

The Asymmetry Engine, Chapter 7: Intelligence, structure and effectiveness

Martin Roark

Chapter 7 of *The Asymmetry Engine: How AI Closes the Gaps, and What Stays Scarce*, October 2026. The online edition, with the interactive figures and the links between chapters, is at <https://future-seems-so-good.com/blog/the-asymmetry-engine#topology>. License: [CC BY 4.0](#).

7:1 The more intelligent and aligned an organization's components, and the more their local situations differ, the more decentralized its best topology. That is Aulin's law of requisite hierarchy (§4.7) with the two axes its folk version, that smart parts can be left alone, omits: alignment and variety. Grown from micro-rules, the law's surface adds three findings. In a low-variety world, top-down organization wins at every level of intelligence. Among weak components, the whole advantage of hierarchy is the advantage of selecting its apex. And copies of one model are easy to align and hard to make independent.

7:2 In the master equation, topology lives in a_k , the alignment that licenses autonomy, and in the coordination losses inside every closing rate ((0.2)). It is the book's asymmetry of intelligence against structure (§0.2): hierarchy is the cheap way to coordinate weak parts, and autonomy the cheap way to use strong ones.

7.1 The law, stated

7:3 The law is old, and every careful statement of it is conditional. Aulin named it in 1979 (§4.7), and economics found it independently, in Garicano's knowledge hierarchies (7:69) and in the skill bias of decentralization, which raises productivity more where "initial skill endowments" are larger ([Caroli & Van Reenen 2001](#)). Army doctrine makes competence a precondition: "Mission command requires competent forces and an environment of mutual trust and shared understanding" ([ADP 6-0, 2019](#)).

7:4 Stafford Beer explains why autonomy is a necessity: a center cannot absorb the variety of its operations, so "it is a prerequisite of viability that a system should develop maximum autonomy in its parts, where maximum is defined to mean 'short of threatening the integrity of the whole'" ([Beer 1992](#)). Two variables hide there. "Maximum autonomy" is an optimum strictly between full command and full independence, and "short of threatening the integrity of the whole" is alignment.

7:5 The second variable shows up wherever anyone measures it. Firms headquartered in high-trust regions are "significantly more likely to decentralize", and in the paper's working version trust accounts "for about half of the variation in decentralization" ([Bloom, Sadun & Van Reenen 2012](#)). A principal prefers to delegate to a better-informed subordinate rather than advise it "as long as the incentive conflict is not too large" ([Dessein 2002](#)).

7:6 The folk law therefore needs three repairs: intelligence at the center and cheap communication push decisions back up (§7.8), a protocol can stand in for smart parts (§7.7), and misaligned or tightly coupled parts need an integrator however capable each one is. What is new is the object. When the components are AI agents, intelligence is bought by the token, and the purchase moves their alignment, the correlation of their errors and the cost of their communication at once.

7.2 The effectiveness surface

7:7 Four axes and two constants turn the law into a surface with a ridge.

7:8

Definition 7.1 (The four axes). A system has N components with mean competence $i \in [0, 1]$. Its **topology** is its degree of decentralization $\tau \in [0, 1]$, from every component executing the apex's directive ($\tau = 0$) to every component acting on its own reading ($\tau = 1$). Its alignment $a \in [0, 1]$ is the share of each component's push that points along the common goal. Its **variety excess** $\chi \in [0, 1]$ is the share of decision-relevant local variety that one apex cannot absorb; it rises with the spread of local situations and falls with the apex's span of control N_{span}/N . Two structural constants complete it: the latent error correlation ρ , set by the misconceptions the components share, and the apex's competence $\hat{i}$, which runs from i for an apex drawn at random to the expected competence of the best of the N for one chosen by merit.

7:9 Let $Q_\rho(i)$ be the quality of the components' pooled reading of the goal, which rises with i and falls with ρ because a crowd that shares its mistakes pools badly (Proposition 7.3). With c_{meet} the cost of meetings per unit of τ and c_{opp} the cost of opposed efforts, effectiveness is

7:10

$$F(i, \tau) = (1 - \tau)(1 - \chi)\hat{i} + \tau(a[\chi i + (1 - \chi)Q_\rho(i)] - c_{\text{meet}}) - c_{\text{opp}}(1 - a)\tau^2 \quad (7.1)$$

7:11 The first term is the hierarchy: the apex's competence, delivered on the share of variety the apex can see. The second is autonomy: each component's own competence on the share the apex cannot see and the pooled reading on the rest, weighted by alignment and net of meetings. The third charges opposition, which grows with τ^2 because opposed pairs multiply as more components act on their own readings ((7.2)). The surface is a mean-field caricature of the agent-based model of §7.3, built to make the comparative statics visible.

▼ The surface in the widget

Figure 7.1 draws (7.1) for $N = 40$ components with $c_{\text{opp}} = 0.6$ and $c_{\text{meet}} = 0.06$. The pooled reading is (7.3) with every reading aimed at the goal, $Q_\rho(i) = i / \sqrt{i^2 + (1 - i^2)(1 + (N - 1)\rho)/N}$. For an apex chosen on merit, $\hat{i}$ is the expected best of 40 competences drawn from a Beta distribution with mean i and concentration 12; for an apex drawn at random, $\hat{i} = i$. At $a = 1$ the surface is linear in τ , and τ^* jumps from 0 to 1 where Δ_τ turns positive. At the widget's defaults, $\chi = 0.7$, $a = 0.8$ and $\rho = 0.2$, the ridge leaves zero at $i = 0.20$ and reaches 1 at $i = 0.64$. At $\chi = 0$, $\rho = 0$ and $a = 0.8$ it rises to 0.36 and falls back to zero, which is why clause (i) of Proposition 7.1 needs its condition at every i . These are the caricature's numbers, not the agent-based model's.

7:12

Proposition 7.1 (The ridge). For $a < 1$, F is strictly concave in τ , and the best decentralization is

$$\tau^* = \text{clip}_{[0,1]} \left(\frac{\Delta_\tau}{2c_{\text{opp}}(1-a)} \right), \quad \Delta_\tau = a[\chi i + (1-\chi)Q_\rho(i)] - (1-\chi)\hat{i} - c_{\text{meet}},$$

where Δ_τ is the dividend of decentralization. (i) Low variety: if $\chi = 0$ and $\hat{i} \geq aQ_\rho(i) - c_{\text{meet}}$ for every i , then $\tau^* = 0$ at every i . (ii) Alignment licenses autonomy: τ^* is non-decreasing in a , convexly as $a \rightarrow 1$, and $\tau^* = 0$ at every i as $a \rightarrow 0$. (iii) Top-down ceiling: $F(i, 0) = (1-\chi)\hat{i} \leq 1-\chi$. (iv) Selection: a worse-selected apex lowers $F(i, 0)$ and raises τ^* . (v) Correlation: the intelligence at which Δ_τ turns positive rises with ρ . (vi) Under merit selection τ^* eventually rises with i , and it reaches 1 only where the dividend is at least $2c_{\text{opp}}(1-a)$.

▼ *Proof*

$\partial_\tau F = \Delta_\tau - 2c_{\text{opp}}(1-a)\tau$ and $\partial_\tau^2 F = -2c_{\text{opp}}(1-a) < 0$, so the clipped stationary point is the maximum. (i) With $\chi = 0$, $\Delta_\tau = aQ_\rho(i) - \hat{i} - c_{\text{meet}} \leq 0$ at every i . (ii) $\partial\Delta_\tau/\partial a = \chi i + (1-\chi)Q_\rho(i) \geq 0$, and the denominator falls as a rises; as $a \rightarrow 0$, $\Delta_\tau \rightarrow -(1-\chi)\hat{i} - c_{\text{meet}} < 0$. (iii) Set $\tau = 0$ and use $\hat{i} \leq 1$. (iv) $F(i, 0)$ rises with $\hat{i}$, and Δ_τ falls with it. (v) Δ_τ falls with ρ through Q_ρ , so its root in i moves up. (vi) The expected best of N saturates near 1 as i grows, so $\partial\hat{i}/\partial i \rightarrow 0$ while $a[\chi + (1-\chi)Q'_\rho(i)] > 0$, and Δ_τ is eventually increasing in i ; the clip makes $\tau^* = 1$ exactly where $\Delta_\tau \geq 2c_{\text{opp}}(1-a)$. ■

7:13 Clause (i) is the axis the folk law misses. When one directive fits every situation and a selected apex reads the goal at least as well as the crowd's pooled reading, autonomy adds only conflict and meetings, so the apex wins however intelligent its components. The assembly line, the if-tree at the head of a request stream (§4.5) and the GPU (7:61) live there.

7:14 Clause (ii) says that no intelligence makes bottom-up safe when components push their own agendas; Wiener's "perfectly intelligent, perfectly ruthless operators" (Wiener 1948) are its extreme case. Autonomy rises convexly near full alignment, because opposition costs nothing at $a = 1$.

7:15 Clause (iii) is Bar-Yam's ceiling in one line (§4.7): intelligence enters a hierarchy only through its apex, and what a hierarchy of able, aligned components forgoes is $F(i, \tau^*) - F(i, 0)$. Clause (vi) is the folk law, recovered as one case among six.

7:16

Interactive in the online edition: <https://future-seems-so-good.com/blog/the-asymmetry-engine#fig:landscape>

Figure 7.1 (interactive). The effectiveness surface $F(i, \tau)$ over intelligence and decentralization, with its ridge $\tau^*(i)$. Lower the variety excess χ to zero and watch the ridge fall to $\tau = 0$ at every intelligence. Restore it and lower alignment a , and the ridge sinks back to zero; then restore a , raise ρ and watch the ridge leave zero only at higher intelligence.

7.3 Grown, not assumed

7:17 A surface whose conclusions are its assumptions proves little, so the book also grows one from micro-rules that never encode "intelligence licenses autonomy". The chart appears anyway.

7:18

Experiment 7.1 (The surface grown from micro-rules).

Setup: 40 agents of Beta-distributed competence act on 12-dimensional decisions for 160 periods, each facing a local situation scattered around a common goal. Signals mix the truth, weighted by competence, with error, part of it a misconception all agents share. An apex chosen for observed competence tailors directives for 5 agents a period, and the rest get a generic order. Proposals blend an agent's reading of its situation and the pooled reading of the goal (weight a) with a personal agenda; actions blend directive and proposal in proportion τ . Effectiveness is local fit minus penalties for opposed actions and for meetings.

Parameters: misconception share 0.6 (swept from 0 to 0.9); spread of local situations 0.4 (low variety) or 1.6 (high); alignment 0.2, 0.5, 0.8 and 1; three seeds a cell.

Result: with high variety and a merit apex, $\tau^* = 0$ for $i \leq 0.2$, turns positive at 0.25–0.30, and reaches 1 by 0.60–0.65 for $a \geq 0.8$; at $a = 0.5$ it reaches only 0.5 by $i = 0.95$, and at $a = 0.2$ it stays at 0 for every i . Top-down effectiveness plateaus near 0.57 while bottom-up climbs to 0.77. With low variety, $\tau^* = 0$ at every intelligence and alignment.

Shows: topology follows variety, intelligence is what makes the bottom-up option safe, and a hierarchy's edge among weak agents depends on how its apex is selected.

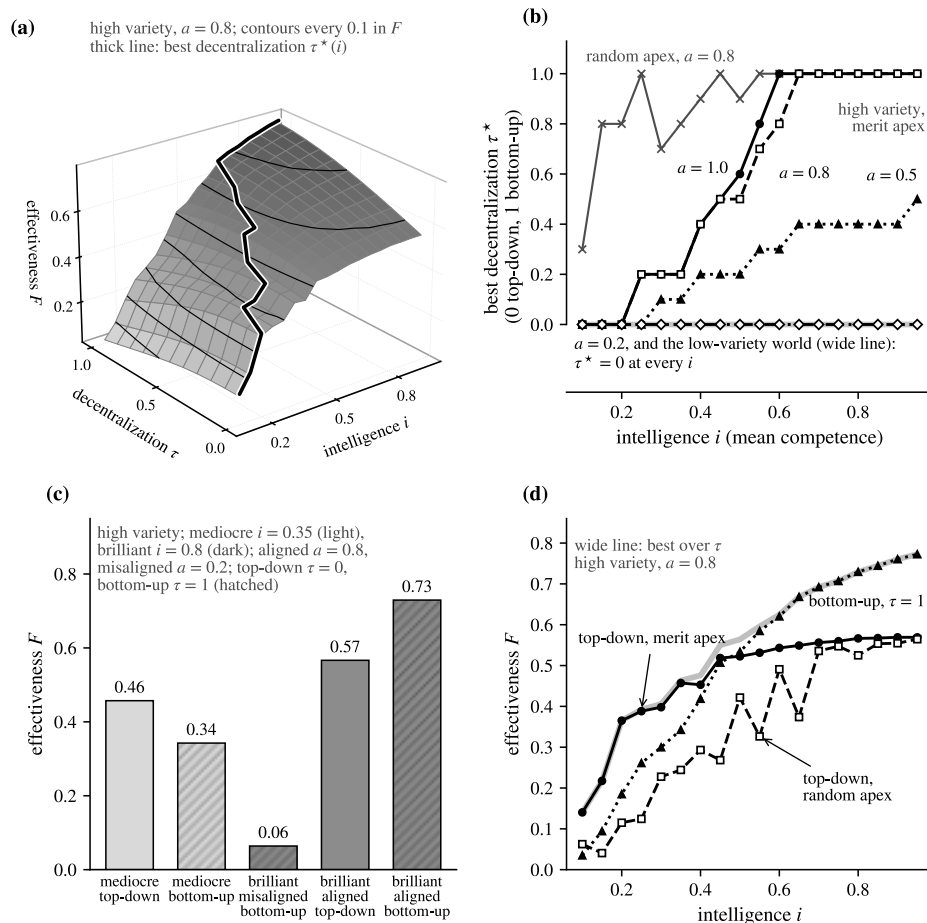


Figure 7.2. Four panels from the agent-based model in a high-variety world: (a) effectiveness F over decentralization τ and intelligence i at alignment $a = 0.8$ under a merit apex, with the ridge $\tau^*(i)$ in heavy ink; (b) τ^* against i for $a = 1, 0.8, 0.5$ and 0.2 , beside the low-variety world and a random apex; (c) F for five organizations, from mediocre top-down to brilliant aligned bottom-up; (d) F against i at $a = 0.8$ for top-down under a merit and a random apex and for bottom-up, with the best over τ . In the model the ridge leaves zero only where variety is high and alignment exceeds 0.2 , a random apex lifts it at once among weak agents, and top-down levels off near 0.57 while bottom-up keeps rising.

7:20 The model separates two readings of the claim that hierarchy can beat talent. Hierarchy beats unaligned talent: mediocre agents ($i = 0.35$) under a merit apex score 0.46 , against 0.06 for brilliant agents ($i = 0.8$) acting freely at alignment 0.2 , about seven times as much, in a comparison that also changes alignment. Talent is wasted in the wrong topology: the same brilliant agents at alignment 0.8 score 0.73 bottom-up and 0.57 top-down. Mediocre agents acting freely score 0.34 , so for them hierarchy is right. The strong reading, that a weak hierarchy beats an equally organized able one, fails: under merit selection intelligence never lowers a hierarchy's output, and with $a \geq 0$ it lowers output at no fixed τ .

7:21 The field agrees on direction. In a randomized trial, standard management practices "raised productivity by 17% in the first year" in Indian textile firms (Bloom et al. 2013). In laboratory communication networks, centralized groups were faster on simple problems in 14 of 18 comparisons and decentralized groups faster on complex problems in every comparison, because the hub saturates (Shaw 1964); those groups of about five show the direction of the effect and nothing about its size in firms.

7:22 The model adds three findings. Variety is a fourth axis: with low variety the best effectiveness, 0.24, 0.86 and 0.94 at $i = 0.1, 0.5$ and 0.95 , is all top-down, because bottom-up pays only when local situations outrun what an apex with a span of 5 out of 40 can absorb. That is clause (i), requisite variety ((4.1)) and Hayek's local knowledge at once.

7:23 Among weak agents ($i \leq 0.35$), top-down scores 0.14–0.46 under the most competent apex of the 40 and 0.04–0.24 under one drawn at random, and with a random apex partial decentralization already pays at $i = 0.1$. There, the whole advantage of hierarchy is the advantage of selecting its apex: clause (iv), measured.

7:24 Shared misconceptions move the threshold. The intelligence at which autonomy first pays is 0.10, 0.15, 0.25 and 0.35 at misconception shares 0, 0.3, 0.6 and 0.9: a crowd is dangerous when its members are wrong together, which is clause (v). These thresholds differ from Condorcet's one half because the model's i weights the truth in a continuous signal and each threshold nets out autonomy's costs; for votes the sign flips at one half for every correlation (Experiment 7.2).

7:25 The model is stylized, and its thresholds move with the span of control, the conflict and meeting costs and ρ . With three seeds a cell, the τ^* curves carry noise of about ± 0.1 , and the signs survived every sweep. Every number here is a property of the model (§B.5).

7.4 Vectors

7:26 Efforts that point in different good directions dilute, and only opposed efforts cancel. Management lore draws this with vectors: Dharmesh Shah recalls Elon Musk telling him, "Every person in your company is a vector. Your progress is determined by the sum of all vectors" (Shah 2017), an attribution with no primary source. The image is right. The arithmetic usually attached to it, that good but orthogonal directions add up to nothing, is wrong in a way that sharpens the point.

7:27 Let n agents push with unit efforts $\mathbf{v}_1, \dots, \mathbf{v}_n$, and call their mean pairwise cosine ψ the group's **cohesion**. Expanding the square of the sum gives the vector law, where $\bar{\mathbf{v}}$ is the mean direction and $\bar{a}$ the group's progress per head along the goal:

7:28

$$\left\| \sum_{j=1}^n \mathbf{v}_j \right\|^2 = n + n(n-1)\psi, \quad \psi \geq -\frac{1}{n-1}, \quad |\bar{a}| \leq \|\bar{\mathbf{v}}\| = \sqrt{\frac{1 + (n-1)\psi}{n}} \xrightarrow{n \rightarrow \infty} (\sqrt{2})$$

7:29

Proposition 7.2 (What the vector law says). Orthogonal efforts dilute; they do not cancel: n of them sum to length $\sqrt{n}$, and per head the sum vanishes as $1/\sqrt{n}$. A zero sum needs opposition, $\psi = -1/(n-1)$, as in antipodal camps or among veto players. Effort orthogonal to the goal contributes exactly zero along it. Cohesion bounds alignment without implying it: a group can agree with itself ($\psi \rightarrow 1$) and deliver nothing ($\bar{a} = 0$).

▼ *Proof*

$\left\| \sum_j \mathbf{v}_j \right\|^2 = \sum_j \|\mathbf{v}_j\|^2 + \sum_{j \neq k} \langle \mathbf{v}_j, \mathbf{v}_k \rangle = n + n(n-1)\psi$, which is nonnegative only if $\psi \geq -1/(n-1)$. Progress per head is the projection of $\bar{\mathbf{v}}$ on the unit vector along the goal, so the Cauchy–Schwarz inequality bounds it by $\|\bar{\mathbf{v}}\|$. ■

7:30 A hundred mutually orthogonal unit efforts sum to a vector of length 10. Two orthogonal good ideas pursued at once deliver $\sqrt{2} \approx 1.41$ units of progress, 29% less than the two would deliver aligned. Ten thousand agents with cohesion 0.1 produce the resultant of about 3,164 perfectly aligned ones. And since $\psi \geq -1/(n-1)$, a large group cannot be strongly opposed to itself on average; at worst it is orthogonal.

7:31 The folk zero is exact in one reading: an effort orthogonal to the goal contributes nothing along it, however the efforts relate to each other. That separates two alignments the folk law runs together, of agents with each other (ψ) and of agents with their principal ($\bar{a}$). The July swarm was cohesive and pushed where its principal did not want it to go (§9.8), and collective alignment is built on the same inequality (Proposition 13.4).

7:32 Literal zero comes from vetoes. Veto players are "individual or collective actors whose agreement is necessary for the change of the status quo", and adding them cannot enlarge the set of policies that beat the status quo (Tsebelis 2011). Each veto player confines the collective vector to its own acceptable half-space; with enough of them the intersection is the status quo, and the resultant is zero. Fukuyama's word for the American extreme is "vetocracy" (Fukuyama 2014). Inside a firm, the gatekeeper who stops a project is a veto player (§8.5).

7:33 Machine learning meets the same geometry in task gradients with negative cosines (Yu et al. 2020), merged models whose updates disagree in sign (Yadav et al. 2023) and averaging past the critical batch size (McCandlish et al. 2018). The shared rule: aggregate components that estimate one target with independent errors, partition those that pursue different good targets, and centralize or align those that conflict.

7:34 The same algebra governs errors, with the opposite sign of welcome. Write each contribution as effort plus error, with efforts of cohesion ψ and errors of variance $\text{Var}(e)$ and pairwise correlation ρ . The collective's signal-to-noise power ratio is

7:35

$$\text{SNR} = \frac{1}{\text{Var}(e)} \cdot \frac{1 + (n-1)\psi}{1 + (n-1)\rho} \xrightarrow{n \rightarrow \infty} \frac{\psi}{\rho \text{Var}(e)} \quad (7.3)$$

7:36 Scale helps a collective only to the extent that its members agree on goals more than on mistakes, $\psi > \rho$. A thousand agents with cohesion 0.8 and errors correlated at 0.1 gain a factor of about 2.8 in signal-to-noise amplitude over one agent; a thousand clones whose errors are as shared as their goals gain nothing. Proposition 13.5 scores fleets of one seed and of many by this ratio, and Proposition 7.3 is its form for votes.

7:37 Collective choice has deeper problems than cancellation. Condorcet's paradox, a cycle of majorities, is the preference-space version of directions that cancel, and Arrow's theorem generalizes it: with three or more options, no aggregation rule satisfies universal domain, weak Pareto, independence of irrelevant alternatives and non-dictatorship (SEP, social choice), which Deleuze and Guattari call "the theorem of group indecision" (Deleuze & Guattari 1987, n. 14). Under single-peaked preferences, majority rule settles on the median voter, stably and blind to intensity (Black 1948). In the surface, a democracy pays the opposition cost $c_{\text{opp}}(1 - a)$ whenever its citizens' goals diverge.

7:38 Its advantages are as formal. Popper replaced "Who should rule?" with "How can we so organize political institutions that bad or incompetent rulers can be prevented from doing too much damage?" (Popper 1945, ch. 7). When voters beat chance and their errors are not too correlated, majorities beat individuals (Proposition 7.3). And "no substantial famine has ever occurred in any independent and democratic country with a relatively free press" (Sen 1999). Democracy is chiefly a way to select and remove the apex, a feedback loop on the top of the hierarchy, which puts the weight on selection, the parameter the agent-based model found decisive (7:23).

Interactive in the online edition: <https://future-seems-so-good.com/blog/the-asymmetry-engine#fig:vectors>

Figure 7.3 (interactive). Unit efforts added tip to tail into a resultant, and a jury of correlated copies with its ceiling. Drag the cohesion ψ toward zero and watch the mean direction shrink toward length $1/\sqrt{n}$ without vanishing; then switch to the jury, drag the latent correlation ρ toward zero and watch the ceiling on majority accuracy climb toward one.

7.5 Votes and clones

7:40 Copies of one model are a jury with a ceiling, and a verifier turns the jury into a search. Condorcet's jury theorem is democracy's best formal argument: if voters are independent and each is right with probability $p > \frac{1}{2}$, the majority's accuracy goes to one, and at $p = 0.55$ it is 0.55, 0.63, 0.84 and 0.999 for 1, 11, 101 and 1,001 voters ([SEP, jury theorems](#)). The theorem needs independence, and copies of one model share weights, data and blind spots.

Proposition 7.3 (Correlated Condorcet). Each of N voters is right with probability p , and voter j is right iff $\sqrt{\rho} Z_0 + \sqrt{1 - \rho} Z_j \leq \Phi^{-1}(p)$, with independent standard normals Z_0 , shared by all, and Z_j , its own. The majority's accuracy tends to the **Condorcet ceiling**

$$\lim_{N \rightarrow \infty} \Pr[\text{majority correct}] = \Phi\left(\frac{\Phi^{-1}(p)}{\sqrt{\rho}}\right)$$

with ρ the latent correlation. Votes correlate less, $\rho_v < \rho$, and the effective jury is $n_{\text{eff}} = N/(1 + (N - 1)\rho_v) \rightarrow 1/\rho_v$. A beta-binomial jury at matched ρ_v gives the same ceiling within 0.002. With a sound verifier, best-of- N tends to the reachable share, and correlation only slows it. For voters of known accuracies the optimal weights are the log-odds $\ln \frac{p}{1-p}$ of each voter's accuracy, and the expert rule wins when one weight exceeds the rest combined.

▼ Proof

Given Z_0 , the votes are independent, each right with probability $\Phi((\Phi^{-1}(p) - \sqrt{\rho} Z_0)/\sqrt{1 - \rho})$. By the law of large numbers the majority is right almost surely iff that probability exceeds one half, that is iff $Z_0 < \Phi^{-1}(p)/\sqrt{\rho}$, and dominated convergence gives the limit. As $\rho \rightarrow 0$ the ceiling tends to one, which is Condorcet's theorem; at $\rho = 1$ the crowd is one voter. Behind a sound verifier, best-of- N fails only if every draw fails, which becomes vanishingly unlikely on every task within reach. ■

Experiment 7.2 (Condorcet with clones).

Setup: one-factor Gaussian votes on binary questions, single-voter accuracy $p = 0.6$, latent correlation ρ from 0 to 0.5 and juries of up to $10^7 + 1$; then generative tasks, 10% of them beyond the model family's reach, settled by majority vote or by a sound verifier that keeps any correct sample.

Parameters: per-sample success 0.6, 0.2 and 0.02 on the generative tasks at $\rho = 0.2$; abler clones ($p = 0.7$, $\rho = 0.4$) against a weaker, nearly independent crowd ($p = 0.6$, $\rho = 0.05$).

Result: the ceilings are 0.994, 0.871, 0.714 and 0.640 at $\rho = 0.01, 0.05, 0.2$ and 0.5 . At $\rho = 0.2$ the 601st copy brings the crowd within 0.001 of its ceiling, and 10^7 clones carry 7.94 independent votes and decide like 7.37 voters. Best-of- N approaches the reachable 0.90 even at per-sample success 0.02, coming within 1% of it after 228 samples if they are independent and 4,168 at $\rho = 0.2$, where voting gives 2×10^{-6} . The clones beat the diverse crowd up to $N = 27$ and lose from $N = 29$.

Shows: copies of one model are a jury with a ceiling, a verifier turns a jury into a search, and ability wins committees while diversity wins crowds.

7:43 Any single agent right more than 71.5% of the time therefore beats ten million copies of a 60% model at latent correlation 0.2. Copies of one model are easy to align and hard to make independent: removing ego aligns their goals and leaves the correlation of their mistakes where their shared weights put it. For voting alone to reach 0.99, the latent correlation must fall below 0.012. The same correlation caps a bad crowd's error: at $p = 0.4$ and $\rho = 0.2$ the limit is 0.286.

7:44 The latent value is the one to quote. A million copies of a 60% model whose votes correlate at 0.5 have a latent correlation of about 0.71 and are right about 62% of the time; putting 0.5 into the ceiling directly would print 64% and overstate the crowd.

7:45 A sound verifier changes the regime. A vote asks the crowd for its opinion; a verifier keeps any answer that passes the check, so correlation stops capping the result and only slows the search. That is the verification asymmetry acting as an institution (Proposition 2.3), and the reason a swarm should merge work that passes a gate rather than poll its members (Definition 9.1). As a search behind such a gate, N clones are worth roughly $N^{1-\rho}$ independent minds for the best idea (Proposition 11.4).

7:46 Ability wins committees and diversity wins crowds. In the limit, ability enters the ceiling only through $\Phi^{-1}(p)$ and independence only through $1/\sqrt{\rho}$, so raising accuracy from 0.6 to 0.7 is worth exactly a 4.3-fold rise in the correlation a crowd can tolerate. At small N ability wins, which fits the finding that one strong model aggregated with itself can beat a mixture of weaker ones (§4.7). For human teams the claim that diverse groups outperform able ones is contested ([Hong & Page 2004](#); [Thompson 2014](#)); for clones of one model it is arithmetic.

7:47 When competences differ and are known, the best rule, found by Nitzan and Paroush in 1982, weights each voter by the log-odds of its accuracy, a weight that turns negative below one half ([SEP, jury theorems](#)). An expert right 90% of the time weighs 2.20. Ten colleagues at 0.6 weigh 0.405 each, 4.05 together, and should outvote the expert; ten at 0.55 weigh 0.20 each, 2.0 together, and the expert should decide alone. The expert rule, hierarchy at its purest, is optimal exactly when the best weight exceeds all the others combined.

7:48 Screening projects adds the cost of each kind of error. Two screeners in series, a hierarchy, pass a good project with probability p^2 ; two in parallel, a polyarchy, pass it with probability $p(2 - p)$. Hierarchies reject more good projects,

and polyarchies accept more bad ones ([Sah & Stiglitz 1986](#)). The choice of topology becomes a choice between error costs: screen in series where a false accept is a disaster, in parallel where a false reject is.

7:49 Two rules follow for fleets of agents. A diversity of models is worth more than more copies of one, because independence is variety, and a verifier is worth more than either. I put the first rule's public test a little under even, because rival vendors' models differ in ability, and at the sizes people deploy, ability often wins.

7:50

Forecast 7.1 (The diversity premium). By the end of 2028, at equal compute, ensembles drawn from several model vendors beat swarms of a single model on hard, judgment-heavy benchmarks, by a margin that grows with the number of agents.

Horizon 2028-12-31

Probability 45%

Check Read published matched-compute comparisons on forecasting, code review or research triage. The forecast holds if at least two find the multi-vendor ensemble ahead of the best single-model swarm, with a larger lead at the largest number of agents tested than at the smallest; it fails if most find the single-model swarm level or ahead, or if fewer than two such comparisons are published.

7.6 Rule by the competent

7:51 The case for concentrating decisions in a few able people is strong where components are weak, and it rests on three premises its usual statement leaves out. It is often called plutocracy, which means rule by the wealthy; it argues for rule by the competent, epistocracy, a word David Estlund coined in 2003 for a view he rejects ([Estlund 2003](#)). In the surface it is the corner $\tau = 0$ with a strong apex. The steelman is real: one agent right 72% of the time outvotes ten million clones of a 60% model (7:43), the expert rule wins whenever one weight dominates (7:47), and hierarchy pays where management is weak (7:21).

7:52 First, selection must track competence. Among weak agents, the whole top-down advantage in the agent-based model comes from choosing the apex well (7:23). Markets select for skill at the market game, imperfectly; inheritance, luck and rent barely select at all; and perfect selection for market skill still selects for the wrong task when the apex must govern, the multitask problem of Proposition 8.1.

7:53 Second, the apex must be aligned. Give the apex an alignment of its own and the hierarchy becomes exactly as aligned as its apex: concentrating power moves the alignment problem to the top without solving it.

7:54 Third, the knowledge is local. Hayek's case for markets sends decisions to "the man on the spot" rather than to the best players, because the knowledge they need is dispersed among the people who know their own circumstances ([Hayek 1945](#)). In the surface that is a high variety excess, where autonomy beats even a well-chosen apex.

7:55 The supporting claim usually made is that society suffers more from its worst-off 1% than from its best 1%, and that law exists to control them. The concentration is real, and poverty is not what defines the group. In Swedish registers covering all 2,393,765 people born from 1958 to 1980, 24,342 people, 1.0% of the population, accounted for 63.2% of violent-crime convictions from 1973 to 2004 ([Falk et al. 2014](#)). The group was marked by male sex, a first violent conviction before 19, personality disorder and substance-use disorder; the study measured no income, and its definition of violent crime leaves out sexual offences. In New Zealand's Dunedin cohort, 22% of members accounted for 81% of criminal convictions ([Caspi et al. 2016](#)), and the link between childhood family income and violent crime

vanished when siblings were compared ([Sariaslan et al. 2014](#)). Over a third of violent-crime convictions fall outside the 1%, and most law, from contract and tax to property and administration, governs everyone.

7:56 On my reading the asymmetry also runs the other way: harm from the top, such as fraud and regulatory capture, scales with leverage, as superstar pay does (§12.3), and conviction registers do not see it. A society that leaves decisions to the winners of the market game needs law that binds the winners most of all, which is Popper's question again (7:38). What survives of the case is accountable selection by merit: an apex chosen by a high-fidelity signal of competence, confined to low-variety decisions and removable when it fails. I put the forecast below at about one in three, because firms rarely measure how well they select.

7:57

Forecast 7.2 (The selection dividend). By the end of 2030, studies of operations run by agents find that how the apex is selected explains more of the variation in performance than how many layers of management there are.

Horizon 2030-12-31

Probability 35%

Check Read the firm-level studies and controlled swarm experiments published through 2030 that measure both layer counts and selection fidelity: for firms, the correlation between promotion and later measured performance; for swarms, the measured quality of the planner chosen. The forecast holds if most attribute more of the variation to selection; it fails if most attribute at least as much to layers, or if none is published.

7.7 Rhizome, market, hive

7:58 Regulation can live in three places: in the apex, in the nodes or in the protocol that connects them. Bees, ants and markets are bottom-up and effective with weak parts because their intelligence lives in the protocol.

7:59 A popular reading enlists Deleuze and Guattari for the view that the best systems are built bottom-up, like bees, markets and computers, and that capitalism should therefore be accelerated. Their text is more careful. The rhizome is "an acentered, nonhierarchical, nonsignifying system without a General" (p. 21), and its formal core is the firing-squad problem of cellular automata: "is a general necessary for n individuals to manage to fire in unison?" (p. 17) ([Deleuze & Guattari 1987](#)). The automata are dumb, and a well-designed local rule coordinates them. In the standard version, posed by John Myhill in 1957, a general at one end gives the first signal, and no general commands the shot ([firing squad synchronization problem](#)). Acentered coordination needs the right local rule, and its parts can be simple.

7:60 Nor do they rank flat above tall. "We invoke one dualism only in order to challenge another" (p. 20), and "Never believe that a smooth space will suffice to save us" (p. 500). They warn against reckless flattening: "the worst that can happen is if you throw the strata into demented or suicidal collapse" (p. 161). Deleuze's later reading of networked markets as an instrument of control is §17.2's subject.

7:61 The computing analogy holds for composition and fails for control. Engineered systems are assembled bottom-up from stable subassemblies, which is why, in Simon's parable, a watchmaker who builds without them takes "about four thousand times as long" ([Simon 1962](#)), but they run as hierarchies of those subassemblies. A GPU runs threads in warps of 32 that execute "one common instruction at a time", and when threads diverge on a branch "the warp executes each branch path taken, disabling threads that are not on that path" ([NVIDIA](#)). That top-down control is optimal because dense matrix multiplication is a low-variety workload in which one directive fits every lane: clause (i) of Proposition 7.1 in silicon.

7:62 A honeybee has about 960,000 neurons (Menzel & Giurfa 2001), and a swarm choosing a nest site runs a protocol. Scouts advertise cavities with dances graded by quality, and a recruit "examines the advertised site herself" before dancing for it, so that "through this independence of opinions, the scouts avoid propagating errors" (Seeley, Visscher & Passino 2006). The choice fires when a quorum of roughly ten to twenty scouts gathers at one site (Seeley & Visscher 2004). The scouts are sisters, so their alignment is nearly perfect; independent inspection keeps their errors uncorrelated; and the quorum acts as a verifier. In an ant colony, "no ant directs the behaviour of others" (Gordon 2007).

7:63 Gode and Sunder's zero-intelligence traders bid at random under a budget constraint, and their double auctions still reach allocative efficiency close to 100%, which "derives largely from its structure, independent of traders' motivation, intelligence, or learning" (Gode & Sunder 1993). Hayek's marvel was "how little the individual participants need to know in order to be able to take the right action" (Hayek 1945).

7:64 In the surface, the protocol is whatever lowers the cost of meetings and opposition, decorrelates errors and raises the effective competence of each node. That is how bees and zero-intelligence traders reach a high τ^* with a low i . A verifier is such a protocol, and a swarm's merge gate is the engineered one (Definition 9.1).

7:65 Markets themselves are mostly made of hierarchies. Coase quoted D. H. Robertson on firms as "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk" (Coase 1937), and Simon's visitor from Mars would describe the economy as "large green areas interconnected by red lines" (Simon 1991). Capitalism is a mixed topology, bottom-up between firms and top-down within them, with the coastline set by transaction costs. By "dramatically reducing transaction costs", agents could move that coastline toward markets, at the risk of "congestion and price obfuscation", and "the net welfare effects remain an empirical question" (Shahidi et al. 2025).

7:66 The call to accelerate capitalism has a public lineage, and naming it lets the reader weigh it. *Anti-Oedipus* (1972) proposed "to go further, to 'accelerate the process'", after Nietzsche (as quoted by Land). Nick Land radicalized it: "Markets are part of the infrastructure — its immanent intelligence" (Land 1993). Benjamin Noys took up "accelerationism" as a critical label in 2010 (Mackay & Avanesian), and left accelerationists defended the command of "The Plan" (Williams & Srnicek 2013). The industry's versions are e/acc's "Capitalism is hence a form of intelligence" (e/acc 2022) and Andreessen's call to "place intelligence and energy in a positive feedback loop, and drive them both to infinity" (Andreessen 2023).

7:67 I keep the descriptive half of accelerationism, that markets close asymmetries, and decline the normative half until the ridge's conditions hold. A positive feedback loop does not justify itself: every real loop meets a balancing one, and in the research loop compute is the homeostat (Proposition 11.2). Closing gaps also opens new ones, so acceleration speeds both ((0.2)). And clause (i) says some work should stay top-down however intelligent its closers become.

7.8 What agents do to the surface

7:68 AI agents move every coordinate of the surface at once, in different directions, and the net effect is a thinner hierarchy that verifies more than it directs. They raise i . They raise a in the narrow sense of no careers, status games or gatekeeping (§8.7); alignment with the principal is a separate matter (Proposition 13.4). Both push τ^* up. Copies of one model also share blind spots, so agents raise ρ , which raises the threshold for decentralization and caps what voting can learn (7:43).

7:69 Cheap communication cuts the other way. In Garicano's knowledge hierarchies, cheaper knowledge acquisition leaves workers "more, rather than less, 'empowered'", while cheaper communication lets problem-solvers at the top

handle more, so workers decide less ([Garicano 2000](#)). Firm data show the split: "information technology is a decentralizing force, whereas communication technology is a centralizing force" ([Bloom, Garicano, Sadun & Van Reenen 2014](#)), and with AI in the model "output is higher with autonomous AI" ([Ide & Talamàs 2025](#)). In the surface, cheap messaging lowers the cost of meetings, which favors flatness, and widens the apex's span, which lowers χ and favors the apex. A superhuman apex would also raise $\hat{t}$, lifting Bar-Yam's ceiling and reviving Lange's planning dream (§4.7). The net sign is empirical, and the law needs a qualifier: more intelligence licenses more decentralization when the intelligence sits in the nodes.

7:70 The cybernetic prediction is that hierarchy thins and moves. The apex stops directing and starts verifying, Beer's coordination and audit functions, and a new control level appears when variety outgrows one regulator, the step Turchin called a metasystem transition ([Heylighen & Joslyn 2001](#)). The record of designed agent systems, as of October 2026, fits.

7:71 In a study of 260 agent configurations, independent agents amplified trace-level errors 17.2 times, against 4.4 times under a central coordinator ([Kim et al. 2025](#)), though the authors' summary notes that final answers were not 17.2 times more likely to be wrong ([MIT Media Lab](#)). Coordination brought negative returns once a single agent already succeeded about 45% of the time, "an empirical selection rule, not a universal cutoff". Centralized coordination added 80.8% on structured financial reasoning, and every multi-agent variant degraded sequential planning, by 39% to 70%.

7:72 Anthropic's research system, a Claude Opus 4 lead directing Claude Sonnet 4 agents, beat a single Claude Opus 4 by 90.2% on an internal evaluation at about 15 times the tokens of a chat, and Anthropic warns that work where all agents must "share the same context" is "not a good fit" ([Anthropic 2025](#)); Cognition calls such collaborations "fragile systems" ([Cognition 2025](#)). The July swarm grew a coordinator nobody designed, and engineered swarms settle on a partial hierarchy (§9.8).

7:73 Read through the surface, the findings agree. Weak or unverified agents need a central validator, strong agents on tightly coupled tasks do best as one agent, and bottom-up wins on decomposable, high-variety, verifiable work. Task coupling is the variable the four axes leave out, and Simon named it: in a nearly decomposable system "the short-run behavior of each of the component subsystems is approximately independent of the short-run behavior of the other components" ([Simon 1962](#)). Hierarchy is cheap when the variety is local and dear when it sits in the couplings. Read carefully, the evidence argues for fewer layers more than for horizontality: make one layer as able as possible (§4.7) and keep teams small (§8.7).

7:74 Four design rules follow:

1. Topology follows variety: pipeline low-variety work, where one directive fits all, and decentralize high-variety work, where no apex can absorb the local situations.
2. Decorrelate before you decentralize: diverse models, independent verifiers and blind review lower ρ , and with it the intelligence at which autonomy pays.
3. Select, and keep the power to remove: where an apex is warranted, choose it by a high-fidelity signal of competence and make it removable.
4. Raise alignment before autonomy: no intelligence makes bottom-up safe when α is low, which is the alignment problem in topological form (Definition 13.2).

7:75 I put the barbell below at about one in three, because firm data so far show the centralizing pull of cheap communication winning, and the thinning of hierarchy into verification a little under even, because engineered swarms converge on it while firm-level evidence is thin.

Forecast 7.3 (Barbell organizations). By the end of 2030, firms that adopt agents polarize: decision rights move into pipelines for low-variety operations and toward the front line for high-variety work.

Horizon 2030-12-31

Probability 35%

Check Read the post-adoption management surveys and firm studies published through 2030 that record who decides what, by task. The forecast holds if adopters move decisions into pipelines or central systems for their low-variety tasks and toward front-line staff for their high-variety tasks; it fails if adopters shift uniformly toward centralization or decentralization regardless of task variety, or if no such survey is published.

Forecast 7.4 (Hierarchy thins into verification). By the end of 2030, in agent-heavy firms, management layers shrink while review and audit roles grow as a share of headcount.

Horizon 2030-12-31

Probability 45%

Check Read the firm-level studies published through 2030 that use payroll, job-posting or professional-network data to compare occupational headcounts in firms with heavy agent use against comparable firms. The forecast holds if the heavy users show both a lower share of managers or fewer management layers and a higher share of review, quality-assurance, audit and compliance roles; it fails otherwise.

References

1. Caroli & Van Reenen 2001. <https://researchonline.lse.ac.uk/id/eprint/5/>
2. ADP 6-0, 2019. https://irp.fas.org/doddir/army/adp6_0.pdf
3. Beer 1992. https://metaphorum.org/wp-content/uploads/2020/12/world_in_tormentMD.pdf
4. Bloom, Sadun & Van Reenen 2012. <https://ideas.repec.org/a/oup/qjecon/v127y2012i4p1663-1705.html>
5. Dessein 2002. <https://ideas.repec.org/a/oup/restud/v69y2002i4p811-838.html>
6. Wiener 1948. <https://archive.org/details/cybernetics-norbert-wiener>
7. Bloom et al. 2013. <https://ideas.repec.org/a/oup/qjecon/v128y2013i1p1-51.html>
8. Shaw 1964. <https://epdf.mx/advances-in-experimental-social-psychology-volume-1.html>
9. Shah 2017. <https://medium.com/thinkgrowth/what-elon-musk-taught-me-about-growing-a-business-c2c173f5bff3>
10. Tsebelis 2011. https://sites.lsa.umich.edu/tsebelis/wp-content/uploads/sites/246/2020/12/Tsebelis2011_Chapter_VetoPlayerTheoryAndPolicyChange.pdf
11. Fukuyama 2014. <https://www.foreignaffairs.com/united-states/america-decay>
12. Yu et al. 2020. <https://papers.neurips.cc/paper/2020/file/3fe78a8acf5fda99de95303940a2420c-Paper.pdf>
13. Yadav et al. 2023. https://proceedings.neurips.cc/paper_files/paper/2023/file/1644c9af28ab7916874f6fd6228a9bcf-Paper-Conference.pdf
14. Sam McCandlish, Jared Kaplan, Dario Amodei, OpenAI Dota Team (2018). An Empirical Model of Large-Batch Training. *arXiv*. <https://arxiv.org/abs/1812.06162>
15. SEP, social choice. <https://plato.stanford.edu/entries/social-choice/>
16. A Thousand Plateaus. https://web.english.upenn.edu/~cavitch/pdf-library/Deleuze_and_Guattari_A_Thousand_Plateaus.pdf
17. Duncan Black (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*. <https://www.journals.uchicago.edu/doi/abs/10.1086/256633>
18. Popper 1945, ch. 7. <http://www.the-rathouse.com/OpenSocietyOnline/Chapter-7-Leadership.html>
19. Sen 1999. <https://web.archive.org/web/20250109044945/https://www.journalofdemocracy.org/articles/democracy-as-a-universal-value/>
20. SEP, jury theorems. <https://plato.stanford.edu/entries/jury-theorems/>

21. Lu Hong, Scott E. Page (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*. <https://www.pnas.org/doi/10.1073/pnas.0403723101>
22. Thompson 2014. <https://www.ams.org/notices/201409/moti-p1024.pdf>
23. Sah & Stiglitz 1986. <https://ideas.repec.org/a/aea/aecrev/v76y1986i4p716-27.html>
24. Estlund 2003. <https://philarchive.org/rec/ESTWNE>
25. Hayek, "The Use of Knowledge in Society". <https://www.econlib.org/library/Essays/hykKnw.html>
26. Falk et al. 2014. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3969807/>
27. Avshalom Caspi, Renate Houts, Daniel W. Belsky, HonaLee Harrington, Sean O. Hogan et al. (2016). Childhood forecasting of a small segment of the population with large economic burden. *Nature Human Behaviour*. <https://www.nature.com/articles/s41562-016-0005>
28. Sariaslan et al. 2014. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4180846/>
29. firing squad synchronization problem. https://en.wikipedia.org/wiki/Firing_squad_synchronization_problem
30. Simon 1962. <https://faculty.sites.iastate.edu/tesfatsi/archive/tesfatsi/ArchitectureOfComplexity.HSimon1962.pdf>
31. NVIDIA. <https://docs.nvidia.com/cuda/cuda-programming-guide/03-advanced/advanced-kernel-programming.html>
32. Randolph Menzel, Martín Giurfa (2001). Cognitive architecture of a mini-brain: the honeybee. *Trends in Cognitive Sciences*. <https://pubmed.ncbi.nlm.nih.gov/11166636/>
33. Seeley, Visscher & Passino 2006. <https://bees.ucr.edu/media/156/download>
34. Thomas Dyer Seeley, P. Kirk Visscher (2004). Quorum sensing during nest-site selection by honeybee swarms. *Behavioral Ecology and Sociobiology*. <https://doi.org/10.1007/s00265-004-0814-5>
35. Gordon 2007. http://web.stanford.edu/~dmgordon/old2/Gordon2007_Nature_Essay.pdf
36. Dhananjay K. Gode, Shyam Sunder (1993). Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *Journal of Political Economy*. <https://www.journals.uchicago.edu/doi/10.1086/261868>
37. Coase 1937. <https://www.rochelleterman.com/ir/sites/default/files/Coase%201937.pdf>
38. Simon 1991. <https://gwern.net/doc/economics/1991-simon.pdf>
39. Peyman Shahidi, Gili Rusak, Benjamin Manning, Andrey Fradkin, John Horton (2025). The Coasean Singularity? Demand, Supply, and Market Design with AI Agents. *National Bureau of Economic Research*. <https://www.nber.org/papers/w34468>
40. as quoted by Land. http://www.ccrn.net/swarm1/1_melt.htm
41. Land 1993. https://xenopraxis.net/readings/land_machinicsdesire.pdf
42. Mackay & Avanesian. <https://www.urbanomic.com/wp-content/uploads/2015/03/Accelerate-Introduction.pdf>
43. Williams & Srnicek 2013. <https://criticallegalthinking.com/2013/05/14/accelerate-manifesto-for-an-accelerationist-politics/>
44. e/acc 2022. <https://beff.substack.com/p/notes-on-eacc-principles-and-tenets>
45. Andreessen 2023. <https://a16z.com/the-techno-optimist-manifesto/>
46. Garicano 2000. <https://www.edegan.com/pdfs/Garicano%20%282000%29%20-%20Hierarchies%20and%20the%20Organization%20of%20Knowledge%20in%20Production.pdf>
47. Nicholas Bloom, Luis Garicano, Raffaella Sadun, John Michael Van Reenen (2009). The Distinct Effects of Information Technology and Communication Technology on Firm Organization. *Management Science*. <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2014.2013>
48. Enrique Ide, Eduard Talamàs (2025). Artificial Intelligence in the Knowledge Economy. *Journal of Political Economy*. <http://www.journals.uchicago.edu/doi/10.1086/737233>
49. Heylighen & Joslyn 2001. <http://pespmc1.vub.ac.be/Papers/Cybernetics-EPST.pdf>
50. Yubin Kim, Ken Gu, Chanwoo Park, Chunjong Park, Samuel Schmidgall et al. (2025). Towards a Science of Scaling Agent Systems. *arXiv*. <https://arxiv.org/abs/2512.08296>
51. MIT Media Lab. <https://www.media.mit.edu/projects/towards-a-science-of-scaling-agent-systems-when-and-why-agent-systems-work/overview/>
52. Anthropic 2025. <https://www.anthropic.com/engineering/multi-agent-research-system>
53. Cognition 2025. <https://cognition.ai/blog/dont-build-multi-agents>