Introduction


2. As Malcolm Gladwell puts it, “I think that the task of figuring out how to combine the best of conscious deliberation and instinctive judgment is one of the great challenges of our time.” Gladwell, *Blink: The Power of Thinking Without Thinking* (New York: Little, Brown, 2005), 269.


Chapter 1


5. Ibid., 648–52.


31. University of Chicago researchers Lubos Pastor and Pietro Veronesi argue that “learning is facilitated by the existence of vast quantities of financial data, but it is also hampered by the large amount of randomness pervading financial markets.” They describe how valuation models that incorporate Bayesian updating and uncertainty can explain seemingly irrational phenomena, like the high valuations of technology stocks (a rational function of uncertainty, in their view) or the subsequent crash (when investors learn of the true profitability of the firms). They point out that the stocks of younger firms have higher volatility and prices than those of older firms; a phenomenon that they conjecture reflects greater initial uncertainty about the business models, which diminishes as investors learn about the firms over time. Learning may explain why individual investors trade excessively (it helps them learn about their stock-picking skills) or why venture capital firms’ performance is persistent (they may learn valuable information from their investments). See Pastor and Veronesi, “Learning in Financial Markets,” *Annual Review of Financial Economics* (2009), and Andrew W. Lo, “Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis,” *The Journal of Investment Consulting* 7:2 (2005).
32. This terminology comes from Philip Tetlock, who studied the forecasting accuracy of a panel of political experts. Tetlock uses the term *calibration* to refer to the degree to which subjective probabilities (forecasts) were aligned with objective frequencies. *Discrimination* measures the ability of forecasters to do better than predicting base rates by assigning high probabilities to events that actually occur and near-zero probabilities to events that do not. See Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005), 47–49.

33. J. Edward Russo and Paul J. H. Schoemaker, “Managing Overconfidence,” *Sloan Management Review* 33:2 (1992), 10–11. The authors argue that in order to learn from experience (and thus improve calibration and forecasting accuracy), forecasters must get feedback on and be held accountable for their predictions.

34. Christopher Cherniak describes the “minimal rationality” that characterizes human cognition as evolving in a social context, where people vet what they know not only against their own beliefs but also against the shared knowledge of the community. “Doctrines like ‘Ignorance of the law is no excuse’ indicate that the agent is held to a standard of care that includes responsibility for collecting information by consulting appropriate experts in the community” (Cherniak, *Minimal Rationality* [Cambridge: MIT Press, 1986], 115).

35. Risk premiums cannot be directly observed. They can be estimated as the difference between equity returns and U.S. Treasury bond returns, but equity risk premiums have fluctuated across different periods, and there has been enormous variability in the risk premium from year to year. See Bradford Cornell, *The Equity Risk Premium: The Long-Run Future of the Stock Market* (New York: Wiley, 1999).

36. The concept of the “null hypothesis” is commonly used in statistical hypothesis testing. Bayesian statisticians prefer to think of new evidence modifying the “prior distribution” (or theory) before conducting the experiment. For the Bayesian point of view, see Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach*, 3rd ed. (Chicago: Open Court, 2006).


Chapter 2

1. Nassim Nicholas Taleb draws a distinction between totally unpredictable Black Swans and what he calls “near-black” or “gray” swans that may be predictable. He also acknowledges that “severe Black Swans” may occur when surprises could have been predicted but people misunderstood the sources of uncertainty or “lack imagination” in their forecasts. Taleb, The Black Swan: The Impact of the Highly Improbable (New York: Random House, 2007), 36–37.


3. John C. Hull, Options, Futures, and Other Derivatives, 5th ed. (Upper Saddle River, NJ: Pearson/Prentice Hall, 2003), 392–94. In this example, we assumed no rate of growth in the stock price.


Chapter 3

4. Some researchers have argued that overconfident traders are more aggressive than rational traders at exploiting mispricings caused by liquidity or noise traders; thus they expect to find in long-run equilibrium that overconfident traders make up a substantial fraction of the trading population. Other reasons cited in the academic literature for overconfident traders to persist include the tendency of rational traders to cut back on their own trades when they realize that overconfident traders are buying or selling aggressively. See Guo Ying Luo and David A. Hirshleifer, “On the Survival of Overconfident Traders in a Competitive Securities Market,” Journal of Financial Markets 4:1 (January 2001).
11. Ibid.
15. For example, during the Cold War, roughly 80% of information about the Soviet Union came from classified sources and only 20% from open sources. Today, 80% is open source. The data are taken from Mark M. Lowenthal, Intelligence: From Secrets to Policy, 3rd ed. (Washington, DC: CQ Press, 2006), 102. The importance of open-source information is one of the key themes of Gregory F. Treverton, Reshaping National Intelligence for an Age of Information (Cambridge, UK, and New York: Cambridge University Press, 2003). For example (p. 10), “The more-open world is blurring the distinction between collection and analysis. The best looker is not a spymaster, much less an impersonal satellite, but someone steeped in the substance at hand—in short, an analyst.” Also, see Robert M. Clark, Intelligence Analysis: A Target-Centric Approach, 2nd ed. (Washington, DC: CQ Press, 2007), 87, 124.


Chapter 4


2. This estimate is based on 10^{25} atoms in a kilogram of matter and the ability to store 1,024 bits of information in the magnetic interactions of the protons of a single molecule containing 19 hydrogen atoms, which has apparently been demonstrated by researchers at the University of Oklahoma. See Ray Kurzweil, The Singularity Is Near: When Humans Transcend Biology (London: Penguin Books, 2005), 131.


4. For example, in one study, researchers found a correlation between public opinion about the problem of illegal drugs and the number of media stories on the subject. See Thomas Kida, Don’t Believe Everything You Think: The 6 Basic Mistakes We Make in Thinking (Amherst, NY: Prometheus Books, 2006), 178.


6. The theoretical minimum cost of computing is quite low, especially if you could operate the computer at a temperature close to absolute zero and harness futuristic technologies like reversible and quantum computing. However, even with these technologies, there is in practice a limit to how powerful a computer could become, because transient fluctuations in energy at the subatomic level would introduce occasional mistakes into the calculations. Thus the device would need some kind of routine to check for and correct errors. The amount of energy dissipated in the processing of a single bit of information is kT, where k refers to Boltzman’s constant and T to the ambient temperature in degrees Kelvin. See Kurzweil, Singularity


9. The “halting problem” and intractability are some of the problems that lead philosophers to consider minimal rationality as a characteristic of human cognition, rather than the perfect rationality assumed in classic economics. See Christopher Cherniak, Minimal Rationality (Cambridge: MIT Press, 1986), 77–81.


13. Simon, Sciences of the Artificial, 51–58. Similarly, “the main reason to focus attention on a select set of target hypotheses is to economize the acquisition of new data . . . [and] to confine inferences to pertinent regions of the network,” making propagation of new information through the entire network unnecessary (thus economizing on computational time). See Judea Pearl, Probabilistic Reasoning in Intelligent
14. Neuroscientists view our ability to pay attention as a limited resource that
one allocates to an object or a location in order to enhance the speed or accuracy of
sensory perception. Experiments with monkeys and humans that involve the moni-
toring of a random selection of visual cues show that the subjects’ ability to system-
atically focus their attention at locations was consistent with the probability of the
cues appearing there. See Paul Glimcher, *Decisions, Uncertainty, and the Brain: The

15. Van Hecke cites one study where researchers observed that law students failed
to consider both sides of an issue before writing an essay that required them to do
so. In another, a researcher studying Fortune 100 employees identified failure to
stop and think as a greater barrier to intelligent behavior than either limited motiva-
tion or ability. See Madeleine L. Van Hecke, *Blind Spots: Why Smart People Do


1999), 175.

1997), 353.

19. If 50,000 chunks of knowledge define expert knowledge (of the sort acquired
over a decade’s study), then it would take about 16 dichotomous tests to identify the
relevant information. At about 10 milliseconds per test, the act of identification
would take less than 200 milliseconds. Apparently it takes in practice a few hun-
dred milliseconds to 1 or 2 seconds for memory to produce results. See Simon, *Sciences of the Artificial*, 88, 90.

decision making and a case study.

21. William Duggan, *Strategic Intuition: The Creative Spark in Human Achieve-

22. See Simon, *Sciences of the Artificial*, 193–200, for a discussion of the trial-
and-error approach to searching and the near decomposability of certain kinds of
problems.


25. Rob Johnston, *Analytic Culture in the U.S. Intelligence Community: An Eth-

26. David Hirshleifer and Ivo Welch, “An Economic Approach to the Psychology
of Change: Amnesia, Inertia, and Impulsiveness,” Yale Cowles Foundation Discus-
27. Because of the “combinatorial explosion” entailed in trying to search through all branches of a decision tree, searching must be “selective, that is, partial and risky” (Haugeland, Artificial Intelligence, 178).


Chapter 5


2. In a report downgrading CIT to neutral on the heels of the announcement, I wrote that the transaction rationale was unclear: “The real issue at CIT, in our view, is the company’s high cost of capital and low returns on equity. Until CIT improves its funding costs, we think it will have trouble competing with bank-owned finance companies, not to mention the likes of GE Capital, especially given the thin margins characteristic of vendor finance and other commercial finance businesses. We don’t see Tyco’s purchase of CIT as providing any benefit to CIT’s funding costs. In fact, CIT’s corporate debt spreads widened 2 bps yesterday on news of the deal, while TYC’s were unchanged. CIT’s corporate debt is ranked by the fixed-income rating agencies several notches better than Tyco’s (A1/A+ versus Baa1/A-). Moody’s has indicated it will review CIT’s ratings for possible downgrade, although Standard & Poor’s has reaffirmed its current ratings. CIT’s recent 5-year bond issuance (’06 maturity) is currently trading at 165 bps over Treasuries, 10 bps lower than Tyco’s 10-year debt (’11 maturity) at 175 bps.” See Kenneth A. Posner and Athina Meehan, “The CIT Group: Downgrading to Neutral,” Morgan Stanley Dean Witter Research, March 14, 2001.


7. For examples of studies on this topic, see Max Bazerman, Judgment in Managerial Decision Making, 6th ed. (Hoboken, NJ: Wiley, 2006), 35–36.
8. Tavris and Aronson detail many of these bad decisions in *Mistakes Were Made*.


11. The other explanation comes from prospect theory, according to which people are willing to take greater risk to avoid a loss than they are to sustain a gain. See Andrea Frazzini, “The Disposition Effect and Under-Reaction to News,” *The Journal of Finance* LXI.4 (August 2006): 2017–46.


16. Some practitioners have come to the same conclusion. For example, the CIA conducted case studies on how its analysts gauge diagnostic power when they update their subjective probability assessments. In a declassified study conducted in the early 1970s, a group of analysts was polled on their changing views of the odds of hostilities between the Soviet Union and China. The study revealed that the analysts’ judgments about probabilities were inconsistent with their assessment of the diagnostic power of new information—a troubling inconsistency. See Charles E. Fisk, “The Sino-Soviet Border Dispute: A Comparison of the Conventional and Bayesian Methods for Intelligence Warning,” originally published in *Studies in Intelligence*, 1972, reprinted in *Inside CIA’s Private World: Declassified Articles from the Agency’s Internal Journal 1955–1992*, ed. H. Bradford Westerfield (New Haven, CT: Yale University Press, 1995).


Chapter 6


5. Sallie Mae, Q3 2005 Earnings Call Transcript, CallStreet, October 20, 2005.


7. Ibid., 7, 8.


10. For those unfamiliar with the term, a *deferred tax asset* reflects the value of tax deductions that have run through the generally accepted accounting principles (GAAP) income statement but have not yet been taken in the tax returns. Freddie had established reserves against future credit losses in its financial statements under GAAP, even though it would not deduct those reserves for tax purposes until it wrote off the loans. If a company cannot demonstrate that it will have sufficient taxable income in future periods to use those tax deductions, then its auditors may force the
company to write off a portion of the deferred tax asset, a move that would impair its capital. For this reason, regulators limit the size of deferred tax assets that banks can count as capital. The regulatory guidelines for Freddie and Fannie Mae did not address deferred tax assets, but the companies were still vulnerable to their auditors’ determination.


12. Undoubtedly, part of the reason for the deferred tax write-down was the auditors’ concern that the government would have the company take actions to benefit public policy, such as reducing its guarantee fees, rather than maximize profits. Whether the company would have been forced to write off its deferred tax asset if it had remained private is an open question.


Chapter 7

1. In computational theory, *mapping* is another term for a mathematical function that produces an output for one or more inputs. See Michael Sipser, Theory of Computation (London: Springer, 2007), 15.


Chapter 8


Chapter 9


13. Graham’s image of Mr. Market is “probably the most brilliant metaphor ever created for explaining how stocks can become mispriced” (Jason Zweig, commentary to chapter 8 of Benjamin Dodd’s *The Intelligent Investor* [New York: Harper-Business, 2005], 213).


19. It was thought that MBNA’s largest investors were interested in monetizing their stakes, so perhaps management had to push for growth as aggressively as possible if it wanted a chance to maintain the company’s independence.


22. “The growing fashion in risk management, supported by the Basel Committee on Banking Supervision, is a move away from discretionary judgments about risk and a move to more quantitative and market-sensitive approaches” (Avinash Persaud, “Sending the Herd off the Cliff Edge: The Disturbing Interaction Between Herding and Market-Sensitive Risk Management Practices,” Bank for International Settlements Papers, No 2 [2000]).


25. For a quick summary, see Sam Peltzman, “Regulation and the Natural Progress of Opulence,” 2004 Distinguished Lecture, AEI-Brookings Joint Center for Regulatory Studies, American Enterprise Institute for Public Policy Research,

