BREAKOUT
GROWTH PARTNERS

Bounded Rationality in the Age of AI

- Strategic decisions are made by people working under limits, which Simon called bounded rationality: too little time, too much to hold in mind, and an early stop once an option is good enough.
- I group those limits into five recurring failures: directional bias, inconsistent noise, an inside view that ignores the base rate, cognitive load that truncates the problem, and organizational dynamics that scale individual limits up.
- These limits are persistent constraints on managerial judgment; training alone will not eliminate them, so sound process and organizational design remain essential.
- AI does not remove these limits; it moves where some of them sit. Under controlled conditions, AI can augment effective memory, standardize some judgments, widen the search for options, and prompt alternative perspectives.
- It leaves the judgment bounds, the risk appetite, the trade-offs, and the ownership of the call untouched, and it adds new bounds of its own: automation bias, anchoring to the model's framing, and a shared machine narrowing everyone's thinking.
- Used well, AI shifts the boundary of bounded rationality outward; used carelessly, it relocates the error. The discipline is knowing which limits a model lifts and which it leaves.

Even highly capable teams often fall into recognizable traps when making strategic decisions. They face time constraints, struggle to juggle numerous complex elements simultaneously, and operate within organizations where politics and incentives subtly restrict the range of options considered. Under this pressure, people tend to satisfice: they stop searching for alternatives as soon as one meets the minimum threshold, rather than continuing to look for the optimal choice.[1,2]

Herbert Simon described this phenomenon as bounded rationality. Preferences form as we go; the search for alternatives stops early, and reasoning runs up against the limits of time, attention, and computation.[1] Later, Kahneman and colleagues refined the understanding of judgment failures, distinguishing between two types: bias, the systematic error that pushes an entire group in the same wrong direction, and noise, the scatter among people who should agree but don't.[3]

The limits do not go away, so the response has to be built into the process rather than left to good intentions. Simon's proposed remedy was procedural rationality: when you cannot reason your way to the best answer directly, a sound method gets you close.[4] This paper sets out what the limits are and where they come from; the Decision Quality overview addresses the process that tackles them. I have kept the two separate on purpose, so this one can go deeper on the why.

Simon was an early pioneer of artificial intelligence and advocated for its role in decision-making, so AI's involvement in this domain is not new. What has changed is the nature of AI itself. While the symbolic systems of Simon's era expanded computational power and memory, the very limits he highlighted, today's generative AI goes further. It widens the range of perspectives considered and tackles unstructured problems that earlier systems could not address. Although AI does not remove human limitations, it can alter where some of them sit: it broadens the search for options, helps teams consider more aspects of a problem at once, improves access to base rates, and provides a shared basis for reasoning, which reduces the variability in individual judgments. However, AI also brings its own set of constraints, which I will discuss later. Before asking what AI might change, we must first understand the boundaries of human judgment.

Understanding bounded rationality

Simon's point was that making the optimal choice is impossible in practice. To truly optimize, you would need to consider every possible alternative, fully understand all potential outcomes, and have the time and computational resources to evaluate them, conditions that never exist in reality.[1] As a result, managers adopt a different approach: they set a target, search for options until one meets that standard, and then stop. Simon called this satisficing, and it is adaptive rather than a sign of laziness. Stopping at "good enough" is often the rational choice under real-world limitations.[1]

This perspective shifts our understanding of what counts as a good decision. Simon's later distinction is important: substantive rationality, choosing the absolute best action given perfect information, is unattainable, so the focus turns to procedural rationality: using a sound method to make decisions within existing constraints.[4] The rest of this section is about what those limits are and where they come from. I group them into five categories that tend to reinforce one another. The grouping is a working frame drawn from several literatures rather than Simon's original categories.

Five limits on management judgment

- **Bias sends judgment in a consistent direction** — These errors are systematic. A first number anchors the estimates that follow, and recent or vivid events feel likelier than they are.[5] Outcomes are judged against a reference point rather than on their own terms, and because the pain of a loss outweighs the pleasure of an equal gain, managers are more likely to take risks to avoid losses than to pursue matching gains.[6] A team in trouble doubles down; the same team sitting on a gain turns cautious.

- **Noise is the scatter that bias hides** — Where bias skews judgment in a single direction, noise is the unwarranted variability in decision-making: two equally qualified people arrive at different conclusions on the same issue, and even the same person may reach different judgments from one day to the next. Kahneman, Sibony, and Sunstein found this phenomenon in settings where consistency is taken for granted, yet the range of judgments was much broader than organizations expected.[3] Noise is often hidden because of averaging the group's judgments, but the variance is real, and it carries a cost.
- **The inside view crowds out the base rate** — When forecasting, managers tend to base their estimates on the unique details of the situation, such as the specific team, market, and plan involved. Lovallo and Kahneman explain how this "inside view" leads to over-optimism because it neglects the "outside view," which draws on actual outcomes from similar past initiatives.[7] Flyvbjerg observed the same effect in a study of hundreds of infrastructure projects: forecasts grounded in each project's particulars often proved too optimistic, while referencing a comparison class of similar projects provided a more accurate prediction.[8]
- **Cognitive load truncates the problem** — Strategic decisions involve more elements than working memory can hold. Miller's well-known estimate put the limit near seven items, though he meant it loosely; later work tightened it to about four.[9,10] Regardless, four to seven items is a fraction of what a strategic problem contains. Typically, the most abstract factors are lost first: vivid, recent details stay top of mind while slow-moving structural ones fade, so the team reasons over a simplified, skewed version of the problem.
- **Organizational dynamics scale the limits up** — Decisions are rarely made by a single individual. March and Simon characterized organizations as structures that constrain rationality by directing attention and filtering information before it even reaches the decision-maker.[11] Where a manager sits then shapes what they see: Ireland and colleagues discovered that executives at various levels interpret the same evidence differently, meaning that information entering the process is already colored by positional bias.[12] Layer on ambiguous decision rights, hierarchical status that silences junior voices, and incentives that dictate which options are safe to propose, and the group can become less rational than any one member. Bias and noise don't cancel out in a group; conformity can lead everyone to align behind the same mistake.[3]

These five factors describe the realities of management decision-making. This raises two key questions: what disciplines these limits, and what can shift them? Decision Quality addresses the first, while AI is increasingly shaping the second. The rest of this paper will explore how.

What AI moves, and what it leaves

The limits discussed earlier were once seen as unchangeable aspects of human judgment, and for decades, the main remedy was simply to improve the decision-making process. That is changing. My argument here is narrow and, I think, defensible: AI does not eliminate these limits. It adds processing capacity that the human decision-maker lacked, shifting where several of the limits sit. Some constraints that were once fixed are now more flexible, while the ones that matter most remain unchanged. At the same time, AI introduces its own set of limitations. The situation deserves careful examination.

Beginning with the limits that are easiest to shift. Noise is the clearest example: a model held at fixed settings does not drift the way a tired or distracted person does, which is the consistency Kahneman, Sibony, and Sunstein prescribe against scatter.[3] That steadiness is real but conditional, since sampling settings, prompt wording, and a version update between Friday and Monday all move the output. The inside view also relaxes, since AI can reduce the effort required to identify candidate reference classes and assemble outside evidence, provided those cases and data are independently verified. This breadth is exactly what AI excels at. Likewise, position-bound perception, the constraint Ireland described where a manager's view is shaped by their organizational seat, is weakened.[12] AI is not anchored to any one role. When prompted to analyze an issue from the perspectives of finance, sales, and the customer, it can reveal insights that a single viewpoint would miss. While these perspectives are valuable, they are not substitutes for evidence from the stakeholders themselves.

However, each of these benefits comes with a caveat. Just because a model is consistent doesn't mean it's accurate; low noise with high bias is a confident, repeatable error. A reference class a model constructs can be fabricated as easily as it is sourced, and both may appear convincing. Similarly, the perspectives it offers are inherently plausible, regardless of whether they represent any real stakeholder. AI provides real advantages, but they require careful scrutiny.

Cognitive load can be both alleviated and exacerbated. Strategic problems often arrive in a state of clutter, and teams spend much of their energy managing that chaos instead of understanding the core issue. AI models can help by synthesizing and structuring information, reducing wasted effort and freeing up mental resources for deeper reasoning. In Sweller's terms the clutter is extraneous load, capacity spent on the presentation of a problem rather than on the problem itself; his work concerns learning tasks, but the mechanism carries over.[13] That is a meaningful gain. However, the same solution can create a new problem: people may delegate too much of their thinking to the tool, and because the decision to offload rests on a self-assessment that is often wrong, a team leans on the model past the point where it should still be reasoning for itself.[14] Consequently, the mental capacity the model freed goes unused, and the reasoning to build a real grasp of the problem gets handed over too, and the understanding never forms. So the discipline is to use the model to organize the problem and still reason through it yourself.

Bias is another area where AI has a double edge. On the one hand, a model can act as a built-in devil's advocate, offering counterarguments and surfacing disconfirming evidence that teams might otherwise overlook. Assigning that role measurably shifts a group's search away from evidence that merely confirms the favored plan.[15] Early research suggests that AI-mediated

anonymity may also help surface underrepresented views while reducing the dissenter's interpersonal exposure, although whether it reproduces the effects of genuine dissent remains an open question.[16] On the other hand, AI introduces new forms of bias. Automation bias describes our tendency to place too much trust in machine-generated results and to stop questioning, especially when the output is articulate and the stakes are high.[17] A model's first framing anchors the discussion as firmly as any human first number, so a team that asked AI to widen its options can be narrowed by the answer it gets back. Furthermore, when everyone relies on the same model, individual differences in judgment can be replaced by a new kind of conformity, as all reasoning flows from a single source.

The boundaries that determine the final decision remain unchanged. Setting the risk appetite, weighing growth against margin against reputation, choosing which objective takes precedence when they conflict, owning the call afterward: these are judgments about values, and a model has no authority to make them. AI can provide input to inform the decision, but it cannot make the decision itself.

AI eases the constraints of memory, consistency, reach, and perspective. It does not shift the boundaries of judgment and ownership, and it adds limits of trust, anchoring, and offloading that a disciplined team now has to manage. The process itself remains essential; what changes is the set of issues the process must now address.

Bottom line

Managers satisfice, anchor on early information, interpret problems through their organizational position, and lose factors they cannot keep active in mind. These are persistent features of managerial judgment and organizational life, not habits that training alone can eliminate. The reliable response has always been a process built to expect them, which is the work Decision Quality does. What has changed is that AI now moves some of these limits: it extends memory, reduces variability, broadens reach, and loosens the grip of any single perspective. However, it does not alter the judgments that actually resolve a decision, the risk appetite, the trade-offs, and ownership. And it introduces its own challenges: misplaced trust and the risk of a single model shaping everyone's thinking. In my observation, the teams that excel at decision-making are those that understand precisely where their own limits lie and have learned exactly which boundaries a model can move and which it leaves in place.

References

1. Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1), 99–118.
2. Simon, H. A. (1957). *Models of Man: Social and Rational*. Wiley.
3. Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*. Little, Brown Spark.
4. Simon, H. A. (1976). From substantive to procedural rationality. In S. J. Latsis (Ed.), *Method and Appraisal in Economics* (pp. 129–148). Cambridge University Press.
5. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
6. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
7. Lovallo, D., & Kahneman, D. (2003). Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review*, 81(7), 56–63.
8. Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. *Project Management Journal*, 37(3), 5–15.
9. Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
10. Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114.
11. March, J. G., & Simon, H. A. (1958). *Organizations*. Wiley.
12. Ireland, R. D., Hitt, M. A., Bettis, R. A., & De Porras, D. A. (1987). Strategy formulation processes: Differences in perceptions of strengths/weaknesses and environmental uncertainty by managerial level. *Strategic Management Journal*, 8(5), 469–485.
13. Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
14. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688.
15. Schulz-Hardt, S., Jochims, M., & Frey, D. (2002). Productive conflict in group decision making: Genuine and contrived dissent as strategies to counteract biased information seeking. *Organizational Behavior and Human Decision Processes*, 88(2), 563–586.
16. Lee, S., Kim, M., Hwang, S., Kim, D., & Lee, K. (2025). Amplifying minority voices: AI-mediated devil's advocate system for inclusive group decision-making. In *Companion Proceedings of the 30th International Conference on Intelligent User Interfaces* (pp. 17–21).

17. Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410.

BREAKOUT
GROWTH PARTNERS

Kerr Manson
(847) 602-4275
kerr.manson@breakout-growth.com
www.breakout-growth.com