

What You Need to Know Before You Expand

The evidence questions that matter most when a pilot is ready to grow

FOR

executive directors, program directors, board members, county administrators, and anyone deciding whether to take a small success bigger

INCLUDES

a scaling-readiness checklist, a pilot-to-scale evidence comparison table, a say/cannot-say table, and guidance on when to bring in help

The moment

Your housing navigator pilot has been running for fourteen months in two neighborhoods. The numbers are encouraging: participants found stable housing faster than the waitlist comparison group, and retention in supportive services improved. Your executive director presented the results to the board last week. A county commissioner called this morning asking whether the program could expand to six more sites by next fiscal year.

Everyone is excited. The staff who built this pilot have worked hard. The families who were served are real people with real outcomes. And the question on the table feels like good news, which makes it harder to slow down and ask the uncomfortable version: What do we actually know, and is it enough to justify doing this six times bigger?

This is one of the highest-stakes evidence moments an organization faces. Scaling a pilot makes a much larger bet based on what you learned from a much smaller one. The evidence that justified the pilot may be too weak to justify the expansion, and the conditions that made the pilot work may change at scale.

Before expanding, identify the assumptions and operating conditions on which the larger commitment depends.

The common mistake

The most common mistake at this moment is treating pilot results as proof that the program works at scale. Organizations take a set of promising early numbers, present them as settled evidence, and build expansion plans around them.

Here is what that usually looks like: the pilot served 60 families across two sites with a handpicked team. Results were strong. Leadership concludes that expanding to 360 families across twelve sites will produce the same results. The expansion plan uses the pilot data as its justification.

The pilot data answers a different question than the one the expansion raises. The pilot showed that the program can work under the conditions tested. The expansion asks whether it will work with more sites, new staff, different neighborhoods, and potentially different populations.

Three specific gaps tend to get overlooked:

The team gap. Pilots are often run by the people who designed the program – experienced, motivated, deeply committed. Expansion requires hiring new staff who may not share that background. What the pilot showed about the model may partly reflect what it showed about those particular people.

The context gap. The neighborhoods, partner agencies, referral networks, and community conditions at the pilot sites may not exist at the new sites. A housing navigator program that worked because of a strong relationship with a local landlord network may not replicate in a neighborhood where those relationships do not exist yet.

The fidelity gap. At pilot scale, the program director can observe every case, catch problems early, and make adjustments in real time. At six times the scale, that level of oversight is not possible. The question becomes whether the program model is clear and structured enough to produce good results without the founder looking over every shoulder.

Organizations that skip these questions do not always fail. But when scaling does fail, the post-mortem almost always traces back to one of these three gaps.

What to prepare

Before you present an expansion plan to leadership or a partner agency, you need to be clear about four things.

1. What the pilot evidence actually shows. State the results precisely. How many people were served? Over what time period? What was measured? What was the comparison – a waitlist, a historical baseline, a matched group, or no comparison at all? Be honest about the design's strength. A before-and-after comparison with 40 participants provides useful evidence with more uncertainty than a quasi-experimental study with a matched comparison group. [How Much Evidence Is Enough?](#) provides a proportionality framework for matching the evidence standard to the stakes.

2. What conditions made the pilot work. List the specific factors that contributed to the results. These might include: the experience level of staff, the size of caseloads, the quality of referral partnerships, the characteristics of the population served, the

intensity of supervision, or the availability of complementary services. Be concrete. “Good staff” is not a condition you can replicate. “Staff with at least two years of housing case management experience, carrying no more than 15 active cases” is.

3. Which of those conditions will change at scale. Go through the list from step two and mark which conditions you can maintain at the new sites and which you cannot. This is the honest part. If the pilot worked partly because the program director personally supervised every case and that will not be possible at twelve sites, that is a condition that changes at scale. Name it.

4. What you will monitor during expansion. Scaling without a learning plan is scaling blind. Before you expand, identify three to five indicators you will track from the first month – not just final outcomes, but early signals that the program is running as designed. If mentor-match completion rates drop below 70 percent at a new site, you want to know in month two, not month twelve.

A ready-to-use checklist: scaling readiness

Bring this to the meeting where the expansion decision is being discussed. Work through it together before committing.

QUESTION	YOUR ANSWER	IMPLICATIONS
How many people did the pilot serve, and over what period?		Smaller samples and shorter periods mean more uncertainty about whether results will hold.
What was the comparison? (Waitlist, baseline, matched group, no comparison?)		The weaker the comparison, the less confident you can be that the program caused the results.

QUESTION	YOUR ANSWER	IMPLICATIONS
Can you name three specific conditions that made the pilot work?		If you cannot name them, you may not know what you are replicating.
Which of those conditions will be different at the new sites?		Each difference is a risk. Not a reason to stop, but a reason to plan.
Will new staff deliver the program? If so, what training and supervision will they receive?		New staff without adequate preparation is the single most common reason scaled programs underperform.
What is your plan for monitoring implementation quality during the first six months?		If you have no plan, you will not know whether poor results reflect the model or the rollout.
What would make you pause the expansion?		Defining a pause threshold before you start is much easier than deciding mid-crisis.
Have you documented the program model clearly enough that someone who was not part of the pilot could deliver it?		If the model lives in people's heads, it will drift at scale.

What not to say

INSTEAD OF THIS	SAY THIS
"The pilot proved the program works."	"The pilot produced promising results under specific conditions. We believe those results can translate to new sites with the right preparation."
"We are ready to scale."	"We have identified what made the pilot successful and have a plan to maintain those conditions at new sites."

INSTEAD OF THIS	SAY THIS
"The data speaks for itself."	"Here is what the data shows, here is what it does not show, and here is how we are addressing the gaps."
"We just need to do the same thing in more places."	"Scaling requires adapting the model to new contexts. We have mapped the differences and have a plan for each one."
"If it worked there, it will work here."	"We expect the model to translate, and we have built in early monitoring so we can adjust if conditions at new sites require it."

When to get help

Outside support is worth considering when:

- The expansion involves a significant increase in budget or population served, and the pilot evidence rests on a pre-post comparison or descriptive data alone. A methodologist can help you understand what the pilot evidence supports and design a learning plan for the expansion phase.
- You need to present the pilot results to a policymaker or oversight body that will ask tough questions about causation, and you want to be prepared with honest, defensible language.
- The expansion moves into communities or populations that differ meaningfully from the pilot sites, and you need help designing an adaptation strategy that preserves the core model while allowing for local conditions.
- Your organization has never managed a multi-site program before, and you need guidance on implementation monitoring — the systems that let you catch problems early rather than discovering them in the year-end evaluation.

- The pilot produced mixed results — some outcomes improved, others did not — and you need help interpreting what can scale and what needs redesign first. [How to Learn From a Program That Did Not Work as Hoped](#) provides a diagnostic framework.