Sound Reasoning

Great outcomes don't prove great decisions—great reasoning does. In strategic choices, luck can flatter a weak process and bad breaks can punish a strong one. The lever leaders control is the *quality of the reasoning* that connects framing, alternatives, information, and values to a defensible choice. In Decision Quality (DQ), **Sound Reasoning** is the fourth link: it makes the argument logical, the probabilities honest, the math transparent, and the conclusions traceable—*before* you know the outcome.[1] This paper follows the same series structure. It builds directly on earlier links—**Appropriate Frame**, **Creative Alternatives**, **Relevant & Reliable Information**, and **Values & Trade-offs**—so that downstream **Commitment** rests on solid logic.

The theory in brief (why reasoning determines strategy)

Strategic decisions unfold in uncertainty. Executives work with incomplete facts, noisy signals, and competing objectives—yet must still reach defensible conclusions. The difference between intuition-led storytelling and genuine strategic reasoning lies in making the logical machinery visible and testable. When reasoning is explicit, the same inputs yield the same conclusions regardless of who runs the analysis.[1-3]

The challenge is that humans don't naturally reason this way. We jump from data to recommendations without stating how the evidence supports the claim—what logicians call the missing "warrant."[4] We craft compelling narratives about *this* project, *our* market, *our* unique advantages while systematically ignoring what happened to the last ten companies that told similar stories. Planning fallacy is pervasive—large projects frequently run over budget and behind schedule; reference classes help correct these errors.[5,7,14] Meanwhile, confirmation bias turns analysis into ammunition for positions already taken, and hidden-profile effects mean critical unshared information stays buried unless processes explicitly draw it out.[5-9]

The solution isn't more analysis but better-structured reasoning. The **Mediating Assessments Protocol (MAP)** offers a practical synthesis: decompose big strategic calls into a handful of decision-critical judgments (e.g., "price realization at month 12," "partner conversion by Q3"). Define each on a common scale, estimate independently with ranges and base rates, then combine transparently to rank options. This approach dampens narrative sway, makes disagreements visible on dimensions that matter, and improves calibration—turning procedural rationality from abstraction into executable discipline.[5,6,10]

Field evidence confirms the payoff. Studies consistently link procedural rationality and comprehensiveness to more effective strategic choices. Teams using structured dissent identify significantly more critical assumptions and achieve higher decision-quality ratings than consensus-only groups. When reasoning follows explicit protocols—with warrants stated, uncertainty quantified, and logic tested adversarially—decisions become genuinely auditable rather than superficially justified.[10,11]

From theory to practice:

These six moves operationalize decades of findings on how leaders reason well under uncertainty.

A. Make the argument explicit (claims → evidence → warrant → conclusion)

Why it works.	What good looks like.
Toulmin's argument model reveals that most strategic disagreements occur at the **warrant**—the reasoning that connects evidence to claim. Making warrants explicit surfaces hidden assumptions—the locus of many strategic disagreements. Adding **rebuttals** forces acknowledgment of boundary conditions where the logic fails.[4,9]	A one-page **Reasoning Chain**: (i) decision criteria; (ii) top **claims**; (iii) **evidence with sources**; (iv) the **warrant** that connects evidence to claim; (v) known **rebuttals/limits**; and (vi) the **qualified conclusion** ("...provided that..."). Keep the chain aligned with the agreed frame, option set, and value model.

B. Quantify uncertainty (ranges and probabilities, not point guesses)

Why it works.	What good looks like.
Point estimates mask uncertainty and feed overconfidence. 90% confidence intervals **often** capture far fewer outcomes than intended, evidence of overconfidence that range-based estimates can mitigate.[5–6,13] Ranges enable Bayesian updating, reveal true confidence, and focus attention on the variables that actually swing the decision. In practice, a small handful of variables typically drives most of the value variance—tornado analysis makes this visible.[8]	Ranged estimates for each decision-swing variable; a brief **confidence note** (why this spread). A **tornado chart** of top five sensitivities. **Flip points**—the minimal changes that reverse the winner.

C. Anchor the inside view with base rates (reasoning-specific use)

Why it works.	What good looks like.
Many projects exceed budgets and targets are frequently missed; reference class forecasting helps by forcing explicit deviation rationales.[5,7,14] Reference class forecasting **can materially reduce** planning errors by anchoring estimates to outcomes from comparable cases.[5,7,14] This enforces probabilistic coherence in the reasoning step.[7,8,13]	A **Base-Rate Box** beside each pivotal estimate: reference class (≥10 cases), median & 10th/90th percentile, your estimate, and a one-line rationale for any material deviation; show how moving to the base rate changes the ranking.

D. Separate facts, assumptions, and opinions (so challenge hits the logic)

Why it works.	What good looks like.
Mixed categories hide logical vulnerabilities. Research on hidden profiles shows tagging helps surface unshared, decision-critical information that often remains hidden in groups. Clear categorization shows where testing and dissent should focus.[6,9]	A **Source Table** for decision-critical inputs labeled **Fact / Supported assumption / Opinion** with source & date; flag any **single-source linchpins** for verification or a fast test.

E. Test the logic adversarially (before commitment)

Why it works.	What good looks like.
Compared with consensus-only discussion, DA/DI and premortems surface more disconfirming evidence and failure paths, improving decision quality ratings. The mechanism: legitimized challenge disrupts confirmation bias and forces search for disconfirming evidence.[11,12]	A **Challenge Pack** attached to the deck: (i) a tight DA brief aimed at the **top two warrants**; (ii) an **A vs. B** dialectic with 3–5 discriminators; (iii) a **premortem** with top failure modes and checks; plus a short **"what changed"** note.

F. Close with decision math and rules (so action is automatic)

Why it works.	What good looks like.
Expected value calculations (or multi-attribute value models) convert reasoning into choice and prevent post-hoc rationalization. Pre-specified decision rules—linked to observable evidence—create commitment devices that overcome present bias and political pressure. MAP-style independent assessments improve both accuracy and buy-in.[2,3,5]	Make the **EV formula** explicit: $EV(option) = \Sigma \Pr(scenario_i) \times Value(scenario_i)$ or use a **Value Scorecard** (0–100 value scores × weights). Combine via a MAP-style roll-up; document the **trade-off sentence**, leading indicators, and **stop/surge rules** ("If price realization <80% at M6, pause and re-vote.").

BREAKOUT
GROWTH PARTNERS

Exhibit — Reasoning Audit Sheet

Use the Reasoning Audit Sheet to document the decision, align the team on what would change the choice, and create an auditable trail for follow-ups.

REASONING AUDIT (page 1)

Decision Snapshot

Decision: _____ Date: _____
Type: ☐ Type 1 ☐ Type 2
Decision criteria: _____

Reasoning Chain (claims → evidence → warrant → rebuttal)

Claim	Evidence (source/date)	Warrant (how evidence implies claim)	Rebuttals / limits
1			
2			
3			

Critical Assumptions

Assumption	Would-Have-To-Be-True (for this to work)	Owner	Test/Validation
1			
2			
3			

Uncertainty Ranges

Variable	P10	P50	P90	Confidence Note (why this spread?)
1				e.g., vendor dispersion high; early sample; triangulated by X
2				
3				

REASONING AUDIT (page 2)

Base Rate Analysis

Our Estimate	Reference Case	Base rate (P10/Med/P90)	Delta & rationale
1		/ /	
2		/ /	
3		/ /	

Top 5 Tornado Drivers

#	Flip point
1	
2	
3	
4	
5	

Impact on ranking if moved to base rate:

Stop/Surge Triggers

If _____ < _____ by [date], then _____
If _____ > _____ by [date], then _____

What changed since last review:

Practical limitations (and how to work with them)

- **Computational complexity & interactions.** Additive EV/score models can miss second-order effects (e.g., ramp × price). Test the few plausible interactions among top drivers and use small scenario slices where needed.[2,3]
- **Group dynamics.** Loud voices and consensus pressure compress ranges and mute dissent. Collect estimates independently, poll silently, rotate a devil's-advocacy role, and record an authored dissent paragraph in the memo.[11]
- **Cultural resistance to probability.** Teams may resist ranges and qualifiers. Standardize P10–P50–P90, track calibration over time, and normalize "we don't know yet" paired with explicit review triggers.[5,6]
- **False precision & model overreach.** Clean spreadsheets can mask fragile warrants and omissions. Show ranges and flip points, include confidence notes, and audit the warrant—not just the numbers. Document what is not modeled—and why.[4–6]
- **Time pressure.** Analysis can sprawl or stall. Time-box work on decision-swing variables and move when VOI falls below the cost of delay; convert residual uncertainty into review triggers.[2,5]

Generative AI as scaffold (not substitute)

Where AI helps. Use AI to reduce extraneous cognitive load so humans can focus on warrants and judgment. AI excels at: drafting **argument structures** (claims → evidence → warrants → rebuttals); converting point estimates into **P10-P50-P90 ranges**; assembling **reference classes** and base-rate boxes; producing **tornado charts** and flip points; flagging **contradictions** across sources; and supporting **MAP roll-ups** on common scales. This accelerates the mechanics of reasoning while preserving human judgment on what matters.[5-7,11]

Where AI doesn't replace you. Don't outsource **values**, **risk posture**, or **confidence calls**. AI cannot set thresholds, decide when evidence is "good enough," or resolve political stakes. Watch for **hallucination** on facts—require source-tagging (asserted vs. cited) and verify decision-critical claims with primary sources. Treat AI outputs as **claims to be tested**, not conclusions.[2,3,11]

Four right-sized prompts:

1. **Reasoning Chain builder.** "From this memo and evidence pack, draft *claims → evidence (with sources) → warrants → rebuttals*; flag missing warrants and any single-source linchpins. Mark which items map to mediating assessments on a 0–100 scale."[4–6,11]

2. **Base-rate finder.** "Suggest ≥10-case reference classes for these estimates; report medians and P10/P90; compute the delta to our inside view and list one falsifiable reason to deviate for each."[5,7,8]

3. **MAP + sensitivity aide.** "Given these mediating assessments, output a tornado list and the minimal changes (**flip points**) that reverse the winner; show how the ranking changes if we move to the base rate."[5,6]
4. **Challenge Pack generator.** "Draft a devil's-advocacy critique aimed at the *top two warrants*, an *A vs. B* dialectic with 3–5 discriminators, and a pre-mortem list with checks/mitigations; summarize *what would change the choice*."[11–13]

Bottom line

Logically Correct Reasoning is the bridge from analysis to action. Make warrants explicit, quantify uncertainty with ranges, tether inside-view estimates to base rates, separate facts from assumptions, stress-test the logic with designed dissent, and close with clear decision math and rules. Do that, and your choices become **auditable, adaptable, and defensible**—the hallmark of Decision Quality—and the right platform for **Values & Trade-offs** and **Commitment** to stand on.

References

1. Spetzler, C., Winter, H., & Meyer, J. (2016). Decision Quality: Value Creation from Better Business Decisions. Wiley.
2. Keeney, R. L., & Raiffa, H. (1976). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley.
3. Hammond, J. S., Keeney, R. L., & Raiffa, H. (1998). The hidden traps in decision making. Harvard Business Review, 76(5), 47–58.
4. Toulmin, S. (1958). The Uses of Argument. Cambridge University Press.
5. Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before you make that big decision... Harvard Business Review, 89(6), 50–60.
6. Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). Noise: A Flaw in Human Judgment. Little, Brown.
7. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291.
8. Morgan, M. G., & Henrion, M. (1990). Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press.
9. Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making. Journal of Personality and Social Psychology, 48(6), 1467–1478.
10. Dean, J. W., & Sharfman, M. P. (1996). Does decision process matter? Academy of Management Journal, 39(2), 368–392.
11. Valacich, J. S., & Schwenk, C. R. (1995). Devil's advocacy and dialectical inquiry effects on decision making. Organizational Behavior and Human Decision Processes, 63(2), 158–173.
12. Klein, G. (2007). Performing a project premortem. Harvard Business Review, 85(9), 18–19.
13. Tetlock, P. E. (2005). Expert Political Judgment. Princeton University Press.
14. Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. Project Management Journal, 37(3), 5-15

BREAKOUT
GROWTH PARTNERS

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