

The Leadership Evaluation Reference Guide

Nine Situations. The Right Questions. Better Decisions.

This guide is organized around the situations you actually face as a leader, not evaluation methodology. Find the situation. Ask the questions. Coach the conversation.

How to use this guide:

1. Before a meeting with a program director, find the matching situation
2. Scan the questions. Pick 2-3 that fit your specific context
3. Use the coaching tip to frame how you raise them
4. Refer to the appendix for the complete five-category framework

Situation 1

A Director Proposes a New Program

Your program director brings you a proposal for a new initiative. They are excited. They have a plan, a timeline, and a budget request. They want your approval.

Before you approve, you need to understand the bet they are placing. Every new program is a hypothesis: ***we will do [activity] and it will produce [outcome] for [population]***. That hypothesis needs to be explicit before resources are committed.

Questions to Ask

What specific problem or need is this program designed to address?

What observable change do you expect to see in students as a result?

How will we know if this program is working versus just being popular?

What is the hypothesis? What are we betting will happen?

Is there an existing program that already addresses this need?

What would success look like at 6 months? At 2 years?

How will we evaluate this, and when should we first check?

COACHING TIP: Do not ask 'how will you measure success?' That leads to satisfaction surveys. Instead ask 'what change should we see in students, and how would we know if it did not happen?' That leads to evaluable outcomes.

KEY INSIGHT: The time to define what success looks like is before the program launches, not after. If the director cannot name the expected change, the program is not ready for investment.

Situation 2

A Team Member Is Starting or Mid-Evaluation

Your staff member is working through the evaluation process. They come to you with a draft evaluation question, or they are deciding what part of the program to examine. They want your input.

Your coaching matters most at three specific moments. If you wait until the final briefing to engage, you have missed your three best opportunities to shape the work.

At the Evaluation Question Stage

Is this the right question for the decision we actually face?

If we answer this question, does it change something we would do?

Are we asking one clear thing, or three things jammed together?

Can we actually answer this with evidence we can collect?

At the Component Selection Stage

Are we looking at the right part of the program?

If this component were broken, would that explain the pattern we are seeing?

Can we actually change this component based on what we find?

At the Findings Stage

Does this give us what we need to make a decision?

Is the finding about an outcome or just an activity?

What does this evaluation not tell us?

COACHING TIP: You do not need evaluation expertise to ask these questions. You need leadership judgment. Your role is to ensure the evaluation is pointed at a real decision, not to evaluate the methodology.

Situation 3

A Team Member Presents Findings

Your staff member delivers an evaluation briefing. It includes a finding, a root cause, and a recommendation. You need to know how to engage with it productively.

Strong evaluation findings have three parts. Check that all three are present and connected:

Element	What It Answers	Red Flag If Missing
Finding	What happened?	You have description, not evaluation
Root Cause	Why did it happen?	The recommendation will target symptoms
Recommendation	What should we do?	You have insight but no action path

Five Questions to Ask When Receiving Findings
Is the finding about an outcome or an activity?
Does the root cause point to something we can change?
Does the recommendation target the cause, not the symptom?
What would we do differently if the finding were reversed?
What does this evaluation not tell us?

COACHING TIP: Question 4 is the most powerful. If the answer is 'nothing would change,' the evaluation did not inform a real decision. Question 5 prevents overconfidence. Name the blind spots before you act.

Situation 4

A Program Director Resists Evaluation Findings

Your program director built this program. They believe in it. The evaluation shows it is not producing the intended outcomes. They push back: 'The evaluation does not capture what this program really does.'

This is the hardest coaching conversation and the most important one. The goal is not to override the director or to concede. It is to hold both perspectives and use questions to move from belief to evidence.

Questions That Validate Without Conceding

I hear you that students value this experience. What outcome do you think it produces that we did not measure?

If the program is working in ways the evaluation missed, can we name that outcome and test it next cycle?

The evaluation found the design is not transferring skills. Could we add a transfer mechanism rather than defending the current design?

What would a redesigned pilot look like that keeps what works and fixes what does not?

If we had designed this evaluation together, what would you have wanted to measure?

KEY INSIGHT: The director is not wrong that the program has value. The evaluation is not wrong that the intended outcome is not occurring. Both can be true. The leadership skill is holding that tension and asking questions that move toward evidence rather than choosing a side.

COACHING TIP: If you lead with the data, you get defensiveness. If you lead with curiosity, you get a conversation. Start with 'help me understand' before 'the evidence shows.'

Situation 5

Budget Season: A Director Requests Increased Funding

Your program director submits a budget request for additional funding. They cite growing participation, strong satisfaction, and increased demand. They want more money.

Budget requests are often supported by activity data (how much we do) rather than outcome data (what changes as a result). Your job is to shift the conversation from volume to value.

Questions to Ask Before Approving
What evidence do we have that this program is producing the outcomes it was designed for?
Is the increased participation reaching the students who need it most, or the ones who are easiest to reach?
What is the cost per student, and is it proportionate to the impact?
If we invest more, what specifically will change for students that is not happening now?
Are there lower-cost ways to achieve similar outcomes?
What would we need to see in an evaluation to justify this investment level?

KEY INSIGHT: Growing participation is not evidence of effectiveness. A program can serve more students every year and still not produce the intended change. Before increasing investment, ask what the current investment is producing.

COACHING TIP: Frame it as partnership, not interrogation: 'I want to support this request. Help me build the case. What evidence can we point to that shows this investment is paying off for students?'

Situation 6

A Legacy Program Has Never Been Questioned

A program has been running for years, possibly decades. It continues because it has always existed. No one has asked whether it is still relevant or effective. You suspect it may be running on momentum rather than evidence.

Legacy programs are the hardest to examine because they are woven into institutional identity. The questions need to honor the program's history while opening the door to honest assessment.

Questions to Surface Strategic Drift
What problem was this program originally designed to address?
Is that problem still relevant today, or has the landscape shifted?
If this program disappeared tomorrow, what would we actually lose?
Who are the students this program serves today? Are they the students who need it most?
Has this program ever been evaluated, or has it been monitored?
When was the last time we fundamentally reexamined its design?
If we were starting from scratch today, would we build this program the same way?

COACHING TIP: The most powerful question for legacy programs is the last one. It separates the program's purpose (which may still be valid) from its design (which may be outdated). You can honor the mission while questioning the method.

KEY INSIGHT: Programs often persist long after the original need has changed. Continuation is not the same as effectiveness. The absence of bad news is not the same as the presence of good evidence.

Situation 7

A Director Reports Strong Satisfaction and Participation

Your program director presents a dashboard or report showing high satisfaction scores, growing attendance, and positive participant feedback. They conclude the program is working.

This is the Greenfield scenario. The data looks good on the surface. But satisfaction is not an outcome. Attendance is not impact. Before accepting the conclusion, test the assumptions.

Questions That Test the Assumptions
What changed for students as a result of this program? Not what they experienced, but what changed.
Satisfied compared to what? Is there a baseline or comparison?
What decision would change if satisfaction were lower?
Is participation growing because the program is effective, or because the eligible population grew?
Are we reaching the students who need this most, or the ones who are most likely to show up?
What is the repeat participation rate? Are students coming back, or coming once and leaving?

The Four Assumption Traps

The Assumption	The Question That Tests It
High attendance = reaching the right students	Who is actually participating vs. who needs it?
High satisfaction = students are benefiting	Satisfaction measures the experience, not the outcome
Growing participation = program is effective	Is the participation rate growing, or just the raw number?
The program name = what it does	Has anyone defined what this looks like in observable terms?

COACHING TIP: Do not dismiss the satisfaction data. Validate it, then extend: 'These numbers tell us students value the experience. What we do not know yet is whether the experience is producing the change we designed it for. Those are different questions.'

Situation 8

A Program Is Underperforming and You Must Decide

Evidence suggests a program is not producing its intended outcomes. You face a decision: redesign, scale back, or end the program. The stakes are high.

This decision requires distinguishing between three different problems. Each one leads to a different action:

The Hypothesis Is Wrong	The Design Is Wrong	The Implementation Is Wrong
The program cannot produce this outcome regardless of design	The right outcome, but the wrong approach to achieving it	The right design, but it is not being delivered as intended
Action: End or fundamentally reconceive	Action: Redesign the program components	Action: Fix delivery, training, or resources

Questions Before Deciding
Is the hypothesis wrong, or is the design wrong? These lead to very different decisions.
Are we seeing failure across the board, or only for certain populations or components?
Is the program being delivered as it was designed?
What would a redesigned version look like? Could we pilot it before making a permanent decision?
What would we lose if we ended this program? Are there outcomes we are not measuring?
What do we owe the students currently in the program?

KEY INSIGHT: The evaluation tells you the current design is not working. It does not tell you the outcome is impossible. Do not confuse a design failure with a hypothesis failure. The first is fixable. The second is not.

Situation 9

Annual Review or Strategic Planning: Reviewing the Portfolio

You are reviewing your full portfolio of programs during annual planning or strategic review. You need to assess which programs to expand, maintain, redesign, or sunset.

A portfolio review requires a different lens than evaluating a single program. You are looking for patterns across programs, not just within them. Use the five-category framework as a portfolio scan.

Portfolio-Level Questions
Which programs have defined, evaluable outcomes? Which are running on assumed effectiveness?
Where are we investing the most resources, and do we have proportionate evidence of impact?
Are there overlapping programs addressing the same need?
Which programs are reaching the students who need them most? Which are mostly serving students who would succeed regardless?
Which programs have been evaluated in the last three years? Which have never been evaluated?
If we had to cut 20% of the portfolio, what evidence would we use to decide?
What institutional priorities have shifted since these programs were designed?

Quick Portfolio Diagnostic:

For each program, ask:	Yes	No or Unsure
Does it have a defined, evaluable outcome?	Evaluation ready	Start here first
Has it been evaluated in the last 3 years?	Evidence exists	Prioritize for evaluation
Do we have evidence of student change?	Impact demonstrated	Operating on assumption
Does it align with current institutional priorities?	Strategically aligned	Examine for drift
Is the cost proportionate to the evidence of impact?	Resource justified	Investigate further

COACHING TIP: A portfolio review is not about cutting programs. It is about directing evaluation resources toward the programs where evidence is most needed and decisions are most urgent. Start with the highest-cost, lowest-evidence programs.

Appendix: The Complete Leadership Question Framework

The five categories below contain the full set of leadership questions. In the situations above, questions are selected and reassembled for specific contexts. This appendix provides the complete reference.

1. Purpose & Alignment — *Are we solving the right problem?*

- What problem or need was this program originally designed to address?
- Is that problem still relevant today?
- How does this program contribute to institutional priorities?
- If this program disappeared tomorrow, what would we lose?
- Are there overlapping programs addressing the same need?

2. Outcomes & Impact — *Is the program producing meaningful outcomes?*

- What outcomes are we expecting from this program?
- What evidence do we have that those outcomes are occurring?
- Are the outcomes meaningful for students?
- Are the outcomes sustained over time?
- Which students benefit most and least?

3. Program Design & Effectiveness — *Which parts are actually working?*

- Which components of the program contribute most to outcomes?
- Are there parts of the program that add little value?
- Are we delivering the program as intended?
- What assumptions about the program design might be wrong?
- If we redesigned this program today, what would we do differently?

4. Resource Use & Value — *Is the program worth the resources invested?*

- How much time, staff effort, and funding does this program require?
- Are the outcomes proportionate to the resources invested?
- Are there lower-cost ways to achieve similar outcomes?
- Which programs in our portfolio deliver the highest impact per resource invested?
- Are there programs that continue mainly due to history rather than results?

5. Learning & Improvement — *What are we learning and how are we improving?*

- What have we learned from evaluating this program?
- What improvements have been made based on evidence?
- What questions about this program remain unanswered?
- What evidence would help us improve the program?
- How often should we re-examine this program's design?

Remember, you do not need to become an evaluator

**Your goal is to be a good questioner and coach of the people
who do**

Start with purpose. Demand outcomes. Coach the process. Act on findings.