

The Asymmetry Engine, Chapter 6: The energetics of attention

Martin Roark

Chapter 6 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#attention>. License: [CC BY 4.0](#).

6:1 People ration attention. The brain's 20 watts barely change with hard thinking, so avoiding thought saves almost no energy; what deliberation costs is felt, and a chooser keeps a default whenever the bits needed to beat it cost more than they are expected to gain. A nudge attenuates a decision's variety and passes its absorption to the architect, whose objective decides whether the attenuator serves the chooser or captures them. Empires and churches regulated what people call peace in the same way, and every asymmetry of the book is a budget of variety.

6.1 The mammal's budget

6:2 Thinking hard barely moves the brain's energy meter. The brain is about 2% of body weight yet "accounts for about 20% of the oxygen and, hence, calories consumed by the body," and "this high rate of metabolism is remarkably constant despite widely varying mental and motoric activity" ([Raichle and Gusnard 2002](#)). That is a budget of about 20 W ([Balasubramanian 2021](#)). Task-evoked increases are "often <5%" ([Raichle 2009](#)): about a watt, under one kilocalorie per hour of hard thought, and cortical computation itself takes only about 0.1 W ([Levy and Calvert 2021](#)). Avoiding thought saves almost nothing.

6:3 Energy shaped the brain's architecture instead. Spikes are so costly that Lennie argues energy limits, "possibly to fewer than 1%, the number of neurons that can be substantially active concurrently," which "explains ... the need for mechanisms of selective attention" ([Lennie 2003](#)). Attention is the brain's answer to its own power bill. The body paid too: by the expensive-tissue hypothesis, our large brain was financed by a smaller gut, made possible by "high-quality, easy-to-digest food," above all meat ([Aiello and Wheeler 1995](#)).

6:4 Behavior economizes attention without letting up. People pick the less demanding of two courses even when unaware of the difference ([Kool et al. 2010](#)), and 170 studies across 29 countries lead their authors to conclude that "mental effort is inherently aversive" ([David, Vassena and Bijleveld 2024](#)). The currency is not glucose: ego depletion, the account in which effort draws down a fuel, failed a 23-lab preregistered replication with $d = 0.04$ ([Hagger et al. 2016](#)). Felt effort reads better as an opportunity-cost signal over scarce executive capacity ([Kurzban et al. 2013](#)), while energy is saved where it is literal: walkers re-optimize their gait within minutes for savings under 5% ([Selinger et al. 2015](#)). What people ration in thought is attention, priced by what else it could do.

6.2 Attention, measured

6:5 What has shortened is the time people spend on one screen before switching. Gloria Mark's in-situ studies of knowledge workers found about two and a half minutes in 2004, 75 seconds around 2012 and a mean of 47

seconds, median 40, in recent years ([Mark, APA interview](#); [Mark et al. 2016](#)), with shorter focus going with lower productivity. Most of the decline came before short video spread, and Mark now says "we seem to have reached a steady state." The measure is switching behavior from small samples, and it says nothing about capacity; collective attention turns over faster too ([Lorenz-Spreen et al. 2019](#)).

6:6 The case against short video is correlational. A meta-analysis of 71 studies and 98,299 participants links short-form-video use to poorer attention ($r = -.38$), and its authors warn that the results "should be interpreted cautiously given the predominance of cross-sectional designs"; a longitudinal experiment they cite "showed no significant change following SFV use" ([Nguyen et al. 2025](#)). The "8-second attention span, shorter than a goldfish's" is a myth, a 2015 Microsoft Canada marketing figure credited to "Statistic Brain" for which the BBC found no evidence ([BBC](#)). What holds is that digital workers switch focus about three times as often as in 2004.

6:7 Herbert Simon stated the economics in 1971: "a wealth of information creates a poverty of attention" ([Simon 1971](#)). Foraging theory reads short dwell times without pathology. Charnov's marginal value theorem says a forager should leave a patch when its marginal yield falls to the habitat's average rate, so the optimal stay shortens as travel between patches gets cheaper ([Charnov 1976](#)), and people foraging for information follow the same rule ([Pirolli and Card 1999](#)). A short-video feed puts the next patch one swipe away, each decision worth about one bit, so short dwell is what an optimal forager would choose there. The welfare question is whose objective sets the patches (Proposition 6.2).

6:8

Forecast 6.1 (Dwell time follows switching cost). Through 2030, median dwell time per item tracks the cost of reaching the next item, whatever the content's length: it falls as switching gets cheaper, and an added delay or confirmation lengthens it.

Horizon 2030-12-31

Probability 70%

Check Read the experiments published from 2027 through 2030 that vary only the cost of switching, such as a delay or a confirmation step, and measure dwell time per item. The forecast holds if most of them find that the added cost lengthens median dwell time; it fails if most find no effect, or if none is published.

6.3 A budget for deliberation

6:9

Definition 6.1 (Attention budget). A chooser faces decisions $j = 1, \dots, J$ in a period, each with a default. Finding decision j 's best option means resolving H_j bits of uncertainty at a utility cost c_{bit} per bit, for an expected gain $G_j^{\text{att}} \geq 0$ over the default. The chooser's **attention budget** is the number of bits H_{att} it can resolve in a period, the channel capacity of a human regulator.

6:10

Proposition 6.1 (When defaults win). (i) Alone, a decision keeps its default iff $G_j^{\text{att}} \leq c_{\text{bit}} H_j$. (ii) Under the budget, the chooser deliberates exactly on the decisions with $G_j^{\text{att}} / H_j > c_{\text{bit}} + p_{\text{att}}^{\text{sh}}$, where the shadow price of attention $p_{\text{att}}^{\text{sh}} \geq 0$ is the lowest at which those decisions fit in H_{att} ; it rises as more decisions compete, and so does the share settled by default. (iii) If deliberation can be partial, at a cost c_{nat} per nat of information between the state and the choice, option i is chosen with probability

$$\Pr[i] = \frac{q_i e^{u_i / c_{\text{nat}}}}{\sum_k q_k e^{u_k / c_{\text{nat}}}}$$

where u_i is its payoff and q_i its prior probability of being chosen, so choices collapse onto the prior as attention gets dear ([Matějka and McKay 2015](#); locals q_i, u_i, c_{nat}).

▼ *Proof*

Part (ii) is a fractional knapsack whose Karush–Kuhn–Tucker conditions rank decisions by gain per bit and admit them until the budget binds; adding decisions above the threshold can only raise the multiplier. Part (iii) is Matějka and McKay's theorem for entropy-costly attention, building on [Sims 2003](#). ■

6:11 The proposition is Simon's poverty of attention as an inequality: an information-rich world raises the shadow price of a bit of deliberation, and every decision whose gain per bit falls below it goes to whoever set the default. Defaults are rational inattention.

6:12 Defaults are the strongest nudges on record. Automatic enrollment raised 401(k) participation from 37% to 86% ([Madrian and Shea 2001](#)). Effective organ-donor consent runs 4.25–27.5% in opt-in countries and 85.9–99.98% in opt-out ones, though actual donations rose only 16.3% ([Johnson and Goldstein 2003](#)): defaults move stated choices more than outcomes. Across 58 studies defaults average $d = 0.68$ ([Jachimowicz et al. 2019](#)), while nudges as a class shrink to $d \approx 0.04$ after correction for publication bias ([Maier et al. 2022](#)). Defaults also carry endorsement, so a sticky default is not proof of inattention alone. The speed default of \$5.6 is a special case.

6:13

Definition 6.2 (Nudge). Let O^* be the chooser's best option and o_{def} a default kept with probability p_{keep} . The architecture leaves the chooser to resolve, on average,

$$(1 - p_{\text{keep}}) H(O^* \mid O^* \neq o_{\text{def}}) \text{ bits}, \quad \Delta H = H(O^*) - (1 - p_{\text{keep}}) H(O^* \mid O^* \neq o_{\text{def}})$$

A **nudge** is a default that removes these ΔH bits, the variety it attenuates: Beer's attenuation applied to a person. It moves their absorption to the architect, who must carry the model of the chooser.

6:14 By requisite variety, an outcome's variety cannot fall below what the regulator leaves unabsorbed, so a default leaves a decision's variety intact and the architect computes what the chooser no longer does. Thaler and Sunstein require that a nudge be "easy and cheap to avoid" ([Thaler and Sunstein](#)), the rule that keeps the escape channel open.

6:15

Proposition 6.2 (Whose purpose). Let an architect set personalized defaults from its model $\hat{\vartheta}$ of the chooser's values ϑ , to maximize its own objective. If that objective is the chooser's utility, a better model raises the chooser's welfare on both counts: more deliberation absorbed, and defaults nearer to the best option. Otherwise a better model still absorbs more deliberation, and it also lets the architect pick, among the defaults the chooser will not escape, those that serve the architect.

▼ *Proof*

By Proposition 6.1 (i), the chooser escapes a default only when the expected loss exceeds $c_{\text{bit}}H_j$; below that bound the architect's objective alone sets it, and a sharper $\hat{\vartheta}$ widens the set of defaults that trigger no escape. ■

6:16 The better the model, the more the regulated system serves the regulator's purpose; in Stafford Beer's words, "the purpose of a system is what it does" ([Beer 2002](#)). A feed optimized for time spent saves its users deliberation and spends their time, the case Thaler calls sludge ([Thaler 2018](#)). Engagement optimization is misalignment at the scale of one person (§13.8).

6:17

Forecast 6.2 (Agents as personal choice architects). By the end of 2030, a personal AI agent that sets defaults on its users' behalf (subscriptions, purchases, settings) reports more than 100 million monthly users, and field data show its personalized defaults overridden less often than one-size defaults in the same domain.

Horizon 2030-12-31

Probability 25%

Check Read company disclosures for user counts and published field studies, by the company or by researchers, for override rates. The forecast holds if both conditions are met by the horizon; it fails otherwise.

6.4 Peace and the helmsman

6:18 People say they want to save time and keep their peace, and the two demands are one: fewer corrections per unit of life. Peace on this reading is the absence of error, with essential variables held inside their bounds at low regulatory cost, which is homeostasis as a person lives it (§6.5).

6:19 W. Ross Ashby built peace as a machine. His homeostat, finished on 16 March 1948, coupled four units whose essential variables had to stay in bounds; when one strayed, a stepping switch gave its unit new random parameters, searching 390,625 combinations in all, until the whole held still ([Cariani 2009](#); [Ashby](#)). Ashby called the property ultrastability: one loop regulates, and a second resets the first when regulation fails. Struggle is the second loop searching, and peace is its rest.

6:20 Empires regulated peace on their subjects' behalf. Juvenal's tenth satire says the Roman people, who once handed out *imperium*, *fascies*, *legiones*, *omnia*, now anxiously long for two things only, *panem et circenses*, bread and circuses ([Juvenal, Satire X](#)). The bread was an institution: in 2 BC those "in receipt of public grain ... comprised a few more than 200,000 persons" ([Res Gestae](#)). Grain and games held the city's essential variable, public order, and attenuated political variety to two channels. Juvenal's complaint is that the regulator replaced the agency it served.

6:21 The Church was the longer-lived regulator. Before print, Benedict Anderson notes, Rome won every war against heresy "because it always had better internal lines of communication than its challengers" ([Anderson](#)). Michel Foucault traced modern government to the Christian pastorate, an art of "taking charge of men collectively and individually throughout their life," where the Greek pilot "does not govern the sailors; he governs the ship" ([Foucault](#)). Its marriage rules reshaped European kinship ([Schulz et al. 2019](#)), and in 2026 it addressed the machine, objecting to an AI whose morality is set by a few (§13.7).

6:22 Such institutions are what Stafford Beer called attenuators. Out of everything people might do, they select the few things they should do and spare each person the cost of choosing; the spared cost is the peace, and a liturgical calendar that synchronizes millions without messages is regulation at scale. Describing what an institution regulates says nothing about the truth of its faith, and the analogy mixes levels and can hide coercion, belief and contingency; religions as recipes for alignment are §13.7's subject.

6:23 The vocabulary carries the same idea. Greek *kybernētēs* is a steersman, from *kybernan*, "to steer or pilot a ship" ([Etymonline](#)), and Latin borrowed it as *gubernare*, "originally 'to steer, to pilot,'" whence govern, governor and government ([Etymonline](#)). Ampère in 1834 named *cybernétique* the "étude des moyens de gouvernement" ([CNRTL](#)). Maxwell's "On Governors" (1868) analyzed the feedback of steam engines, and in 1948 Wiener, crediting that paper, chose "the name Cybernetics, which we form from the Greek κυβερνήτης or steersman" ([Wiener, via Language Log](#)).

6:24 The chain closes in software. Kubernetes, the container system Google open-sourced in June 2014 ([Google](#)), takes its name from the Greek "meaning helmsman or pilot" ([Kubernetes](#)), and its 2019 documentation called the word "the root of governor and cybernetic" and its mechanism control processes that "continuously drive the current state towards the provided desired state" ([Kubernetes, 2019](#)): a negative-feedback loop over a fleet. Its co-founder Tim Hockin says there is "no magic behind the name" ([The Changelog](#)), so the lineage is a rhyme. It still says what steering is: sensing error and correcting it.

6:25

Forecast 6.3 (Bread and circuses, computed). By the end of 2030, recipients of large cash-transfer programs vote, volunteer and join civic groups less than comparable controls: where AI displaces work, transfers plus abundant AI entertainment settle into a low-participation homeostasis.

Horizon 2030-12-31

Probability 20%

Check Read the evaluations published from 2027 through 2030 of cash-transfer programs with at least 1,000 recipients that compare turnout, volunteering or civic membership with controls. The forecast holds if most of them find recipients lower on these measures; it fails if most find participation unchanged or higher, or if none is published.

6.5 It is all cybernetics

6:26 Read as method, the claim that it is all cybernetics commits the book to a grammar of feedback, regulation and variety. Rosenblueth, Wiener and Bigelow defined its negative form as one in which "the signals from the goal are used to restrict outputs which would otherwise go beyond the goal," and its positive form as one that "adds to

the input signals, it does not correct them"; "all purposeful behavior may be considered to require negative feedback" (Rosenblueth, Wiener and Bigelow 1943).

6:27 Negative feedback that holds a variable near a set point is **homeostasis**, Walter Cannon's word for the processes "which maintain most of the steady states in the organism"; he insisted that it "does not imply something set and immobile, a stagnation. It means a condition—a condition which may vary, but which is relatively constant" (Cannon 1932). Positive feedback amplifies "an insignificant or accidental initial kick" (Maruyama 1963). Variety counts the states a regulator must tell apart, and requisite variety is its conservation law.

6:28 In that grammar every asymmetry of the book is a budget of variety. A verifier must have the variety of the generator's errors (§4.6). A model must have the variety of its task, and more is not worth paying for (Definition 3.1). A brake must act faster than the loop it governs can grow (Proposition 4.3), which makes latency a loop delay. Each object of the book then has a cybernetic role:

6:29

Object	Cybernetic role	Home
Verification	the sensor: how cheaply error can be measured	Chapter 2
Compute and gigawatts	the regulator's power	Chapter 3, Chapter 10
Variety	the regulator's repertoire	Chapter 4
Latency	loop delay, which bounds steering	Chapter 5
Attention	the channel capacity of the human regulator	§6.3
Topology	where the regulators sit	Chapter 7
Recursion	a positive loop on intelligence itself	Chapter 11
Alignment	the setpoint: whose essential variables are held	Chapter 13
Inertia	the time constant of the plant	Chapter 15

6:30 Two warnings keep the table honest. Wiener denied that markets regulate themselves (4:48). And in the research loop the negative feedback that bounds the runaway is physical: compute is the homeostat (Proposition 11.2).

6:31 Every loop in the table ends at a human, and AI amplifies everything except the human channel. Ashby saw the opening: intellectual power can be amplified (4:12). When the human channel cannot absorb the variety, regulation must be delegated, and Proposition 6.2 says delegation is safe only when the delegate's objective is ours. Wiener saw the cost: "The future offers very little hope for those who expect that our new mechanical slaves will offer us a world in which we may rest from thinking. Help us they may, but at the cost of supreme demands upon our honesty and our intelligence" (God & Golem, Inc.). Peace is a well-regulated loop, and the book's questions are who holds the helm, what setpoint it steers toward and how fast the loop can turn.

6:32 The attention economy is the grammar's simplest positive loop: engagement yields data, data a better model of the user, and the model a more engaging feed, checked by fatigue, time and regulation; Figure 11.1 draws the same loop for research. Closing the labor gaps of the coming decade frees hours, and that loop is waiting for them.

6:33

Forecast 6.4 (Freed hours are recaptured). Through 2030, people whose paid hours fall spend more of the freed hours on screen media (television, streaming, games and leisure computer use) than on education, care, civic life and in-person company combined, so closing labor gaps regenerates attention gaps.

Horizon 2030-12-31

Probability 60%

Check Compare people whose paid hours fell with similar people in the American Time Use Survey, using the latest published analysis or the microdata available at the horizon. The forecast holds if the screen-media categories absorb more of the freed hours than education, caring for others, volunteering and in-person socializing combined; it fails otherwise.

References

1. Marcus E. Raichle, Debra A. Gusnard (2002). Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences*. <https://www.pnas.org/doi/10.1073/pnas.172399499>
2. Vijay Balasubramanian (2021). Brain power. *Proceedings of the National Academy of Sciences*. <https://pubmed.ncbi.nlm.nih.gov/34341108/>
3. Raichle 2009. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6665302/>
4. William B. Levy, Victoria G. Calvert (2021). Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proceedings of the National Academy of Sciences*. <https://www.pnas.org/doi/10.1073/pnas.2008173118>
5. Lennie 2003. <https://www.cell.com/current-biology/fulltext/S0960-9822%2803%2900135-0>
6. Aiello and Wheeler 1995. <https://gwern.net/doc/algernon/1995-aiello.pdf>
7. Wouter Kool, Joseph T. McGuire, Zev B. Rosen, Matthew Botvinick (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology General*. <https://pubmed.ncbi.nlm.nih.gov/20853993/>
8. Louise David, Eliana Vassena, Erik Bijleveld (2024). The unpleasantness of thinking: A meta-analytic review of the association between mental effort and negative affect. *Psychological Bulletin*. <https://pubmed.ncbi.nlm.nih.gov/39101924/>
9. Martin S. Hagger, Nikos L. D. Chatzisarantis, H. Alberts, Calvin Octavianus Anggono, Cédric Batailler et al. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science*. <https://journals.sagepub.com/doi/10.1177/1745691616652873>
10. Kurzban et al. 2013. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3856320/>
11. Selinger et al. 2015. <https://www.cell.com/current-biology/fulltext/S0960-9822%2815%2900958-6>
12. Mark, APA interview. <https://www.apa.org/news/podcasts/speaking-of-psychology/attention-spans>
13. Gloria Mark, Shamsi T. Iqbal, Mary P. Czerwinski, Paul Johns, Akane Sano (2016). Neurotics Can't Focus. *CHI Conference on Human Factors in Computing Systems*. <https://dl.acm.org/doi/10.1145/2858036.2858202>
14. Lorenz-Spreen et al. 2019. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6465266/>
15. Lan Nguyen, Jared Walters, Siddharth Paul, Shay Monreal Ijurco, Georgia E Rainey et al. (2025). Feeds, feelings, and focus: A systematic review and meta-analysis examining the cognitive and mental health correlates of short-form video use. *Psychological Bulletin*. <https://pubmed.ncbi.nlm.nih.gov/41231585/>
16. BBC. <https://www.bbc.com/news/health-38896790>
17. Simon 1971. <https://gwern.net/doc/design/1971-simon.pdf>
18. Eric L. Charnov (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*. <https://doi.org/10.1016/0040-5809%2876%2990040-X>
19. Peter L. T. Pirolli, Stuart Card (1999). Information foraging. *Psychological Review*. <https://doi.org/10.1037/0033-295X.106.4.643>
20. Matějka and McKay 2015. <https://ideas.repec.org/a/aea/aecrev/v105y2015i1p272-98.html>

21. Christopher A. Sims (2003). Implications of rational inattention. *Journal of Monetary Economics*. <https://doi.org/10.1016/S0304-3932%2803%2900029-1>
22. Madrian and Shea 2001. https://www.nber.org/system/files/working_papers/w7682/w7682.pdf
23. Johnson and Goldstein 2003. <https://www.dangoldstein.com/papers/DefaultsScience.pdf>
24. Jachimowicz et al. 2019. <https://www.cambridge.org/core/journals/behavioural-public-policy/article/when-and-why-defaults-influence-decisions-a-metaanalysis-of-default-effects/67AF6972CFB52698A60B6BD94B70C2C0>
25. Maier et al. 2022. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9351501/>
26. Thaler and Sunstein. <https://archive.blogs.harvard.edu/nudge/what-is-a-nudge/>
27. Stafford Beer (2002). What is cybernetics? *Kybernetes*. <https://doi.org/10.1108/03684920210417283>
28. Richard H. Thaler (2018). Nudge, not sludge. *Science*. <https://www.science.org/doi/10.1126/science.aau9241>
29. Cariani 2009. <http://petercariani.com/uploads/1/2/3/9/123936871/cariani-2009-ashbyhomeostat.pdf>
30. Ashby. https://archive.org/stream/designforbrainor00ashb/designforbrainor00ashb_djvu.txt
31. Juvenal, Satire X. <https://www.thelatinlibrary.com/juvenal/10.shtml>
32. Res Gestae. https://droitromain.univ-grenoble-alpes.fr/Anglica/resgest_engl.htm
33. Anderson. <https://www.thetcdkarchive.com/library/benedict-anderson-imagined-communities>
34. Foucault. [https://www.asc.uw.edu.pl/wp-content/uploads/2022/03/Foucault-Security_Territory_Population_Lectures_at_the_Co... - Pg_147-231.pdf](https://www.asc.uw.edu.pl/wp-content/uploads/2022/03/Foucault-Security_Territory_Population_Lectures_at_the_Co...-Pg_147-231.pdf)
35. Jonathan Schulz, Duman Bahrami-Rad, Jonathan P. Beauchamp, Joseph Henrich (2019). The Church, intensive kinship, and global psychological variation. *Science*. <https://doi.org/10.1126/science.aau5141>
36. Etymonline. <https://www.etymonline.com/word/cybernetics>
37. Etymonline. <https://www.etymonline.com/word/govern>
38. CNRTL. <https://www.cnrtl.fr/definition/cybern%C3%A9tique/nom>
39. Wiener, via Language Log. <https://languagelog.ldc.upenn.edu/nll/?p=27973>
40. Google. <https://opensource.googleblog.com/2014/06/an-update-on-container-support-on.html>
41. Kubernetes. <https://kubernetes.io/docs/concepts/overview/>
42. Kubernetes, 2019. <https://web.archive.org/web/20190530141916/https://kubernetes.io/docs/concepts/overview/what-is-kubernetes/>
43. The Changelog. <https://github.com/thechangelog/transcripts/blob/master/podcast/the-changelog-250.md>
44. Rosenblueth, Wiener and Bigelow 1943. https://archive.org/download/wiener-1938/RosenbluethEtA11943_0_djvu.txt
45. Cannon 1932. <https://raw.githubusercontent.com/peatysharing/bibliography/main/Walter%20Cannon/1932%20-%20Walter%20Cannon%20-%20The%20Wisdom%20Of%20The%20Body.pdf>
46. Maruyama 1963. <http://pespmc1.vub.ac.be/books/Maruyama-SecondCybernetics.pdf>
47. God & Golem, Inc.. <https://mitpress.mit.edu/9780262730112/god-and-golem-inc/>
48. American Time Use Survey. <https://www.bls.gov/tus/>