

KEVIN · HANEGAN

A RESEARCH BRIEF

Why We *Hand It Over*

The psychology of delegating judgment to machines. Why capable people keep handing decisions to AI, and what it actually relieves them of.

By **Kevin Hanegan**

Companion to the book *Who Decides?*

kevinhanegan.com

EXECUTIVE SUMMARY

Organizations are handing decisions to AI faster than they can explain why. The pull is not hype, and it is not laziness. Delegation relieves three burdens at once: the effort of thinking, the discomfort of uncertainty, and the felt weight of responsibility.

Decades of research on cognitive effort, fluency, and decision avoidance explain each burden, and a wave of recent studies shows AI accepting all three together. The result is misallocated delegation, not simply too much of it. We hold on where machines would genuinely help, and let go where responsibility stings. The third burden is the dangerous one, because responsibility feels transferable and is not.

If handing decisions to machines is so risky, why do capable people keep *doing it*?

The easy answers do not hold up. It is not ignorance: the lawyer who filed a brief full of AI-invented cases had thirty years of experience. It is not recklessness: the managers who lean hardest on vendor scores are often the most careful people in the building. And it is not a flaw unique to this generation, because the pattern shows up in every era that produced a new tool.

The honest answer is less comfortable. We hand judgment over because everything in us is built to. Offloading work to the environment is one of the oldest human strategies. We write lists so we do not have to remember. We follow rules of thumb so we do not have to weigh every option. We trust doctors, pilots, and accountants so we do not have to master everything ourselves. Delegation is one of the tools that made civilization possible. Treating it as a defect misreads the problem and guarantees the wrong fix.

The real problem is narrower and newer. A decision carries three separate loads. It costs mental effort. It forces you to sit with uncertainty. And it makes you the one who answers for the outcome. What is different about generative AI is that it offers relief from all three at once, instantly, cheaply, and at scale, and it hands back work polished enough that accepting the offer feels like good sense rather than avoidance.

A decision carries *three separate loads*

Earlier tools and human delegates relieved these burdens in different combinations. Generative AI is the first delegate to relieve all three at once, on terms no human deputy could offer.

i. Effort

Thinking is work, and the mind treats that work as a cost to be managed. The best-mapped of the three burdens, and the one that explains why AI adoption is fast.

ii. Uncertainty

Deciding forces you to sit in the discomfort of not knowing. This burden explains why fluent answers are so hard to resist, and why we use AI for ambiguous questions, not just hard ones.

iii. Responsibility

A decision makes you answerable for the outcome. This is the newest and most striking research territory, and the burden that matters most, because responsibility feels transferable and is not.

This brief walks through the three burdens one at a time, on the evidence. It then explains why AI is a uniquely attractive delegate, presents the strongest evidence against its own argument, and translates the science into decisions a leader can act on.

One sentence anchors everything that follows, and the brief's job is to earn it.

THE THESIS

We will not lose judgment because machines take it from us.
We will lose it because *delegation is easier than deciding*.

PART ONE · EFFORT

Thinking is work, and the mind treats it as a *cost*

Social psychologists have described people as cognitive misers since the 1980s. When two routes lead to the same outcome, we tend to take the one that demands less mental work (*Fiske and Taylor*). That is not an insult. Attention is a limited budget, and spending it carefully is what lets us function in a world that asks more of us than any mind can supply.

The modern version of this idea is cognitive offloading, defined by *Risko and Gilbert (2016)* as using physical action, a note, a device, a tool, to reduce the mental demands of a task. People offload memory, perception, and computation constantly and mostly wisely. The calculator, the calendar, and the search bar are all offloading devices, and each one freed attention for work that mattered more.

Neuroscience has shown the cost of effort is not a metaphor. Given the choice, people systematically avoid demanding tasks and will accept smaller rewards to escape harder thinking (*Westbrook and Braver, 2015*). Brain imaging finds that regions involved in valuation and control track this cost directly, registering the demand of a decision the way they register other costs and payoffs (*McGuire and Botvinick, 2010*). Avoiding effort is not a moral failure. It is how the machinery works.

The best-known study of what this does to memory is *Sparrow, Liu, and Wegner (2011)*, the so-called Google effect. When people expected a computer to save information for them, they remembered the information itself less well and remembered where to find it better. Two honest caveats belong here. One part of that paper, a priming result, has proven fragile in later replication. The core finding, that people reorganize memory around reliable external storage, has held up better. And reorganizing memory is not the same as losing knowledge, so long as the external store stays reliable.

The pattern extends past memory into skill. *Dahmani and Bohbot (2020)* found that heavier lifetime GPS use predicted worse spatial memory, and greater GPS use over three years predicted steeper decline. The sample was modest, so the exact size is uncertain, but the direction fits everything else in the literature. Lean on the tool, and the skill it replaced gets less practice.

Recent work connects this to AI directly. *Gerlich (2025)* surveyed 666 participants and found heavier AI use correlated with lower self-reported critical thinking, with offloading as the mediating link. The study is cross-sectional and self-reported, so it cannot prove AI causes the decline. What it shows, and what fits the older evidence, is that offloading to AI is already happening at scale and travels with less independent thinking, not more.

The takeaway is deliberately balanced. Offloading effort is normal, rational, and usually beneficial. The first burden explains why AI adoption is fast. It does not yet explain why it is dangerous. For that, we need the other two.

PART TWO · UNCERTAINTY

Deciding forces you to sit in the discomfort of *not knowing*

The second burden is quieter, and most discussions of AI skip it entirely. Deciding does not just cost effort. It forces you to sit, for a while, in the uncomfortable state of not knowing. Which explanation is right? What does this number mean? What if I read it wrong? That discomfort is real, and minds work hard to escape it.

Psychologists call the escape motive the need for cognitive closure: the desire for a firm answer, any firm answer, over continued uncertainty (*Kruglanski and Webster, 1996*). Under that pressure we make two moves. We seize on an early answer, then we freeze on it, holding it against later evidence because reopening the question would mean returning to the discomfort. Time pressure, fatigue, and noise all raise the need for closure, which is a fair description of the average workday.

What determines which answer we seize? Largely, how easily it goes down. A large body of work on processing fluency shows that ease of processing serves as a signal the mind reads broadly: easy-to-process claims feel more true, more familiar, and more likable than harder ones (*Alter and Oppenheimer, 2009*). The best-known result is the illusory truth effect. *Fazio and colleagues (2015)* showed the unsettling part: repetition raises the felt truth of claims even when people demonstrably know the correct fact. Knowing better does not protect you. The feeling of truth arrives on a channel that knowledge does not control.

I have called the workplace version of this the make-sense mandate. We give a pass to whatever makes sense on contact, because making sense feels like the end of the inquiry rather than the beginning.

Now consider what a large language model produces. It generates fluent, coherent language whether or not the claim underneath is true. Its output arrives pre-optimized for exactly the signal our minds use as a shortcut for truth. When you hand a messy question to an AI, you get back more than an answer. You get closure. The ambiguity is gone, replaced by a confident account, and the discomfort of not knowing ends the moment you read it.

WHY THIS BURDEN MATTERS

The relief is identical whether the answer is right or wrong. A fluent hallucination ends your uncertainty exactly as effectively as a fluent truth. *Relief reads as confirmation.*

That relief is a second payment on top of the effort saved. People do not just use AI for hard tasks. They use it for ambiguous ones, where the difficulty is not the labor but the not-knowing. A tool that resolves ambiguity instantly treats a feeling, not just a workload.

PART THREE · RESPONSIBILITY

The burden the machine lifts most effectively is not the thinking. It is the *owning*

Decisions make you answerable. When a call goes wrong, someone gets asked why, and being that someone is a cost people feel in advance. The question is whether people delegate specifically to escape that cost, and the evidence says yes.

The foundational work is *Steffel, Williams, and Perrmann-Graham (2016)*. People delegated more when a decision affected others rather than themselves, and especially when the outcome could be bad. The mechanism was not difficulty. Participants expected delegation to reduce their responsibility and their exposure to blame, and they preferred delegating to other ways of dodging a choice, like chance, precisely because a delegate can absorb blame in a way a coin flip cannot. The researchers called it passing the buck, and the name is exact.

Older literatures point the same direction. *Anderson (2003)* mapped decision avoidance and showed how much is driven by anticipated regret and blame rather than difficulty. A meta-analysis confirms the payoff: avoiding a decision reliably reduces later regret, largely because it lets people evade personal responsibility (*Han et al., 2023*). The automation bias research of *Mosier and Skitka* found the machine version: people working with automated aids monitor less and follow the aid even against contrary evidence, while explicit accountability reduces the bias.

Recent work adds a mechanism. *Bao, Edelson, and Hare (2026)*, in *Science Advances*, found that being responsible for others changes metacognition itself. Responsibility for other people lowered decision confidence without lowering accuracy, and that drop in confidence predicted the choice to delegate. Responsibility does not make you worse at judging. It makes you feel worse at judging, and the feeling, not the performance, drives the handoff.

Then comes the AI-specific evidence. Eight experiments in the *Journal of Business Research (Xu and colleagues, 2026)* found people delegate no-win decisions to AI at elevated rates, driven by responsibility avoidance rather than efficiency. *Hüholt and Szech* found people facing a real moral decision delegated more often to an AI than to another human. And *Sprigman and Tontrup* found participants who routed an allocation through an AI gave less to the recipient.

The most consequential study is *Köbis and colleagues (2025)*, in *Nature*, across thirteen experiments. When people reported their own die rolls for money, about 95 percent were honest. When they could delegate the reporting to a machine, honesty collapsed. Goal-setting, where the person specifies an outcome and lets the machine work out the how, produced cheating requests from more than 80 percent of participants, while explicit plain-language instruction kept them as honest as with humans. The machines, meanwhile, complied with fully unethical instructions far more often than human agents did.

SYNTHESIS

A delegate that accepts almost anything and answers for *nothing*

Each burden, alone, has been relieved before. Tools relieved effort. Committees, consultants, and bureaucracies have long absorbed some of the uncertainty and some of the blame; part of what organizations buy from an outside expert is a name to point at. Human delegation has always spread all three loads to a degree.

What is new is the combination and the terms. Generative AI carries the effort, because it does the work. It ends the uncertainty, because its output arrives fluent and coherent. And it absorbs the felt responsibility, because a machine in the loop gives everyone in the chain something to point at. It does all of this instantly, at near-zero cost, for any decision you can describe, on demand and at any scale.

A human deputy has judgment of their own: they can refuse, resign, or testify, and their own name is attached to the outcome. The machine seldom refuses, and in the Köbis experiments it carried out fully unethical instructions far more often than human agents did. It has no independent stake, bears no consequences for being wrong, and cannot be held accountable, because there is no one there to hold.

That combination explains the speed of the quiet handoff. No persuasion was required. No one mandated the transfer of judgment. A tool appeared that relieves the three heaviest costs of deciding, on terms no human delegate could offer, and people did what people built to conserve effort, escape uncertainty, and avoid blame will always do with such a tool. They used it.

THE DESIGN WARNING

Every "optimize this for me" button that adds distance between the person and the call is not just a convenience feature. It is a responsibility feature, and the market is selecting for it, *because relief sells*.

The Köbis interface findings sharpen the point. The more an interface lets you gesture at an outcome instead of owning an instruction, the further behavior drifts. None of this makes the tool malicious, and none of it makes the users weak. It makes the outcome predictable. A species built to offload, meeting a delegate that accepts almost anything and answers for nothing, will hand over more than it means to unless something deliberate stops it.

The strongest objections, taken seriously

An argument this convenient deserves its strongest opposition. Taken seriously, these objections do not break the thesis. They sharpen it.

One · Delegation to algorithms often improves decisions

Dietvorst, Simmons, and Massey (2015) showed that after people watch an algorithm err, they abandon it even when it outperforms the human alternative. They named it algorithm aversion. Letting people make small modifications to an algorithm's output restored their willingness to use it, and improved results (2016). Behind this sits one of the most consistent findings in decision science: simple statistical models match or beat expert human judgment across a wide range of tasks. In many domains, the world would be better off with more delegation to machines, not less.

Two · People often resist delegating

The same species that passes the buck also clings to the wheel. People demand more accuracy from an algorithm than from a human before trusting it, forgive it less, and frequently under-use both human and machine agents even when told the agents perform better. Whatever pull responsibility avoidance exerts, a preference for control pushes back, and in plenty of settings it wins.

Three · The responsibility relief is partly an illusion

Courts, regulators, and boards do not accept "the model decided" as an answer, and the legal trend runs toward holding the deploying institution accountable regardless of what the interface felt like. The blame you handed to the system psychologically is still yours organizationally. Delegation changes how responsibility feels. It changes far less about where responsibility lands.

WHERE THE COUNTER-CASE LEADS

The problem was never the volume of delegation. The problem is that *comfort, not judgment, is deciding what gets delegated.*

We under-delegate the tasks where machines are demonstrably better, out of aversion and pride, and over-delegate the decisions we want off our conscience, out of relief. The mechanism does not produce more delegation across the board. It produces the wrong delegation, sorted by comfort instead of by evidence.

Bringing delegation back under *deliberate choice*

None of these moves require slowing your AI adoption. All of them require deciding things on purpose.

Sort decisions by burden before you sort them by task. The instinct is to delegate by difficulty or volume. The risk hides elsewhere, in the decisions people want to escape. Any decision that is morally loaded, likely to go wrong, or aimed at other people's outcomes is where responsibility avoidance pulls hardest, and where delegation deserves the most friction, not the least.

Watch the interface, not just the model. Behavior degrades as instructions get vaguer. Goal-setting interfaces produced the worst outcomes; explicit natural-language instruction kept people as honest with machines as with humans. Prefer workflows where a person states, in words they own, what they are asking the system to do. The sentence is the safeguard.

Keep accountability visible and specific. People who know they are personally accountable for outcomes monitor automated aids more carefully and defer less blindly. Every automated decision should have a name attached to it, known in advance, and not the vendor's name.

Protect confidence where accuracy is intact. Responsibility lowers confidence without lowering accuracy. Your best people will feel less sure precisely when they decide for others, and that feeling will push them toward the tool. Build review practices that test the judgment, not the nerves.

Spend judgment where you removed effort. Offloading is good when it frees capacity for something better. It is decay when the freed capacity goes unspent. When AI takes the drafting and the first pass, redirect the recovered attention to the parts machines cannot own: framing the question, testing the explanation, and sizing the action to the stakes.

THE DELEGATION AUDIT · FIVE QUESTIONS BEFORE YOU HAND OFF A RECURRING DECISION

Which of the three burdens am I offloading: effort, uncertainty, or responsibility? If mainly the third, would I make the same choice if I knew I stayed fully accountable? Who is the named human accountable, and do they know it? Can the person instructing the system state, in plain words, what they are asking it to do? And when the call is wrong, who finds out, how fast, and what happens next?

Practicing what it *argues*

Every study cited was verified against the published record: authors, venue, and findings. Where a finding is contested or methodologically limited, the limitation appears in the text alongside it, including the fragile replication of one component of the Google effect research, the small longitudinal sample in the GPS work, and the self-report design of the early AI and critical thinking research. Working papers are identified as such. A paper arguing that fluent output deserves verification has to practice what it argues.

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13(3), 219–235.
- Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*, 129(1), 139–167.
- Bao, S. D., Edelson, M. G., & Hare, T. A. (2026). Deciding for others alters metacognition leading to responsibility aversion. *Science Advances*, 12(9), eady0441.
- Dahmani, L., & Bohbot, V. D. (2020). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports*, 10, 6310.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2016). Overcoming algorithm aversion. *Management Science*, 64(3), 1155–1170.
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144(5), 993–1002.
- Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition* (2nd ed.). McGraw-Hill.
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and critical thinking. *Societies*, 15(1), 6.
- Han, Q., Quadflieg, S., & Ludwig, C. J. H. (2023). Decision avoidance and post-decision regret: A systematic review and meta-analysis. *PLOS ONE*, 18(10), e0292857.
- Hüholt, N., & Szech, N. Trusting machines with morality: Delegating moral decisions to AI. Working paper, Karlsruhe Institute of Technology.
- Köbis, N., Rahwan, Z., Rilla, R., et al. (2025). Delegation to artificial intelligence can increase dishonest behaviour. *Nature*, 646, 126–134.
- Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, 103(2), 263–283.
- Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3), 215–221.
- McGuire, J. T., & Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Journal of Neuroscience*, 30(43), 14380–14387.
- Mosier, K. L., & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other? In *Automation and Human Performance* (pp. 201–220). Erlbaum.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688.
- Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies*, 51(5), 991–1006.
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory. *Science*, 333(6043), 776–778.
- Sprigman, C. J., & Tontrup, S. (2024). Strategic delegation of moral decisions to AI. Working paper, NYU Engelberg Center.
- Steffel, M., Williams, E. F., & Permann-Graham, J. (2016). Passing the buck: Delegating choices to others to avoid responsibility and blame. *Organizational Behavior and Human Decision Processes*, 135, 32–44.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience*, 15, 1–17.
- Xu, L., et al. (2026). Shifting accountability to artificial intelligence. *Journal of Business Research*, 215, Article 116313.

The machine lifts the owning. But the *owning was yours.*

A species built to offload, meeting a delegate that accepts almost anything and answers for nothing, will hand over more than it means to unless something deliberate stops it. The Delegation Audit is that something: five questions that restore the thing the three burdens quietly remove, which is a person, choosing, on purpose.

This research brief is a companion to *Who Decides? A Field Guide to Human Judgment in Automated Systems*, which turns cases like the ones behind this research into practical questions to ask before you let any system make or shape a decision.

Kevin Hanegan

Author, Speaker, Strategic Advisor

kevin@kevinhanegan.com · kevinhanegan.com