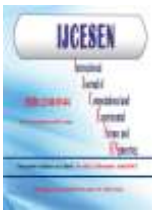
International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 7039-7052
http://www.ijcesen.com



Copyright © IJCESEN

ISSN: 2149-9144

Research Article

From Experimentation to Standardization at a Startup: A Strategic Framework for Scaling Operations through Project Portfolio Governance

Vivek Udaykumar*

Head of Project Management, Lima- USA
* **Corresponding Author Email:** vive2k@gmail.com - **ORCID:** 0000-0002-5247-225X

Article Info:

DOI: 10.22399/ijcesen.5532
Received : 05 September 2025
Revised : 25 October 2025
Accepted : 27 October 2025

Keywords

project management;
startup scaling;
entrepreneurial experimentation;
portfolio management;
benefits realization;
strategic alignment

Abstract:

Scaling a startup requires a transition from founder-led experimentation to repeatable delivery without losing the ability to revise uncertain assumptions. This paper examines how project management can support that transition by connecting product development, financial capacity, and market execution. It develops a conceptual framework from selected research on entrepreneurial experimentation, management control, agile delivery, benefits realization, and organizational learning. The framework treats startup scaling as a portfolio of bounded investments whose continuation depends on customer evidence, operating readiness, and resource availability. It distinguishes project outputs from customer outcomes and business benefits, and explains how governance should change as a venture progresses from discovery to repeatability and expansion. Practical applications include project selection, rolling-wave planning, staged funding, dependency management, risk ownership, and performance measurement. A hypothetical software startup illustrates how these practices can redirect investment toward an onboarding bottleneck before expanding customer acquisition. The paper also proposes a phased implementation approach and a research design for testing the framework. No original survey or causal analysis is reported. The central argument is that appropriately tailored project management can make growth decisions more explicit, coordinated, and reversible, while excessive process can impose costs that outweigh its benefits. Sustainable scaling therefore depends on the quality of decisions and realized benefits rather than the quantity of projects completed.

1. Introduction

Startups often begin with a small group making decisions through frequent, informal conversations. Founders can negotiate with customers, redirect development, and resolve conflicts without elaborate reporting structures. As the venture adds employees, products, partners, and customers, those arrangements become harder to sustain. A promise made during a sales call can change engineering priorities, delay a deployment, and create support costs that were absent from the financial forecast. Scaling therefore creates a coordination problem alongside the commercial challenge of finding demand. Project management becomes useful when it makes those relationships visible and helps the organization decide what to undertake, how to deliver it, and when to stop.

This paper defines startup scaling as the deliberate expansion of a venture's capacity to deliver and capture customer value through increasingly repeatable processes. Growth refers to an increase in activity, revenue, customers, or resources. Scaling, as used here, additionally requires attention to whether operating capacity and economic performance can support that increase. A startup can grow while delivery quality deteriorates or cash requirements rise faster than its ability to finance them. This definition is analytical rather than a universal classification: labor-intensive and capital-intensive ventures may need substantial resource growth even when their expansion is well managed.

Project management provides a way to organize temporary initiatives that change the business. Examples include validating a new segment, automating onboarding, introducing a pricing model, integrating a distribution partner, or preparing a product for a larger deployment. Product management determines which customer problems and product outcomes deserve attention.

Operations sustains recurring delivery after a change has been introduced. Portfolio management chooses among competing investments. These responsibilities can sit with the same person in a small startup, but distinguishing the decisions prevents a completed feature from being mistaken for a successful growth strategy.

The starting point for this paper is the alignment of product design, finance, and market positioning. That relationship remains important, particularly for AI-enabled products whose technical performance, service costs, and commercial promises are closely connected. The present argument extends it by asking how alignment is maintained in daily work. A strategy statement cannot allocate a scarce engineer between two urgent projects or determine whether a pilot result warrants further funding. Those decisions require explicit ownership, useful evidence, a schedule for review, and a way to adjust commitments.

The paper addresses three questions. First, which project management practices can help a startup expand without multiplying unresolved execution problems? Second, how should those practices change with uncertainty, organizational complexity, and financial exposure? Third, how can founders assess whether project management improves scaling rather than merely increasing administrative activity? The contribution is a practical conceptual framework linking investment decisions to delivery and benefits. It is intended for founders, project and program managers, accelerator teams, and researchers studying the organization of entrepreneurial growth.

2 Method and Scope

This is a conceptual paper based on a targeted narrative synthesis of published research and selected professional guidance. Sources were chosen for their direct relevance to five mechanisms: entrepreneurial learning, startup formalization, adaptive delivery, benefits realization, and coordination under uncertainty. Peer-reviewed studies provide empirical context, while the Scrum Guide and the National Institute of Standards and Technology AI Risk Management Framework provide clearly identified practice guidance. Publisher records, institutional repositories, and official guidance pages were used to verify bibliographic information and the specific findings summarized here. The review is selective and does not claim exhaustive coverage, a formal search protocol, or meta-analytic estimates.

The synthesis separates evidence from the paper's proposed applications. Research on a particular practice does not establish that the full framework causes startup success. Likewise, a professional framework specifies an approach but does not by itself demonstrate its effectiveness. Sections that recommend funding gates, project records, meeting cadences, or implementation sequences present a managerial design derived from the synthesis. These recommendations require local adaptation. The numerical case is hypothetical and exists to show decision logic, not to provide a benchmark or an observed company result.

No original respondents, interviews, experiments, or financial observations form part of this study. The paper does not estimate statistical relationships or report new performance effects. Its contribution is conceptual development followed by a proposed empirical evaluation. Empirical findings from the literature are described within their original settings, and their relevance to startup project management is treated as an inference. Evidence about established organizations is not assumed to apply unchanged to new ventures. Likewise, estimates in the illustrative case depend on explicitly stated assumptions and cannot validate the framework. These boundaries establish the basis on which readers should assess the argument and identify where further research is required.

The framework is most immediately applicable to knowledge-intensive ventures where product changes and commercial experiments can be tested incrementally. Hardware, biotechnology, regulated services, and physical operations can also use its decision principles, although their lead times, testing requirements, and investment reversibility differ. The discussion does not assume that every startup should pursue rapid expansion. Maintaining a smaller operation, postponing a market entry, or discontinuing a project may be the more defensible decision when customer evidence or financial capacity is insufficient.

3 Research Foundations

3.1 Entrepreneurial learning and investment discipline

Entrepreneurial uncertainty limits the usefulness of plans that assume a stable market and a known route to profitability. Sarasvathy (2001) distinguishes predictive reasoning from an effectual logic that emphasizes control over available means and commitments under uncertainty. For the present framework, this supports a modest conclusion: early project plans should define what the venture can afford to test and learn before treating distant forecasts as reliable. It does not imply that forecasting is unnecessary or that every founder follows the same decision process.

Camuffo et al. (2020) studied a scientific approach to entrepreneurial decision-making in a randomized trial involving 116 Italian startups. The intervention taught participants to articulate predictions and test hypotheses. The reported results support the relevance of disciplined inquiry to entrepreneurial choices. A later replication involving 759 firms across four trials found effects on idea termination and a more nuanced pattern of strategic changes (Camuffo et al., 2024). These studies support explicit assumptions and decision criteria, while cautioning against interpreting experimentation as an instruction to pivot continuously or as a guarantee of revenue growth.

Koning et al. (2022) examine experimentation through the adoption of A/B testing in technology startups. Their analysis connects experimentation with learning and subsequent venture performance. Its relevance here is the organizational use of evidence, rather than any universal performance multiplier. A startup with very few customers, long purchase cycles, or substantial spillovers between users may not be able to run informative A/B tests. Interviews, paid pilots, observed workflows, and staged operational trials can provide different forms of evidence, each with limitations that the project record should acknowledge.

3.2 Formalization and adaptive delivery

Davila and Foster (2007) investigate management control systems in 78 early-stage startup companies. Their work makes formalization relevant to the study of entrepreneurial growth and challenges the assumption that all management systems are inherently incompatible with startups. However, management control is broader than project management, and the study does not establish that adopting a particular project tool causes scaling. The implication developed here is that ventures should introduce coordination mechanisms in response to observable decision and execution needs.

Serrador and Pinto (2015) analyze 1,002 projects across industries and countries and report positive relationships between agile use and measures of project success. The study supports considering adaptive delivery, but its sample is not restricted to startups and should not be treated as a startup-specific causal test. The Scrum Guide describes an iterative framework with defined accountabilities, goals, inspection, and adaptation (Schwaber & Sutherland, 2020). Together, these sources help distinguish adaptive work from the absence of planning. They do not require every venture or every project to use Scrum.

Professional guidance also emphasizes tailoring the management approach to the organization and project. Whitaker (2014) describes methodology development as an iterative process of assessment, design, and improvement. Its practical relevance is the need to examine whether a process is used and whether it helps decisions. The supporting convenience survey should not be interpreted as causal evidence that greater formalization always improves startup performance.

3.3 Benefits and organizational learning

Serra and Kunc (2015) examine benefits realization practices and report positive relationships with strategic project success. Their work is useful because a delivered output can fail to create its intended business benefit. For startups, a billing integration might work technically while customers still abandon setup. The application proposed in this paper is to retain a named benefits owner after delivery, with a scheduled review of adoption and economic effects. This extends project evaluation beyond whether the work finished on time.

Edmondson (1999) finds an association between psychological safety and learning behavior in a study of 51 manufacturing work teams. Although that setting differs from a startup, the underlying concern is relevant to project reporting: teams need to disclose mistakes and uncertainty. This paper proposes that founders make early escalation credible by responding to bad news with investigation and decisions. A reporting process loses value when employees believe that revising an estimate or questioning a customer commitment will be treated as disloyalty.

Taken together, the literature supports examining mechanisms rather than prescribing a single maturity model. Experimentation informs which commitments are justified; coordination practices help deliver them; benefits management checks whether they mattered. The missing empirical step is a direct evaluation of how these mechanisms interact across startup stages. The framework below is intended to make that evaluation possible while offering practitioners a coherent set of decisions to test in their own organizations.

4 A Project Management Framework for Startup Scaling

4.1 Connecting strategy to accountable work

The proposed framework begins with a scaling constraint. Founders identify what currently limits the venture's ability to serve more customers sustainably: uncertain demand, slow onboarding,

low retention, unreliable delivery, weak distribution, or another specific condition. They then define an outcome that would demonstrate improvement. A project is authorized only when its proposed output has a plausible connection to that outcome, an accountable owner, and a resource requirement the venture can accommodate. This connection is revisited when new evidence changes the original rationale.

For example, a startup may identify delayed activation as its immediate constraint. Its desired outcome is that suitable customers reach a useful first result sooner. Potential projects include simplifying configuration, changing customer training, or improving data import. Each addresses the same outcome through a different mechanism. The team should compare those alternatives before committing to a larger platform rebuild. This avoids equating the most technically ambitious proposal with the highest-value response to the business problem.

Every authorized project should have a short charter that records the customer or operational problem, intended outcome, scope boundary, principal assumptions, project owner, benefits owner, and next investment decision. It should also identify the resource envelope and the evidence needed to continue. The charter is a decision record rather than a ceremonial document. If nobody uses a field to allocate resources, coordinate work, or evaluate results, the team should question whether that field deserves to remain.

The project owner coordinates delivery and makes dependencies visible. The benefits owner remains responsible for checking whether the change improves the business after implementation. A founder or functional leader acts as sponsor when a decision exceeds the project owner's authority. Separating these accountabilities is useful even when one person initially holds all three. As hiring proceeds, the startup can transfer responsibility without losing the original purpose of the work.

4.2 Matching governance to the stage of uncertainty

The framework distinguishes discovery, repeatability, and expansion as decision contexts rather than fixed company ages or funding rounds. In discovery, the important question is whether the problem and proposed response justify further commitment. During repeatability, the venture tests whether it can deliver the promised value across comparable customers without relying on exceptional founder effort. During expansion, it evaluates whether additional volume, segments, or locations can be supported economically and operationally. A single company may have projects in all three contexts at once.

Discovery projects should have bounded experiments, learning questions, and explicit expenditure limits. Repeatability projects need baseline process data, acceptance criteria, and a plan for transferring work into operations. Expansion projects usually require more explicit capacity forecasts, cross-functional dependencies, and contingency planning. The amount of governance should also rise with potential harm and irreversible commitments. A small venture deploying a high-consequence system may need substantial assurance even before it has repeatable revenue.

4.3 Maintaining an evidence and decision cycle

The framework operates through a recurring cycle of selecting work, testing assumptions, delivering a change, observing outcomes, and reallocating resources. At each review, the responsible decision-maker can continue, narrow, pause, redirect, or stop the project. These choices should be recorded with the evidence available at the time. A decision log helps later teams distinguish a reasonable choice under uncertainty from an execution mistake and prevents retrospective narratives from hiding why a project expanded beyond its original purpose.

This cycle requires a distinction between acceptable uncertainty and missing accountability. A team may legitimately be unable to forecast adoption precisely. It should still know who will obtain customer evidence, when the evidence will be reviewed, and how much the next test can cost. Similarly, a technically uncertain task can have an accountable investigation plan even when its completion date is expressed as a range. Project management contributes by structuring the next decision without pretending to eliminate uncertainty.

5 Selecting and Sequencing the Scaling Portfolio

5.1 Identifying the current constraint

Portfolio selection starts by asking where additional effort is most likely to improve the venture's ability to deliver value. If customers do not return after initial use, increasing acquisition may enlarge an expensive replacement cycle. If demand is credible but implementation queues are growing, delivery capacity may deserve attention before another campaign. This is a diagnostic question, not a universal rule about the order of startup activities. The team should examine segment-level evidence because an aggregate metric can conceal a healthy

segment alongside an unsuitable one.The startup can compare proposed projects using strategic relevance, strength of evidence, expected benefit, time to learn, resource demand, and dependency exposure. A short written rationale is often more informative than a score with unsupported numerical precision. Where scoring is useful, leadership should explain the weights and test whether plausible changes alter the ranking. A project should not appear compelling merely because its sponsor used an optimistic revenue estimate while another sponsor used a cautious one.
Mandatory work requires separate treatment. Contractual delivery, essential maintenance, and necessary risk controls may consume capacity even when their immediate revenue contribution is difficult to quantify. They should be visible in the portfolio rather than hidden as interruptions. Management can then see how much capacity remains for experiments and expansion. The appropriate allocation depends on the venture's commitments and exposure; a universal percentage would obscure the trade-offs that the portfolio review is meant to resolve.

5.2 Accounting for shared resources and dependencies

Two individually attractive projects can form an infeasible portfolio when both require the same specialist or external approval. The project manager should maintain a dependency map that identifies the prerequisite, responsible party, expected completion range, and consequence of delay. Dependencies that affect several projects deserve particular attention because their failure can invalidate multiple schedules at once. A shared data migration, for example, may determine the feasibility of customer onboarding, analytics improvements, and a product launch.
Work-in-progress limits provide a practical response to overcommitment. Leadership decides how many significant initiatives the available team can advance meaningfully and makes an explicit trade-off before starting another. The limit should be adjusted using observed waiting time, unfinished work, and operating demands. An urgent request that exceeds capacity must displace, delay, or reduce something else. This makes the cost of urgency visible to the person requesting it.
Sequencing also has informational value. A short demand test may determine whether a much larger integration is worth building. An operational pilot may reveal whether expansion requires another hire or a different process. Projects that resolve expensive uncertainties early can deserve priority even if they do not immediately increase revenue. The portfolio record should state which later decision depends on the learning, so that the experiment has a concrete strategic purpose.

5.3 Defining stop and continuation criteria

Before work begins, the team should specify evidence that would support continuation and evidence that would challenge the investment. These criteria may include customer behavior, delivery reliability, cost exposure, or a deadline for obtaining an external commitment. They should not be interpreted mechanically when measurement is weak. If a pilot result is inconclusive, the decision-maker must weigh the cost and value of another test rather than silently declaring success or failure.
Stopping a project should release its people, budget, and related obligations. A nominally paused initiative can continue consuming attention through maintenance, vendor contracts, or customer expectations. Closure therefore includes communicating the decision, preserving useful learning, and assigning any remaining responsibilities. This makes termination a managed portfolio decision. It also reduces the pressure to keep weak projects active merely because no one has planned how to end them.

6 Planning and Delivering Under Uncertainty

6.1 Choosing the delivery approach

The appropriate delivery method depends on the stability of requirements, reversibility of changes, external dependencies, and consequences of failure. Iterative approaches suit work where frequent feedback can improve the solution. More predictive planning is useful when a sequence is constrained by physical installation, procurement, or an externally fixed event. A hybrid approach can place iterative product development inside a broader plan for contracts, capacity, and launch readiness. The project manager should explain the choice in relation to the work.
Where a startup uses Scrum, it should respect the framework's defined accountabilities rather than adding a project manager who informally overrides team decisions. The project manager may coordinate external commitments and portfolio dependencies while product priorities remain with the Product Owner and the team manages its delivery work. A company that selects only a few Scrum practices should describe its approach accurately. The name of the method is less useful than clarity about who can decide what and how feedback changes the plan.

6.2 Using rolling planning and explicit scope

Rolling-wave planning specifies near-term work in more detail while leaving distant work at a level consistent with current knowledge. A startup might commit to a short discovery or delivery interval and maintain a broader forecast for the quarter. The near-term plan includes acceptance criteria, dependencies, owners, and an estimate of available capacity. The longer horizon states intended outcomes and major assumptions. The exact intervals should match customer feedback cycles and operational constraints.

Scope should be described in terms of the behavior or capability the project must enable. An onboarding project could require a target customer to connect a supported data source, complete setup, and obtain a usable result without founder intervention. This is more testable than an instruction to improve the onboarding experience. The team should also record exclusions, such as unsupported integrations, so that commercial discussions do not quietly expand the delivery obligation.Acceptance criteria and a definition of done address different questions. Acceptance criteria determine whether a particular deliverable meets its intended requirements. A shared definition of done establishes the quality conditions that apply across relevant work, such as testing, documentation, monitoring, or support readiness. Neither establishes that the business benefit has occurred. That judgment requires observation after customers or staff use the change in practice.

6.3 Managing estimates and change

Early estimates should expose uncertainty rather than conceal it behind a single date. The team can describe a plausible completion range and the assumptions that would move work toward either end. As delivery data accumulates, forecasts should be updated using comparable work and current constraints. A range is not useful if its basis remains unexplained. Stakeholders need to know whether uncertainty comes from technical discovery, another team's availability, a supplier, or customer participation.When a change request arrives, the project owner assesses its effect on the intended outcome, schedule, cost, dependencies, and risk. Small reversible changes within the agreed envelope can remain with the team. Changes that alter the customer promise, require additional funding, or displace another initiative should go to the designated sponsor. The purpose is a timely trade-off, not a lengthy approval ritual. The decision and its consequences should be visible to everyone whose commitments are affected.

Delivery reviews should inspect working outputs and unresolved constraints. A demonstration can expose misunderstandings that a status report misses, while a short dependency discussion can resolve a blockage that the team cannot remove independently. Retrospectives should end with a small number of owned changes and a later check of whether those changes helped. Repeated discussion without follow-through adds meeting time without improving execution.

7 Connecting Projects to Financial Capacity

7.1 Funding the next justified commitment

Project funding should reflect the maturity of the evidence. Discovery may justify a limited test, repeatability may justify process investment, and expansion may justify greater capacity. Each funding decision should state what has been learned, what remains uncertain, and why the next commitment is proportionate. This prevents a preliminary concept budget from becoming an assumed entitlement to full implementation. It also makes finance a participant in planning rather than a reviewer brought in after the scope is fixed.

The project budget should separate cash expenditure from internal effort. Existing salaries may not create additional cash outflow when staff move between projects, but that effort still has an opportunity cost. Contractors, tooling, travel, recruitment, infrastructure, and supplier deposits can change near-term cash requirements. The financial view should also show recurring operating costs that begin after launch. A project can fit its development budget while introducing a service model the business cannot afford to sustain.

The baseline should identify committed spending, discretionary future spending, and contingency. Committed spending cannot always be removed simply by stopping a project. Contingency should correspond to identifiable uncertainty rather than function as hidden scope. When a risk consumes contingency, the team should update both the forecast and the decision record. If the remaining exposure exceeds available capacity, leadership needs to adjust scope, timing, or the portfolio.

7.2 Linking runway and milestones

Cash runway is a planning estimate of how long available cash can support expected net outflows. Dividing unrestricted cash by a stable monthly net burn gives a simplified estimate, but an actual startup forecast should reflect changing collections, payroll, supplier payments, and one-time obligations. A project milestone matters financially

when it changes those flows or produces evidence needed for another decision. Completing a prototype is not equivalent to collecting revenue or securing financing.For an illustrative venture with $900,000 in cash and monthly net burn of $100,000, the simplified runway is nine months. If a project adds an immediate $60,000 payment and $10,000 of monthly spending without new collections, the same approximation becomes $840,000 divided by $110,000, or about 7.6 months. This calculation omits timing variation and is not a financial forecast. It demonstrates why project authorization should consider the company's remaining decision time as well as the project's nominal budget.Funding assumptions should have explicit scenarios. A base case can reflect current expectations, while a downside case considers slower collections or delayed commercialization. Prospective investment should not be treated as available cash before it is secured. If the venture depends on a future financing event, the portfolio should identify which commitments can be reduced if that event occurs later than expected. This enables adjustment while options still exist.

7.3 Assessing the economics of expansion

Expansion projects should connect expected customer value to the cost of acquiring and serving those customers. Useful measures vary by business model, but may include contribution margin, acquisition payback, implementation effort, returns, and retention. Contribution margin must use an explicit definition of variable cost. Comparisons are misleading when one analysis includes support and infrastructure while another omits them. Early customer lifetime value estimates are especially uncertain when retention histories are short.The project manager should work with finance and commercial owners to model a range of plausible outcomes. The question is whether the decision remains reasonable under changed assumptions, not whether a spreadsheet produces an attractive point estimate. For AI products, usage-sensitive inference, human review, and rework can materially affect delivery economics. Customer growth can therefore increase both revenue and service exposure. The project plan should include the instrumentation needed to observe those costs after launch.

8 Organizing People and Stakeholders for Scale

8.1 Clarifying authority without creating bottlenecks

As a startup hires specialists, ambiguity can grow around decisions that founders previously made informally. The company should state who owns product priorities, approves additional expenditure, accepts operational readiness, and communicates external commitments. A responsibility matrix is useful when several functions contribute to the same deliverable, but it should remain short enough to guide actual decisions. Each important decision needs a clear accountable owner; assigning several people joint accountability can leave no one responsible for resolution.

Founder involvement should concentrate on choices that require strategic authority, such as changing the target market or accepting a material financing exposure. Routine decisions should move to the people with relevant information within agreed boundaries. Delegation requires access to that information and confidence that a reasonable decision will not be reversed without explanation. If every priority dispute returns to the founder, the project manager can document the pattern and propose a clearer decision rule.

Cross-functional teams can reduce handoff problems when the people needed to produce an outcome work together. However, a startup may not have enough specialists to dedicate them to each initiative. Shared experts need visible commitments and an agreed priority order. A project should not assume full availability merely because a person's name appears on its staffing plan. Capacity planning should account for operational responsibilities, leave, onboarding, and the time needed to support colleagues.

8.2 Managing communication and adoption

Stakeholder communication should support decisions rather than reproduce every task update. A project sponsor needs the intended outcome, forecast, major risks, and decisions requiring authority. A delivery team needs current priorities and dependencies. Customers need accurate expectations and a clear route for feedback. These audiences can use the same underlying evidence while receiving different levels of detail. This reduces the risk of parallel reports that disagree about project status.

Commercial commitments require particular discipline. Sales and partnership teams should know which capabilities are available, which are being tested, and which remain proposals. A commitment outside the approved scope should trigger a visible feasibility and capacity decision. Product and engineering teams should likewise understand the consequence of missing a promised date. Alignment becomes practical when both sides

can see the cost of a change before communicating it externally.Internal adoption deserves its own plan. A new workflow may require training, changes to incentives, data cleanup, or revised responsibilities. Launching the software does not complete that transition. The project owner should identify affected users, arrange a pilot, and collect evidence of actual use. The operations owner should accept the handover only when support responsibilities and unresolved issues are clear. Adoption reviews then determine whether the intended process improvement has occurred.

8.3 Building reliable reporting behavior

A startup needs reporting that makes uncertainty discussable. Teams should be able to identify a weak assumption, a slipping dependency, or a customer complaint before the issue becomes a crisis. Leadership can support this by distinguishing an unfavorable result from negligent work. A carefully designed experiment that rejects an idea may have protected the venture from a larger loss. Conversely, optimistic reporting that conceals known problems should not be rewarded because it temporarily preserves confidence.

The project manager can reinforce this behavior by asking for evidence and explaining how reports affect decisions. Risk status should reflect actual exposure, not the desired impression for an investor meeting. When estimates change, the discussion should examine the new information and available responses. Over time, a decision history can improve forecasting and help new employees understand why certain trade-offs were made.

9 Managing Risk Quality and Operational Readiness

9.1 Integrating risk into delivery

Risk management should address events that could prevent the project from achieving its outcome or create unacceptable harm. A useful risk record states the cause, possible event, consequence, owner, response, and trigger for action. An assumption is something the plan currently treats as true; an issue is a condition that has already occurred. Keeping these distinctions clear helps teams decide whether they need more evidence, preventive work, or immediate recovery.

The record should remain focused on exposures that affect decisions. For example, a partner's uncertain integration date may require a fallback launch plan, while concentration in one customer may affect project funding. Likelihood and impact categories can help prioritize attention, but multiplying subjective scores should not be mistaken for a precise expected loss. Some low-frequency events deserve special treatment because their consequences are severe or difficult to reverse. Quality requirements should be established before testing begins. They may address reliability, accessibility, data integrity, usability, and service performance. The team should state who can accept residual defects and which defects prevent release. This authority should correspond to the consequence of the decision. A marketing deadline alone does not determine whether a high-impact defect is acceptable, and a project manager should not assume authority held by a qualified specialist or accountable executive.

9.2 Addressing AI and data dependencies

AI-enabled startups require particular attention to the system surrounding a model. Sculley et al. (2015) identify sources of technical debt in machine learning systems, including data dependencies and interactions that can increase maintenance complexity. The implication for project planning is to account for monitoring, testing, integration, and upkeep alongside model development. A prototype that performs well on a demonstration set may still lack the infrastructure and controls needed for dependable customer use.The NIST AI Risk Management Framework organizes risk activities around Govern, Map, Measure, and Manage and treats governance as cross-cutting (Tabassi, 2023). This paper uses that voluntary framework as a reference for structuring AI project questions, not as a certification or a claim of legal compliance. Startup-specific release criteria might examine intended use, evaluation coverage, foreseeable misuse, human oversight, and response arrangements. Those criteria require domain expertise and must reflect the consequences of the actual application.For an AI feature, the project should establish which outputs require human review and how incorrect results are detected and handled. Evaluation data should represent the intended use sufficiently to support the release decision, with known limitations recorded. Vendor changes, model updates, and changing input data should trigger appropriate reassessment. The project should also identify who owns incidents after handover; otherwise, a delivery milestone can leave an operational risk without an accountable owner.

9.3 Preparing for operation and recovery

Readiness reviews should ask whether the business can support the change after release. Evidence may

include tested monitoring, trained staff, a support process, realistic capacity, and recovery arrangements. For changes that can be reversed, the team should define how to roll back and what information must be preserved. Where reversal is difficult, the pilot and approval process should be more demanding. This connects delivery method to exposure rather than applying an identical checklist to all projects.

Operational handover should identify remaining defects, maintenance needs, vendor obligations, and benefit review dates. The receiving owner needs sufficient documentation to operate the change without continuous dependence on the original project team. An early post-launch review can test whether demand, cost, and support assumptions remain valid. Its purpose is to resolve emerging problems before the startup increases volume further.

10 Measuring Project Contributions to Scaling

10.1 Distinguishing delivery from business results

A startup should assess project performance at several levels. Outputs describe what was delivered. Outcomes describe how customers or staff behave differently because of the change. Benefits describe the value the venture expects from those outcomes. An automated setup tool is an output, faster activation is an outcome, and reduced service effort or improved retention may be benefits. The relationship among them is a hypothesis that needs evidence, not an automatic sequence.

Delivery measures remain useful. Schedule variation, blocked work, defects, and budget forecasts help teams coordinate commitments. However, counting completed tasks does not establish customer value, and comparing teams by story points can create misleading incentives because estimation scales differ. The startup should select a small set of measures linked to the project's purpose and interpret them with context. A rise in reported defects may reflect improved detection rather than deteriorating quality.

10.2 Establishing measurement discipline

Each metric needs a definition, denominator, observation window, data source, and accountable owner. Activation might mean connecting an account in one business and completing a meaningful transaction in another. Without a stable definition, the team cannot tell whether performance changed or the reporting method changed. Cohort analysis can help distinguish newly acquired customers from established customers and expose whether a change benefits the intended segment.

Baseline data should be gathered before the intervention where feasible. If several changes occur together, the benefits review should acknowledge that attributing the result to one project is difficult. Comparison groups, staggered rollouts, or carefully designed experiments can strengthen interpretation when appropriate. Qualitative evidence can explain a pattern but should not be presented as a numerical estimate of impact. A useful review states both the observed change and the limits of the inference.

Targets also need guardrails. Reducing onboarding time is not a success if it produces incorrect configurations or shifts work onto customers. Increasing conversion may be undesirable if it attracts customers the venture cannot retain economically. The benefits owner should therefore identify a primary outcome alongside relevant quality and economic measures. Management can then evaluate trade-offs explicitly rather than reward a single number that encourages local optimization.

10.3 Evaluating the cost of project management

Project management itself consumes scarce resources. The startup should periodically examine meeting time, reporting effort, decision delays, and duplicated documentation. These costs should be considered alongside practical improvements such as fewer unresolved dependencies or faster correction of weak investments. A new template is justified when it supports a recurring decision more reliably than the previous approach. It should not be retained simply because a mature organization uses something similar.

A small company can begin with a shared portfolio view, concise project charters, a decision log, and a risk record. Additional tools become useful when the existing arrangement no longer supports coordination. The software chosen should fit the workflow and ownership model. Purchasing a project platform before agreeing on decision rights can automate ambiguity rather than resolve it.

11 A Hypothetical Application in a Software Startup

11.1 Defining the problem and alternative investments

Consider a hypothetical startup that sells an AI-assisted reporting service to business customers. All

quantities in this example are invented for illustration. The company has 40 paying accounts at $1,000 per month, yielding $40,000 in monthly recurring revenue before adjustments. Its founders are considering a larger acquisition campaign. However, implementation staff report a growing setup queue and frequent manual corrections. The project manager frames the decision as whether the business can absorb more customers reliably before increasing acquisition expenditure.

A preliminary capacity calculation estimates 80 implementation hours available per month and an average of eight staff hours per new account. The implied onboarding capacity is ten accounts per month, assuming those hours are usable and effort remains stable. The proposed campaign is expected, with substantial uncertainty, to produce 15 new accounts monthly. If both assumptions held, the queue would grow by five accounts each month. This does not prove that the campaign should be rejected, but it identifies a dependency the investment decision must address.

The team compares three responses: hire additional implementation capacity, simplify onboarding, or reduce the campaign's initial scale. Hiring may increase capacity more directly but adds recurring expenditure and training time. Automation may reduce effort but depends on technical feasibility and customer behavior. A smaller campaign limits immediate exposure while preserving learning. Leadership authorizes a bounded onboarding project and a limited acquisition test, reserving the larger campaign decision until capacity evidence is available.

11.2 Designing the project and its review

The project charter defines the intended outcome as reducing staff effort for a clearly specified customer segment while maintaining setup quality. Its scope covers one common data source and the associated configuration workflow. A product lead owns the customer experience, an engineering lead owns technical delivery, and the customer success lead owns the post-launch benefit review. The project manager coordinates dependencies and maintains the forecast. The founder retains authority over additional cash expenditure and changes to the commercial commitment.

The team first observes setup sessions and separates avoidable product friction from necessary customer-specific work. It then tests a simplified workflow with a small pilot cohort. Acceptance criteria cover successful data import, accurate configuration, and a documented recovery path. The pilot records staff effort, elapsed activation time, correction work, and customer feedback. These observations are reviewed together because a reduction in visible setup effort could otherwise hide increased support work after activation.

Suppose the pilot records an average of four staff hours per account with no observed increase in corrections. Under the same simple capacity assumption, 80 available hours would support 20 accounts monthly. The team should still avoid treating that figure as a proven operating limit. Pilot customers may have cleaner data, staff may provide unusual attention, and the sample may be too small to reveal infrequent problems. A second staged rollout can test whether the improvement persists under more representative conditions.

11.3 Connecting the outcome to the economic decision

Assume the onboarding project requires $12,000 in incremental cash expenditure and 120 hours of existing staff time. These amounts should remain separate in the decision record. At a planning value of $50 per hour, saving four hours across 15 monthly onboardings would represent $3,000 of monthly capacity value. Dividing $12,000 by $3,000 gives a four-month simple payback on that capacity valuation. It is not a cash payback unless the released capacity reduces actual expenditure or enables incremental contribution that would otherwise be unavailable.

The benefit may instead take the form of shorter queues, greater resilience, or avoiding a future hire. Finance should model the mechanism the company actually expects and avoid counting the same saved hours both as cost reduction and as capacity for additional customers. Likewise, the existing $40,000 in recurring revenue should not be attributed to the onboarding project. Only an appropriately supported incremental effect belongs in its benefit assessment.

The final expansion decision considers repeatability, customer quality, cash exposure, and operating readiness together. The company may increase the campaign gradually, retain a temporary manual fallback, or delay expansion if the second rollout exposes substantial rework. This case illustrates the proposed framework's central mechanism: project management makes the assumptions connecting acquisition, delivery capacity, and finance inspectable before the company increases its commitments. It does not establish that automation is always preferable to hiring or that the hypothetical results would occur elsewhere.

12 Implementing the Framework Over Ninety Days

12.1 The first thirty days

A startup can introduce the framework without establishing a large project management office. During the first month, leadership should inventory active initiatives and identify their owners, expected outcomes, major commitments, and next decisions. This inventory often reveals projects that compete for the same people or continue without a clear business rationale. The objective is a usable portfolio view, not a complete historical reconstruction of every task the company has performed.The team should select one scaling constraint and establish its current baseline. It should also agree on decision rights and a concise charter format for a small number of important projects. Existing meetings can be adapted to review decisions and dependencies, avoiding a parallel reporting structure. The person coordinating this work needs access to leadership and operational information, whether their title is project manager, operations lead, or founder.

12.2 Days thirty one through sixty

The second month should test the approach on a limited set of projects. Teams define continuation criteria, make resource assumptions explicit, and record changes to scope or funding. The project manager examines where work waits and whether unresolved dependencies require sponsor action. Finance participates in the review of recurring cost and cash implications. The benefits owner establishes how outcomes will be measured after delivery.This period should also test the administrative burden. If a report repeats information already available in the work system, it can be removed or automated. If an approval adds delay without changing the decision, authority can move closer to the team. Feedback should come from the people doing the work as well as leadership. A framework that looks clear to executives but requires excessive manual reconciliation is unlikely to remain useful.

12.3 Days sixty one through ninety

By the third month, the startup should have enough experience to compare the intended process with actual use. It can examine whether priorities are clearer, dependencies are resolved earlier, and benefit reviews influence investment decisions. It should retain practices that support those outcomes and revise those that do not. Ninety days is an implementation horizon, not a scientifically established threshold for organizational maturity or a guarantee of measurable financial improvement.
Expansion of the project management function should follow demonstrated need. Multiple interdependent initiatives may justify program coordination, while a larger portfolio may require a dedicated portfolio review. A formal project management office becomes useful only if its services address recurring problems at an acceptable cost. The objective is a repeatable way to make and execute growth decisions, with enough flexibility to respond when the business changes.

13 Research Propositions and Evaluation Design

13.1 Propositions arising from the framework

The first proposition is that the usefulness of project governance depends on its fit with uncertainty and coordination requirements. More formal planning may improve execution when dependencies and commitments are substantial, while imposing unnecessary burden on a small exploratory task. The predicted relationship is therefore contingent rather than uniformly positive. Researchers would need to measure the actual practices and their use, not simply whether a company employs a project manager.The second proposition is that explicit learning criteria improve resource allocation by making continuation and termination decisions more responsive to evidence. The proposed mechanism is better selection of subsequent commitments, rather than more experiments completed. The third proposition is that benefits ownership helps connect delivered outputs with realized operational or commercial value. The fourth is that disclosure and decision behavior moderate these effects: accurate project information is less useful when leaders do not act on it or when staff cannot safely raise concerns.These propositions are not findings. They identify relationships that could explain why similar tools produce different results across startups. Alternative explanations include founder experience, financing conditions, market demand, technical quality, and investor involvement. Stronger ventures may also adopt management practices earlier, creating reverse causality. An empirical study must address these possibilities before claiming that project management caused superior scaling performance.

13.2 A proposed longitudinal study

A future study could follow startups over multiple periods, recording the introduction and actual use

of project practices alongside customer, operating, and financial outcomes. Sampling should vary by sector, organizational size, and business-model maturity. The unit of analysis could combine projects nested within ventures, allowing researchers to separate project-level practices from company-level conditions. Sample size should be justified through a power analysis linked to the proposed model rather than selected for convenience.Data collection should combine project records, operational measures, and interviews with more than one function where possible. A founder's assessment of alignment should not be the sole source for both the explanatory variable and the performance outcome. Researchers should document metric definitions, changes in measurement, and the availability of historical baselines. Firms that close or abandon initiatives should remain part of the analysis where feasible, because studying only survivors would distort the evaluation of scaling decisions.

Analysis could examine changes within firms while accounting for stable differences between them and common time effects. A staggered introduction of practices might support stronger comparisons if its assumptions are defensible. Even then, time-varying confounding and selective adoption require attention. Qualitative process evidence could help determine whether better coordination actually explains an observed change. Reported uncertainty, robustness checks, and transparent model specifications would be necessary for credible interpretation.Missing data should be investigated rather than routinely replaced with overall means. The appropriate response depends on why values are missing and on the analysis being performed. Construct reliability should also be distinguished from validity: consistent survey items do not guarantee that a measure captures effective governance. Researchers should avoid combining unrelated indicators into a single maturity score unless its meaning and measurement properties are justified.

Table 1 *Governance priorities across scaling contexts*

Context	Primary decision	Minimum useful evidence	Typical project controls
Discovery	Whether to continue testing	Observed problem; customer response; cost of next test	Hypothesis record; spending limit; review date
Repeatable delivery	Whether delivery can be repeated	Comparable customer outcomes; process effort; recurring costs	Acceptance criteria; dependency ownership; operational handover
Expansion	Whether additional volume is supportable	Capacity forecast; cohort economics; service quality under load	Staged rollout; integrated schedule; contingency and benefits reviews

Note. Proposed framework. Contexts can coexist within one startup; these are not universal maturity thresholds.

Table 2 *Measures for evaluating a scaling project*

Level	Illustrative measure	Interpretive question
Delivery	Accepted output; completion forecast; unresolved defects	Can stakeholders rely on the delivery commitment?
Customer outcome	Activation rate; time to first useful result; cohort retention	Is the intended customer behavior changing?
Operating benefit	Staff effort per account; queue length; rework	Can the venture serve more customers reliably?
Economic benefit	Incremental contribution; service cost; cash effect	Does the change improve the venture’s economics?
Learning	Assumption tested; decision changed; remaining uncertainty	Did the evidence improve the next investment choice?
Guardrail	Error rate; complaints; customer effort	Did improvement create an unacceptable trade-off?

Note. Suggested measures, not validated scales. Select measures appropriate to the business model and define their denominators and observation windows.

14 Discussion and Limitations

The framework reframes strategic alignment as an ongoing set of project and investment decisions. Product design determines what can be delivered, market evidence indicates what customers value, and finance constrains the timing and scale of commitments. Project management connects these domains through accountable work and reviewable evidence. Its contribution is strongest when coordination problems would otherwise prevent information in one function from changing decisions in another.

Several limits remain. The underlying literature covers different populations, methods, and definitions of performance. Evidence from established organizations or broad project samples may not transfer directly to resource-constrained startups. The framework has not been tested as an integrated intervention, and its implementation recommendations represent reasoned design choices. The hypothetical case simplifies customer heterogeneity, cost behavior, and capacity variation to make the decision process visible.

The framework also depends on leadership behavior. A founder who overrides priorities without acknowledging consequences can undermine an otherwise clear process. Excessive measurement can divert attention from customers, while aggressive targets can encourage selective reporting. The appropriate response is periodic review of both outcomes and administrative burden. A startup should be willing to simplify its project management practices when the coordination value no longer justifies their cost.

15 Conclusion

Project management can support startup scaling by translating uncertain growth ambitions into bounded commitments, coordinated delivery, and observable benefits. The proposed framework connects project selection with evidence about demand, financial capacity, and operational readiness. It preserves the importance of product, finance, and market alignment while explaining how that alignment can be maintained through ownership, planning, risk decisions, and post-delivery review.

Start with a clear scaling constraint, a small portfolio of justified projects, and explicit decision criteria. Increase governance as dependencies and exposure grow. Although the framework requires empirical testing, it offers founders a practical structure for assessing whether project investments create the capacity to serve more customers reliably and sustainably.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Use of AI Tools:** The author(s) declare that no generative AI or AI-assisted technologies were used in the writing process of this manuscript.

References

1. Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science, 66*(2), 564–586. https://doi.org/10.1287/mnsc.2018.3249
2. Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., & Spina, C. (2024). A scientific approach to entrepreneurial decision-making: Large-scale replication and extension. *Strategic Management Journal, 45*(6), 1209–1237. https://doi.org/10.1002/smj.3580
3. Davila, A., & Foster, G. (2007). Management control systems in early-stage startup companies. *The Accounting Review, 82*(4), 907–937. https://doi.org/10.2308/accr.2007.82.4.907
4. Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44*(2), 350–383. https://doi.org/10.2307/2666999
5. Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science, 68*(9), 6434–6453. https://doi.org/10.1287/mnsc.2021.4209

6. Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–263. https://doi.org/10.5465/amr.2001.4378020
7. Schwaber, K., & Sutherland, J. (2020). *The Scrum Guide: The definitive guide to Scrum: The rules of the game.* https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-US.pdf
8. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems, 28*, 2503–2511. https://research.google/pubs/hidden-technical-debt-in-machine-learning-systems/
9. Serra, C. E. M., & Kunc, M. (2015). Benefits realisation management and its influence on project success and on the execution of business strategies. *International Journal of Project Management, 33*(1), 53–66. https://doi.org/10.1016/j.ijproman.2014.03.011
10. Serrador, P., & Pinto, J. K. (2015). Does Agile work? A quantitative analysis of agile project success. *International Journal of Project Management, 33*(5), 1040–1051. https://doi.org/10.1016/j.ijproman.2015.01.006
11. Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1
12. Whitaker, S. (2014). *The benefits of tailoring: Making a project management methodology fit* [White paper]. Project Management Institute. https://www.pmi.org/learning/library/tailoring-benefits-project-management-methodology-11133