

Beyond “Prompt In, Answer Out”: A Reading List for Agentic Work

Recommended Reading from Tech Founders

2026-08-07

Beyond “Prompt In, Answer Out”: A Reading List for Agentic Work

By Recommended Reading from Tech Founders • August 7, 2026

The strongest recommendations move from prompt-in/answer-out toward systems that specify work, track state, and make progress measurable; the surrounding reading asks how to direct ambition and attention.

The most useful diagnosis: agents are workflows, not chatbots

“Some obvious & non-obvious reasons I think AI agents may not have really been widely adopted yet, even though the tech is ready” — *X thread* · Creator: @BadCapitalVC · Recommended by: Aaron Levie.

Levie calls the post “a great place to start” for understanding real-world agent adoption. Its central distinction is that an agent is a process you set up and steer, not a chatbot that simply returns an answer; prompting is therefore closer to writing a specification, including a definition of “done.” [1]

The thread extends that diagnosis into a practical checklist: delegation, technical setup, agent-to-agent and agent-to-colleague handoffs, asymmetric trust when an agent can send or edit something, and the absence of a public job-to-be-done. [2] Levie adds that the payoff requires new data flows, cross-organizational context, and redesigned human review—not merely putting an agent on top of an unchanged workflow. [1]

Why it matters: This is the clearest adoption filter in the set. When assessing an agent product, ask whether it defines the work, supplies the context, manages handoffs, and makes review safe; a better prompt alone is not the transition Levie is describing.

A benchmark for long-horizon agency

Factorio Learning Environment (FLE) v0.3.0 — *evaluation environment / benchmark* · **Creator:** Factorio Learning Environment project · **Recommended by:** Tobi Lütke.

Tobi’s endorsement is deliberately provocative: “Let’s all agree that this is the correct and final eval for agi.” [3] The release describes FLE 0.3.0 as an environment for testing agents on long-term planning, reasoning, and world modelling. [4] Its lab-play benchmark gives an agent fixed resources and a production target, has it write Python against the FLE API, and limits the run to 64 steps; targets are 16 units per minute for solids and 250 for fluids. [4]

The useful difficulty is operational rather than exam-like. FLE says its model ranking is closer to an economically oriented benchmark than to static exam benchmarks, while also warning that many successful agents still shuttle resources manually or use chests and belts as buffers. A 60-second holdout period is used to reduce that shortcut. [4]

Why it matters: FLE gives Tobi’s claim a concrete basis: it tests whether an agent can maintain a model of a changing environment and construct robust multi-step automation, not just produce a plausible answer. Treat “final eval” as Lütke’s standard, not settled consensus; the release itself presents a constrained lab-play setting with known measurement limitations.

A serious reading path into agency: Spinoza

Three biographies of Spinoza, by Nadler, Goldstein, and Stewart — *book set* · **Recommended by:** Garry Tan. (*The talk names the authors but does not supply full titles or links.*)

Tan calls these three biographies “the best biographies about the man.” He says an agent read roughly 1,500 pages and produced a chronology, a map of disagreements among the biographers, and cited quotations that he used to build his talk. [5] He also tells prospective founders that they could do well to learn from Spinoza, whose concept of joy he summarizes as an increase in one’s power of acting. [5]

Why it matters: This is a substantive humanities recommendation with a direct founder lens: the biographies provide a way into a thinker Tan uses to ask whether a tool, decision, or practice increases a person’s capacity to act. The recommendation is also unusually specific about how to read a difficult corpus—compare biographies, preserve disagreements, and retain the quotations that survive the comparison.

Priority and attention

“First and Second Things” — *essay* · **Creator:** C.S. Lewis · **Recommended by:** David Perell.

Perell points to Lewis’s essay for a paradox: putting the highest thing first can bring the second thing with it, while putting the second thing first can cost both. [6] He connects that idea to the warning that results are rarely achieved by obsessing over results themselves, and that a less outcome-dependent identity can create more freedom to perform. [7]

Why it matters: It is a compact decision rule for ambition: keep the activity or value that makes the outcome possible ahead of the outcome as an identity test. (*Perell’s post did not provide a direct link to the essay; the link above is a reading pointer.*)

Tim Ferriss offers a related, more personal set of cues. He says seven quotes shaped his thinking and changed his behavior over the previous year, and that he revisits them often. [8] Two of the book-derived recommendations are:

- ***Notes of a Native Son*** — *book* · **Creator:** James Baldwin. Ferriss selects Baldwin’s observation that people may cling to hatred because abandoning it would force them to face pain. [8] **Why it matters:** It is a prompt to examine the emotional function of a story or grievance before treating it as a settled belief.
- ***The Art of Possibility*** — *book* · **Creators:** Rosamund and Benjamin Zander. Ferriss highlights the instruction: “I am here today to cross the swamp, not to fight all the alligators.” [8] **Why it matters:** It directs attention toward completing the central passage rather than spending the day responding to every obstacle.

Sources

1. X post by @levie
2. X post by @BadCapitalVC
3. X post by @tobi
4. Factorio Learning Environment
5. Garry Tan: “Personal AGI Is How You Stay Under Your Own Power”
6. X post by @david_perell
7. X post by @MVP_Mindset
8. X post by @tferriss