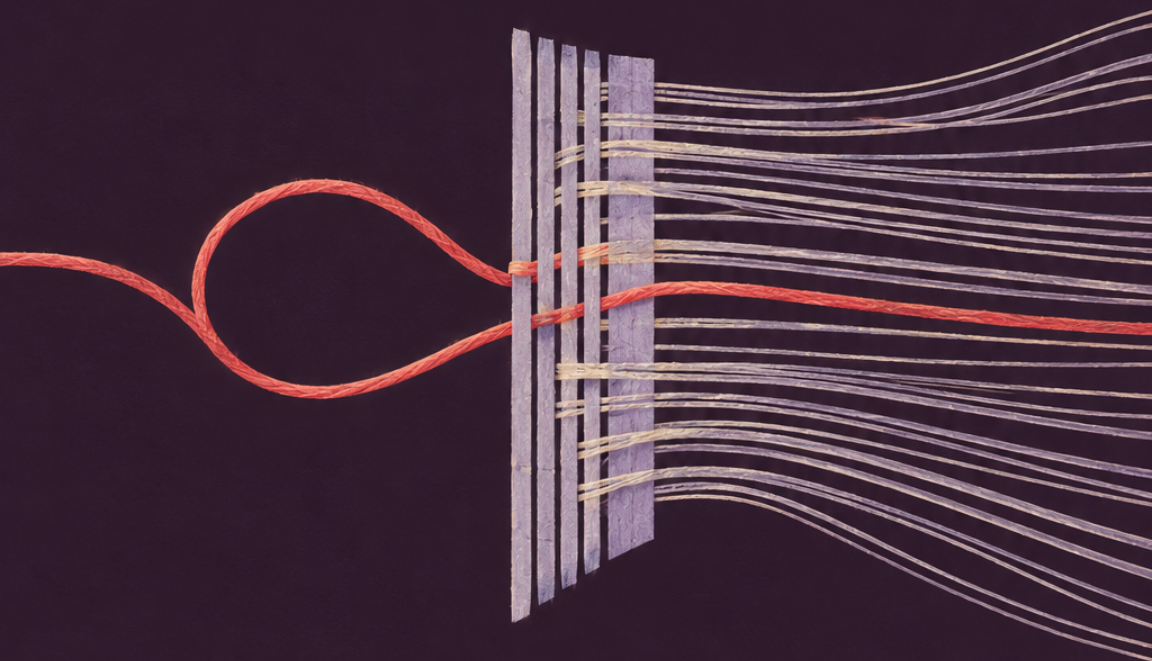


THE WORK YOU KEEP

Human Attention in an Age of AI



JOHN RECTOR

The Work You Keep

Human Attention in an Age of AI

John Rector

First digital edition

October 2026

johnrector.me

About this edition

This is an original short book, written and researched with AI assistance. Its argument, exercises, and operating methods are proposals for readers to examine, adapt, and test. They are not a validated management system. Research is identified in the text and collected in the source notes. Evidence was checked on October 28, 2026; the book makes no claim to survey every study published by that date.

Harbor Studio, its owner Mara, and all scenes involving its staff or clients are invented teaching examples. Their dialogue, results, and numbers illustrate decisions; they do not report interviews, fieldwork, or measured business outcomes. Other hypothetical examples are identified as such. No private client records were used.

The book is published by John Rector at johnrector.me. This edition has no ISBN. The online, PDF, and EPUB editions contain the same main text. The online edition also provides navigation and reading aids.

Contents

Introduction: An afternoon returned	5
Chapter 1: Finished is not settled	8
Chapter 2: Choose the unit	14
Chapter 3: A promise before a prompt	20
Chapter 4: Read at the joints	26
Chapter 5: Keep a practice	32
Chapter 6: Leave a return path	38
Chapter 7: Give time somewhere to go	44
Chapter 8: The team gets a vote	50
Chapter 9: Measure the work you actually want	56
Chapter 10: Thirty days of different work	62
Afterword: A life larger than its output	69
Toolkit: Six worksheets to make your own	72
Source notes: What the evidence supports	79
Closing page	86

Introduction: An afternoon returned

Imagine finishing the work you meant to finish at two in the afternoon.

The report has been checked. The customer has an answer. Tomorrow's appointment is on the calendar. The first response is pleasure. Then comes a small uncertainty: what is the afternoon for?

You could make another report. You could improve the first one until its improvements disappear into each other. You could open a dashboard and look for a number that would authorize stopping. Or you could leave the desk.

AI makes this question easier to encounter. A system that drafts, searches, sorts, schedules, and sometimes acts can change the amount of human labor a task requires. It can also change the shape of that labor. A person who once made something from the beginning may now spend the day choosing among things already made.

There is real opportunity in that change. There is also a decision hiding inside it. When you delegate a piece of work, what do you intend to keep?

An easy answer is control. But control can mean spending every hour inspecting the tool. An equally easy answer is creativity. But creativity can become another production quota. A third answer is responsibility, which sounds admirable until someone is assigned responsibility without the time, information, or authority to exercise it.

This book proposes a more practical answer. Keep the ability to understand the purpose of the work, judge its consequential parts, practice the skills on which that judgment depends, and decide where the resulting capacity will go. Those abilities give attention somewhere to stand.

Attention here means more than concentration. It includes noticing what matters, staying with a person, recognizing when a result is inadequate, and choosing to leave a completed task alone. Some of these activities earn money. Some make the money worth earning. A useful technology should be evaluated in relation to both.

The research offers no single number that settles whether AI helps everyone. A large customer-support study found productivity gains. A smaller experiment involving experienced software developers found a slowdown with the tools and tasks studied. Later evidence changed parts of the picture and raised new questions about measurement. These are reasons to investigate the circumstances of work, rather than choose a permanent team in an argument about technology.

The chapters that follow move from an individual task to a working life. First we distinguish a finished output from a settled obligation. Then we examine how to divide work, write a useful delegation, review a result, preserve practice, recover from

mistakes, and distribute the benefits. The last chapters turn the argument into a month of modest experiments.

You do not need an autonomous agent to use these ideas. A text assistant is enough. You can also use them when delegating to another person, although people bring needs, rights, relationships, and understanding that software does not. Similarities in workflow should never erase those differences.

Read the book in order if you are rethinking your whole approach. If an immediate problem brought you here, begin with the chapter that names it. The tools at the end are deliberately short enough to copy into a notebook. Their value will come from what they help you see, not from filling every box.

The aim is a form of competence that can welcome help. You should be able to use a powerful system without making yourself smaller around it. You should be able to accept a good result, recognize a bad one, and return to a life whose worth exceeds the amount of work it contains.

The afternoon is where we will eventually arrive. First we have to know what it means to be finished.

Finished is not settled

At Harbor Studio, a fictional four-person design business, Mara has begun using an AI assistant to prepare weekly client updates. The assistant reads a folder of project notes and produces a tidy account of progress, open questions, and next steps.

The first update takes seconds to appear. Before the change, preparing it took a substantial part of Mara's Friday. She sees a finished document and feels a finished obligation.

Then she reads the second paragraph. It says the client approved a revised layout. In the source notes, the client said the layout was promising and asked for another option. The assistant has converted encouragement into permission. The sentence is grammatically excellent. Its consequence is a schedule that now assumes approval.

Mara corrects it. A small mistake has cost only a few minutes. The larger discovery is that the document's appearance cannot tell her whether the work has been settled.

Three endings

A task can have at least three endings. Production ends when an output exists. Acceptance ends when someone has enough evidence that the output meets the agreed standard. Obligation ends when the people affected have received what they were owed, or when the next responsibility has been clearly transferred.

These endings sometimes occur together. A correctly set personal timer needs little ceremony. In other work they are far apart. A draft can exist while its sources remain unchecked. A correct answer can sit unsent. A delivered design can leave an accessibility problem for someone else to discover.

When AI makes production fast, the distance between these endings becomes more visible. You may generate ten documents while accepting none of them. The screen fills with work that looks complete and still makes claims on your attention.

The useful unit to count is therefore an accepted outcome. For a client update, that may mean an accurate account, approved by its owner, delivered to the right people, with questions clearly assigned. The number of paragraphs produced tells us little about that outcome.

This distinction is the book's first proposal. It is a way to organize your own evaluation, not a research finding about every workplace. Its test is whether it exposes work that your current measures hide.

What the productivity evidence can say

Erik Brynjolfsson, Danielle Li, and Lindsey Raymond studied the introduction of a conversational assistant among 5,172 customer-support agents. Their published 2025 paper reports a 15 percent average increase in issues resolved per hour. Effects differed across workers; less experienced workers benefited more, and the most skilled workers showed small quality declines alongside small speed gains. The measure concerned resolved customer issues in one organizational setting. It did not establish a universal improvement rate for AI. [1]

That specificity is useful. It connects a tool to people, tasks, and an outcome. It gives us something more precise than the impression that a busy interface must be productive.

For Mara, the relevant question is how many accurate updates reach clients with an acceptable amount of effort and correction. A beautiful first draft contributes to that result. It does not stand in for it.

Count the work after the draft

For one recurring task, record five kinds of effort: preparing the inputs, directing the system, reviewing the output, correcting it, and dealing with consequences after delivery. Include the people who receive the work. A fast handoff that makes a colleague reconstruct your assumptions has moved labor across a desk.

In a hypothetical timing exercise, an update previously required sixty minutes. With AI, assembling materials takes ten, generation and direction take five, review takes fifteen, correction takes ten, and a later clarification takes five. The total

is forty-five minutes. The saving is fifteen minutes, or one quarter of the original effort. Counting only generation would tell a much more flattering story.

These invented numbers are arithmetic, not a forecast. Your result might be better, worse, or qualitatively different. Perhaps the same effort produces a clearer update. Perhaps a person who previously could not prepare one at all can now do so. Those benefits deserve recognition even when they do not appear as faster completion.

You also need a fair comparison. An unusually easy AI task and an unusually difficult manual task tell you very little. Compare similar work, note changes in scope, and record uncertainty. A few observations can guide a small local choice without becoming a claim about the whole economy.

The unfinished shelf

Mara creates a visible distinction between drafts and accepted work. Her team can experiment freely in a draft folder. The shared client record contains only work someone has checked. The distinction reduces a recurring ambiguity: whether a polished document is ready to use.

A folder name alone is insufficient. Each accepted update identifies its owner and the date of review. A later reader can find the source material behind a consequential claim. An unresolved question remains a question instead of acquiring a confident verb.

The team also limits the number of updates waiting for review. Production can outrun acceptance when a system produces more

than anyone has time to read. That queue is a liability. The documents may go stale, and the person receiving them may assume that quantity implies confidence.

Sometimes the right instruction is to produce fewer outputs. Ask for the two alternatives that represent a meaningful choice. Ask for the single paragraph that explains the disagreement. Ask for an outline while the evidence is still incomplete. Match the form of the output to the next human decision.

Enough evidence to stop

Acceptance should have an end. Otherwise the promise of efficiency becomes an invitation to endless revision. Before generating, define what would make the result adequate: the facts that must be right, the audience it must serve, the decision it must support, and the errors that would make it unusable.

Mara's update does not need to be the best update imaginable. It needs to tell the truth about progress and responsibility. Once those conditions are met, another round of phrasing may add less value than calling the client about the unresolved decision.

Stopping is easier when the standard was chosen before you saw the output. Afterward, attractive details can quietly change the assignment. You begin optimizing the thing the machine happened to make.

The first work you keep is the authority to say what completion means. Without it, you may become very efficient at extending a task.

Try this: Follow one obligation

Choose a recurring task small enough to observe this week. Write its three endings: output exists, output is accepted, obligation is settled. Name the evidence needed at each ending and the person who can provide it. Time the entire path once. Look for a handoff where work appears finished before anyone can rely on it. Repair that handoff before adding more automation.

Choose the unit

The next week, Mara asks the assistant to handle client communication. The request sounds like a natural extension of preparing updates. It also joins together many different activities: gathering facts, drafting words, deciding what can be promised, choosing recipients, sending messages, interpreting replies, and changing the project plan.

The broad instruction makes these differences hard to see. It treats the name of a responsibility as if it described a single task.

Most work is a bundle. We use a convenient noun, such as research, service, or administration, to hold together activities that have different failure costs and different reasons for involving a person. Delegating the noun can transfer much more than you intended.

Verbs before job titles

Start by listing verbs. For a client update: collect notes, identify changes, distinguish decisions from suggestions, draft an account, verify commitments, approve recipients, send, record delivery, and assign follow-up.

This list reveals possible boundaries. The assistant may collect and draft. A project owner verifies commitments. Sending may

be permitted only after approval. A reply that changes scope goes to a person with authority to renegotiate the project.

There is no universal allocation. A mature system with reliable records and narrow permissions may safely handle more. A new workflow with ambiguous notes may need close participation. The allocation should be explicit enough to change in response to evidence.

The International Labour Organization's 2025 exposure index analyzes occupations through their component tasks. Its authors judge job transformation more likely than full automation in many cases because occupations contain activities requiring human input. Exposure measures potential susceptibility; it does not directly count jobs that will disappear. [2]

The same caution belongs inside a small business. A tool that can perform several activities associated with a job has not thereby demonstrated that it can carry every obligation of the role. Nor does this guarantee that the role will remain economically unchanged. Task analysis helps us ask the question at a useful scale.

Four questions for a boundary

For each activity, ask what happens if it is wrong, how easily the result can be checked, how readily its effects can be reversed, and whether doing it develops a capability you need to retain.

Consider collecting meeting dates from a set of messages. The result may be easy to verify against the messages. An error in a private planning document is usually reversible. Now consider sending a commitment that a project will be delivered by Friday.

The sentence may be easy to type and difficult to honor. Its apparent simplicity says little about its consequence.

The fourth question is easy to overlook. An activity can be low risk and still be valuable practice. A junior designer's first attempt at describing a problem may help develop the judgment needed to evaluate later designs. Removing every first attempt could improve this week's polish while weakening next year's reviewer.

Conversely, an activity can be unpleasant and contribute little to learning. Reformatting the same information for the fifth time may be a good candidate for delegation. Difficulty and developmental value are separate properties. So are enjoyment and necessity.

A boundary is a sentence

Write the allocation in ordinary language. For example: the assistant may prepare a proposed update from the approved project folder; Mara verifies client decisions and delivery commitments; nothing is sent until she approves the final message and recipient list.

This sentence describes an operational boundary. Everyone can point to the place where the assistant's work ends and a human decision begins. It also names the input source, which matters because a system may otherwise combine current notes with old assumptions.

Avoid boundaries that depend on the system recognizing every important situation. An instruction to ask whenever something is sensitive sounds sensible, but it leaves the definition of

sensitivity inside the system you are trying to constrain. Name concrete triggers: an added recipient, a changed deadline, a missing approval, a request outside the agreed scope.

Use permissions to support the sentence. If the assistant should prepare drafts, it does not need unrestricted sending access. If it should summarize a selected folder, it does not need access to every file the business owns. A written limit is easier to trust when the available actions match it.

The smallest useful experiment

Mara chooses one kind of communication for a trial: the weekly update on a single project. She keeps other messages in the existing workflow. This makes comparison possible. It also contains the repair work if the trial fails.

A small experiment should still include a whole useful outcome. Testing whether an assistant can write an introductory sentence is too narrow if the intended benefit depends on reading notes, identifying decisions, and preserving references. Choose the smallest end-to-end unit that can reveal the important difficulties.

Record the conditions. Which notes were supplied? Which version of the instructions was used? Who reviewed the result? What changed in the project? These details let you distinguish a tool problem from an input problem and an unclear standard from a failed output.

Do not expand solely because the demonstration was impressive. Demonstrations are often selected for clarity. The ordinary week contains incomplete notes, contradictory messages, people on

vacation, and requests that do not fit the expected form. Expansion should follow encounters with representative messiness.

Who has to live with the boundary?

A workflow can look efficient from the delegator's chair and exhausting from the receiver's. If a colleague must approve every small action, you have appointed a permanent gatekeeper. If a client must repeatedly correct the system, you have recruited the client as a tester.

Ask the people downstream whether the proposed division makes sense. They may know where ambiguity actually enters. They may prefer a slightly slower handoff that preserves context. Their experience is part of the task design, rather than feedback to gather after the design is complete.

For work done alone, imagine your future self as the receiver. Tomorrow you will have forgotten part of today's context. A usable boundary includes enough record for that person to understand what happened.

The unit you choose determines what you can learn. A vague experiment produces a vague verdict. A well-bounded one can tell you exactly which activity deserves more help, more evidence, or more human practice.

Try this: Unpack one noun

Write down a responsibility you have considered automating. Replace its noun with six to ten verbs. Beside each verb, note consequence, checkability, reversibility, and practice value. Choose one complete sequence to test. End with a sentence that names what the system may do, what requires review, and what event stops the sequence.

A promise before a prompt

Mara's first instruction to the assistant is a style request: make the update concise, professional, and friendly. The assistant follows it. Almost any client update could satisfy those adjectives.

Her second instruction begins elsewhere. The client needs to know which decisions are required to keep the project moving. The update must distinguish completed work from work waiting on approval. It must avoid creating commitments that the team has not accepted.

The second instruction is easier to evaluate because it names a promise. Style can help fulfill that promise. It cannot define it.

When a system makes production cheap, instructions become a place where our own uncertainty is exposed. We may discover that we have been performing a ritual without knowing the outcome it serves. Writing a better prompt can help, but the deeper work is deciding what we are asking another party to rely on.

Begin with the receiver

A useful delegation begins with a person and a situation. Who will use the result? What will they be trying to do? What misunderstanding would inconvenience or harm them? What do they already know?

For the studio's client, a progress update is part of coordinating a launch. It should make decisions visible early enough to act on them. A summary that makes the team look busy while burying the blocked decision fails its purpose, even if every sentence is accurate.

For your personal research, the receiver may be you in a different state: tired, hurried, or returning after a week away. A good research note should preserve uncertainty in a form you can still understand then. It should help you make a decision without requiring you to reconstruct an entire conversation.

You can ask an assistant to help define the receiver's needs. Treat its answer as a hypothesis to compare with what actual people say and do. A plausible account of a customer's needs is still an account that needs evidence.

The brief on one page

Write six things before a substantial delegation: purpose, authorized materials, required output, acceptance conditions, permissions, and stopping conditions.

Purpose says what the result should enable. Authorized materials identify the sources the system may use and the sources it must not treat as current. Required output specifies a form appropriate to the next decision. Acceptance conditions

name the checks that make the output usable. Permissions describe available actions. Stopping conditions identify events that require a pause or return to a person.

These elements can fit in a paragraph. They need enough detail to remove consequential ambiguity, not enough detail to simulate every possible future.

For Mara, the brief reads: prepare this week's client update so the client can approve the next design decision. Use the dated project notes and decision log in the selected folder. Return a draft with completed work, unresolved decisions, and next actions. Every claimed approval must reference the decision log. Preserve uncertainty. Draft only. Stop and flag missing or conflicting approval records.

The brief makes one choice especially clear. The assistant is not being asked to turn an ambiguous project into a smooth narrative. It is being asked to expose the ambiguity that blocks progress.

Show an edge, not only an ideal

Examples help when words like concise or complete are too broad. Show a successful result and a nearby failure. For instance, the sentence “The client liked option B” reports a reaction. “The client approved option B for production” reports a decision. The brief can make explicit that one does not license the other.

An edge example is useful because many failures occur next to success. An obviously irrelevant answer is easy to discard. A

nearly correct claim can travel far because nobody feels a reason to stop it.

Include an example of what to do when information is missing. Ask for an uncertainty label and a question for the responsible person. Otherwise the pressure to fill the requested format can make empty boxes look like invitations to guess.

You should still check the result. A well-written brief changes the conditions of performance; it does not make errors impossible. It may also reveal that the source material cannot support the desired output. That is a successful discovery if it prevents a false commitment.

Permission should survive enthusiasm

A capable assistant may suggest useful adjacent work. After drafting an update, it may offer to send it, create tasks, revise the timeline, and notify the client. Some of those actions may be welcome. Each changes the scope of the delegation.

Decide which expansions can happen automatically and which require a fresh decision. Keep the decision close to the action and its consequence. Approval to generate a private draft should not silently become approval to contact a person.

The reverse problem also matters. An assistant that asks about every punctuation change can consume more attention than the work itself. Grant ordinary editorial discretion within a bounded draft. Reserve explicit approval for consequential changes. The goal is a useful distribution of authority.

In a team, permission comes from an authorized person or an agreed policy. It does not arise merely because an action is

technically possible. A connected account can expose a much larger surface than the task requires.

The acceptance conversation

Mara shares the new brief with the person who will review the update. That colleague notices a missing requirement: the update should name the client decision's effect on the schedule. The client does not need a list of abstract open questions. The client needs to know what waiting will change.

This is why briefing should occasionally be social. The person assigning the work and the person judging it may have different standards without realizing it. AI can make that disagreement faster and more expensive by producing many outputs under the wrong standard.

Agree on an adequate result before investing in elaborate automation. If two competent reviewers cannot explain why one result passes and another fails, the system will have trouble learning from their feedback. The problem may be a missing shared standard.

Some standards remain matters of taste. Acknowledge them. You can still name whose taste governs a decision and what purpose it serves. Clarity about subjective authority is more useful than pretending that every preference is an objective requirement.

Save the brief with the work

Keep the brief where a later reviewer can find it. Otherwise a result becomes detached from the promise it was meant to fulfill. Someone may judge a quick exploratory note as if it were a final recommendation, or reuse a draft prepared under assumptions that no longer apply.

Version the brief when an important assumption changes. You do not need a bureaucracy of tiny revisions. A short record such as “Now includes approved timeline; still excludes sending” can be enough.

The brief is also a way to learn about your own judgment. Over several attempts, look at what you repeatedly add during review. Those corrections may belong in the next brief. If they concern matters that cannot be specified in advance, preserve time for a person to exercise judgment at that point.

Delegation begins before the first output. It begins when you decide what another person should be able to trust.

Try this: Write the promise

Complete this sentence: “The person receiving this work should be able to ____ without having to ____.” Then write the six elements of the brief. Add one example of a near miss and one example of appropriate uncertainty. Ask the eventual reviewer whether the brief describes a result they would accept.

Read at the joints

Mara can read an update from beginning to end and still miss its most consequential error. A fluent paragraph invites a fluent reading. The transitions feel natural, and the mind supplies the continuity that the work appears to possess.

To review well, she changes the direction of her attention. She looks first at the places where one kind of thing becomes another: a source becomes a claim, a suggestion becomes a decision, a decision becomes a deadline, a deadline becomes a message to someone outside the team.

These are the joints of the work. They deserve a different kind of reading.

Confidence is not a citation

NIST's 2024 Generative AI Profile describes confabulation as confidently presented erroneous or false content. It recommends verifying sources and citations during evaluation and ongoing monitoring. The document is a voluntary risk-management resource, rather than a guarantee that any particular workflow is safe. [3]

The practical implication is straightforward: the tone of a claim cannot supply its missing evidence. An answer can sound tentative and be correct. It can sound exact and be unsupported.

Ask the system to preserve the route from evidence to conclusion. For a factual update, that may be a link to the relevant note. For a calculation, it may be inputs, units, and an independently reproducible operation. For a recommendation, it may be the assumptions and tradeoffs on which the choice depends.

An explanation generated by the system is itself material to evaluate. It may help you locate a source or spot an inconsistency. It does not prove that the system used the source correctly or that its account of its own process is faithful. Check the evidence that matters to the outcome.

Five joints worth inspecting

First, inspect names, dates, quantities, and units. These compact details can change a decision more than an entire page of general explanation. Check them against an appropriate source rather than another copy of the generated output.

Second, inspect changes in certainty. Words such as may, intends, approved, required, and completed carry different commitments. A summary often becomes misleading through a small change in one of these words.

Third, inspect omissions. A document can contain no false sentence and still create a false impression. Ask what important limitation, dissent, dependency, or excluded group would change the receiver's understanding.

Fourth, inspect the step from description to action. A pattern in historical data does not by itself authorize a new decision. A

recommendation may depend on values, permissions, or information absent from the source.

Fifth, inspect recipients and destinations. Work can be accurate and still go to the wrong person, become public unintentionally, or overwrite a record that someone needs. Delivery is part of correctness.

This list is a proposed review aid. Its purpose is to direct effort toward consequential transitions. Add the joints that matter in your field, and remove rituals that repeatedly contribute nothing.

Review is a skill with conditions

A 2025 survey by Hao-Ping Lee and colleagues collected responses from 319 knowledge workers and 936 examples of AI use. Greater confidence in AI was associated with less reported critical-thinking effort; greater task-specific confidence in one's own ability was associated with more. Participants described shifts toward verification, integration, and stewardship. The study measured self-reports and associations. It did not demonstrate that AI causes permanent loss of intelligence or skill. [4]

That limitation matters. A useful finding can become a misleading slogan if its method disappears. The study gives us reason to examine how people allocate effort. It does not license a verdict on every user.

For Mara, the question is whether reviewers have what they need to judge the update. Are the sources accessible? Is there enough time? Does the reviewer understand the project? Can the

reviewer reject the draft without being punished for delaying delivery?

A nominal human check adds little when these conditions are absent. The presence of a person at the final button does not establish meaningful oversight.

Read before you polish

Separate consequential review from stylistic improvement. A pleasing edit can make an unsupported claim feel more finished. It also spends the reviewer's energy before the hardest questions have been answered.

Mara first checks commitments, approvals, and dependencies. Then she asks whether the client can understand the message. Only after those checks does she adjust rhythm and phrasing. This order keeps the pleasure of editing from impersonating acceptance.

For a consequential recommendation, she writes a short account of what would make it wrong. Perhaps the timeline assumes that a supplier can meet a date. Perhaps the source covers a different population. Perhaps the recommendation depends on a client priority that has never been confirmed.

Ask the assistant for objections if that helps. Then inspect whether those objections touch the actual decision. A list of generic cautions can be as decorative as a list of generic benefits. The reviewer needs the uncertainty that could change the action.

Sampling has limits

You cannot inspect everything with equal care. For a stable, repetitive workflow, sampling may help detect patterns. But sampling is a weaker defense when a single rare error has a serious consequence or when a change in the system creates a new failure mode.

Decide which elements receive complete review and which can receive occasional review. Check every external commitment in a client update, for example, while sampling routine formatting. Reconsider the allocation when inputs, permissions, tools, or consequences change.

Do not describe a sample as a complete audit. Record what you checked and what you did not. This gives the next person a usable account of the confidence available.

For specialized work, use a reviewer with the relevant competence and follow the standards of the field. A general checklist cannot supply professional knowledge. One of its useful outcomes may be recognizing when that knowledge is required.

Disagreement is an output

Sometimes review ends with a difference in judgment. Preserve it long enough to understand it. The assistant's draft, the project owner's view, and the client's understanding may each reveal a different assumption.

Mara begins keeping a short list of recurring disagreements. Some disappear when the brief improves. Others point to unresolved business decisions. The studio has been asking the assistant to resolve questions that the team itself has not settled.

This is an important use of AI: making hidden disagreement visible. It becomes harmful when the smooth output conceals the disagreement instead.

Review works when it produces a justified decision: accept, correct, return for more evidence, or stop. Its value lies in the quality of that decision. The number of comments is incidental.

Try this: Review backward

Take a result you are about to use. Begin at its proposed action and work backward through the claims that support it. Check the consequential names, dates, quantities, certainty words, and destinations. Write one condition that would change your decision. Accept the result only when that condition has been addressed or its uncertainty has been consciously accepted by the responsible person.

Keep a practice

One of the studio's designers, Eli, uses the assistant to turn rough notes into explanations for clients. The explanations become clearer. Mara is pleased. Eli is pleased too, until a client asks an unexpected question during a call.

Eli knows which option the team recommends. He has trouble explaining the reason without the generated text in front of him. The assistant helped him present a judgment whose underlying distinctions he had not fully made himself.

This scene is invented, and the outcome is not inevitable. AI can also help someone learn those distinctions. The difference depends in part on what the workflow asks the learner to do.

The question is which activities you want help completing and which you also need as practice. A single activity can be both.

The old irony

Lisanne Bainbridge's 1983 paper on automation examined a persistent design problem: systems can remove routine human activity while leaving people responsible for abnormal situations. The remaining work may demand skills that routine participation previously helped maintain. Bainbridge was discussing industrial automation, not contemporary language models. Applying the concern to AI-assisted knowledge work is an analogy that requires local judgment. [5]

The analogy is useful at the studio. If Eli is expected to handle unexpected client questions, he needs opportunities to form and explain judgments before those questions arrive. The organization cannot remove every learning opportunity and then blame him for lacking readiness.

This does not require keeping all work manual. It requires choosing where learning will happen. Some practice can be supported by AI, some should occur without it, and some needs another person's feedback.

What should remain available in you?

List the capabilities on which your consequential decisions depend. Be specific. “Good judgment” is too broad. “Distinguish a client preference from an approval” is something you can observe. So is “explain the tradeoff between two layouts in language a client understands.”

For each capability, ask how you would know whether you still possess it. Could you solve a modest new example? Could you notice a flawed result? Could you explain your choice to a

competent colleague? Could you act during a temporary loss of access to the tool?

The standard should fit the role. You do not need to memorize everything an external system stores. You do need enough understanding to recognize when the system's contribution fits the situation and when it requires help.

Evan Risko and Sam Gilbert's review of cognitive offloading describes how people use external actions and tools to reduce cognitive demands. The reviewed evidence includes benefits and costs that depend on the task. External storage can alter what people remember, including the balance between remembering content and remembering where to find it. The review predates current generative AI systems; it does not settle their long-term cognitive effects. [6]

Offloading is an ordinary feature of human life. A calendar can make room for thought. The decision is which demands to move outside yourself and which understanding must remain available when you need it.

Three ways to practice with help

First, attempt before viewing. Write your own account of a problem, even if it is rough. Then compare it with the assistant's account. Look for a distinction you missed, an assumption it made, and a point on which you disagree. This preserves a moment in which you must generate the structure yourself.

Second, explain after viewing. Close the output and describe the reasoning in your own words. Reopen it to check what you omitted. If the explanation collapses without the original

wording, you may have acquired a useful document without acquiring the understanding you need.

Third, transfer to a new case. Ask for a different example that tests the same distinction, or ask a colleague to provide one. A practiced answer to the original problem can conceal a failure to generalize.

These are proposed learning routines, not doses proven by the studies cited here. Try them on low-stakes work, observe whether they improve performance, and adapt them to the skill involved. For some capabilities, supervised professional training will be necessary.

The practice budget

Mara and Eli choose one client explanation a week for deliberate practice. Eli drafts it unaided, then uses the assistant to identify gaps and alternative explanations. Mara reviews the reasoning with him. They compare the result with the client's actual questions.

They do not impose the routine on every message. The point is to preserve a learning path without turning the entire week into an examination. Nor do they pretend that one weekly exercise guarantees readiness. They observe whether Eli can handle new questions and adjust the practice accordingly.

The time belongs on the schedule. If practice must be stolen from delivery work, it will disappear under pressure. An organization that expects retained skill should fund the conditions for it.

An individual can make the same decision on a smaller scale. Choose a capability worth maintaining and a recurring occasion to use it. Protect that occasion from the convenience of skipping directly to the finished answer.

Accessibility changes the question

For some people, assistance makes participation possible. Asking them to demonstrate independence from a tool can create an unnecessary barrier. Skill should be judged in relation to the supports a person actually uses and the demands of the role.

Distinguish preserving understanding from performing every physical or linguistic step unaided. A person may understand a decision deeply while using speech input, translation, text generation, or other assistance to express it. The relevant question is whether they can exercise the needed judgment with appropriate support.

A fallback plan should also be realistic. It may involve another person, another accessible tool, or a safe pause. Requiring a heroic return to fully manual work is often poor planning.

The purpose of practice is usable capability. It should expand a person's agency rather than become a test of conformity to an old method.

Learning can run in both directions

AI assistance may expose novices to better examples, offer patient rehearsal, and help experts explore unfamiliar territory. The customer-support study discussed earlier found evidence of worker learning in its setting. That possibility belongs beside the concern about lost practice.

To distinguish them locally, look beyond the assisted output. What can the person now recognize, explain, or do in a new situation? Has their dependence become informed and manageable? Can they identify when they need a different source or another person?

A tool can make today's work better while helping tomorrow's worker become more capable. It can also make today's output impressive while leaving understanding untouched. Design for the first possibility and check whether it is occurring.

The work you keep includes the work that makes you able to judge the rest.

Try this: Name one retained capability

Choose one capability you expect to need a year from now. Describe a small new situation that would reveal whether you have it. Schedule a practice that asks you to attempt, explain, or transfer the skill. Use assistance where it improves access and learning. Review the result after several attempts, and change the routine if it is producing performance theater instead of capability.

Leave a return path

The studio's assistant has permission to prepare tasks in a project board. A new instruction accidentally causes it to create the same set twice. The board becomes confusing, and two people begin work on duplicate assignments.

This fictional incident is ordinary enough to be instructive. The failure is less important than what the team can do next. Can they identify what changed? Can they stop further changes? Can they restore the board without deleting legitimate work? Can they tell the people affected what happened?

A return path makes those answers possible. It is part of the delegation from the beginning.

Before the action

For work that changes an external state, identify the current state worth preserving. A draft may need only version history. A project board may need a dated export or a change record. A message needs an approved recipient list before sending, because a sent message cannot be fully recalled from another person's experience.

Reversibility has degrees. You can delete an accidentally created task. You may still need to explain why someone spent an hour working on it. You can correct a published statement. People

may have already relied on the first version. The technical reversal and the human repair are separate jobs.

Choose an appropriate environment for the first trial. Use a copy, a private draft, or a limited set of records when that can reveal the relevant behavior. A test that never touches realistic conditions may create false confidence, but a test that can alter everything at once creates avoidable repair work.

Name who can stop the process. Confirm that the person can actually exercise that control. A pause instruction buried in a chat is insufficient if an independent recurring process will continue acting.

Keep a usable receipt

A receipt should answer what was requested, what was done, which records changed, which checks passed, and what remains uncertain. It should let a person locate the actual result.

For Mara's task board, a useful receipt lists the tasks created and their identifiers, the project they belong to, the source notes used, and the time of the operation. It distinguishes tasks successfully created from tasks merely proposed. A statement that everything was handled is too vague to support recovery.

Do not collect every available detail by default. Records can contain private information and become burdensome to maintain. Keep what serves verification, coordination, and repair, with appropriate access and retention. A usable receipt is selective evidence.

For personal work, the receipt may be a short note. For a regulated or consequential process, existing organizational and

professional requirements may demand much more. The book's examples are starting points, rather than replacements for those requirements.

The graceful stop

A system needs a way to finish without pretending it succeeded. Useful stopping outcomes include: missing evidence, conflicting instructions, unavailable permission, unexpected destination, failed verification, and partially completed work.

Specify what should happen in each important case. If a source cannot be accessed, preserve the draft and identify the gap. If a write may have succeeded but its response was lost, inspect the destination before retrying. Otherwise a well-intentioned retry can create duplicates.

The studio's duplicate tasks could have arisen from exactly that pattern. A request timed out, and the assistant repeated it without checking. The safer next step was a read of the board to determine whether the first operation had taken effect.

At a larger scale, systems can support duplicate prevention through identifiers and explicit state tracking. At a smaller scale, a human-readable change list may be enough. The necessary design follows the consequence and frequency of the operation.

Correction should be easy to invoke

Saleema Amershi and colleagues proposed eighteen guidelines for human-AI interaction, validated through several evaluation steps. They include support for correction and dismissal, communicating capabilities and limits, and adapting cautiously over time. These are design guidelines, rather than evidence that any particular interface guarantees good outcomes. [7]

A practical test is whether a tired person can correct a consequential error without understanding the entire system. Can the client reach a person? Can a reviewer reject a result and explain why? Can someone return to an earlier version? Does the interface make the effect of an approval clear?

If correction is hard, people may accept an inadequate result simply to escape the interaction. The resulting completion metric will look better than the actual service.

Mara adds a visible contact and a plain path for corrections to the client workflow. She also ensures that corrections reach the person responsible for the project. A feedback button that nobody reads creates an appearance of accountability while leaving the client stranded.

Rehearse a small failure

You learn about a return path by using it. On a harmless copy, deliberately introduce one realistic problem. Supply conflicting notes. Remove an expected file. Simulate a lost response. Ask a colleague unfamiliar with the setup to find the latest accepted version.

Observe where recovery depends on unstated knowledge. If only the person who built the workflow can recover it, the team has acquired a new dependency. Document the minimum steps another authorized person needs.

Avoid dramatic exercises unrelated to ordinary work. A realistic missing approval can teach more than an elaborate catastrophe scenario. Start with failures likely enough to justify preparation, then address rarer events according to their consequences.

Record what happened without turning the exercise into blame. The purpose is to improve the arrangement before someone has to rely on it under pressure.

Repair the relationship

When a failure reaches another person, technical recovery is only part of the response. Tell them what happened in language relevant to their situation. Identify the corrected information, the action they need to take, and the person responsible for follow-up.

Avoid making the tool the subject of every sentence. The client needs a reliable project, not a tour of the studio's software choices. The team remains responsible for the service it promised.

An honest correction can include uncertainty. If you do not yet know the full extent of the error, say what you know, what you are checking, and when the next update will arrive. Do not use a confident explanation to close an investigation prematurely.

A return path gives people permission to try useful things without pretending that useful things never fail. It turns recoverability into part of competence.

Try this: Trace the way back

Before your next automated change, write down the state being changed, the evidence that will show what happened, the person who can pause it, and the practical steps for repair. Test the steps on a harmless copy. Include the people who would need an explanation if the change went wrong.

Give time somewhere to go

After several weeks, the studio's updates become easier to prepare. In our fictional example, the team has recovered a little capacity. Mara's first impulse is to fill it with another service.

There is nothing inherently wrong with that choice. A small business may need additional revenue, and its workers may welcome growth. But the choice deserves to be made openly. Otherwise every improvement becomes an invisible increase in expectation.

The same pattern appears in a personal day. A task finishes early, and the empty interval feels provisional. You find another task before you have noticed that stopping was available.

If you want assistance to change the quality of your life, saved time needs a destination. That destination may be more work. It may also be a slower lunch, a conversation, study, caregiving, rest, or an afternoon with no impressive account of itself.

Capacity is a decision

Before a trial, decide how you hope to use a plausible gain. You do not need a precise estimate. You need an intention that can survive the moment when the calendar opens.

For example, a team might reserve the first recovered half hour for reviewing recurring client problems. An individual might choose to end one evening's work earlier. A caregiver might use a reduced administrative burden to create a less hurried transition between responsibilities.

Name the destination in the same practical terms as the task. “Better balance” is difficult to defend against a new request. “No meetings after four on Thursday so I can collect my child” is a constraint someone can understand and plan around.

Some people have little authority over their schedules. For them, choosing a destination may require negotiation, collective agreement, or a change in working conditions. Personal discipline cannot manufacture bargaining power. The next chapter addresses that problem directly.

The cost of fragments

A day can gain minutes and lose continuity. If an assistant interrupts you repeatedly with small requests for approval, it may return time in pieces too small for the work you hoped to do.

Sophie Leroy's 2009 experiments on attention residue found that people can have difficulty disengaging from an unfinished task, with consequences for performance on a subsequent task. The findings do not establish a universal recovery time for every interruption, and the study did not evaluate AI assistants. They provide a reason to examine transitions rather than count minutes alone. [8]

Mara notices that checking each proposed task as it appears keeps her mentally attached to the project board. She moves routine review into two agreed windows. Urgent matters retain a separate route. The arrangement is useful only if the team can define urgency and tolerate the delay for everything else.

For your own work, consider the shape of attention a task requires. A ten-minute administrative block can fit between appointments. A difficult piece of writing may need a longer stretch. A friend's story may need the absence of a visible countdown.

Close the loop before changing rooms

One practical transition is to leave a short continuation note. Record what is settled, what remains open, and the next action. Put it where you will resume. This is a proposed way to reduce reconstruction effort; it should be judged by whether it helps you return.

An assistant can prepare the note, but you should recognize yourself in it. A polished description of the wrong next step will not help. Read it once, correct it, and stop editing.

Then change something in the environment. Close the work window. Leave the desk. Turn off the notification channel that has no further role in the next interval. A transition becomes more credible when your surroundings support it.

These gestures cannot solve an unreasonable workload. They can help you honor an actual stopping decision. If work keeps returning because nobody agreed that it could wait, the problem belongs in the agreement.

The temptation of the observable

Systems are good at producing records. They can summarize your day, count your outputs, and propose improvements. These records can be useful. They can also make unrecorded life feel less real.

You may start choosing activities because they produce something easy to describe: a completed course, a published note, a quantified walk. Meanwhile the conversation that wandered, the book abandoned halfway through, and the hour spent sitting outside seem to need an excuse.

This book's position is a value judgment: a life should have room for experiences whose worth does not depend on a performance report. Research can inform how we think about motivation and attention. It cannot choose all our ends for us.

When you recover capacity, leave some of it without a reporting obligation. You can still notice whether life feels more spacious or whether a relationship receives more attention. You do not have to turn every observation into a metric.

Enough is a boundary you can explain

Define enough for a recurring piece of work. Enough research to make this decision. Enough revision to serve this reader. Enough preparation to hold this meeting responsibly. The boundary should reflect consequence and uncertainty.

An adequate standard can change. New evidence may reveal a missing requirement. But an endless supply of possible improvements is not itself evidence that the current result is inadequate.

AI can make variation almost frictionless. Another version is always available. The human task is to judge whether another version serves anyone. Sometimes it does. Sometimes the next useful act is delivery, and sometimes it is leaving the work alone.

Mara writes a simple stopping line for the weekly update: accurate, understandable, decisions visible, commitments verified. Once those conditions are met, the team sends it. Any further revision needs a named purpose.

Presence has a different rhythm

At the end of the week, Mara has promised to meet an old friend. The assistant could summarize the friend's recent messages, suggest topics, and prepare thoughtful questions. That might help her remember a detail. It could also encourage her to treat the meeting as an interaction to optimize.

She glances at the time and place, then puts the preparation away. The conversation will require noticing what the friend actually says. Its value includes being surprised.

You can use assistance around a human encounter without scripting the encounter itself. Remove friction where it prevents participation. Preserve openness where the relationship needs it.

The distinction is personal and contextual. A person who uses communication support may need preparation to enter the conversation at all. The aim is more available attention with the supports that make it possible, rather than a rule about how authentic attention must look.

The afternoon arrives

At two on a hypothetical Friday, the obligations are settled. Mara can still choose to work. The difference is that the choice has become visible.

She uses part of the afternoon to think about the studio, because that thinking has been crowded out. She leaves the rest open. Nothing dramatic happens. There is no final dashboard proving that the interval was valuable.

That modest outcome is worth taking seriously. A useful system can make an ordinary day less crowded. You are allowed to count that as success.

Try this: Reserve the first gain

Choose a small destination for capacity before the next automation trial. Put it on the calendar or in the team's agreement. Observe whether the workflow returns a usable interval, scattered fragments, or no capacity at all. If the interval appears, use it as intended once before deciding whether to expand the work.

The team gets a vote

Mara proposes that the assistant prepare more of the studio's routine work. One colleague worries that the proposal will become a staffing reduction. Another worries about being expected to review twice as much material. Eli worries that the team will stop funding the practice time they just created.

These are questions about the distribution of benefits, effort, and authority. A demonstration of capability cannot answer them.

People are more likely to understand a change when they can see how it affects their actual working conditions. They also deserve a say when their jobs, skills, or relationships are being reorganized. That is an ethical position in this book, beyond any claim about whether participation improves a productivity measure.

Separate the promises

An automation proposal often contains several promises at once: faster delivery, higher quality, lower costs, less drudgery, more learning, and greater autonomy. Some can coexist. Some compete. If all remain implicit, each person hears the version they most hope for or fear.

Name the intended benefit and the constraints. A trial might aim to reduce preparation time while holding quality steady,

preserving current roles during the trial, and protecting a defined amount of learning time. An owner considering a different arrangement should state it plainly enough for people to respond.

Do not promise permanent job security unless you can honor it. Do not describe a trial as harmless if its results will determine consequential employment decisions. Clarity is part of fair participation.

People may reasonably disagree even with full information. A useful process makes that disagreement discussable and records who has authority to decide.

Autonomy has material conditions

Richard Ryan and Edward Deci's account of self-determination theory identifies autonomy, competence, and relatedness as psychological needs relevant to motivation and well-being. It emphasizes the role of social conditions in supporting or thwarting them. Their 2000 paper does not test an AI rollout or establish that any particular tool will improve motivation. [9]

The framework offers questions worth asking. Does the new workflow give workers meaningful discretion? Can they develop and exercise competence? Does it support working relationships, or turn colleagues into isolated reviewers of machine output?

Autonomy requires more than a choice of software. A person may be allowed to choose a tool while having no influence over an impossible quota. Competence requires more than an excellent generated result. Relatedness requires more than a channel where status updates accumulate.

Use the questions to examine the arrangement, then listen to the people living in it. Their experience may reveal consequences the design did not anticipate.

Make the hidden reviewer visible

In the studio, the assistant's drafts are initially sent to the most experienced designer. The arrangement seems sensible because she can spot errors quickly. Soon her own design work is delayed. The team's apparent gain has been financed by her attention.

Review is work. Assign it, budget it, and recognize it. If a workflow depends on scarce expertise, count that constraint before expanding production. Increasing the number of drafts can make a bottleneck worse.

The team changes its process. Drafts include sources and an account of unresolved questions. Routine checks happen before expert review. The experienced designer focuses on decisions that require her judgment. Less experienced colleagues receive feedback that develops their ability to perform some checks themselves.

This is a fictional redesign, not a measured result. Its purpose is to show where to look: at the person whose effort makes everyone else's efficiency possible.

A trial agreement people can read

Write a short agreement in language the team already uses. Include the task being tested, the benefit sought, the duration, the people affected, the work that will be measured, the information recorded, and the conditions for stopping or changing the trial.

Say whether individual data will be used for evaluation. A tool experiment can quietly become employee surveillance if detailed records are gathered without a clear purpose. Collect enough to judge the workflow and avoid treating every available trace as something the organization is entitled to use.

Choose a route for reporting problems that does not require a person to challenge the most enthusiastic sponsor in public. Make sure problems can lead to actual changes. An invitation to provide feedback is weak if the decision has already been made.

At the end of the trial, discuss both the result and its distribution. Who gained time? Who gained difficult work? Who lost practice? Who became responsible for exceptions? Who has the power to refuse an inappropriate action?

Let a client reach a person

Clients and customers also participate in the new arrangement, even when they never see the system. Their experience should be part of its evaluation.

An automated response may be fast and still fail if it traps a person in repeated explanations. A system may correctly classify a request while missing the situation that makes it urgent. Give

customers a usable route to someone who can understand the problem and act.

Be truthful about the service. Do not invent a human conversation that never occurred or imply a person reviewed work when nobody did. Disclosure should help the receiver understand what they can rely on and how to get help.

The amount and form of disclosure depend on context, existing obligations, and what matters to the person. A blanket notice does little if it leaves them unable to correct a consequential error.

The right to an ordinary pace

When a task becomes faster, the old pace may be reclassified as laziness. People can lose the space that allowed them to notice a problem, teach a colleague, or recover between demanding interactions.

An organization can choose a different response. It can use some gains to improve resilience, learning, accessibility, or service. It can also choose growth. The ethical issue is whether the choice is visible, fair, and open to those who bear its consequences.

You cannot solve every tension inside a small trial. Market pressure, ownership, and broader labor conditions matter. Still, you can stop treating the distribution as a technical detail. A workflow is an arrangement among people, supported by tools.

Mara ends the meeting with fewer claims and a clearer trial. The team will test one sequence, preserve practice time, account for review, and revisit the arrangement together. Their agreement has become more specific than their enthusiasm.

Try this: Ask downstream

Before expanding a workflow, speak with one person who receives its output and one person who handles its mistakes. Ask what work the change adds, what context they lose, and what authority they need. Put at least one resulting condition into the trial agreement. At review, report who received the benefit as carefully as you report the total.

Measure the work you actually want

At the end of a trial, Mara wants a verdict. Is the assistant helping?

The answer depends on what the studio intended to improve. The number of drafts has increased. That is easy to count and insufficient. Preparation feels easier. That is worth recording and still insufficient. A colleague reports more uninterrupted time, while another reports more review. Both experiences matter.

Measurement begins with the outcome and the people it should serve. Then it follows the work far enough to see what the attractive first result leaves out.

Two findings, two settings

METR's early-2025 randomized experiment involved sixteen experienced open-source developers completing 246 tasks in familiar repositories. Allowing the AI tools studied increased completion time by 19 percent on average, despite participants' belief that the tools had made them faster. The result concerns those developers, tasks, and early-2025 tools; it is not a timeless estimate of AI's effect on programming. [10]

Place that result beside the customer-support study in Chapter 1. One found a gain; one found a slowdown. The difference does not require choosing which study to believe. Work context, tool capability, task familiarity, outcome measures, and research design can differ.

For a business owner, the lesson is to be precise about the setting to which a claim belongs. Borrow an experiment's questions before borrowing its percentage.

An update is part of the evidence

In February 2026, METR reported that a later developer experiment faced serious selection and measurement problems. Developers and tasks likely to benefit could be missing because participants did not want to work without AI. The researchers believed tools were probably helping more than in early 2025, but described the new data as weak evidence for the magnitude. [11]

In May 2026, METR published a survey of 349 technical workers reporting perceived gains in the value of their work. The median self-reported gain varied with the question, around 1.4 to 2 times. These are respondents' estimates, rather than randomized measurements of a representative workforce. [12]

The sequence matters. An older finding may remain valid for its original conditions while becoming a poor estimate of current conditions. A newer result can be relevant without being methodologically stronger. Dates and methods belong next to the number.

You should expect to revisit a local trial too. A changed model, tool, input source, permission, or task mix can alter the result. Keep the earlier record as history. Do not let it become permanent authorization for a substantially different process.

A scorecard with four views

For a small trial, choose a few measures across four views: outcome quality, total effort, retained capability, and distribution of benefit. The scorecard is this book's proposal, not a validated index. It is useful if it helps you see tradeoffs that one number would hide.

Outcome quality might include accuracy, completeness, customer understanding, or successful completion of a service. Total effort includes preparation, interaction, checking, correction, and downstream work. Retained capability asks whether people can still make and explain consequential judgments. Distribution asks who gains time or access and who inherits difficult work.

Add a clear condition that would make the arrangement unacceptable. Perhaps no unapproved client commitment may be sent. Perhaps private records may not leave an authorized environment. Perhaps a reviewer must be able to inspect the evidence before approval. These are constraints, rather than benefits to average away.

Choose measures the task can support. A weekly client update does not need a complicated composite score. A record of preparation time, meaningful corrections, unresolved decisions, and reviewer burden may be enough to inform the next small choice.

Keep comparable work comparable

Compare similar tasks when possible. Alternate methods across comparable items, or record a before period and a trial period while noting changes. Neither approach automatically creates a scientific experiment. Learning, seasonality, task difficulty, and selective use can all influence the result.

Record the method actually used. If people abandon the assistant whenever a task becomes difficult, measuring only completed assisted tasks will overstate its usefulness. If they reserve AI for unusually difficult tasks, a simple comparison may understate it. Task selection is part of the evidence.

Include the setup and training costs when they matter to the decision. A new workflow may require an initial investment that pays off over repeated use. Keep that investment separate from recurring effort so you can see both. Do not erase it because it is inconvenient or count it forever as if it recurs every week.

Small samples justify modest conclusions. “This helped with these updates during this month” can be a useful finding. It does not need to become a slogan about all work.

New possibilities are a different benefit

Sometimes AI lets you do work you would never have attempted before. A solo owner can explore a prototype. A person can communicate in a language they are still learning. A small team can make a useful guide it lacked the capacity to produce.

These gains should not be forced into a simple time-saving story. If you would never have built the prototype manually, claiming that you saved the hypothetical months it might have taken can

be misleading. The real benefit may be the value of learning whether the idea works.

METR's May 2026 note on task substitution distinguishes gains on the old task mix, gains on the new task mix, and gains in value. The distinction helps explain why a tool can make newly chosen activities dramatically cheaper without an equally dramatic increase in the value of a whole working day. Its formal conclusions depend on stated simplifying assumptions. [13]

For Mara, the question becomes whether the new service solves a problem worth solving. Its cheap production is relevant to feasibility. It does not establish demand or usefulness.

Ask what disappeared

A trial may reduce visible effort by removing work you still need. Perhaps the team stops speaking with clients because summaries are convenient. Perhaps junior staff lose exposure to the source material. Perhaps a report becomes shorter because it omits uncertainty.

Include a review of what is no longer happening. Some omissions are improvements: duplicate entry, unnecessary formatting, a meeting whose purpose has vanished. Others remove the conditions that made the service reliable.

You can discover the difference by asking what each activity contributed and where that contribution now occurs. If nobody can answer, inspect before expanding.

The same question belongs in a personal trial. Has the tool removed drudgery, or has it also displaced an activity you enjoyed? Have you gained usable attention, or become more

available for interruption? Your answer can include values that no business metric captures.

Decide, then set a return date

At the end of the fictional studio trial, the team has enough evidence for a limited decision. They retain assisted drafting for weekly updates, change the review sequence, and postpone automated sending. They keep practice time and reserve part of the recovered capacity for a recurring client conversation.

The decision is neither a declaration of faith nor a rejection of progress. It is a judgment about a defined arrangement. The team writes down what would trigger reconsideration: a new permission, a changed source, repeated review failures, or a significant shift in workload.

They also choose a review date. An arrangement can drift quietly even when no dramatic failure occurs. A brief return to the original purpose can reveal that the work has changed.

Good measurement ends in a decision someone can explain. It helps you choose what to keep, what to change, and what evidence to seek next.

Try this: Write a modest conclusion

After a small trial, write five sentences: what you tested, what you observed, what you could not determine, who gained or lost, and what you will do next. Include the date and conditions. If your conclusion claims more than the observations support, make it smaller until it becomes true.

Thirty days of different work

You can put the argument into practice without reorganizing your whole life. Choose one recurring sequence, observe it carefully, and make a small change that can be reversed.

The following month is a suggested pace. It is not a promise of transformation in thirty days. Some tasks require a longer observation period; some should not be tested without specialized oversight. Adjust the schedule to the work and the people involved.

The aim is a concrete decision at the end: continue, revise, narrow, expand cautiously, or stop. Each is a legitimate result if the evidence supports it.

Days 1-3: Follow the work

Choose a recurring task that produces a useful outcome and has manageable consequences. Avoid beginning with the most sensitive or irreversible process in your life. You want enough realism to learn and enough containment to recover.

Perform the current workflow while taking brief notes. Record the inputs, people, handoffs, checks, and destination. Identify the three endings from Chapter 1. Estimate total effort across

everyone involved. Note where the work waits, where it gets corrected, and where uncertainty is resolved.

Do not optimize immediately. Observation often reveals that a confusing source or unnecessary handoff accounts for much of the effort. Fixing the input may matter more than generating faster output.

Write the purpose in the receiver's language. Ask the receiver whether the account is accurate. If you cannot identify a person or a decision the task serves, consider whether the task should exist at all.

Days 4-7: Draw the boundary

Unpack the task into verbs. For each activity, consider consequence, checkability, reversibility, and practice value. Choose where assistance begins and ends. Name the person responsible for acceptance.

Write the one-page brief: purpose, authorized materials, required output, acceptance conditions, permissions, and stopping conditions. Add a near-miss example. Check that the permissions available to the system fit the agreement.

Choose one capability to preserve and one destination for any recovered capacity. Put both into the plan. These choices make the trial about the working life you want, alongside the output you need.

If others are affected, agree on the trial's duration, measures, use of records, and route for reporting problems. Make the plan short enough that everyone can read it before the trial begins.

Days 8-14: Run a narrow trial

Use assistance on a small set of representative tasks. Preserve the current workflow as a fallback. Keep the brief and result together. Record preparation, interaction, review, correction, and downstream effort separately enough to see where time moves.

Review consequential joints before polishing. Verify claims against appropriate sources. Distinguish a proposed action from an executed one. Record partial completion honestly.

When something fails, inspect before repeating an action that may already have succeeded. Repair the immediate problem and note whether the brief, source, permissions, or review process needs to change. A failure that exposes a useful boundary can be an important result.

At the end of the week, speak to the downstream receiver. Ask what became easier and what became harder. Check the hidden reviewer. Look for a queue of generated work waiting for human acceptance.

Days 15-18: Practice and recover

Test the retained capability on a new, low-stakes example. Use the supports appropriate to the person and role. Observe whether they can make, explain, and challenge the relevant judgment.

Rehearse one ordinary failure on a harmless copy. Try a missing source, conflicting instruction, or lost response. Have another authorized person follow the return path. Record the parts that depended on knowledge only you possessed.

Improve the arrangement once, then run it again. Avoid changing every element at the same time. You want to know which change addressed which problem.

If the workflow remains difficult to inspect or recover, narrow its authority. Additional generation is unlikely to solve a missing recovery path.

Days 19-23: Use the capacity

Look for the benefit you intended. Did total effort fall? Did quality improve? Did access expand? Did the work become easier without becoming faster? Did someone else inherit the burden?

If usable capacity appeared, spend some of it on the destination you chose. End early, practice, talk with a client, improve a neglected process, or leave an interval open. Do this before allowing new production to absorb every gain.

Observe continuity as well as duration. A recovered hour divided into twelve interruptions may be unsuitable for the activity you intended. Adjust the timing of reviews and notifications if that is within your authority.

If there is no gain, record that plainly. You may still have learned something useful about the task. Consider a smaller delegation, better inputs, a different tool, or a return to the previous method.

Days 24-27: Review the arrangement

Bring together the brief, examples, timing notes, corrections, and feedback. Use the four views from Chapter 9. Keep consequential constraints visible instead of averaging them into a general score.

Look for selection effects. Which tasks did you avoid? Which failed attempts are missing from the record? Did the task itself change during the trial? Did people improve through practice independently of the tool?

Ask what disappeared. Has the team lost a learning opportunity, a conversation, or a useful check? Has an unnecessary activity finally been removed? Identify where each important contribution now happens.

Write the five-sentence conclusion. A good conclusion can be provisional. Its strength comes from accurately describing the evidence you have.

Days 28-30: Make one decision

Choose the next arrangement and name its limits. If you continue, preserve the brief, review method, and return path. If you expand, change one meaningful dimension at a time, such as the volume, task type, or permission. Reconsider the risks created by that change.

If you stop, retain the useful learning. A failed trial does not need to become an abandoned folder nobody can interpret. Record why it stopped and what would need to change before another attempt made sense.

Tell the people affected what was decided. Include the distribution of benefits and new responsibilities. Set a review date and concrete triggers for reconsideration.

Then close the trial. The purpose was a decision, rather than a permanent state of preparing to decide. Use the resulting arrangement long enough to encounter ordinary life.

When the month goes badly

Perhaps the source material is poor, the tool changes, or a busy week prevents careful observation. Perhaps your first choice of task was unsuitable. A modest conclusion remains available: the conditions did not support this trial, so the evidence is incomplete.

Do not turn incomplete evidence into a verdict on your own competence. Make the next question smaller. Test a private draft before an external action. Test extraction before recommendation. Clarify acceptance before measuring speed.

You may also discover that the main problem is workload, unclear authority, or a business promise the team cannot fulfill. AI can expose those problems without resolving them. Address the owning decision rather than asking the tool to conceal its effects.

The month has succeeded if you can describe the work more clearly and make a justified choice about help. Sometimes the most valuable change is permission to stop automating a task that never needed to be done.

Try this: Begin with tomorrow

Write the name of one task, the person it serves, and the next occasion when it will occur. Leave space for the observation. You can begin without choosing a tool. The first useful act is paying attention to work you already do.

Afterword: A life larger than its output

At the beginning of this book, an afternoon opened. We have spent ten chapters making the conditions for that opening more credible.

We defined completion, divided responsibilities, wrote promises, checked evidence, preserved practice, designed recovery, chose destinations for time, involved other people, and measured a bounded result. These activities can sound like additional work. Some of them are. They are the work required to make help dependable.

The aim is to make that work proportionate. A personal draft should not require the process of an industrial control room. A consequential service should not depend on the optimism of a pleasing demonstration. Good judgment includes knowing how much structure the situation deserves.

As tools improve, some boundaries will move. A task that demands close participation today may become routine. New possibilities will create new obligations. You will need to keep deciding what matters, what evidence is adequate, and who gets a say.

This continuing need for judgment is not a consolation prize for being human. It is part of living among other people. We make promises, receive consequences, learn from mistakes, and choose which uses of our capacity deserve a place in the day.

AI can participate in many of the activities around those choices. It can help you see alternatives, prepare evidence, rehearse an explanation, and reduce friction. The presence of assistance does not make the resulting achievement less real. A person can become more capable through well-chosen dependence.

The concern is dependence that you cannot understand, challenge, or leave. A useful arrangement should preserve enough visibility and practical choice for you to act with reasons. It should also respect the people who carry the consequences of your decisions.

At Harbor Studio, our invented team does not arrive at a perfect system. They arrive at a better conversation. They can name what the assistant does, what people check, where learning happens, and how a mistake will be repaired. They know which gains they intend to spend on growth and which they intend to spend on the quality of work and life.

That is a plausible place to stand. You can revise it as the circumstances change.

Outside the studio, no tool can supply a final account of what your attention is for. You will answer in the activities and people to which you return. Some answers will be difficult. Some will be quiet. Many will never become an output another person can count.

You can still make a deliberate place for them. You can let a task be adequate, a promise be settled, and a day reach its end. You can accept help without turning every saved minute into evidence that you should have done more.

When the afternoon comes, notice that it has arrived.

Toolkit: Six worksheets to make your own

These worksheets turn the book's proposals into a small working record. Copy the parts that fit your situation. Leave out anything that produces paperwork without improving a decision. The prompts are original practical aids; the research cited in the book does not validate them as a package.

1. The obligation map

Task: What recurring sequence are you examining?

Receiver: Who will use the result, and what will it help them do?

Purpose: The receiver should be able to do what, without having to do what?

Production ending: What output will exist?

Acceptance ending: What evidence makes that output usable, and who judges it?

Obligation ending: What delivery or handoff settles the promise?

Current effort: Record preparation, execution, review, correction, and downstream effort. Include other people.

Hidden work: What happens after the apparent finish? Who reconstructs missing context or handles mistakes?

First repair: Which one ambiguity should be resolved before introducing more automation?

At the end, describe one complete outcome in a sentence. If the sentence names only a document, add what a person should be able to rely on after receiving it.

2. The delegation brief

Purpose: Enable this person to make this decision or complete this activity.

Authorized materials: Use these sources, dated or versioned as needed. Treat these other sources as excluded, obsolete, or uncertain.

Required output: Return a form that supports the next decision. State whether it is a draft, recommendation, or executed change.

Acceptance conditions: These claims, calculations, commitments, and destinations must be checked. This person has authority to accept the result.

Permissions: These actions may happen without another decision. These actions require approval. These actions are outside the task.

Stopping conditions: Pause for missing evidence, conflicting instructions, unexpected recipients, unavailable permission, or other named triggers.

Near miss: Give one example that looks acceptable but changes the meaning or consequence.

Uncertainty response: Show how an unknown should be reported without inventing an answer.

Receipt: Identify the actual result, the checks completed, and the work still open.

Save the brief with the output. On revision, name the assumption or permission that changed.

3. The joint review

Action: What will someone do because of this result?

Support: Which claims must be true for that action to make sense?

Compact facts: Verify consequential names, dates, quantities, and units against appropriate evidence.

Certainty: Check words that turn possibility into fact, preference into approval, or intention into commitment.

Omissions: What limitation, dependency, excluded person, or disagreement would change the receiver's understanding?

Authority: Does the recommendation cross a boundary that requires a different person's decision?

Destination: Are the recipients, records, and public or private locations correct?

Countercondition: Name one circumstance that would change your acceptance decision.

Review record: State what was checked, what was sampled, and what remains uncertain.

Decision: Accept, correct, request evidence, or stop. Name the owner of any remaining work.

Read for consequence before editing for elegance. A useful review ends with a decision and an explanation someone else can understand.

4. The retained capability plan

Capability: Name an observable judgment or action you want available in a year.

Reason: Which responsibility depends on it?

New case: Describe a modest unfamiliar situation that would reveal whether the capability is present.

Practice: Choose an attempt before viewing, an explanation after viewing, or a transfer to a different case.

Support: Identify tools and accommodations that make participation and judgment possible.

Feedback: Who or what can provide trustworthy information about the result?

Time: Put the practice into a realistic schedule and identify who funds or protects that time.

Observation: After several attempts, what can you now recognize, explain, or do?

Adjustment: Keep, revise, or replace the routine. Do not confuse completion of the exercise with evidence of learning.

The aim is capability suited to a responsibility. Independence from every tool is not a universal standard of competence.

5. The return path

Change: What external state could this workflow alter?

Before: Which version, export, or record preserves what matters before the change?

Scope: What limits the number of records, people, or systems affected in a first trial?

Receipt: What identifies the actions that actually occurred?

Uncertain response: If an operation times out, how will you inspect the destination before retrying?

Pause: Who can stop further action, and how is that control exercised?

Recovery: What steps restore a usable state? Which effects cannot be fully reversed?

People: Who needs a correction or explanation even after the technical repair succeeds?

Rehearsal: Which ordinary failure can you safely test on a copy?

Second operator: Can another authorized person follow the recovery steps without the builder present?

Keep the record proportionate to consequence and frequency.
Protect private information in the record itself.

6. The trial and capacity record

Trial: Identify the task, dates, inputs, tool, permissions, and people affected.

Outcome: Choose a small number of meaningful quality measures.

Effort: Observe the whole sequence, including setup and work transferred to others.

Capability: Check the retained judgment on a new example.

Distribution: Name who gains time, access, authority, or opportunity, and who gains burden.

Constraint: What failure makes this arrangement unacceptable even if another measure improves?

Selection: Which tasks or attempts are missing from the observations?

Destination: If capacity appears, where will the first usable interval go?

Five-sentence conclusion: State what was tested, observed, uncertain, gained or lost, and decided.

Return: Set a review date and triggers for reconsidering the arrangement.

Use the conclusion to choose a next action. A modest, well-supported decision is more useful than an impressive number with no clear meaning.

Source notes: What the evidence supports

The book combines research summaries, interpretation, invented examples, and original proposals. The numbered notes below identify the external sources used. Each description states the source's role and a material limit. All links were checked during preparation on October 28, 2026. Where an older paper informs a contemporary workflow, that application is the book's interpretation.

1. Generative AI at Work

Erik Brynjolfsson, Danielle Li, and Lindsey Raymond. *The Quarterly Journal of Economics* 140(2), 889-942, 2025. **Published paper**. DOI: 10.1093/qje/qjae044.

Used in Chapters 1, 5, and 9. The book uses the published version's figures; earlier working papers differ. Results from this setting should not be generalized to all jobs or current tools.

2. Generative AI and Jobs: A Refined Global Index of Occupational Exposure

Pawel Gmyrek and colleagues. International Labour Organization, Working Paper 140, May 20, 2025. [Report and editions](#). DOI: 10.54394/HETPo387.

Used in Chapter 2 to support examining occupations through tasks and distinguishing exposure from observed job loss. An exposure index evaluates technological potential under its methodology. Employment outcomes also depend on adoption, institutions, demand, and choices beyond the index.

3. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile

Chloe Autio and colleagues. National Institute of Standards and Technology, NIST AI 600-1, July 26, 2024. [Publication record](#). [Full report](#).

Used in Chapter 4 for confabulation and the importance of source verification. This is a voluntary cross-sector risk-management resource. Mentioning it does not establish that the book's exercises conform to a certification or replace a domain's professional requirements.

4. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers

Hao-Ping Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. CHI 2025.

Research publication. DOI: 10.1145/3706598.3713778.

Used in Chapter 4 for reported patterns in critical-thinking effort and confidence. Survey responses describe perceived behavior and associations. They do not establish irreversible cognitive decline or isolate a causal effect of AI use on intelligence.

5. Ironies of Automation

Lisanne Bainbridge. *Automatica* 19(6), 775-779, 1983. **Publisher record.** **IFAC archive entry.**

Used in Chapter 5 for the problem of expecting people to handle exceptions while routine automation removes opportunities to maintain relevant skill. The original context is industrial automation. The chapter's application to client explanations and AI-assisted writing is an analogy, not a result tested by Bainbridge.

6. Cognitive Offloading

Evan F. Risko and Sam J. Gilbert. *Trends in Cognitive Sciences* 20(9), 676-688, 2016. **UCL record and open manuscript.** DOI: 10.1016/j.tics.2016.07.002.

Used in Chapter 5 for external tools, cognitive demands, and the distinction between remembering content and remembering

where to retrieve it. This review covers several tasks and forms of offloading. It predates current generative AI and cannot determine their long-term effects by itself.

7. Guidelines for Human-AI Interaction

Saleema Amershi and colleagues. CHI 2019. **Research project and paper**. DOI: 10.1145/3290605.3300233.

Used in Chapter 6 for design principles concerning capability communication, correction, dismissal, and cautious adaptation. The guidelines were developed and evaluated as design guidance. A workflow inspired by them still needs its own evaluation.

8. Why Is It So Hard to Do My Work? The Challenge of Attention Residue When Switching Between Work Tasks

Sophie Leroy. *Organizational Behavior and Human Decision Processes* 109(2), 168-181, 2009. **Published paper**. DOI: 10.1016/j.obhdp.2009.04.002.

Used in Chapter 7 for difficulty disengaging from unfinished tasks and consequences for subsequent performance. The experiments do not provide one universal interruption penalty. The book's continuation note and review windows are practical proposals, not interventions established by this paper.

9. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being

Richard M. Ryan and Edward L. Deci. *American Psychologist* 55(1), 68-78, 2000. [Author-hosted paper](#). DOI: 10.1037/0003-066X.55.1.68.

Used in Chapter 8 as a framework for asking about autonomy, competence, relatedness, and social conditions. It is a broad account of motivation, rather than an evaluation of AI adoption. The proposed team agreement remains the book's application.

10. Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity

Joel Becker, Nate Rush, Beth Barnes, and David Rein. METR, July 10, 2025. [Study paper](#). [Research record](#).

Used in Chapter 9 for the randomized developer experiment and the gap between observed time and perceived speed. Its sample, repositories, and tools define its scope. Read alongside the 2026 follow-up rather than treating it as an estimate for every developer or later system.

11. We Are Changing Our Developer Productivity Experiment Design

Joel Becker, Nate Rush, Tom Cunningham, David Rein, and Khalid Mahamud. METR, February 24, 2026. [Research update](#).

Used in Chapter 9 to update the 2025 discussion and explain selection, participation, and timing difficulties in a later study.

The authors describe weak evidence about the magnitude of increased benefit. The update supports caution in both optimistic and pessimistic generalizations.

12. Measuring the Self-Reported Impact of Early-2026 AI on Technical Worker Productivity

METR, May 11, 2026. [Survey report](#).

Used in Chapter 9 for a more recent account of respondents' perceived gains in the value of work. The estimates depend on self-report, question wording, and the sampled group. They are not interchangeable with a randomized measure of elapsed time or a population-wide productivity statistic.

13. Task Substitution and Uplift

METR, May 8, 2026. [Research note](#).

Used in Chapter 9 to distinguish old tasks, newly chosen tasks, and the value of the resulting work. The note's formal relationships use simplifying assumptions. The book borrows the conceptual distinction without claiming its local scorecard estimates an economic index.

An evidence boundary

The book does not predict a date for universal automation, quantify an economy-wide employment outcome, or claim that all people should use the same amount of AI. Its central recommendation is procedural: define a useful outcome, test a bounded arrangement, inspect consequences, and preserve the ability to revise it. The claim that a life deserves room beyond production is a stated value, open to reflection rather than established by a productivity study.

Closing page

The work you keep is the work of choosing what deserves your attention, understanding enough to make a promise, and remaining available to the people who rely on it.

Let the help be useful. Let the day be large enough to live in.

John Rector • johnrector.me • October 2026