BREAKOUT
GROWTH PARTNERS

Sound Reasoning

- Sound reasoning is the fifth link in Decision Quality. It is where the frame, the alternatives, the information, and the priorities are assembled into a choice, and the last point at which that choice can be inspected before the outcome is known.
- A good outcome does not prove a good decision. Luck flatters a weak process and a bad break punishes a strong one, so it is the reasoning, not the result, that shows whether a decision was sound.
- Reasoning usually fails in the same places: the warrant connecting evidence to conclusion goes unstated, confidence is drawn too narrow, and a vivid story crowds out the base rate. An argument left implicit cannot be tested for any of these.
- Six moves make the reasoning explicit and auditable: state the argument as claims, evidence, warrant, and conclusion; quantify uncertainty in ranges; tie the pivotal estimates to base rates; separate facts from assumptions and opinions; test the logic adversarially; and close with decision math and pre-set rules.
- Match the rigor to the stakes. A one-way door with large consequences earns the full discipline; a reversible bet rarely does.
- AI is fluent at the mechanics of reasoning, which makes it useful for the legwork and risky when handed the judgment. It can structure an argument and assemble base rates; it cannot set thresholds, weigh risk, or own the call, and it commits the very errors the reasoning is meant to expose.

In my experience, the clearest sign of a weak strategic decision is a team that can tell you what it chose but not why. Reasoning is what connects the frame, the alternatives, the information, and the priorities to a choice, and it is often implicit. A good outcome does not make up for reasoning no one ever spelled out. Luck can make a weak process look good, and a bad break can sink a strong one, so the result alone tells you little about the decision behind it. The reasoning is the part you can actually check before the outcome is in.

Sound reasoning is the fifth link in Decision Quality.[1] It is where the argument has to hold together: the logic stated, the probabilities honest, the model transparent, and the conclusion traceable to the evidence and values it rests on. It inherits the work of the earlier links, the frame, the option set, the information, and the priorities and trade-offs, and turns them into a choice. Get the reasoning wrong, and even a sound frame and a strong option set produce a choice no one can defend.

The theory in brief (why reasoning determines strategy)

Reasoning is the step that turns the frame, the options, the information, and the priorities into a choice. Executives reason under uncertainty, with incomplete facts and competing objectives. Most of the time, it stays in someone's head: a recommendation comes with the evidence but not the logic that ties the two together. Write the logic down, and someone else can take the same inputs and see whether they reach the same conclusion. That is the difference between an argument you can test and one you can only take on faith.[1]

The problem that keeps coming up is a gap in that logic. Toulmin's account of argument calls it the **warrant**: the reason a given piece of evidence supports a given claim.[2] In strategy debates, the warrant is usually left unsaid, so what looks like a fight about the facts is often a fight about an inference no one has spelled out. Say the claim is that Germany should be the next market and the evidence is that it is the largest market open to us in Europe: the warrant is the unspoken rule that size decides where to expand first. People can agree that Germany is the biggest market and still disagree on the move, because some would sequence by ease of entry instead, and the numbers never show that. Two failures hide inside an argument left unstated. The first is **overconfidence**: a single-number estimate says nothing about how wrong it could be, and even when people give a range, they draw it too narrow.[3,4] The second is unchecked input, where the argument runs on material no one has tested. The team's own case for this venture, built from its own strengths and the reasons this time is different, crowds out the plain record of how comparable ventures actually turned out, so an optimistic inside story wins over the **base rate**. In the same way, a fact that only one person in the room holds tends to stay with them unless the process asks for it directly.[5] None of it can be checked until the argument is out in the open.

The fix is to make the reasoning explicit and split it into parts you can judge one at a time. The **Mediating Assessments Protocol**, set out by Kahneman, Lovallo, and Sibony, is the cleanest version: decompose a large call into a handful of decision-critical judgments, define each on a common scale, estimate each independently with ranges and **base rates,** and only then combine them to rank the options.[6] Estimating each part on its own keeps the team's overall optimism from tilting every estimate the same way, narrows the argument to the one or two judgments that actually change the answer, and breaks the decision into concrete pieces the team can test and assign.

The evidence for this approach is consistent. Dean and Sharfman found that a more rational, analytical process led to more effective strategic decisions, independent of how well each was implemented.[7] Likewise, controlled studies of group decision-making show that teams that argue through structured dissent uncover more critical assumptions and produce stronger recommendations than teams that seek consensus.[8] Explicit arguments can be scrutinized and improved; implicit ones can only be justified after the fact.

From theory to practice

Each move below pairs a short reason it works with a description of what good output looks like, so the reasoning becomes a visible, auditable link rather than a leap the team takes on trust. The frame, the options, the information, and the priorities are settled when this work starts; reasoning is what assembles them into a choice.

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

Why it works.	What good looks like.
Toulmin's model of argument separates a claim from the evidence offered for it and from the **warrant**, the reason that evidence supports that claim.[2] In strategic argument the warrant is usually the part left implicit, so that is where disagreement hides and where a challenge should land. Stating the rebuttal next to it, the condition under which the logic would fail, keeps the argument honest about its limits.	A one-page reasoning chain: the decision criteria, the leading claims, the evidence behind each with its source, the warrant connecting evidence to claim, the known rebuttals or limits, and a conclusion stated with its conditions attached ("…provided that…"). The chain ties back to the agreed frame, the option set, and the priorities.

B. Quantify uncertainty (ranges, not point guesses)

Why it works.	What good looks like.
A point estimate hides how uncertain it is and invites overconfidence. The pattern is well documented: when people put a 90% range around an estimate, the real value falls outside it far more than one time in ten.[3,4] A **range** states the uncertainty instead of burying it, and it points attention at the few variables whose spread actually moves the decision. Usually a small handful of them drives most of the variance in value, which a tornado chart makes visible at a glance.[9]	A range, not a point, on every variable that can swing the decision, each with a one-line note on why the spread is that wide. A tornado chart of the top five sensitivities. And the flip points: the smallest change in each that would reverse which option wins.

C. Tie the pivotal estimates to base rates

Why it works.	What good looks like.
Inside the argument sit one or two estimates the whole case turns on, and the inside view builds them from the particulars of this venture while ignoring how comparable ones turned out. Pricing those estimates against a **reference class** of similar cases corrects the optimism, the discipline Flyvbjerg traced across hundreds of overrunning infrastructure projects and that Kahneman, Lovallo, and Sibony urge for any large forecast.[3,10] Assembling the reference class is the information link's job; using it to check the pivotal numbers is this one's.	Beside each pivotal estimate, a short base-rate panel: the reference class and how many cases it holds, its median and 10th and 90th percentiles, your own estimate, and one line justifying any gap between them. Show what happens to the ranking when the estimate moves to the base rate.

D. Separate facts, assumptions, and opinions

Why it works.	What good looks like.
An argument is only as testable as its inputs are legible. A case that blends established fact, supported assumption, and informed opinion into one confident voice hides which parts will actually bear weight under challenge, so labeling each input by its standing shows where a test or a dissent should go first. It also draws out the input most likely to be missing: in groups, information only one person holds tends to stay unshared unless the process asks for it, the **hidden-profile effect** Stasser and Titus demonstrated.[5]	A short table of the decision-critical inputs, each labeled fact, supported assumption, or opinion, with its source and date. Anything the case rests on that traces to a single source is flagged for verification or a fast test before commitment.

E. Test the logic adversarially (before commitment)

Why it works.	What good looks like.
A team that has built a case will defend it, and informal debate rarely breaks that grip. Structured dissent does: in controlled comparisons, groups running dialectical inquiry or devil's advocacy surfaced sounder assumptions and reached better recommendations than groups working toward consensus.[8] Aimed at reasoning, the job is not to generate new options, which the alternatives link already did, but to attack the **warrants** the case depends on and force the search for evidence that would sink them. A premortem takes the opposite approach: by assuming the decision has failed, it surfaces weak assumptions and potential failure points that were not previously articulated.[11]	A short adversarial review of the chosen option by people outside the sponsor's reporting line, aimed at the two or three warrants the case most depends on rather than the whole deck. An A-versus-B dialectic with three to five points the two framings genuinely disagree on. A premortem listing the top failure modes with a check against each. And a brief note of what the challenge changed.

F. Close with decision math and rules

Why it works.	What good looks like.
Reasoning has to close on something the team can point to afterward, or it quietly reopens after the meeting. Usually that is a number or a rule: an **expected-value calculation**, or a multi-attribute value model where objectives compete, converts the argument into a ranking and leaves a record of why one option beat the others.[12] Sometimes it is qualitative, where one option dominates on every objective that matters or a hard constraint rules out the rest, and what counts is that the basis is written down either way. Setting **the decision rule** in advance, tied to evidence anyone can observe, is what keeps the choice from being relitigated under pressure later. Combining the separate assessments rather than arguing the whole option at once also improves both the accuracy of the ranking and the team's willingness to stand behind it.[6]	The decision math written out: either an expected value, $EV(\text{option}) = \Sigma \Pr(\text{scenario}_i) \times \text{Value}(\text{scenario}_i)$, or a weighted value score where objectives compete, combined as a transparent roll-up of the separate assessments. One sentence naming the trade-off the choice makes. The leading indicators to watch, and the stop-or-scale rule set before the result is in ("if price realization is below 80% at month six, pause and re-decide").

Exhibit — Reasoning Audit Sheet

Use the reasoning audit sheet to record the decision and the few estimates the case turns on, capture what each rests on, agree what would change the choice, and leave an auditable trail for the follow-ups.

Reasoning Audit Sheet

Decision: _____

_____

☐ Type 1 ☐ Type 2 **Date:** _____ **Owner:** _____ **Version:** _____

PURPOSE: Document decision logic, align team on what would change the choice, and create auditable trail for follow-ups. Makes reasoning explicit, uncertainty quantified, and conclusions traceable.

① REASONING CHAIN (claims → evidence → warrant → rebuttal)

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

② CRITICAL ASSUMPTIONS

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

③ UNCERTAINTY RANGES

Variable	P10	P50	P90	Confidence Note (wh this spread?)
1				
2				
3				

④ BASE RATE ANALYSIS

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

⑤ TOP 5 TORNADO DRIVERS

#	Flip point
1	
2	
3	
4	
5	

Impact on ranking if moved to base rate:

⑥ STOP/SURGE TRIGGERS

If _____ < _____ by [date], then _____

If _____ > _____ by [date], then _____

What changed since last review:

Practical limitations (and how to work with them)

- **Additive models miss how the parts interact** — A weighted additive value model assumes that trade-offs among objectives can be represented separately, which may fail when the value of one outcome depends materially on another. Expected value itself does not require uncertain variables to be independent, but the model must still represent dependencies and joint effects correctly. Use the additive form as a transparent starting point, then test the few interactions among the pivotal drivers that could plausibly change the ranking.[12]
- **Group dynamics compress the range and mute dissent** — The loudest voice narrows the spread everyone reports, and the wish to agree quiets the objection worth hearing.[13] Collect the estimates independently before any group discussion, poll silently, give the devil's-advocate role to a named person rather than hoping it happens, and record an authored dissent in the memo so it survives the meeting.
- **Teams resist thinking in probabilities** — Asked for a range, people give a number and call the range hedging. Standardize on a 10th–50th–90th format so it stops feeling like evasion, track each person's calibration over time so the ranges earn trust, and make "we don't know yet, here is the trigger that will tell us" an acceptable answer.[4]
- **A clean spreadsheet can hide a weak argument** — Precise outputs lend a model authority its inputs have not earned, and a fragile warrant disappears behind the decimal places. Show the ranges and flip points rather than single figures; keep the confidence notes attached; and audit the warrant, not just the arithmetic. Write down what the model leaves out and why.[9]
- **The reasoning can sprawl until it stalls** — Detailed analysis is satisfying and close to limitless, so it expands to fill the time available. Time-box the work to the variables that actually swing the choice, and stop when the value of another round of analysis falls below what the delay costs, which is the question the value of information was built to answer.[14] Convert whatever uncertainty is left into review triggers rather than holding the decision open for it.

Generative AI as scaffold (not substitute)

AI can reason fluently. The output reads like an argument, set out in claims, warrants, and base rates, whether or not the argument holds. This is the step whose job is to catch bad reasoning, and it is the step where AI produces it. It will state a number it made up in the same tone it uses for one it looked up. That is why the AI's reasoning must be validated. Every warrant has to be tested against what the team actually knows, and any number that looks sourced has to be traced to its source before anyone relies on it.

Guard against that, and AI can save real time on the mechanics. It will structure a loose memo into claims, evidence, and warrants; convert point estimates into ranges; assemble a reference class and read the base rate off it; build the tornado chart and find the flip points; and flag where two sources cannot both be right. That is the slow, clerical half of reasoning, and clearing it lets the team focus on whether the argument is any good.

When the evidence is good enough to act on, where the threshold lies, what risks are worth running, who owns the call once it is made: these are all judgments for the team to make.

A few prompts do most of the work. The wording is a starting point; adapt it to your decision and your data.

- To build the argument: “From this memo and evidence pack, lay out the case as claims, the evidence for each with its source, and the warrant connecting them. Flag any warrant that is missing and any claim resting on a single source.”
- To check the pivotal estimates: “Suggest reference classes of at least ten comparable cases for these estimates, report the median and the 10th and 90th percentiles, and show the gap to our own number with one falsifiable reason we might deviate.”
- To find what moves the answer: “From these assessments, give me the tornado list and the smallest change in each that would reverse which option wins, and show how the ranking shifts if each estimate moves to its base rate.”
- To attack the case: “Draft a devil’s-advocacy critique aimed at the two warrants the recommendation most depends on, an A-versus-B dialectic with three to five points the framings genuinely disagree on, and a premortem with a check against each failure mode.”

The guardrails do not change from link to link. The decision owner signs the argument, not the model. Keep a record of what you asked and what the model leaned on, so you can retrace how the argument was built. Work in an approved environment and keep proprietary detail out of the prompts unless it is secured for that. Mark every fact the model supplies as sourced or asserted before it reaches the room. And ask one question before you commit: where did a confident paragraph from the model paper over a warrant no one had established?

Bottom line

Sound reasoning is what turns the upstream work into a choice and is the last chance to check a decision before it’s acted on. When arguments are made explicit, they can be tested for weak logic, overconfident estimates, ignored evidence, or missed objections. Arguments kept in someone’s head can only be defended after the outcome, when it’s too late to fix them. In my experience, the habit that most protects decisions is writing down the reasoning clearly enough for others to spot and address flaws before action is taken.

References

1. Spetzler, C., Winter, H., & Meyer, J. (2016). *Decision Quality: Value Creation from Better Business Decisions*. Wiley.
2. Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press.
3. Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before you make that big decision. *Harvard Business Review*, 89(6), 50–60.
4. Tetlock, P. E. (2005). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press.
5. Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48(6), 1467–1478.
6. Kahneman, D., Lovallo, D., & Sibony, O. (2019). A structured approach to strategic decisions. *MIT Sloan Management Review*, 60(3), 67–73.
7. Dean, J. W., & Sharfman, M. P. (1996). Does decision process matter? A study of strategic decision-making effectiveness. *Academy of Management Journal*, 39(2), 368–396.
8. Schweiger, D. M., Sandberg, W. R., & Ragan, J. W. (1986). Group approaches for improving strategic decision making: A comparative analysis of dialectical inquiry, devil's advocacy, and consensus. *Academy of Management Journal*, 29(1), 51–71.
9. Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press.
10. Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. *Project Management Journal*, 37(3), 5–15.
11. Klein, G. (2007). Performing a project premortem. *Harvard Business Review*, 85(9), 18–19.
12. Keeney, R. L., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Trade-offs* (2nd ed.). Cambridge University Press.
13. Valacich, J. S., & Schwenk, C. (1995). Devil's advocacy and dialectical inquiry effects on face-to-face and computer-mediated group decision making. *Organizational Behavior and Human Decision Processes*, 63(2), 158–173.
14. Howard, R. A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1), 22–26.

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