How to Learn From a Program That Did Not Work as Hoped

When evidence points to change, leaders need to understand what happened before assigning blame

FOR

executive directors, program directors, board members, county administrators, case managers, and anyone facing difficult evidence about a program they care about

INCLUDES

a learning-first response framework, a redesign-versus-sunset decision table, a say/cannot-say table, and guidance on when to bring in help

The moment

You are sitting in a conference room with your leadership team. The evaluation results are on the screen. Your workforce readiness program — the one you spent two years building, the one your staff poured themselves into, the one the board celebrated at last year's gala — did not move the employment numbers. Participants who completed the program were employed at roughly the same rate as a comparable group that did not participate.

The room is quiet in the specific way that rooms get quiet when people are deciding whether to challenge the data, defend the program, or wait for someone else to speak first.

This is one of the hardest moments in program work. Not because the evidence is complicated — sometimes it is, sometimes it is not — but because the emotional stakes are high. People built this program. People believed in it. Participants trusted you with their time. And now the numbers are saying something nobody wanted to hear.

What happens next will determine whether your organization learns from this moment or wastes it. A structured response keeps diagnosis ahead of judgment.

The common mistake

Organizations typically make one of two mistakes at this moment, and both are understandable.

Mistake one: defending the program by attacking the evidence. Someone says the data was collected wrong, the measure was unfair, the comparison group was not really comparable, or the evaluation did not capture the "real" impact. Some of these objections may be legitimate. When every disappointing result is met with a reason the data must be wrong, the organization loses its ability to learn.

Mistake two: treating the result as a verdict and moving immediately to cut or end the program. A board member reads the headline number and says: "This is not working. We need to redirect these resources." That may turn out to be the right conclusion. But jumping to it before understanding the result means you may cut a program that was poorly implemented rather than poorly designed, or end something that was reaching the wrong population rather than failing the right one.

Both mistakes share a common root: they skip the diagnostic step. They jump from the feeling — disappointment, defensiveness, frustration — to a conclusion. The work that needs to happen is in between.

There is a third mistake that is quieter but equally damaging: doing nothing. Acknowledging the result in a meeting, expressing concern, and then letting the program continue unchanged because nobody wants to have the harder conversation. The evidence gets filed. The program runs another year. The same result comes back.

What to prepare

Before the meeting where the evidence will be discussed, prepare three things.

1. A clear, factual summary of what the evidence shows. State the result in plain language. What was measured, for how many people, over what time period, and compared with what? Do not editorialize in either direction — do not soften it and do not dramatize it. “Participants’ employment rate at six months was 34 percent, compared with 31 percent for the comparison group. The difference was not statistically meaningful” is factual. “The program basically failed” is not. Neither is “The program almost worked.”

2. A preliminary diagnosis using five questions. Before the meeting, ask: Was the program implemented as designed? Was the measure valid? Were the right people reached? Were external factors overwhelming? Did the program genuinely fail to produce the change? Come to the meeting with your best assessment of which explanation is most likely and the evidence for that assessment.

3. A set of options, not a single recommendation. Present two or three possible responses, each tied to a different diagnosis. “If the issue is implementation fidelity, here is what we would change. If the issue is population fit, here is what we would change. If the model itself needs revision, here is what that process would look like.” This gives leadership a structured choice rather than a binary defend-or-cut decision.

Anyone reviewing the results should prepare the same way and ask diagnostic questions before reaching a verdict: “Help us understand the result. Walk us through what you think happened and what you would do differently.”

A ready-to-use framework: the redesign-or-sunset decision

After you have completed the diagnostic work, the question becomes: Do we redesign this program or do we end it? This table maps diagnoses to responses and structures the conversation.

DIAGNOSIS	EVIDENCE FOR THIS DIAGNOSIS	LIKELY RESPONSE	BEFORE YOU DECIDE
Implementation was inconsistent — the program did not run as designed	Attendance records, dosage data, staffing gaps, or fidelity checks show the model was not fully delivered	**Redesign: fix the implementation** before judging the model. Strengthen training, supervision, staffing, or delivery systems.	Ask: If we delivered the full model consistently, do we believe the results would be different? What evidence supports that belief?
The measure missed the real change	Staff and participants report meaningful benefits that the outcome measure did not capture	**Redesign: improve the measurement** while continuing the program. Add or change measures to better capture the program's contribution.	Ask: What do participants and staff say changed? Can we find a valid measure for that?
The program reached people it was not designed for	Enrollment data shows participants had barriers or characteristics the model does not address	**Redesign: adjust targeting or add supports.** Refine eligibility, referral pathways, or add complementary services for the barriers participants actually face.	Ask: Who does this model work for? Can we reach those people more consistently?

DIAGNOSIS	EVIDENCE FOR THIS DIAGNOSIS	LIKELY RESPONSE	BEFORE YOU DECIDE
External conditions overwhelmed the program	Broader trends show the same outcome worsening for everyone, not just participants	**Continue with realistic expectations** and seek comparison data that shows the program's relative contribution. Consider partnerships to address the external factor.	Ask: Did participants do better than they would have without the program, even if the headline number is flat?
The model was delivered well, to the right people, in fair conditions, and the result still did not materialize	Implementation was strong, the measure is valid, the population matched the design, and external factors were not extraordinary	**Consider sunsetting** the program or fundamentally redesigning the theory of change. This is the diagnosis that calls for the hardest conversation.	Ask: What would we need to see to believe the model could work? Is there a plausible path, or are we holding on because of attachment rather than evidence?

If the answer is to sunset the program, do it with honesty and respect. Acknowledge what the program attempted. Recognize the staff who built it. Explain what you learned and how it will inform what comes next. A program ended in response to credible evidence demonstrates that the decision process is working.

What not to say

INSTEAD OF THIS	SAY THIS
"The program failed."	"The program did not produce the outcome we measured. We are working to understand why."

INSTEAD OF THIS	SAY THIS
"The evaluation was flawed, so we should disregard the results."	"We have questions about aspects of the evaluation. Here is what it does tell us, and here is where we need more information."
"We will try harder next year."	"We have identified specific changes to implementation, targeting, or design that address the most likely explanation for the result."
"This is just one study — we should not overreact."	"We take this result seriously. Here is what we are doing to understand it and respond."
"We need to move on and redirect resources immediately."	"Before we make a resource decision, we want to complete the diagnostic work so the decision is based on understanding, not just the headline number."

When to get help

Outside support is worth considering when:

- The disappointing result involves a program that serves vulnerable people, and a wrong interpretation could lead to cutting services that are actually helping — or continuing services that are actually not. The stakes of getting the diagnosis right are high enough to warrant independent review.
- Your team is stuck between defending the program and ending it, and neither side is budging. An outside facilitator can create the space for the diagnostic conversation that the internal dynamics are preventing.
- The result will be shared with a board, oversight body, or the public, and you need help framing the finding in language that is honest, defensible, and constructive.

- You believe the model needs fundamental redesign and want help rebuilding the theory of change based on what the evidence revealed, rather than starting from scratch or repeating the same assumptions.
- The program involves multiple partner agencies, and the response to disappointing evidence needs to be coordinated across organizations with different priorities, risk tolerances, and political pressures.