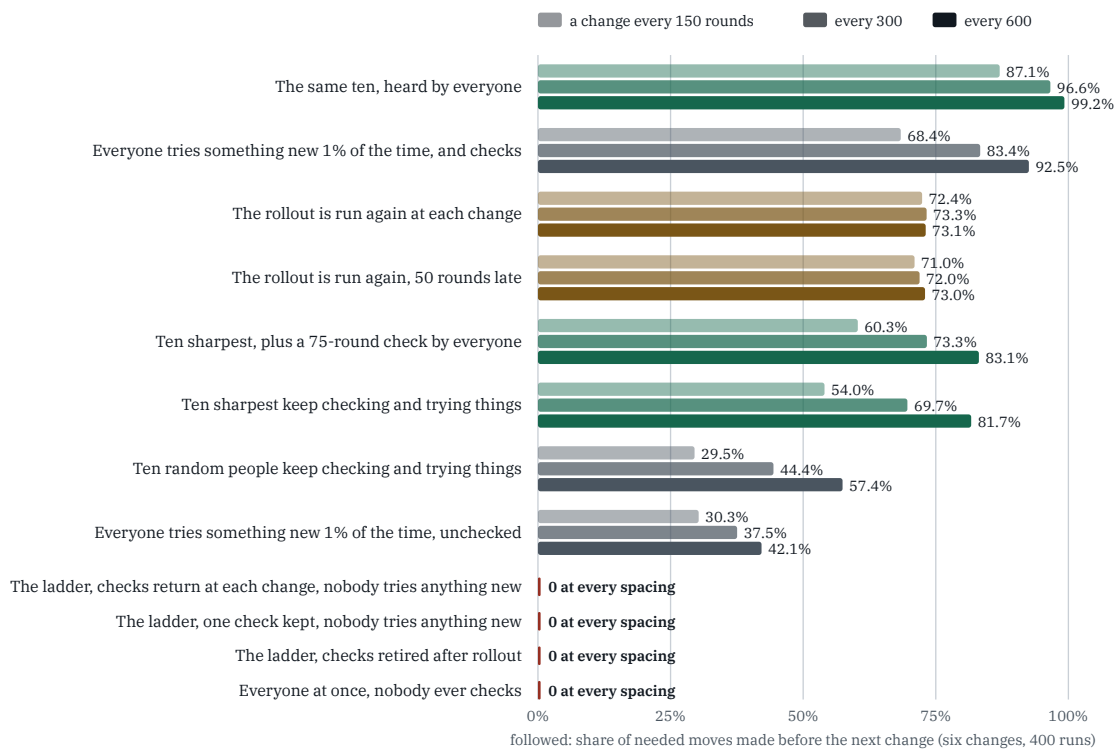


**Figure 12.** Followed: the share of needed moves the organization made before the next change, for every plan, at a change every 150, 300 and 600 rounds. **Simulated; six changes and 400 runs per bar.**

The ten members who read results with the least noise, checking every switch five times and trying a random way in 5% of their updates, brought the organization to the new best way in 54.0% of the needed moves when changes came every 150 rounds, 69.7% at 300 and 81.7% at 600 (Figure 12). The moves they made took a median of 51, 65 and 79 rounds. When the same ten filled 10% of everyone's hearing they followed 87.1%, 96.6% and 99.2%, in a median of 32 to 37 rounds, and with fewer check trials than the ten heard as usual (1,059 against 1,228 per 100 rounds at a change every 300). Ten

members chosen at random, doing the same checking and exploring, followed 29.5%, 44.4% and 57.4%. So the two things that decided spread in Section 6 decide staying current too: who the people are, and whether the organization can hear them. Built into the plan: the organization can tell who reads results most accurately.

The six changes were alike. At a change every 300 rounds the ten followed between 67.4% and 72.2% of needed moves in each of the six cycles, and the ten heard by everyone between 94.7% and 97.5%; nothing in the model wears the ten out or teaches the organization to follow faster. Adding one check by everyone for 75 rounds after each change to the ten who stay raised followed by 6.3, 3.7 and 1.5 points at the three spacings.

### 7.3 A rollout run again is fast, costs more the more often the world changes, and misses small changes

Running the rollout again at each change followed about 73% of the time at every spacing, within a few rounds. It missed the changes that left the organization only a little behind.

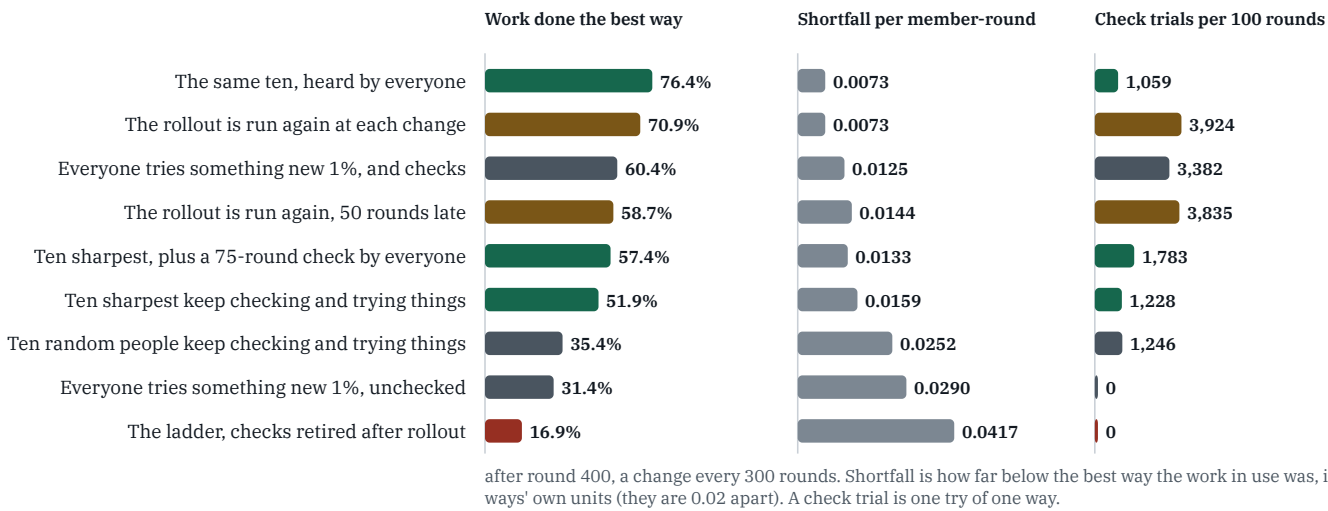


**Figure 13.** What following costs, at a change every 300 rounds: the share of all work done the best way after round 400, the mean shortfall, and the check trials spent. **Simulated; 400 runs per row.**

In the re-run, at each change a fifth of the members pick a way again at random and everyone checks five, two and one times over the next 75 rounds. It followed 72.4%, 73.3% and 73.1% of needed moves at the three spacings. More time between changes did not help it, because it is one shot: by round 75 it is over. It was fast. The moves it made took a median of 13 rounds, against 65 for the ten and 36 for the ten heard by everyone. Because it is fast, it kept more of the work on the best way than its followed share suggests. When changes came every 150 rounds it put 66.1% of all member-rounds after round 400 on the best way, the most of any plan, ahead of the ten heard by everyone at 62.0%. At 300 rounds the order reversed, 70.9% against 76.4%, and at 600 it was 72.9% against 87.2% (Figure 13, Table 6).

**Table 6.** A change every 300 rounds: what each plan followed and what it cost, after round 400. Never and every time are shares of the 400 runs. Shortfall is the mean distance below the best way, in the ways' own units, per member-round. Left is the share of needed moves in which the organization reached the new best way and was off it when the cycle ended.

| Plan                                                                | Followed | Never  | Every time | On the best way | Shortfall | Check trials per 100 rounds | Median rounds to follow | Left  |
|---------------------------------------------------------------------|----------|--------|------------|-----------------|-----------|-----------------------------|-------------------------|-------|
| The same ten, heard by everyone                                     | 96.6%    | 0      | 82.5%      | 76.4%           | 0.0073    | 1,059                       | 36                      | 1.2%  |
| Everyone tries something new 1% of the time, and checks             | 83.4%    | 0      | 36.5%      | 60.4%           | 0.0125    | 3,382                       | 58                      | 3.5%  |
| The rollout is run again at each change                             | 73.3%    | 0.2%   | 16.8%      | 70.9%           | 0.0073    | 3,924                       | 13                      | 1.4%  |
| The rollout is run again, 50 rounds late                            | 72.0%    | 0.2%   | 15.5%      | 58.7%           | 0.0144    | 3,835                       | 64                      | 1.4%  |
| Ten sharpest, plus a 75-round check by everyone                     | 73.3%    | 0      | 19.5%      | 57.4%           | 0.0133    | 1,783                       | 51                      | 5.4%  |
| Ten sharpest keep checking and trying things                        | 69.7%    | 0.2%   | 15.0%      | 51.9%           | 0.0159    | 1,228                       | 65                      | 6.3%  |
| Ten random people keep checking and trying things                   | 44.4%    | 9.5%   | 1.0%       | 35.4%           | 0.0252    | 1,246                       | 96                      | 8.1%  |
| Everyone tries something new 1% of the time, unchecked              | 37.5%    | 10.5%  | 0          | 31.4%           | 0.0290    | 0                           | 108                     | 36.7% |
| The ladder, checks return at each change, nobody tries anything new | 0        | 100.0% | 0          | 17.0%           | 0.0416    | 0                           | none                    | 0     |
| The ladder, one check kept, nobody tries anything new               | 0        | 100.0% | 0          | 16.7%           | 0.0419    | 0                           | none                    | 0     |
| The ladder, checks retired after rollout                            | 0        | 100.0% | 0          | 16.9%           | 0.0417    | 0                           | none                    | 0     |
| Everyone at once, nobody ever checks                                | 0        | 100.0% | 0          | 20.0%           | 0.0394    | 0                           | none                    | 0     |

Its cost is per change. Each re-run spent between 11,707 and 11,804 check trials, about 59 per member, so per 100 rounds it cost 7,804, 3,924 and 1,967 trials at the three spacings, where the ten cost 1,607, 1,228 and 968. That is 4.9, 3.2 and 2.0 times the ten's checking. A re-run that started 50 rounds late followed as often (71.0%, 72.0% and 73.0%) and lost the time: at a change every 150 rounds it put 43.4% of member-rounds on the best way against the on-time re-run's 66.1%.

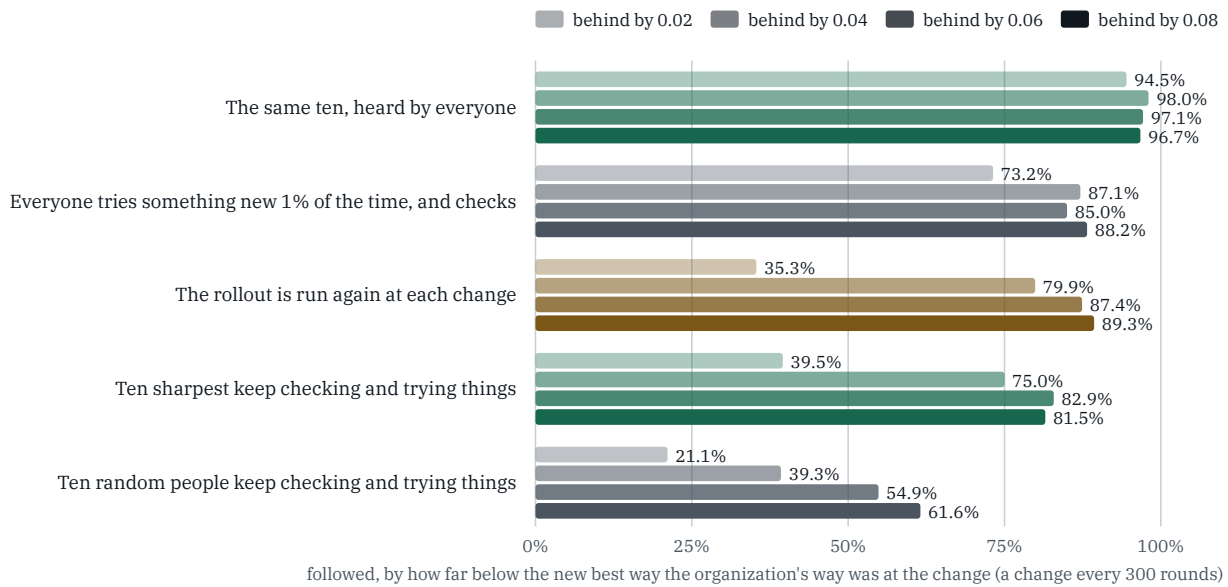


**Figure 14.** Small changes are the ones an organization misses: followed, by how far below the new best way the organization's way was when the world changed, at a change every 300 rounds. **Simulated; six changes and 400 runs; the counts are in Table 7.**

What the re-run missed were the small changes (Figure 14, Table 7). When the change left the organization's way 0.02 below the new best, one step of the scale, the re-run followed 33.6%, 35.3% and 39.3% of the time at the three spacings. When it left the organization 0.04 or more behind, at a change every 300 rounds the re-run followed 79.9%, 87.4% and 89.3% at gaps of 0.04, 0.06 and 0.08. The re-run's checks are the ladder's, made by every member with their own reading noise, and Who Gets It First found that an average member's five-try check let a proposed switch to a worse way through 37.4% of the time. The plans built on people who stay also miss small changes when time is short, and catch them when it is not: at a gap of 0.02 the ten followed 22.8%, 39.5% and 62.6% at the three spacings, and the ten heard by everyone 71.6%, 94.5% and 98.8%. A standing group keeps working on a small difference after a re-run has stopped.

**Table 7.** Followed, by how far below the new best way the organization's way was at the change, with the number of needed moves in brackets. Pooled over six changes and 400 runs.

| Plan and spacing                                        | Behind by<br>0.02 | 0.04        | 0.06        | 0.08        |
|---------------------------------------------------------|-------------------|-------------|-------------|-------------|
| The rollout is run again at each change, every 150      | 33.6% (566)       | 79.8% (574) | 87.5% (567) | 88.8% (564) |
| The rollout is run again at each change, every 300      | 35.3% (552)       | 79.9% (592) | 87.4% (556) | 89.3% (572) |
| The rollout is run again at each change, every 600      | 39.3% (565)       | 81.3% (589) | 86.1% (554) | 86.1% (541) |
| Ten sharpest keep checking and trying things, every 150 | 22.8% (565)       | 50.2% (558) | 72.6% (532) | 73.1% (520) |
| Ten sharpest keep checking and trying things, every 300 | 39.5% (569)       | 75.0% (556) | 82.9% (555) | 81.5% (574) |

| Plan and spacing                                                   | Behind by<br>0.02 | 0.04        | 0.06        | 0.08        |
|--------------------------------------------------------------------|-------------------|-------------|-------------|-------------|
| Ten sharpest keep checking and trying things, every 600            | 62.6% (594)       | 87.8% (574) | 89.8% (569) | 87.2% (572) |
| The same ten, heard by everyone, every 150                         | 71.6% (592)       | 92.8% (596) | 91.3% (589) | 92.8% (556) |
| The same ten, heard by everyone, every 300                         | 94.5% (601)       | 98.0% (607) | 97.1% (595) | 96.7% (581) |
| The same ten, heard by everyone, every 600                         | 98.8% (604)       | 99.7% (603) | 99.2% (606) | 99.3% (585) |
| Everyone tries something new 1% of the time, and checks, every 150 | 42.9% (557)       | 70.3% (576) | 79.2% (559) | 80.9% (564) |
| Everyone tries something new 1% of the time, and checks, every 300 | 73.2% (578)       | 87.1% (590) | 85.0% (567) | 88.2% (576) |
| Everyone tries something new 1% of the time, and checks, every 600 | 88.7% (585)       | 93.1% (594) | 94.7% (619) | 93.6% (561) |
| Ten random people keep checking and trying things, every 150       | 11.7% (504)       | 23.5% (557) | 35.2% (495) | 47.9% (520) |
| Ten random people keep checking and trying things, every 300       | 21.1% (516)       | 39.3% (535) | 54.9% (525) | 61.6% (541) |
| Ten random people keep checking and trying things, every 600       | 39.9% (547)       | 58.4% (567) | 64.7% (539) | 67.5% (505) |

## 7.4 Unchecked exploration finds the best way and loses it

**With everyone trying something new now and then and nobody checking, the organization reached the new best way and then left it in up to half of the cases.**

Everyone trying a random way in 1% of their updates, with no check, followed 30.3%, 37.5% and 42.1% of needed moves at the three spacings and cost no check trials. Its trouble is that it does not hold what it finds. In 22.5%, 36.7% and 50.9% of needed moves most of the organization held the new best way at some point in the cycle and did not hold it when the cycle ended, and the longer the cycle, the more often that happened. With a standing one-try check added, followed rose to 68.4%, 83.4% and 92.5% and reached-then-left fell to 3.6%, 3.5% and 4.6%, at 4,299, 3,382 and 2,776 check trials per 100 rounds, the most of any standing plan. As in Who Gets It First, exploration is what finds the new way and the check is what keeps it. That trade-off has been simulated for a moving target before (March, 1991; Posen & Levinthal, 2012), with the caution that ongoing change is not by itself a case for more exploration.

## 7.5 Under the rule in Who Gets It First

The change in Who Gets It First was narrower: the best and second-best ways trade places, so the organization is never more than 0.02 behind. Repeated six times at 300 rounds, that rule gave the ten 40.6% followed and the ten heard by everyone 92.6%, where the single change in that note gave 37.8% and 92.7% (Table 8). The retired ladder followed 0 of 1,200 needed moves and still spent 50.0% of its member-rounds on the best way, because under this rule two ways take turns being best and the ladder sits on one of them. An organization that never looks can be right half the time if the world happens to swing back, which is a reason to measure whether it follows a change and not only how often it is right.

**Table 8.** The same six changes under the rule in Who Gets It First, where the best and second-best ways trade places every 300 rounds. 400 runs per row.

| Plan                                         | Followed               | Member-rounds on the best way | Check trials per 100 rounds | Median rounds to follow |
|----------------------------------------------|------------------------|-------------------------------|-----------------------------|-------------------------|
| Everyone at once, nobody ever checks         | 0 (0 of 1,728)         | 28.0%                         | 0                           | none                    |
| The ladder, checks retired after rollout     | 0 (0 of 1,200)         | 50.0%                         | 0                           | none                    |
| Ten sharpest keep checking and trying things | 40.6% (640 of 1,576)   | 48.9%                         | 828                         | 118                     |
| The same ten, heard by everyone              | 92.6% (2,088 of 2,256) | 64.8%                         | 1,056                       | 72                      |

## 8 What it means

### 8.1 For the three valleys

On spread, the simulation reproduces the auditors' case for dedicated capacity at the handover (the relationship managers of Section 2) in a setting with no market and no contract; it has no budgets or incentives, so it is silent on who has money at stake. Where the developers paid for the handover out of their own building time, the organization got nothing for it (at the default reach; Section 6.4). Where a small dedicated capacity paid, it returned several times its cost. And the most effective carry slot for slot was the receivers' own attention, which the model prices at nothing and which the audit record and the hospital study say cannot be assumed. That adds a term to the position of Who Gets It First (Atkinson, 2026b). That note found that a few strong builders do more than everyone building only when they are given the capacity to build. They also need a carry, and it should not come out of the same capacity: a first group with a building budget and no carry budget is the first row of Table 2.

On value we add no evidence, because our models take the quality of a way of working as given. The nearest thing in them is a result of that note: with everyone holding the tool and trying each build once,

the organization gained 0.056 by round 200, against 0.279 when ten members did the same amount of building. That is wide use with little gain, and it comes from who builds and how carefully they check, not from the complements and the measurement that this valley is about. What the literature gives is a method and a warning. The method is to count deployment and outcome, not acquisition (Fichman & Kemerer, 1999), and to keep counting after the handover (GAO, 2013). The warning is that an early dip is what a general-purpose technology looks like when it is working (Brynjolfsson, Rock & Syverson, 2021), so a flat reading in the first year does not by itself say which side of the valley an organization is on.

On staying current, the firmest result, true by construction in the model, is that a rollout which ends, with nobody left trying anything new, cannot follow a change at all, however good its checks were, because a check only tests what somebody proposes. After that there are two workable arrangements with different costs. People who stay follow more of the changes the longer they have, catch small changes given time, and depend on being heard. A rollout run again on a trigger costs a fixed amount per change, is fast, does best when changes come often, misses small changes, and has to be told that a change has happened. A vendor's retirement notice is such a trigger and gives about two months. A change in the work gives none. The nearest thing to both that we ran, the ten who stay with everyone checking once for 75 rounds after each change, added a few points to the ten alone.

## 8.2 The roles, reworked

|                                                                                                                                                                                                                      |                                                                                                                                                                                                            |                                                                                                                                                                                       |                                                                                                                                                                                                     |                                                                                                                                                                                                                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Champion</b><br/>the standard model</p> <p>In the single valley<br/>conceptualizes the idea and makes the business case</p> <p>In the three valleys<br/>the builder everyone can hear, not only their team</p> | <p><b>Sponsor</b><br/>the standard model</p> <p>In the single valley<br/>supplies resources to show it is viable</p> <p>In the three valleys<br/>pays for the carry, and not out of the builders' time</p> | <p><b>Gatekeeper</b><br/>the standard model</p> <p>In the single valley<br/>sets criteria and decides</p> <p>In the three valleys<br/>accepts only work the receivers could check</p> | <p><b>Carrier</b><br/>added here</p> <p>In the single valley<br/>relationship managers in GAO's four companies</p> <p>In the three valleys<br/>a few people whose job is taking the work across</p> | <p><b>Checker who stays</b><br/>added here</p> <p>In the single valley<br/>nobody: the standard model ends when the product ships</p> <p>In the three valleys<br/>keeps checking and trying things after the rollout</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**Figure 15.** The three roles of the standard model as they read in the three valleys, and two roles the standard model does not have, added here; the carrier has precedents in change agents, opinion leaders and boundary spanners, and the checker who stays in Tyre and Orlikowski's episodes of renewed adaptation (Sections 4.3 and 4.4). **A proposal, with the sections it rests on in the text; no data.**

Markham and colleagues' three roles carry a project across one valley. In three valleys they read differently, and two are missing (Figure 15). The champion is the builder everyone can hear; in Who Gets It First an audience for the builders did more than an order to adopt their work. The sponsor pays for the carry, and not out of the builders' time (Section 6.1; in the model the exception is a very large audience, Section 6.4). The gatekeeper accepts work that the receivers have been able to check: a carry in which each member tried what they heard before switching harmed no runs even when the builders were chosen at random (Section 6.6), where the unchecked order in that note harmed 8.2% of them. Our position is that two roles are new to the valley model. The first is the carrier, a few people whose job is the handover. It has precedents in the relationship managers that GAO found on both sides of the handover in its four companies, and in change agents and boundary spanners (Section 4.4). In the

model five carriers raised net gain per head from 0.279 to 0.418. The second is the checker who stays. The standard model has no such role, because it ends when the product ships. In the loop, the plans without one followed none of 23,752 needed moves (by construction).

These are readings of a model, and the roles are a proposal. We did not study how organizations staff or fund any of them.

8.3 What to measure to know which valley you are in

Each valley can be hidden by a count that is easy to collect and shown by one that is not (Table 9). The counts in the middle column are the ones in most adoption reports. The counts on the right are, in each case, what an earlier study recommended or what our simulations had to count to see the valley at all. The last column gives each row's status for real organizations.

Table 9. Counts that hide each valley and counts that show it.

| Valley          | A count that hides it                                   | A count that shows it                                                                                                                                               | Why                                                                                                                                                                                                                                          | Status                                     |
|-----------------|---------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|
| Spread          | Builds, pilots and demonstrations completed             | Teams running something another team built; the people and hours assigned to carrying it; the time from first use in one team to use in a second                    | The builders' output is not the constraint (6.1); the carry is, and it has to be somebody's time (6.2, 6.3)                                                                                                                                  | Conjecture; inferred in the model          |
| Value           | Seats, active users, hours reported saved               | An outcome the organization already tracked, before and after, against a comparison group; deployment in the workflow and not acquisition                           | Acquisition and deployment separate (Fichman & Kemerer, 1999); an evaluation in which users reported time savings found no robust evidence of improved productivity, which was not a key aim of it (Department for Business and Trade, 2025) | Conjecture; inferred from others' evidence |
| Staying current | Accuracy at launch; how often the organization is right | At the last model change, the share of workflows that needed to move and did; the date of the last check that failed something; who is assigned to try alternatives | A plan nobody revisits can be right half the time without following anything (7.5); followed is the count that separates the plans (7.1 to 7.3)                                                                                              | Conjecture; inferred in the model          |

9 What contradicts it

The three-valley picture has real evidence against it, and some of our own results cut against the way the author first stated it.

**People adopt fast.** Relative to each technology's first mass-market product, work adoption of generative AI has been faster than the personal computer's, and 27% of employed respondents were using it for work in a given week by late 2024 (Bick, Blandin & Deming, 2026, abstract). A claim that adoption is slow is false for individuals. The valleys are about organizations, and the paper's claims stop there.

**Gains on tasks are real and measured.** In a preregistered experiment with 453 college-educated professionals, access to ChatGPT cut the average time taken on a writing task by 40% and raised output quality by 18% (Noy & Zhang, 2023, abstract). The second valley is not a claim that the tool does not work.

**The second valley may be a delay.** The J-curve reads flat early productivity as investment in complements that will pay later. The S&P 500 study's J in profitability is a different pattern: it runs across firms at different levels of adoption at one time, not through time within a firm, so it does not by itself support the delay reading (Yu et al., 2026). If that reading is right, the second valley is a dip to be funded through, not a death. The two readings predict different things for an organization that measures outcomes over time, and nothing for one that counts seats, which is the argument of Table 9.

**The first valley can be crossed with structure and money.** By 2013 the department that the 2006 audit criticized had twenty transition programs, which reported a transition rate of over 70 percent (GAO, 2013), and six of the seven recommendation items from that audit are closed as implemented. The one that is not is the one about who holds the money.

**Some projects should die.** Branscomb and Auerswald (2002) preferred the image of a Darwinian sea to a valley: the gap is full of contending ideas, and only some of them survive. An organization's pilots are such a sea, and a low rate of spread can be selection working. Our carry model has a version of this: a build is as likely to be worse as better, and the checks screen what spreads.

**Our own results.** The author's position going in was people first. In the loop a process did as well: a rollout run again at each change followed more often than ten people who stay when changes came every 150 rounds (72.4% against 54.0%) and kept more of the work on the best way than any other plan at that spacing. In the carry, the receivers' attention outdid dedicated carriers slot for slot, and where the builders had a large room their own time was enough. And others had already described a loop: Mollick's lab tests again when a new system arrives (Section 4.4).

## 10 Limits

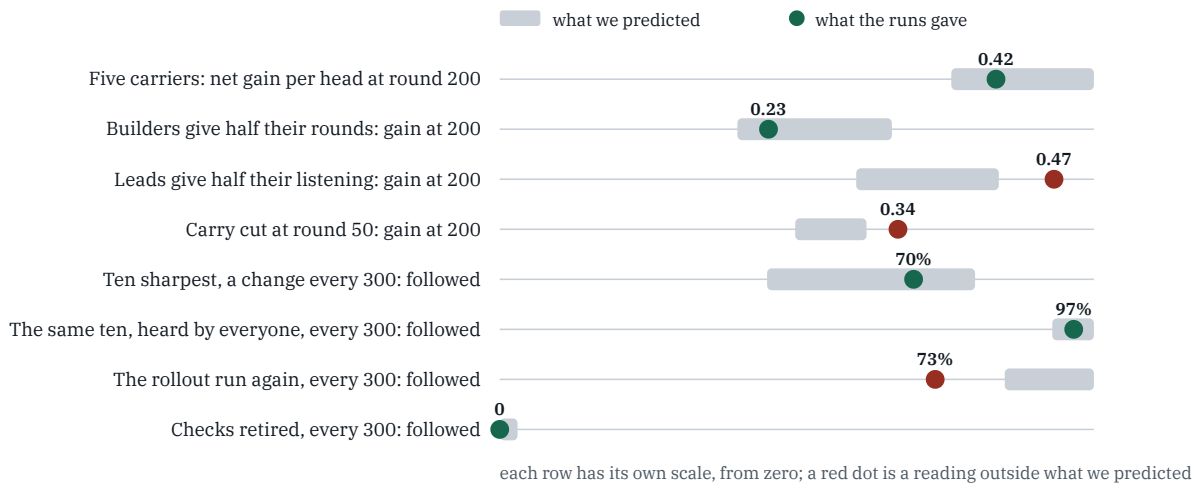


**Figure 16.** Eight of the predictions written into the scripts before the runs, against what the runs gave; each of the eight named a range. A reading outside the range is marked the same way whether the prediction was scored as held in part or as failed. **Simulated; predictions as written in the scripts.**

These are simulations of a simplified model, not of people or of any organization. The reading of spread as hearing slots, and of a carrier as someone who fills them, is ours. A real handover is more than being heard: the work has to be explained, adapted and supported, and none of that is in the model. A carrier relays perfectly unless a run says otherwise, the receivers can try what they hear at no cost, and a lead's attention is free. The only forgetting we ran is members sliding back to the old way one at a time, which the organization repairs by ordinary hearing; a whole team dropping a way of working, or other work competing for the same hearing, is not in the model, and either would make a cut cost more. In the loop the changes are a re-deal of five fixed qualities and no new way of working ever appears; the re-run is told when the world has changed; and the gaps are 0.02 to 0.08 on a scale we chose. Members differ only in how accurately they read results. They are honest, have no incentives, and do not lose skill, which matters here because loss of skill in the people who check is one of the third valley's own mechanisms (Section 4.3); a model with it would follow less. Apart from the carry's reach, no setting was varied: the hearing split, the leads' weight, the deference, the noise, the build spread and the exploration rates are those of Who Gets It First (Atkinson, 2026b), and whether the order of the plans survives other values is untested. A build cannot be abandoned, so the model shows the first valley as a delay, never a death. Rounds are rounds of the model, and we have not mapped them to the months of a vendor's retirement schedule. Counts are not priced: a check trial, a carrier, a lead's attention and a round on a worse way are in different units, and which plan is cheaper depends on prices the model does not have. Each configuration has 400 runs. The second valley is not simulated at all.

The review of earlier work is not systematic, and most of the journal articles were read in abstract only; the disclosure says which, and marks each number taken from one of them. The figures reported by others are theirs: the S&P 500 shares are a language model's scoring of what filings say, the survey shares come from different questions, the endoscopy study is observational, and the retirement record

is one vendor's. The three valleys are the author's proposal, and nothing here shows that organizations fail in these three ways and no others.

Of the 25 predictions written before the runs, 15 held, 8 held in part and 2 failed (Figure 16). The full list and the scoring rule are not reproduced here; below are the failures and five partial results. Both failures were in the loop. We predicted that a rollout run again at each change would follow at least 85% of needed moves and it followed 72.4% to 73.3%; and we predicted that one check by everyone for 75 rounds would add at least 10 points to the ten who stay at a change every 150 rounds, and it added 6.3. In the carry four partial misses ran the same way: more carry kept buying more than we expected. Ten carriers were still above five on both readings, a reach of 128 was well above a reach of 32, the leads' attention came in above our range, and a cut at round 50 left the organization 0.057 above never funded where we predicted a difference of at most 0.03. The fifth was in the sliding-back runs: we predicted that 2% of members sliding back each round would cost the funded organization between 0.03 and 0.12 by round 200, and it cost 0.012, within the noise between like rows of Table 4 (Section 6.5). Our code is not public, so no one else has reproduced these results, and none of this work has been externally peer reviewed.

**Disclosure.** Drafting and literature synthesis were assisted by AI models (Claude, Anthropic) working under the author's direction; the author is responsible for the content. The experiments were designed under the author's direction and written and run by an AI agent (Claude, Anthropic, in Claude Code). The works cited were read by AI reading agents working under the author's direction, to the depth stated below, and quotations and numbers taken from them were checked against the saved text. Every simulation number in the text and tables is filled in by name from the run records when the paper is built, and every chart of simulated results is drawn from those records by a script; the plotted values are recorded beside the scripts. Figures 1, 2, 7 and 15 are drawings. Figures 3 to 6 show numbers reported by the sources named in their captions, typed from those sources and recorded with them. The reports of the Government Accountability Office (2006, 2015) were read in their published PDFs for the pages cited, the 2013 report in its highlights page, and the 2006 report's recommendation status on the office's own page on 3 October 2026. Allen (2026), Weill, Woerner and Sebastian (2024), Mollick (2025) and the Department of Energy's page were read at their own pages; Anthropic's deprecations page was read whole, and its eleven rows were recomputed by a script. Yu et al. (2026) was read at its arXiv PDF for its abstract, method and Table 12, Legara, Jose and Martinez (2026) for their abstract and readiness table, and Branscomb and Auerswald (2002) at its public PDF for the three passages cited. The following were read in abstract only, from their records at OpenAlex, arXiv, PubMed or the NBER, with titles, authors and venues matched: Agrawal, Gans and Goldfarb (2024), Ahuja (2026), Bick, Blandin and Deming (2026), Bonney et al. (2024), Brynjolfsson, Rock and Syverson (2021), Budzyń et al. (2025), Chen, Zaharia and Zou (2024), Cooper and Zmud (1990), Costa (2026), Edmondson, Bohmer and Pisano (2001), Endsley and Kiris (1995), Fichman and Kemerer (1999), Klein and Sorra (1996), Levitt and March (1988), Markham et al. (2010), McElheran et al. (2024), Meyer and Rowan (1977), Noy and Zhang (2023), Parasuraman and Manzey (2010), Repenning (2002), Repenning and Serman (2001) and Tyre and Orlikowski (1994); Auerswald and Branscomb (2003) was read in abstract on its publisher's page. The Department for Business and Trade (2025), Valentini, Hamann and Dorigo (2014) and Soma et al. (2024) were read as described in our earlier notes. A number from a work read in abstract only is

the one its abstract reports, marked "abstract". The works added after the review of 7 October 2026 (Szulanski; March; Posen and Levinthal; Lazer and Friedman; Rogers; Valente; T. J. Allen; Tushman) are cited as prior art at concept level, and no number is taken from them. **Competing interests.** The author owns Wolfberg LLC. The model retirement record in Section 4.3 is Anthropic's, and Anthropic is the vendor of the AI system used to run the experiments and to help draft this note. The work was not funded. **Code and data.** The simulation code, the run log (one record per configuration, with parameters, seeds, metrics and the script's hash) and the figure data are available from the author on request.

## References

- Agrawal, A., Gans, J. S., & Goldfarb, A. (2024). Artificial intelligence adoption and system-wide change. *Journal of Economics & Management Strategy*, 33(2), 327-337.  
<https://doi.org/10.1111/jems.12521>
- Ahuja, P. (2026). *The checking problem: What must be true before AI ships in a regulated firm*. arXiv:2607.28666. [arxiv.org/abs/2607.28666](https://arxiv.org/abs/2607.28666)
- Allen, J. S. (2026, April 3). *Monitoring AI adoption in the US economy*. FEDS Notes, Board of Governors of the Federal Reserve System. [federalreserve.gov/econres/notes/feds-notes/monitoring-ai-adoption-in-the-u-s-economy-20260403.html](https://www.federalreserve.gov/econres/notes/feds-notes/monitoring-ai-adoption-in-the-u-s-economy-20260403.html)
- Allen, T. J. (1977). *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. MIT Press.
- Anthropic. (2026). *Model deprecations*. Claude platform documentation, read 3 October 2026. [platform.claude.com/docs/en/about-claude/model-deprecations](https://platform.claude.com/docs/en/about-claude/model-deprecations)
- Atkinson, B. (2026a). *The fine print on the hive mind: Size, scouting and honesty in a swarm of copiers*. Wolfberg research note, 30 September 2026. [wolfberg.ai/papers/when-the-hive-gets-it-wrong](https://wolfberg.ai/papers/when-the-hive-gets-it-wrong)
- Atkinson, B. (2026b). *Who gets it first: Rollout order and checking in a simulated organization*. Wolfberg research note, 3 October 2026. [wolfberg.ai/papers/who-gets-it-first](https://wolfberg.ai/papers/who-gets-it-first)
- Auerswald, P. E., & Branscomb, L. M. (2003). Valleys of death and Darwinian seas: Financing the invention to innovation transition in the United States. *The Journal of Technology Transfer*, 28(3-4), 227-239. <https://doi.org/10.1023/A:1024980525678>
- Bick, A., Blandin, A., & Deming, D. J. (2026). The rapid adoption of generative AI. *Management Science*. <https://doi.org/10.1287/mnsc.2025.02523>

- Bonney, K., Breaux, C., Buffington, C., Dinlersoz, E., Foster, L. S., Goldschlag, N., Haltiwanger, J. C., Kroff, Z., & Savage, K. (2024). *Tracking firm use of AI in real time: A snapshot from the Business Trends and Outlook Survey* (NBER Working Paper 32319). National Bureau of Economic Research. [nber.org/papers/w32319](https://nber.org/papers/w32319)
- Branscomb, L. M., & Auerswald, P. E. (2002). *Between invention and innovation: An analysis of funding for early-stage technology development* (NIST GCR 02-841). National Institute of Standards and Technology. [nist.gov/system/files/documents/2017/05/09/gcr02-841.pdf](https://nist.gov/system/files/documents/2017/05/09/gcr02-841.pdf)
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333-372. <https://doi.org/10.1257/mac.20180386>
- Budzyń, K., Romańczyk, M., Kitala, D., Kołodziej, P., Bugajski, M., Adami, H. O., et al. (2025). Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: A multicentre, observational study. *The Lancet Gastroenterology & Hepatology*, 10(10), 896-903. [https://doi.org/10.1016/S2468-1253\(25\)00133-5](https://doi.org/10.1016/S2468-1253(25)00133-5)
- Chen, L., Zaharia, M., & Zou, J. (2024). How is ChatGPT's behavior changing over time? *Harvard Data Science Review*, 6(2). <https://doi.org/10.1162/99608f92.5317da47>
- Cooper, R. B., & Zmud, R. W. (1990). Information technology implementation research: A technological diffusion approach. *Management Science*, 36(2), 123-139. <https://doi.org/10.1287/mnsc.36.2.123>
- Costa, F. F. (2026). *The deployment wall: A diagnostic framework and instrument for enterprise AI in the deployment era*. arXiv:2607.29089. [arxiv.org/abs/2607.29089](https://arxiv.org/abs/2607.29089)
- Department for Business and Trade. (2025, August 28). *Microsoft 365 Copilot pilot: DBT evaluation report*. GOV.UK. [gov.uk/government/publications/microsoft-365-copilot-pilot-dbt-evaluation-report](https://gov.uk/government/publications/microsoft-365-copilot-pilot-dbt-evaluation-report)
- Edmondson, A. C., Bohmer, R. M. J., & Pisano, G. P. (2001). Disrupted routines: Team learning and new technology implementation in hospitals. *Administrative Science Quarterly*, 46(4), 685-716. <https://doi.org/10.2307/3094828>
- Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381-394. <https://doi.org/10.1518/001872095779064555>
- Fichman, R. G., & Kemerer, C. F. (1999). The illusory diffusion of innovation: An examination of assimilation gaps. *Information Systems Research*, 10(3), 255-275. <https://doi.org/10.1287/isre.10.3.255>
- Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *Academy of Management Review*, 21(4), 1055-1080. <https://doi.org/10.5465/amr.1996.9704071863>

- Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4), 667-694. <https://doi.org/10.2189/asqu.52.4.667>
- Legara, E. F., Jose, E. D., & Martinez, P. J. (2026). *Beyond model readiness: Institutional readiness for AI deployment in public systems*. arXiv:2605.17203. [arxiv.org/abs/2605.17203](https://arxiv.org/abs/2605.17203)
- Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319-338. <https://doi.org/10.1146/annurev.so.14.080188.001535>
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71-87. <https://doi.org/10.1287/orsc.2.1.71>
- Markham, S. K., Ward, S. J., Aiman-Smith, L., & Kingon, A. I. (2010). The valley of death as context for role theory in product innovation. *Journal of Product Innovation Management*, 27(3), 402-417. <https://doi.org/10.1111/j.1540-5885.2010.00724.x>
- McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. J. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, 33(2), 375-415. <https://doi.org/10.1111/jems.12576>
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340-363. <https://doi.org/10.1086/226550>
- Mollick, E. (2025, May 22). *Making AI work: Leadership, lab, and crowd*. One Useful Thing. [oneusefulthing.org/p/making-ai-work-leadership-lab-and](https://oneusefulthing.org/p/making-ai-work-leadership-lab-and)
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192. <https://doi.org/10.1126/science.adh2586>
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381-410. <https://doi.org/10.1177/0018720810376055>
- Posen, H. E., & Levinthal, D. A. (2012). Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science*, 58(3), 587-601. <https://doi.org/10.1287/mnsc.1110.1420>
- Repenning, N. P. (2002). A simulation-based approach to understanding the dynamics of innovation implementation. *Organization Science*, 13(2), 109-127. <https://doi.org/10.1287/orsc.13.2.109.535>
- Repenning, N. P., & Sterman, J. D. (2001). Nobody ever gets credit for fixing problems that never happened: Creating and sustaining process improvement. *California Management Review*, 43(4), 64-88. <https://doi.org/10.2307/41166101>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

- Soma, K., Bouteiller, Y., Hamann, H., & Beltrame, G. (2024). *The hive mind is a single reinforcement learning agent*. arXiv:2410.17517. [arxiv.org/abs/2410.17517](https://arxiv.org/abs/2410.17517)
- Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17(S2), 27-43.  
<https://doi.org/10.1002/smj.4250171105>
- Tushman, M. L. (1977). Special boundary roles in the innovation process. *Administrative Science Quarterly*, 22(4), 587-605. <https://doi.org/10.2307/2392402>
- Tyre, M. J., & Orlikowski, W. J. (1994). Windows of opportunity: Temporal patterns of technological adaptation in organizations. *Organization Science*, 5(1), 98-118.  
<https://doi.org/10.1287/orsc.5.1.98>
- U.S. Department of Energy. (n.d.). *Adoption Readiness Levels (ARL) framework*. Read 3 October 2026. [energy.gov/technologytransitions/adoption-readiness-levels-arl-complement-trl](https://energy.gov/technologytransitions/adoption-readiness-levels-arl-complement-trl)
- U.S. Government Accountability Office. (2006). *Best practices: Stronger practices needed to improve DOD technology transition processes* (GAO-06-883). [gao.gov/products/gao-06-883](https://gao.gov/products/gao-06-883)
- U.S. Government Accountability Office. (2013). *Defense technology development: Technology transition programs support military users, but opportunities exist to improve measurement of outcomes* (GAO-13-286). [gao.gov/products/gao-13-286](https://gao.gov/products/gao-13-286)
- U.S. Government Accountability Office. (2015). *Defense Advanced Research Projects Agency: Key factors drive transition of technologies, but better training and data dissemination can increase success* (GAO-16-5). [gao.gov/products/gao-16-5](https://gao.gov/products/gao-16-5)
- Valente, T. W. (2012). Network interventions. *Science*, 337(6090), 49-53.  
<https://doi.org/10.1126/science.1217330>
- Valentini, G., Hamann, H., & Dorigo, M. (2014). Self-organized collective decision making: The weighted voter model. In *Proceedings of AAMAS 2014* (pp. 45-52).
- Weill, P., Woerner, S. L., & Sebastian, I. M. (2024, December 19). *Building enterprise AI maturity* (Research Briefing XXIV-12). MIT Center for Information Systems Research.  
[cisr.mit.edu/publication/2024\\_1201\\_EnterpriseAIMaturityModel\\_WeillWoernerSebastian](https://cisr.mit.edu/publication/2024_1201_EnterpriseAIMaturityModel_WeillWoernerSebastian)
- Yu, Y., Fleming, M., Hampton, L., Combemale, C., & Thompson, N. (2026). *AI adoption in S&P 500 firms*. arXiv:2607.08920. [arxiv.org/abs/2607.08920](https://arxiv.org/abs/2607.08920)