

Project Controls in Transformational Projects

AN INTERNATIONAL HANDBOOK



Professor Amgad Badewi



KENT BUSINESS COLLEGE

School of Project Control

Project Controls in Transformational Projects

An international handbook

Professor Amgad Badewi

Copyright © 2026 Professor Amgad Badewi

First edition 2026

Published by Kent Business College

School of Project Control

kentbusinesscollege.com

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means without the publisher's prior written permission, except for brief quotations with appropriate acknowledgement and uses permitted by law.

The documented cases and academic sources are acknowledged within the relevant chapters. Worked teaching examples are distinguished from observed project results.

Contents

Part I Foundations and governance	1
1 The Foundations of Transformational Control	3
2 Principles for Credible Control	17
3 Connecting Governance to the Work	35
Part II Designing the control system	51
4 Authority Ownership and the Human System	53
5 Designing Change That Can Be Measured	69
6 The Economics of Continuing	89
7 Sequencing Change at an Absorbable Pace	107
8 Turning Information into Organisational Learning	131
9 Confidence That Survives Challenge	151
10 Decisions under Uncertainty	177

Contents continued

Part III Applying controls through the transformation lifecycle	205
11 Control as a Repeated Learning Cycle	213
12 From Strategic Ambition to an Investigable Proposal	225
13 Specifying the Future Operating State	237
14 Committing the Next Wave of Investment	249
15 Integrating Delivery without Losing the Purpose	263
16 Adoption Readiness and Operational Stability	277
17 Reforecasting from New Evidence	291
18 Ending the Programme while Protecting Its Value	305
 About the author	 318
 How to cite this book	 319

Using this handbook

Transformational projects ask an organisation to change what it can do, how it works and the value it creates. Their control cannot end with an account of delivered scope, elapsed time or expenditure. It must connect the investment commitment to the operating conditions, behaviour and evidence that make improvement possible. This handbook develops that connection as a practical discipline of measurement, forecasting, professional challenge and authorised action.

The three parts build one coherent argument. Part I establishes the foundations of transformational control and connects its principles to governance. Part II develops the human, design, economic, planning, information, assurance and uncertainty disciplines that make a control system useful. Part III applies those disciplines through a transformation lifecycle, from an investigable strategic proposal to the continuing ownership of value after the programme ends. The lifecycle can repeat and overlap; it is a structure for consequential decisions rather than a promise that change will unfold in a fixed sequence.

The chapters combine academic research, documented cases and worked examples. Source findings retain their context and limitations. The book's original frameworks provide practical ways to organise questions, evidence and responsibilities; their usefulness depends on informed application. Numerical illustrations show their assumptions so that practitioners can adapt the reasoning to their own data, authority and operating conditions. Each chapter's references provide routes back to the underlying research and cases.

Readers can follow the argument from beginning to end or enter through the discipline and lifecycle decision that matters to their current work. The repeated framework names, shared lifecycle and connected teaching examples support movement between chapters. In practice, the most useful reading is active: identify the decision at stake, establish the evidence available, challenge the assumptions, clarify who can act and determine what subsequent observation will show. The purpose is an organisation that can explain and honour its commitments while learning how to create lasting value.

PART I

Foundations and governance

The purpose, principles and authority that make transformational control credible.



KENT BUSINESS COLLEGE

School of Project Control

CHAPTER 01

The Foundations of Transformational Control



IN THIS CHAPTER

Connecting delivery, organisational change and lasting value.

The business must change as well as the project

A transformation earns its name through what becomes different in everyday life. A passenger reaches work more reliably, a patient receives a coordinated service, or an employee resolves a request without entering the same information three times. These changes depend on delivered assets, systems and processes, but also on people using them under ordinary operating conditions. For the project controls professional, this creates a demanding responsibility. A reliable account of expenditure and completed work remains essential, yet it cannot establish that the organisation is becoming capable of producing the intended improvement. This chapter develops a wider purpose for controls: to connect investment promises, delivery evidence and operational results closely enough that responsible people can make timely decisions. It introduces the Transformation Control Loop, a proposed practice framework for examining those connections, and applies it to published UK cases and a fictional service transformation. The argument is practical: control becomes useful when it changes a decision, clarifies a responsibility or tests whether an earlier intervention worked.

Crossrail provides a revealing starting point because its delivery history and its subsequent use require different judgments. The National Audit Office found that reporting before the delayed opening emphasised achieved progress more than the remaining delivery risks. Its 2019 examination also identified a handover schedule assembled from largely separate plans, without a sufficiently detailed integrated sequence of outstanding work and dependencies (NAO, 2019, paragraphs 2.22–2.27). These findings describe more than a problem with presentation. An account of completed construction cannot, by itself, establish when an integrated railway will be ready to operate. Later evidence adds another perspective: Transport for London reported more than 350 million Elizabeth line journeys by its second anniversary in May 2024 (TfL, 2024). That operator statistic indicates substantial use; it is neither a count of unique passengers nor proof of the scheme's total economic benefit. Serious weaknesses in delivery control and substantial subsequent use can coexist. A professional account must preserve both observations.

The book's subject is therefore the connection between delivering change and making change valuable. Its foundations include the programme perspective associated with Managing Successful Programmes, commonly called MSP (AXELOS, 2020), alongside research on project success, benefits management and feedback. The intention is an independent account of project controls, with its own questions, examples and proposed models. Readers may work on a single organisational project, a coordinated programme or a portfolio containing several transformations. The central discipline travels across these settings, although authority and accountability differ. At every level, someone must explain what is expected to improve, what work and adoption are necessary, what the evidence currently supports, and what should happen next. The following pages establish that discipline before later chapters examine governance, organisation, design, justification and delivery in greater depth. They also establish an evidence standard: published observations, interpretation, proposed tools and fictional teaching examples will remain distinguishable throughout the discussion.

What makes transformation difficult to control

A transformational project changes the way an organisation creates or provides value. Replacing an ageing server may mainly preserve an existing service; introducing a shared service can alter responsibilities, information flows, customer access and the meaning of competent performance. The distinction depends on the extent and consequences of the change, rather than its publicity or budget. A relatively small intervention can transform a local service, while an expensive asset replacement may leave operating practices largely intact. Some transformations can be governed as projects with a clear operational owner. Others require a programme because multiple projects and operational activities must remain coordinated as the desired future develops. A programme is consequently more than a large project, and transformation does not automatically require a new organisational layer. The useful question is where interdependent decisions need to be made together. If separate delivery teams can each meet their contracts while the overall service remains unusable, governance must address the connections between them.

Those connections extend beyond technical interfaces. A new booking system may depend on agreed eligibility rules, accurate records, trained staff, an accessible telephone alternative and the capacity to handle exceptional cases. Each element may have a different owner and a different planning rhythm. Software can be released iteratively while building modifications require a fixed sequence and operating teams must maintain a daily service. The integrated plan must make these relationships visible without pretending that every activity has equal certainty. Its purpose is to explain how an acceptable operating state can be reached and sustained. This shifts the scheduling conversation from the date on which a component is supplied to the conditions under which the whole service can work. A milestone becomes stronger when it describes those conditions, identifies who will accept them, and states the evidence required. Calling a milestone “ready” adds little unless readiness has an agreed operational meaning.

Transformation also contains uncertainty about the intervention itself. Teams may understand how to install a system while remaining uncertain about whether customers will use it or whether managers will release the capacity it creates. These are different questions requiring different evidence. An estimate of installation effort cannot resolve uncertainty about behaviour, and a positive staff survey cannot establish technical resilience. The original investment proposal should therefore be treated as a set of connected, revisable claims. Some concern delivery feasibility, others concern adoption, and others concern the relationship between changed behaviour and value. Treating every emerging difficulty as poor execution encourages teams to defend a weak design. Treating every difficulty as unavoidable discovery removes accountability. Controls should make the distinction discussable: what has departed from an agreed commitment, what assumption has proved unreliable, and who is authorised to alter the response? This is how disciplined adaptation can coexist with commitments that mean something.

Research helps explain why completion alone is an incomplete standard. Atkinson’s conceptual discussion of information systems projects broadened evaluation beyond cost, time and quality to include consequences after delivery and stakeholder perspectives (Atkinson, 1999). Shenhar and colleagues identified multiple dimensions of success, including efficiency, customer impact,

organisational results and preparation for the future, with their importance changing over time and with technological uncertainty (Shenhar et al., 2001). These arguments do not make a missed budget irrelevant. They show why different judgments must remain visible instead of being compressed into a convenient verdict. A transformation can meet an approved delivery budget and still impose unacceptable operating costs. It can exceed that budget while creating useful capability whose value deserves separate examination. Responsible governance asks both questions, records the basis of each answer and avoids using one favourable result to erase another unfavourable result. Controls provide the evidence that makes such balanced judgments possible.

Project control as a discipline of evidence and action

The scientific ambition of project controls begins with making claims testable. A forecast should identify its assumptions, the observations that could challenge it and the consequences of being wrong. A performance measure should have a stable definition, an accountable source and a known limitation. An intervention should have an expected effect and a date when that effect will be examined. This does not turn organisations into laboratory systems. It brings disciplined reasoning to situations in which evidence is incomplete and interests may conflict. In engineering, feedback connects measurement with intervention, and its effectiveness depends on matters such as delay, noise and system behaviour (Åström and Murray, 2008). The organisational application here is an analogy, not a mathematical proof. People can change their objectives, dispute authority and respond strategically to measurement. Governance must therefore establish whose objectives count, which constraints cannot be traded away, and who may act when information changes.

Consider the difference between an approved target and a current forecast. A target expresses an authorised intention or constraint; a forecast expresses the best present judgment about what is likely to happen. Pressure to keep them identical damages the information needed to achieve either. A completion date can remain the target while the forecast moves later, prompting an explicit decision about scope, resources, sequence or expectations. Similarly, a baseline preserves the reference against which approved changes and performance are interpreted. Replacing it silently whenever conditions worsen removes the possibility of learning. Keeping it unchanged while suppressing the current forecast is equally unhelpful. A useful report allows the reader to distinguish the original commitment, authorised revisions, present expectation and reasons for movement. Those distinctions are ordinary professional hygiene, but they have ethical force: decision makers should be able to see what has changed without decoding a reassuring colour or reconstructing an old presentation.

Cost and schedule techniques remain central to this account. A credible schedule explains dependencies and the constraints that determine completion; cost control connects authorised scope, commitments, actual expenditure and the estimate of remaining expenditure. Earned value measures can describe progress against an agreed delivery baseline where the work and its measurement are suitably defined. Their interpretation must remain faithful to what they measure. A favourable delivery index does not demonstrate service adoption or a realised benefit. Conversely, promising adoption evidence does not excuse a delivery forecast built on

unverified remaining effort. These streams should inform one another while retaining their definitions. A delayed interface may reduce the period during which a benefit can accrue; a changed service design may alter both operational value and the delivery scope needed to obtain it. The controls professional makes these consequences traceable, so that a decision about one part of the investment does not quietly undermine another.

Forecast quality also requires challenge from outside the immediate delivery narrative. Flyvbjerg, Holm and Buhl's historical study of 258 transport projects found systematic cost underestimation in its sample (2002). Its findings justify examining forecast bias; they do not provide a universal allowance for every transformation or establish the motives of a particular team. Reference class forecasting offers a related discipline: compare the proposal with outcomes from sufficiently similar completed undertakings, while explaining the relevance and limits of the comparison (Flyvbjerg, 2006). For an organisational transformation, useful comparisons may concern migration effort, adoption time or operational disruption, rather than total construction cost. A database is only helpful when definitions, boundaries and context are understood. Internal estimates and external comparisons should be reconciled through an explicit explanation. Where they disagree, the gap becomes a matter for investigation, rather than a choice between whichever number best supports the desired approval.

Uncertainty should be expressed in a form suited to the decision. Where sufficiently relevant data and a defensible model exist, a probability distribution can describe a forecast more honestly than a single number. Where those conditions are absent, explicit scenarios may be more useful than an invented percentage of confidence. A sponsor could examine what happens if adoption develops slowly, if specialist staff become unavailable or if the new service must run alongside the old one for longer. Each scenario should preserve plausible relationships between events: a delay may increase operating overlap costs while postponing benefits, so those effects should not be evaluated in isolation. The aim is to discover which assumptions change the preferred decision and which merely alter its margin. A decision may remain sensible across a wide range of outcomes, or depend critically on an assumption that nobody has tested. Controls should communicate that distinction plainly. Precision becomes useful when it reflects the quality of the evidence; additional decimal places cannot compensate for an undefined population, an omitted dependency or an unsupported behavioural assumption.

The Transformation Control Loop

The Transformation Control Loop proposed in this chapter organises five connected questions: what value justifies the change; what must be delivered and adopted; what evidence deserves confidence; what decision is now required; and whether the resulting action helped. Figure 1.1 presents the loop. Its originality lies in this particular arrangement for professional practice, not in a claim to have discovered feedback, governance or benefits management. It is a synthesis of established ideas, developed here as a tool for structuring inquiry. It has not been empirically validated as a predictor of project success. A reader should use it to expose missing connections, document assumptions and improve the quality of discussion. It is deliberately possible to enter at any point: an unexpected operating result may require revisiting the original value claim, while a funding decision may expose an adoption dependency that delivery reports have omitted. Each movement around the loop should produce a clearer decision or a better specified uncertainty.

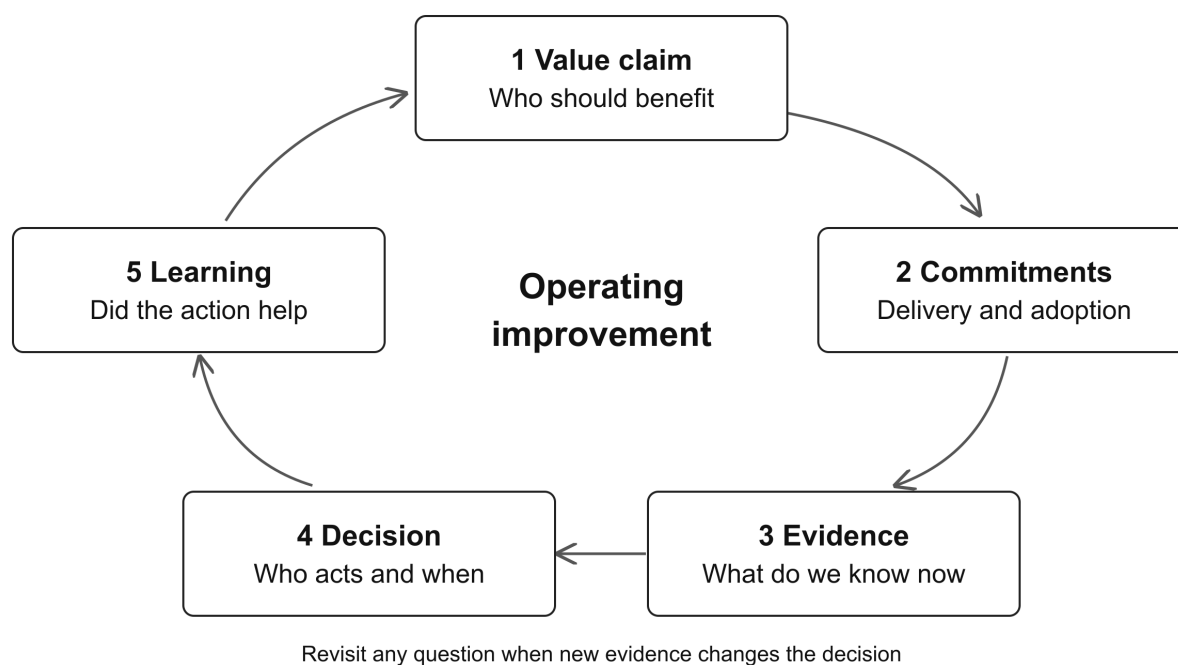


Figure 1.1 The Transformation Control Loop. An original proposed practice framework linking five questions. New evidence may require returning to any question. The loop has not been empirically validated.

The first question establishes a value claim that can survive examination. “Improve efficiency” is an aspiration, whereas reducing avoidable handling time for a defined service population within a specified period is a claim that can be investigated. The claim should identify the present position, the desired change, the affected people, the time horizon and the owner who will judge the result. It should also record possible harms and the conditions under which improvement would cease to be worthwhile. Faster processing that increases incorrect decisions may reduce rather than improve service value. Distribution matters as well: an average improvement can conceal a worse experience for people whose needs are harder to serve. The owner must therefore decide which measures belong together and which thresholds protect acceptable

performance. This is the point where strategy becomes sufficiently precise for controls to help, without pretending that every public or organisational purpose can be expressed as cash.

The second question translates the claim into commitments across the path from output to benefit. An output is something delivered, such as an installed system. Capability is the ability to use the relevant combination of assets, skills, information and arrangements. An outcome is the changed operating condition or behaviour that follows. A benefit is a valued improvement attributable, with appropriate caution, to that change. These distinctions are useful because each transition needs evidence. Installation requires acceptance evidence; capability requires realistic demonstrations; adoption requires observations of use; benefit assessment requires comparison with a defined baseline and attention to other influences. Badewi's study, analysing 200 valid responses through structural equation modelling, supports considering project management and benefits management together when examining success (Badewi, 2016). Its associations do not prove that adding a particular process causes improvement. The practical application proposed here is to connect delivery commitments with named operational responsibilities before either becomes difficult to change.

The third question concerns the strength and relevance of evidence. A report should distinguish an observation from an estimate, an assumption and a forecast. An observed training attendance rate is not an observed competence rate; a forecast transaction volume is not actual demand. Confidence grows when the measure matches the claim, its source is reproducible, its coverage is understood and someone independent enough can challenge its interpretation. This does not mean gathering everything. Measurement consumes effort and can distort behaviour when the easiest count replaces the intended result. The fourth question therefore connects evidence to a decision with a real owner and a useful deadline. A variance requiring sponsor authority should not circulate indefinitely among analysts who cannot alter the commitment. Nor should every minor fluctuation reach the board. Locally agreed tolerances, escalation routes and response times make the distinction workable. Their adequacy depends on the consequences and reversibility of the decision, not simply on a universal percentage threshold.

The final question closes the loop by examining the intervention. If extra staff were assigned to clear migration exceptions, did the backlog fall, were new errors introduced, and can the operating team sustain the result? A decision record should preserve the expected effect, responsible owner and review date so that later reporting can answer those questions. Feedback takes time, and repeated intervention before earlier actions have worked through the system may create confusion. Sterman's inventory simulation research showed how decision makers can misunderstand delayed feedback (1989); applying that insight to transformation is an explanatory analogy rather than a field estimate of project failure. The practical lesson is to examine response delays and accumulated commitments before reacting again. Learning also needs to travel beyond the temporary team. Williams's research on organisational learning from projects highlights difficulties in carrying lessons across project boundaries (2008). A useful control loop therefore changes assumptions, standards or future decisions instead of ending with an archived report.

Putting the loop to work requires a small amount of durable documentation. For each material value claim, the team should be able to locate the relevant delivery commitments, operational owner, measurement definition, latest evidence and next authorised decision. These connections can sit in existing systems; creating a separate register has little merit if it duplicates information and soon becomes inconsistent. The practical test is whether a new reviewer can follow a claim from its stated purpose to its present status without relying on private knowledge. Evidence also has a useful life. A demonstration performed before a major configuration change may no longer establish readiness, while an operating baseline may become misleading after a change in demand. The record should make those changes visible. Review effort can then concentrate on the claims whose importance, uncertainty or exposure has increased. This gives proportionality a concrete meaning: intensive scrutiny follows consequential decisions and fragile assumptions, while stable, low consequence matters receive lighter attention. The loop remains manageable because it organises inquiry around decisions rather than demanding the same reporting burden from every activity.

Reading UK experience through the loop

Returning to Crossrail, the proposed loop directs attention to the transition between reported component progress and evidence of integrated readiness. The NAO's 2021 update described improved indicators of outstanding work, more realistic station sequencing and attention to constrained specialist resources in the revised completion plan. It also found that sponsors lacked a clear overarching benefits management strategy or plan at the time of its examination, although work had begun on one (NAO, 2021, paragraphs 3.19 and 4.14–4.16). These are bounded findings from that reporting period. They suggest two distinct professional questions: whether the remaining delivery sequence is credible, and whether the arrangements for securing and assessing value are adequate. They do not establish that a single scheduling improvement caused recovery. Figure 1.2 uses these two questions to form a Delivery and Value Control Map. The map helps a review team locate its discussion, while the underlying evidence and local governance determine the actual response.

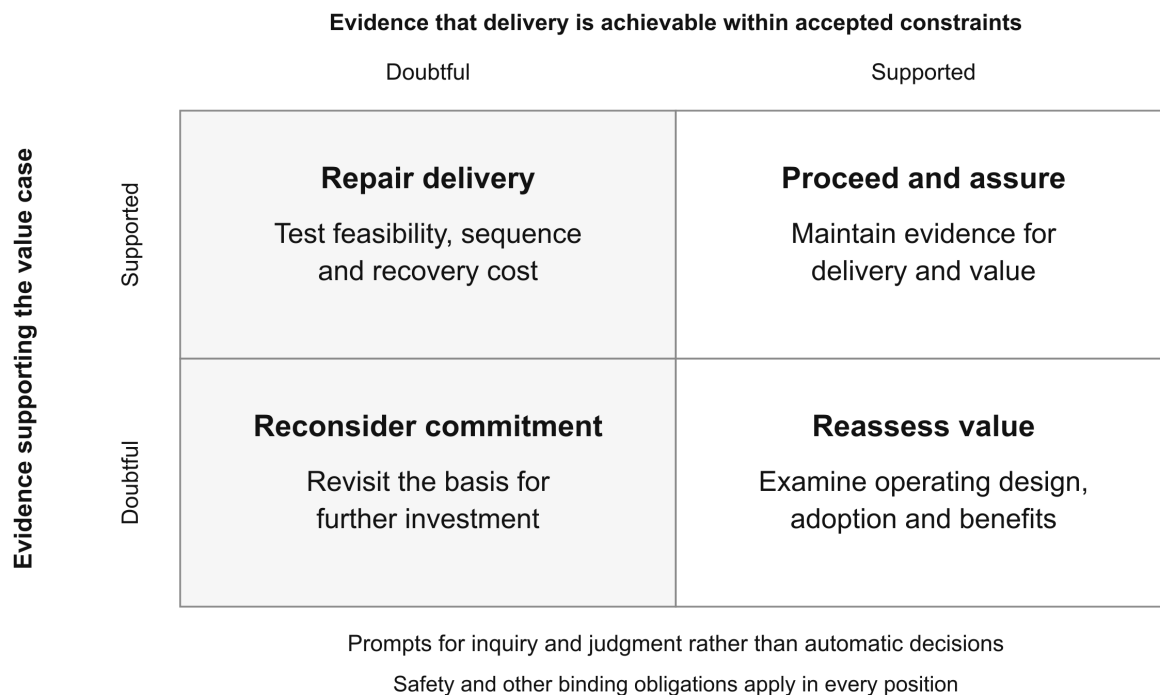


Figure 1.2 The Delivery and Value Control Map. An original decision aid. The delivery dimension concerns evidence of viability within accepted constraints; a credible forecast can still reveal unacceptable delivery.

London 2012 provides a different perspective. The NAO concluded that the Olympic Delivery Authority delivered venues and infrastructure on time and within budget. The same post Games review described continuing responsibilities for coordinating and assuring legacy delivery after the event, involving several organisations (NAO, 2012, paragraphs 3.5 and 2.12–2.13). The positive delivery judgment concerns that defined scope and budget; it should not be expanded into a claim that every original estimate or every legacy ambition was achieved. For the controls professional, the case demonstrates the usefulness of preserving a boundary between successful delivery and continuing responsibility for value. Closure can be appropriate for a temporary delivery organisation while operational owners still have funded obligations. The control loop therefore needs an explicit transfer at that boundary: who will retain the baseline, collect evidence, respond to deteriorating performance and revisit unrealised benefits? Otherwise, the disappearance of the project team can be mistaken for the disappearance of the problem.

The NHS evidence brings measurement itself into view. In its 2013 review of the final benefits statement for programmes previously managed under the National Programme for IT in the NHS, the NAO reported that there had been no comprehensive initial benefits baseline. It examined the compilation approach, without validating individual benefit figures, and noted that 65 per cent of estimated benefits were forecast to arise after March 2012 (NAO, 2013, paragraphs 4, 7 and 45). A statement described as final could therefore still contain substantial forecasts. The professional lesson concerns the status of the evidence, rather than a blanket judgment about every system delivered. A controls report must keep realised results, estimates and future expectations distinguishable. Together, the three cases illustrate different weaknesses that the proposed loop can help readers investigate. They are not a comparative experiment, and

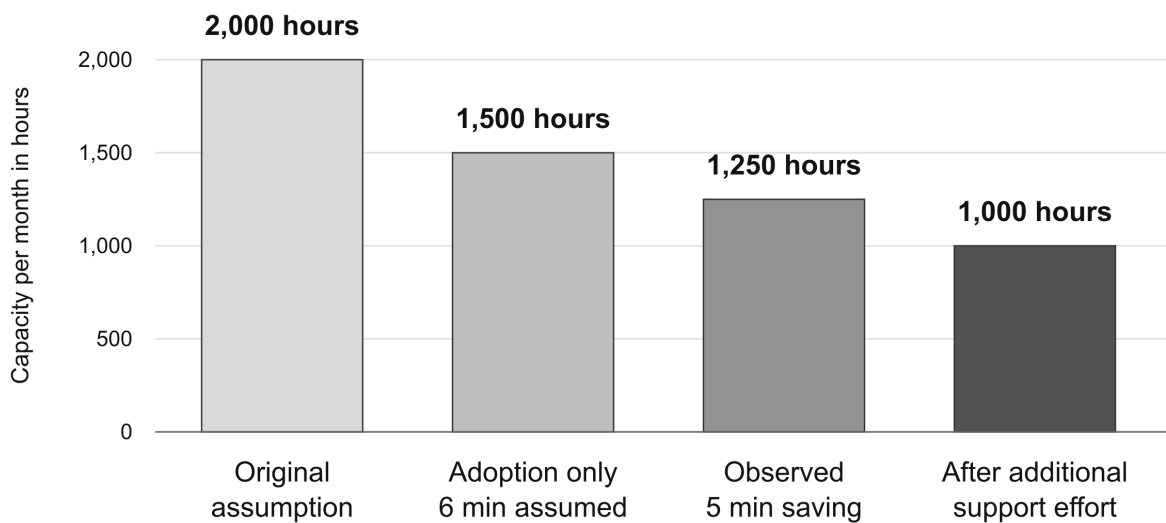
none validates the framework. Their value lies in making specific questions about integration, ownership and measurement harder to overlook.

A worked example from investment promise to operating result

Consider an explicitly fictional UK service organisation introducing a shared digital case management service. All figures in this example are invented for teaching. The investment proposal anticipates 20,000 eligible cases each month and a reduction in average staff handling time from 30 to 24 minutes. At full adoption, with comparable cases and unchanged quality, the arithmetic indicates 2,000 hours of monthly capacity released: 20,000 cases multiplied by six minutes, divided by 60. At an illustrative loaded staff cost of £30 an hour, that capacity has a nominal value of £60,000 a month. These numbers describe a conditional opportunity. They do not show that a payroll budget can fall by £60,000 or that all released time will be productively used. The initial value claim must state whether the organisation intends to reduce overtime, handle additional demand, improve service quality or change staffing expenditure. Each intention requires a different operating decision and a different measure of achievement.

The controls team first establishes a usable baseline. Handling time means active staff effort per case, rather than elapsed waiting time, and the team records how the measure treats repeat contacts, rework and exceptional cases. It separates relevant case types so that a change in complexity is not mistaken for productivity. A sample of current transactions is checked against source records, with its coverage and uncertainty documented. The operational owner agrees that complaint levels and decision accuracy will be reviewed alongside speed. The delivery plan then connects configuration and migration to a realistic demonstration of capability: staff must process representative cases, recover from an interrupted session and route exceptions correctly. A training attendance sheet provides supporting information but cannot satisfy these conditions by itself. The investment owner can now see what acceptance evidence means and which operational conditions must exist before a claim of improved performance deserves confidence. This preparation gives later numbers something dependable to stand on.

After introduction, suppose 15,000 of the 20,000 eligible monthly cases actually use the new process. Among comparable adopted cases, measured handling time averages 25 minutes. The apparent release is now 1,250 hours: 15,000 multiplied by five minutes, divided by 60. Its nominal capacity value is £37,500 at the same hourly rate. Suppose the operating team also records 250 additional hours of monthly support and rework caused by the transition, outside the measured handling time. On this deliberately simplified basis, the net opportunity falls to 1,000 hours, notionally £30,000, before any other relevant costs or effects. Figure 1.3 shows the conversion. The calculation is transparent, but transparency does not establish causation. Changes in demand, staffing mix or case complexity could influence the observation, and the comparison needs further examination. Most importantly, the finance team should recognise a cash saving only when evidence supports an actual reduction in relevant expenditure. Capacity used to reduce a queue can be valuable without becoming cash released.



Fictional data • Capacity value is not an observed cash saving

Figure 1.3 Converting an investment assumption into a qualified capacity estimate. All data are fictional. The 1,500 hour intermediate step holds the original six minute saving constant while applying 75 per cent adoption.

There is a further question before the organisation attributes the whole observed improvement to the project. What would have happened without the change? In the fictional service, handling time might also have fallen because an experienced temporary team joined or because simpler cases formed a larger share of demand. The operational analyst should examine these possibilities using evidence appropriate to the setting. A phased introduction could permit comparison between sufficiently similar groups, provided differences in timing, selection and operating conditions are considered. A time series may reveal whether an apparent improvement simply follows an existing trend. Neither comparison automatically establishes causality, and a complex evaluation may be disproportionate for a small decision. The essential discipline is to state what the available evidence can support and where attribution remains uncertain. The organisation should also avoid double counting: if the same released hours are used to process more cases, their full value cannot simultaneously be claimed as a reduction in staffing expenditure. Different benefits can coexist, but their relationship must be reconciled through the actual use of resources and the outcomes observed.

The Transformation Control Loop now makes the decision concrete. The original value claim remains useful, but adoption and support costs require attention. Investigation finds that exceptional cases are being routed outside the new process because a required approval step is unavailable. This is a fictional finding within the example, rather than an explanation borrowed from a real organisation. The sponsor can authorise a bounded correction, assign the operational owner to test it and retain a separate forecast for the next reporting period. The review date should allow the change to affect a meaningful volume of cases. Before expanding the service, the team checks whether exception handling, adoption and accuracy have improved, whether support effort is falling and whether the additional delivery cost remains justified. A successful intervention would change both the evidence and the decision about expansion. A disappointing

result would require reconsidering the design or value claim, rather than merely replacing the favourable colour on a dashboard. The decision record could state that the sponsor authorises correction of the missing approval step within the contingency, subject to a confirmed cost estimate. The service manager owns acceptance using representative exception cases. Expansion remains conditional on accurate routing and stable service performance across four weeks of observation. The controls lead will report the revised cost forecast and adoption evidence at the next monthly review. Any deterioration in decision accuracy triggers an earlier review with the sponsor and service manager.

Judgment ownership and the next decision

The Delivery and Value Control Map is useful when these conversations become crowded. Where evidence supports delivery within accepted constraints and the operational value case, the task is to continue with proportionate assurance. If delivery viability is doubtful but value remains credible, attention turns to feasibility, sequencing and the cost of recovery. If delivery is viable but value is doubtful, operational design, adoption and the continuing investment case need examination. Where both are weak, sponsors should reconsider the basis for further commitment. These are prompts for inquiry rather than automatic instructions to continue or stop. An accurate forecast can expose unacceptable delivery; forecast credibility alone cannot establish viability. Judgments require stated evidence, not intuition alone. The map also excludes no overriding obligation: legal, safety, accessibility or essential service requirements may constrain the available choices regardless of the quadrant. Its purpose is to prevent a healthy schedule from concealing a weak value case, or a persuasive ambition from concealing an implausible delivery plan.

Professional judgment also depends on the conditions under which evidence can be reported. A team that is punished for changing a forecast has an incentive to delay bad news; a team whose estimates are never challenged has less incentive to improve them. Useful controls require a working agreement that uncertainty can be expressed without abandoning accountability. Analysts should distinguish what they know from what they infer, retain the basis of important numbers and make changes visible. Sponsors must respond to credible warnings with decisions that match their authority. Operational owners must accept responsibility for the conditions within their control rather than leaving every benefit attached indefinitely to the project manager. These responsibilities intersect, but they are not interchangeable. The control loop fails when the information is sound and nobody can act, just as it fails when authority is clear and the evidence is poor. Improving software or reporting frequency addresses neither failure automatically.

Information technology can make this work faster while leaving its hardest judgments unresolved. An automated dashboard can reconcile dates, surface exceptions and preserve the history of a forecast, but somebody must still decide whether the underlying measure represents the intended result. A predictive model trained on past projects may inherit their recording practices and omissions. Its usefulness depends on checking performance against relevant observations, understanding when conditions have changed and retaining a route for professional challenge. These considerations apply equally to an elaborate analytical platform and an ordinary spreadsheet. The first investment in better controls should therefore address a

specific weakness in the decision process. If warnings arrive after commitments become irreversible, improve the timing and ownership of the warning. If the remaining work is unclear, improve its definition and integration. If claimed benefits cannot be traced, improve measurement and attribution. Buying a more sophisticated reporting tool can support those changes once their purpose is clear, but it cannot supply an agreed purpose on behalf of the people responsible for the transformation.

The proposed models themselves should face the same standard of scrutiny. A pilot team could record whether using the loop exposes previously unowned assumptions, shortens the time between a warning and an authorised response, or improves the traceability of benefit claims. Those observations would establish aspects of usability and practice, rather than prove that the model caused better investment outcomes. Stronger evaluation would need a research design that addresses alternative explanations and the differences between participating organisations. Until such work exists, the framework should remain open to amendment. Readers may discover that a question needs greater emphasis in their setting, or that an existing governance arrangement already performs the same function effectively. Such findings are useful because the aim is better control of change, not loyalty to a diagram. Building a professional science requires accumulating evidence about what works, for whom and under which conditions. It also requires the willingness to retire a measure or method when it obscures the decision it was intended to improve.

For a reader beginning this discipline, the first application can be modest. Choose one material investment promise and trace it into the next decision. Identify the intended operating improvement, the delivery and adoption conditions on which it depends, the evidence available today and the person who can act on it. Then examine one gap in that chain closely enough to specify a useful response and a date for checking its effect. This is a more demanding exercise than completing a template because it may reveal disagreement about purpose, ownership or acceptable performance. That disagreement is valuable information when it is made explicit early enough to resolve. The chapters that follow develop the principles and governance needed to sustain this work across an entire transformation. Their starting point is the same as this chapter's: a control system earns its place by helping an organisation make better supported commitments, recognise changing reality and take responsibility for the results of its decisions.

References

- Åström, K. J. and Murray, R. M. (2008). *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press. Author authorised first edition: www.cds.caltech.edu/~murray/FBS/First_Edition.html
- Atkinson, R. (1999). Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management*, 17(6), 337–342. doi: [10.1016/S0263-7863\(98\)00069-6](https://doi.org/10.1016/S0263-7863(98)00069-6)
- AXELOS (2020). *Managing Successful Programmes*. Fifth edition. TSO The Stationery Office. ISBN 9780113316762.
- Badewi, A. (2016). The impact of project management (PM) and benefits management (BM) practices on project success: Towards developing a project benefits governance framework. *International Journal of Project Management*, 34(4), 761–778. doi: [10.1016/j.jiproman.2015.05.005](https://doi.org/10.1016/j.jiproman.2015.05.005)

- Flyvbjerg, B., Holm, M. S. and Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association*, 68(3), 279–295. doi: [10.1080/01944360208976273](https://doi.org/10.1080/01944360208976273)
- Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. *Project Management Journal*, 37(3), 5–15. doi: [10.1177/875697280603700302](https://doi.org/10.1177/875697280603700302)
- National Audit Office (2012). The London 2012 Olympic Games and Paralympic Games: post-Games review. HC 794, 5 December. www.nao.org.uk/wp-content/uploads/2012/12/1213794fr.pdf
- National Audit Office (2013). Review of the Final Benefits Statement for Programmes Previously Managed under the National Programme for IT in the NHS. 6 June. www.nao.org.uk/wp-content/uploads/2013/06/10171-001_NPfiT_Review.pdf
- National Audit Office (2019). Completing Crossrail. HC 2106, 3 May. www.nao.org.uk/wp-content/uploads/2019/05/Completing-Crossrail.pdf
- National Audit Office (2021). Crossrail – a progress update. HC 299, 9 July. www.nao.org.uk/wp-content/uploads/2021/07/Crossrail-a-progress-update-2.pdf
- Shenhar, A. J., Dvir, D., Levy, O. and Maltz, A. C. (2001). Project success: A multidimensional strategic concept. *Long Range Planning*, 34(6), 699–725. doi: [10.1016/S0024-6301\(01\)00097-8](https://doi.org/10.1016/S0024-6301(01)00097-8)
- Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3), 321–339. doi: [10.1287/mnsc.35.3.321](https://doi.org/10.1287/mnsc.35.3.321)
- Transport for London (2024). The Elizabeth line continues to transform travel in London on its two-year anniversary. 24 May. tfl-newsroom.prgloo.com/news/the-elizabeth-line-continues-to-transform-travel-in-london-on-its-two-year-anniversary
- Williams, T. (2008). How do organizations learn lessons from projects—And do they? *IEEE Transactions on Engineering Management*, 55(2), 248–266. doi: [10.1109/TEM.2007.912920](https://doi.org/10.1109/TEM.2007.912920)

CHAPTER 02

Principles for Credible Control



IN THIS CHAPTER

Principles that keep commitments, evidence and judgement credible.

When information fails to change a decision

In June 2015, the executives responsible for Canada's Phoenix pay system cancelled a planned departmental pilot after critical defects and instability emerged. Rescheduling would have delayed implementation. The system subsequently went live in two waves in February and April 2016. Canada's Auditor General later found that readiness dashboards contained calculations or supporting data that the department could not adequately explain, that oversight was not independent of project management, and that an interdepartmental advisory forum lacked formal authority to challenge implementation. Significant known problems were omitted from the final briefing to the responsible Deputy Minister, and the implementation decision was not documented (OAG Canada, 2018, paragraphs 1.39–1.40, 1.74–1.75 and 1.90–1.102). The professional question raised by these findings is uncomfortable. How can a project possess reporting arrangements and still reach a decision that its underlying evidence cannot adequately support? The answer requires attention to the behaviour and authority surrounding the numbers, as well as to the numbers themselves.

Chapter 1 introduced the Transformation Control Loop, connecting value claims, commitments, evidence, decisions and learning. This chapter examines the conditions that make those connections credible. Six principles provide its argument: purpose, traceability, calibration, challenge, accountability and proportionate adaptation. Together they form the proposed Credible Control Framework shown in Figure 2.1. Purpose connects measurement to intended value; traceability preserves the history of commitments and evidence; calibration matches confidence to what is known. Challenge makes contrary evidence discussable, accountability connects judgment to authorised action, and adaptation allows the organisation to learn without quietly abandoning its obligations. These principles reinforce one another. Perfectly recorded data can answer the wrong question, an accurate warning can reach someone unable to act, and an authorised decision can remain ineffective if nobody examines its consequences. The framework is an original synthesis for this handbook, not a validated prediction model. The research that follows supports its components and explains the problems they address, while international cases illustrate the practical meaning of the argument.

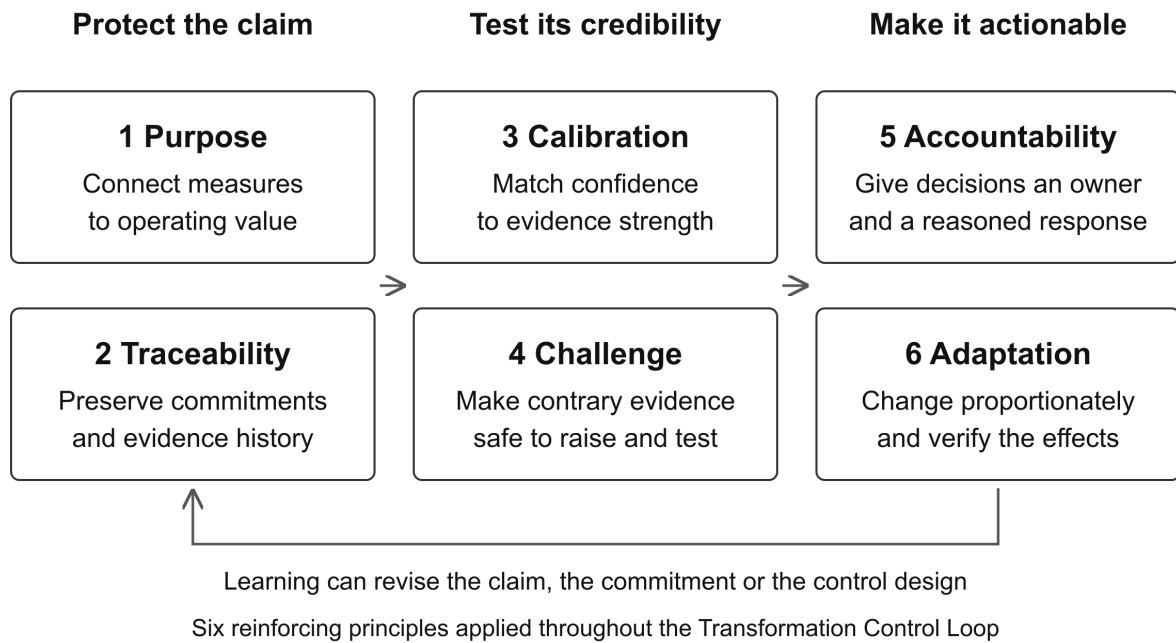


Figure 2.1 The Credible Control Framework. Six reinforcing principles apply throughout the Transformation Control Loop introduced in Chapter 1. This is an original proposed framework, not a validated scoring or prediction model.

The framework does not suggest that six favourable scores can be added into a reliable verdict. A weakness may be decisive even when the other conditions look strong. Well documented evidence is insufficient if it concerns the wrong operating population; an independent challenge is ineffective if it arrives after authority to change direction has disappeared. Nor does credibility require certainty before every action. It requires an honest account of the uncertainty, a response suited to its consequences and responsibility for what follows. A sponsor may reasonably authorise investigation while declining wider deployment, or retain an ambitious target while accepting a less favourable forecast. The relevant test is whether the commitment matches what the evidence can support. This distinction allows the principles to guide judgment under pressure. They do not remove disagreements about value, resources or acceptable exposure, but they help participants identify which disagreement they are having and what kind of evidence or authority could resolve it. Their application begins with the meaning of the performance claim itself.

Begin with the purpose of control

The first principle asks what the control is intended to protect. In transformation, that purpose includes delivering commitments and creating conditions for worthwhile operating change. It cannot be reduced to producing a report that satisfies its recipient. A procurement dashboard showing that purchase orders are processed faster may conceal growing rework or the exclusion of difficult requests. A construction report can show installed quantities without establishing that the resulting system is usable. Ouchi's conceptual account of organisational control relates the suitability of control mechanisms to the ability to measure outputs and understand the process producing them (Ouchi, 1979). Applied here, the implication is that controls should fit the work and the knowledge available. Repetitive transactions may support stable performance measures; exploratory design may initially require evidence of learning and competent judgment. This does not exempt uncertain work from scrutiny. It changes the question from whether a predetermined quantity has been produced to whether the investigation has reduced an important uncertainty enough to justify the next commitment.

Research on goals explains both the value and the danger of numerical targets. Locke and Latham's synthesis connects specific, challenging goals with performance, while emphasising conditions including commitment, ability, feedback and task complexity (2002). For unfamiliar work, learning goals can matter before demanding immediate output. Bevan and Hood's examination of English healthcare targets shows the vulnerability of treating a measured part as though it represented the whole service (2006). Read together, these sources support purposeful measurement rather than either enthusiasm for every target or rejection of targets altogether. A controls professional should ask what behaviour a measure encourages, what it excludes and what complementary evidence protects the underlying purpose. Faster decisions may need to be read alongside accuracy; completed migration records alongside their usability; cost efficiency alongside operating disruption. The selection should remain economical. Adding measures without a decision purpose can create another kind of distortion, in which the organisation devotes its attention to maintaining the measurement system instead of improving the service it was meant to explain.

Purpose also determines whose results belong in the account. A transformation can transfer costs between departments, suppliers or customers while presenting a local saving as an organisational gain. The investment boundary must therefore include the consequences material to the decision. Musawir and colleagues' survey of 333 projects examined relationships among governance, benefits management and project success (2017). Their cross sectional design did not directly measure actual realised benefits, so the study cannot establish that a particular governance intervention produces a specified financial return. It nevertheless gives empirical support to examining delivery governance and benefit responsibilities together. The practical requirement proposed here is to attach each material performance claim to an operating consequence and an accountable owner. This is a test of meaning rather than a demand to monetise everything. Reliability, accessibility, employee workload and continuity may carry value that must be considered explicitly even when it cannot be credibly expressed as cash. Controls become relevant when these considerations enter the decision before commitments make them difficult to address.

Preserve the commitment and qualify the forecast

Traceability begins with a distinction that organisations can find politically inconvenient: an approved commitment and a current expectation are different statements. A target date records an intention; a forecast records a present judgment; a baseline provides an authorised reference. They can move through legitimate decisions, but they should not become interchangeable. A revised forecast does not itself approve additional spending, and a new budget does not erase the original investment proposition. A useful account preserves the initial commitment, approved changes, actual performance and current estimate of remaining work. This allows the reader to distinguish a changed scope from a changed estimate, an execution problem from a missing assumption, and a decision still required from one already made. It also protects learning. If adverse forecasts are repeatedly absorbed into a replacement baseline without explanation, future planners lose the information needed to estimate similar work. Traceability is therefore both a current governance obligation and an investment in the organisation's future judgment.

HealthCare.gov illustrates the consequences of separating a delivery date from the evidence needed to support it. In its examination of selected contracts and task orders, the United States Government Accountability Office found that important requirements were still unknown when development work was commissioned. Oversight arrangements were inadequate, and guidance left uncertainty about authority to approve additional contractor work. The operational readiness assessment moved from March to September 2013; required approvals were not obtained, and launch proceeded without verification that performance requirements had been met (GAO, 2014). These findings concern the examined acquisition arrangements, not every aspect of American health policy. They also do not mean that no testing occurred. Their relevance is more precise: a commitment can remain visible while the evidence that would justify fulfilling it on the announced terms becomes weaker. Credible controls must expose that divergence. A schedule should show the remaining path to accepted operation, including integration and readiness work, rather than treating the public date as evidence that this path is achievable.

Calibration adds a second discipline. Confidence should reflect the strength, relevance and limitations of the evidence, rather than the seniority of the person presenting it. Flyvbjerg distinguishes cognitive bias from the strategic incentives that can influence project accounts (2021). This distinction matters because additional analytical training may help an optimistic estimator while leaving an incentive to conceal unwelcome information untouched. Reference class forecasting offers a way to challenge an internal account using observed outcomes from comparable undertakings (Flyvbjerg, 2008). Its usefulness depends on comparability and on matching the historical decision point to the current one. A forecast based on projects already well advanced cannot automatically calibrate a proposal at initial approval. The controls professional should explain important differences instead of selecting whichever comparison is most reassuring. Where sound probability estimates are unavailable, transparent scenarios are preferable to invented precision. The resulting statement may be less comfortable than a single confident date, but it gives decision makers a clearer understanding of what would need to be true for that date to hold.

Digital systems can preserve this evidence, but their presence does not establish its credibility. Zhang's qualitative study of a software consultancy examined how accountability developed through the ways project participants actually used digital tools (2024). Its findings emphasise shared practices and expectations, rather than a simple effect of making more information visible. This is a bounded organisational study, not a universal comparison of software products. Its practical implication for the present argument is straightforward: a dashboard needs an agreed reporting date, stable definitions, identified owners and a route to its underlying records. A polished chart cannot resolve uncertainty about which transactions were included or who accepted the work. Nor should every user need access to every confidential record. Traceability requires appropriate access and a reproducible account, with relevant restrictions made explicit. When a measure changes, the reason and its effect on comparison should remain visible. Otherwise, digital transparency can become a detailed view of figures whose meaning has quietly moved.

Traceability must also survive consolidation. A programme may combine suppliers' progress reports, financial actuals and operational observations that cover different periods or use different definitions of completion. Adding them together can create an apparently precise account with no consistent measurement basis. The controls team should reconcile reporting dates, scope boundaries, units and acceptance rules before interpreting the aggregate. Costs incurred but not yet invoiced need an agreed treatment; physical progress needs evidence appropriate to the work; and a benefit forecast needs its own assumptions about adoption and operating conditions. Where business units use different currencies, the reporting basis should make conversion effects distinguishable from delivery performance. These are practical expressions of the same principle. A reader must be able to understand what changed and why, rather than attributing every movement to productivity. Reconciliation also exposes the time lag between an event and its appearance in a report. If a decision depends on information that is already materially out of date, the account should say so and identify what must be checked before authorisation.

Make challenge possible and give it a destination

Evidence depends partly on whether people can report what they see. Edmondson's study of 51 manufacturing teams associated psychological safety with learning behaviour (1999). Frazier and colleagues subsequently examined a much broader evidence base through a meta analysis of 136 independent samples (2017). Edmondson and Bransby's later review connects the literature with learning, performance and leadership while identifying continuing research limitations, including extensive reliance on cross sectional evidence (2023). These studies do not establish that a particular meeting format will improve every project. They do establish a substantial research basis for treating the interpersonal conditions of speaking up as relevant to organisational learning. In project controls, the proposed application is to reduce the personal cost of raising a reasoned concern while retaining demanding standards for evidence. A planner should be able to question an unachievable sequence; an operational specialist should be able to describe an untested exception. Their concern still needs examination. Respectful challenge means taking the contribution seriously enough to test it, not accepting every objection as correct or abandoning the requirement to decide.

Formal risk processes can fail before any calculation begins if relevant information never enters them. Kutsch and Hall's qualitative research with IT project managers examined how known risks could be treated as irrelevant (2010). The contribution is a warning about selection, not a claim that every omitted issue was deliberately concealed. A risk register can appear complete because the organisation has defined uncomfortable questions as outside its scope. The controls team should therefore retain a route for questioning those boundaries. Why was an operational dependency excluded? Who decided that a reported problem could not affect the next commitment? What observation would reopen the issue? The same discipline applies to apparently reassuring evidence. A test result deserves questions about its population, operating conditions and exceptions, not merely confirmation that somebody signed the report. Challenge is most useful when it addresses a consequential uncertainty early enough to change action. A ritual invitation for comments immediately before an irreversible approval creates the appearance of participation while leaving little opportunity for investigation.

Accountability supplies the destination for that challenge. Lerner and Tetlock's review shows that accountability can improve or distort judgment depending on its timing, audience and task (1999). Anticipated scrutiny can encourage more careful reasoning, but demanding an explanation after commitment may also encourage defence of a position already taken. The professional response should therefore examine the quality of reasoning before authorisation and preserve the basis for later review. An escalation route needs a named recipient with relevant authority, a response time and a recorded disposition. Consultation alone does not supply those conditions. A controls analyst may identify the implications of an incomplete test, but responsibility for accepting or changing the commitment belongs with an authorised decision maker. Equally, that decision maker cannot make the concern disappear by returning it to the analyst without direction. The record should state what was accepted, rejected, investigated or deferred, why that response was chosen and who remains responsible for its consequences. This gives disagreement a useful organisational form.

Norway's arrangements for major public investment show why challenge must also be designed institutionally. Research on front end governance distinguishes selecting the right investment from managing its subsequent delivery efficiently (Samset and Volden, 2016). Comparative work on six public investment governance regimes identifies different institutional arrangements rather than a universally applicable structure. In Norway, independent concept appraisal, known as QA1, is distinguished from QA2, which reviews the management basis and cost estimate before investment approval. Reviewers assess the evidence while elected authorities retain investment decisions (Volden and Samset, 2017a). A Concept programme evaluation of twenty evaluated projects concerned investments that had undergone QA2; none had passed through QA1 (Volden and Samset, 2017b). Those cases therefore cannot establish that concept stage assurance caused their results. This distinction matters to an international handbook. A review mechanism must be understood in its institutional and historical setting before its outcomes justify adoption elsewhere. The transferable principle is the opportunity for informed challenge to influence commitment; the appropriate institutional arrangement requires local design.

Adapt the response while retaining responsibility

Control can support skilled work or become an obstacle to it. Adler and Borys distinguish enabling from coercive forms of formalisation, examining how procedures can help employees understand and improve their work rather than merely constrain them (1996). The application here is to design control arrangements that make problems diagnosable and correction possible. A procedure should explain its purpose, the evidence required and the legitimate route for dealing with an exception. Requiring every local adjustment to wait for a distant committee can delay learning; allowing local changes without an account of their effects can destroy integration. Proportionate adaptation lies between those failures. The intensity of scrutiny should reflect consequence, uncertainty, interdependence and reversibility. An easily reversible experiment within agreed boundaries may justify delegated authority. A change that affects payment integrity or service continuity may require stronger evidence and wider agreement. This is a design judgment that the organisation should explain, rather than an excuse to describe inconvenient controls as bureaucracy whenever a deadline approaches.

The learning problem has a further dimension. March's theoretical analysis distinguishes exploration of new possibilities from exploitation of existing knowledge and capability (1991). A transformation requires both, although their balance changes with the work. Teams may need to investigate an unfamiliar operating model while applying dependable disciplines to expenditure, access permissions and service continuity. Research by Turner and colleagues using eight IT service project cases examines how projects combine exploitation and exploration through management practice (2016). These sources support considering disciplined delivery and learning together; they do not establish that any activity labelled agile is appropriately controlled. A useful local experiment specifies the uncertainty to be reduced, the resources it may consume and the observations that would change the next decision. A learning exercise without a question can become unbounded activity. A performance commitment imposed before the necessary learning has occurred can encourage superficial compliance. The controls professional helps

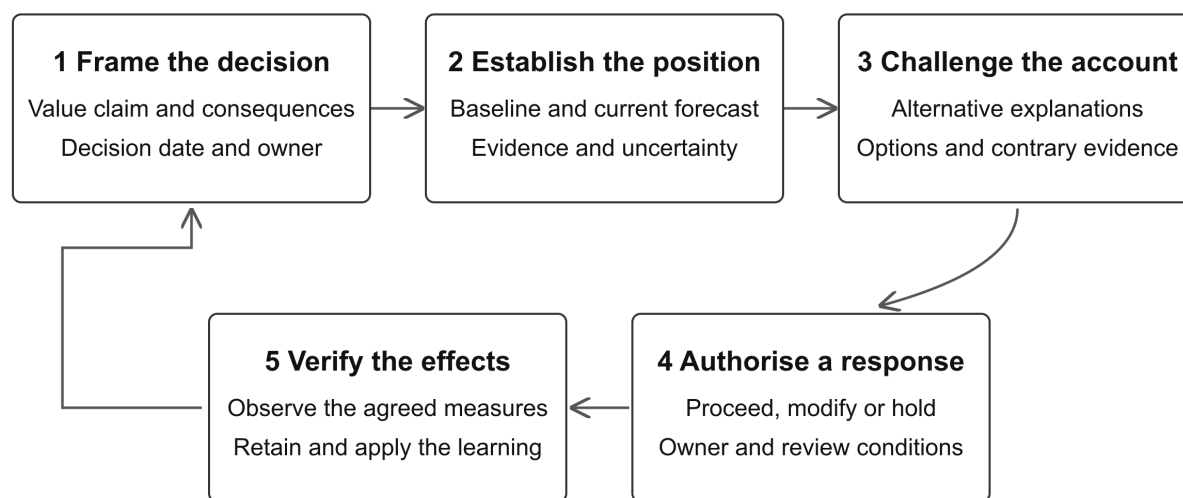
distinguish these situations and ensures that the chosen arrangement remains connected to the wider investment.

Singapore's NEWater programme offers a positive illustration of staged evidence and commitment. PUB's retrospective account records a demonstration plant in 2000, testing and multidisciplinary expert review during 2000–2002, and introduction of NEWater as a national water source in 2003 (PUB, 2013, pp. 40–41). A contemporaneous government statement records acceptance of an expert panel recommendation in September 2002 and a plan for a small initial introduction into reservoirs, followed by progressive expansion (Singapore Government, 2002). These are sponsor and government accounts, not an independent audit proving that a particular governance method caused success. The announced quantities in the 2002 statement were plans at that date. Nevertheless, the chronology illustrates a useful separation: demonstration generates evidence, specialist review examines it, and an authorised decision determines the next exposure to operational use. That separation matters beyond water infrastructure. In a digital service or industrial transformation, a pilot becomes informative when its conditions, observations and consequences for expansion are defined before the pressure to announce completion overwhelms its learning purpose.

Adaptation also requires an honest account of improvement after difficulty. In a later HealthCare.gov review, GAO reported that HHS described improved availability and functionality during enrolment for 2018 coverage. Officials attributed part of the improvement to an operating system upgrade and comprehensive testing. GAO simultaneously identified gaps in performance measurement, including the absence of numerical enrolment targets and incomplete measurement of aspects of consumer experience (GAO, 2018). This supports a qualified account of operational improvement, not a verdict that every programme objective had been achieved. For controls practice, the implication is to preserve distinctions between a repaired technical service, improved user experience and the wider outcomes of the investment. Recovery should be recognised when supported by evidence, while unresolved obligations remain visible. Equally, a troubled beginning should not become a permanent label that prevents learning from subsequent improvements. A credible account can contain progress, uncertainty and disagreement without forcing them into a single favourable or unfavourable colour.

A methodology for credible control

The six principles become practical through a repeatable method of framing, establishing, challenging, authorising and verifying. Figure 2.2 presents this process as an application of Chapter 1's control loop. The framework describes the qualities a control system needs; the method describes how a team can apply those qualities to a particular decision. A Control Principles Charter records the local commitments that make the method workable. These are proposed practice tools, developed for this book. Their value should be assessed through use, and their presence should never be treated as evidence that governance is effective. A short, well supported decision record may serve the purpose better than a large completed template. The method is intended for material commitments and exceptions, with lighter application where consequences are limited. It can be used during investment definition, before a delivery wave, at an operational transition or when new evidence challenges an existing commitment. Its common feature is a decision whose basis deserves to remain understandable after the meeting has ended.



The decision record carries evidence, reasoning, authority and a review trigger

Return to an earlier step whenever new evidence changes the basis for action

Figure 2.2 The Credible Control Method. An original five-stage process for applying the principles to a material decision. New evidence may require returning to an earlier stage; the process does not substitute for delegated authority.

Framing begins by identifying the decision, its owner, the operating value at stake and the consequences of being wrong. The team specifies the date by which a decision is useful and the constraints it must respect. Establishing the position then reconciles the approved commitment with current evidence and forecasts. This includes remaining work, relevant dependencies and known limitations in the information. The two stages should expose ambiguity early. If the question is whether to expand a service, evidence that the first site was technically installed may be insufficient; the decision may require stable operation, competent users and reliable handling of exceptions. A statement that the project is within budget may also conceal an unsupported

estimate of future expenditure. The record should separate what has happened from what is expected and identify which assumptions connect them. It should be possible for a reviewer to explain both the case for proceeding and the main reason that case might be wrong without having to reconstruct the entire project history.

Challenge then examines alternative explanations and credible options. It should ask what evidence would change the recommendation, whether a relevant external comparison exists and whether the information has been selected to defend an earlier decision. Staw's experimental work on escalating commitment showed how personal responsibility for an unsuccessful initial choice could influence further investment in the study's setting (1976). Sleesman and colleagues' meta analysis found several determinants of escalation, including relationships involving sunk costs, personal responsibility and proximity to completion (2012). These findings do not diagnose a particular sponsor or imply that every overrun should be cancelled. They justify comparing prospective options explicitly. Continuing can be sensible if its remaining costs and expected consequences compare favourably with alternatives. The weakness is continuing mainly because expenditure already incurred must somehow be vindicated. The review should therefore consider modification, staging, deferral and stopping alongside continuation, with the same attention to uncertainty in each option. Challenge serves the decision by improving the comparison, rather than rewarding opposition for its own sake.

Disagreement can remain after this examination, particularly when different groups bear different consequences. The response should identify the unresolved question rather than force a premature consensus. A supplier may consider its component complete while operations cannot yet use the integrated service. Both statements can be accurate under their respective definitions. The decision record needs to explain that boundary and the work required to cross it. Where the dispute concerns missing evidence, a bounded test may help. Where it concerns who may accept a consequence, a technical test cannot supply the missing authority. Where it concerns competing values, additional decimal places will not settle the choice. This classification keeps inquiry purposeful and prevents analysis becoming a substitute for judgment. It also preserves a useful distinction between challenging an account and accusing its author. People remain responsible for timely, accurate reporting, while a changed forecast is examined for its basis and implications before conclusions are drawn about competence or conduct. The method can thereby support candid revision without removing professional accountability for the quality of the work.

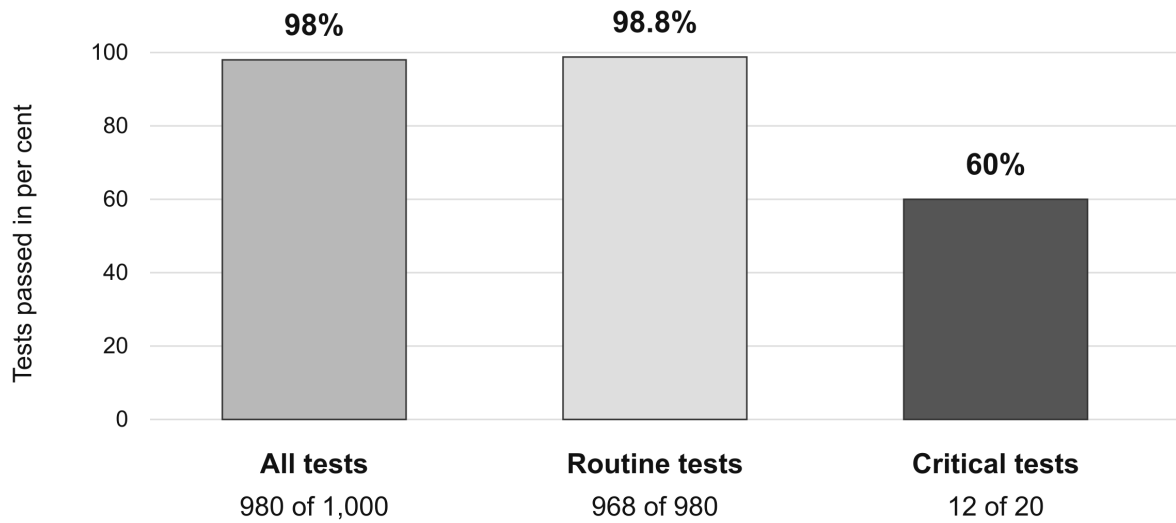
Authorisation converts that comparison into an accountable response. The decision maker records the selected option, the reasoning, any conditions, the responsible owner and the next review trigger. If evidence is insufficient, a bounded investigation or a temporary hold can be an authorised response; uncertainty does not automatically justify either unlimited delay or immediate commitment. Verification then examines whether the agreed action had the expected effect and whether new information changes the investment case. The review interval should reflect how quickly consequences can emerge and how long an intervention needs to work. A weekly meeting cannot create weekly evidence of a benefit that develops over several months. Conversely, waiting for a monthly board cycle may be unacceptable for a rapidly deteriorating operating condition. The resulting learning should update relevant assumptions, practices or commitments through the appropriate authority. This completes the method without confusing a

changed forecast with an approved change, or a completed action with proof that the underlying problem has been solved.

A charter that changes the next commitment

Consider a fictional international manufacturer introducing a shared procurement and payment platform across three business units. All names, quantities and decisions in this example are illustrative. The approved delivery budget is US\$12 million and the target completion date is 30 June. A current assessment forecasts US\$14 million and completion on 31 August. Testing reports that 980 of 1,000 specified tests have passed. A sponsor facing a publicised deadline might be tempted to treat the 98 per cent pass rate as sufficient reassurance and replace the old budget with the latest number. Applying the framework changes the conversation. Purpose asks whether the service can process and authorise payments reliably. Traceability retains the US\$12 million commitment while showing the additional US\$2 million now forecast. Calibration asks what the test population represents. Challenge gives operational and finance specialists a route to question the recommendation. Accountability identifies who can approve additional investment or alter the release. Adaptation considers a limited next step within explicit conditions rather than assuming that only complete launch or total cancellation is possible.

The test breakdown reveals why aggregation matters. Of 980 routine tests, 968 have passed; of twenty critical authorisation and exception tests, only twelve have passed. The combined result is still 980 successful tests out of 1,000, but eight critical failures remain. Figure 2.3 separates these populations. The arithmetic is correct in both views, yet the aggregate is inadequate for the release decision. In this fictional organisation, the agreed release condition requires all twenty critical tests to pass, together with a demonstrated recovery arrangement and an accepted operating responsibility. That is a local condition for the example, not a universal assurance standard. Passing the selected tests would not establish a probability of safe operation or prove that no unknown defect exists. It would establish that defined evidence required for this commitment has been obtained. Each of the twelve failed routine tests also needs a documented disposition, an owner and an acceptance decision. The controls professional must therefore communicate both the result and its boundary, including untested operating conditions, outstanding defects and any material changes made since the tests were performed.



Eight critical failures remain hidden within a 98 per cent aggregate pass rate

Fictional example • Test counts do not establish a probability of safe operation

Figure 2.3 Correct arithmetic can still conceal a material weakness. In this fictional example, the aggregate pass rate is 98 per cent while eight of twenty critical tests remain unsuccessful. Routine failures also require an explicit disposition.

The charter makes these expectations explicit before the decision is taken. Table 2.1 records the six principles as local commitments with an owner and a review consequence. Its purpose is to prevent terms such as transparency or accountability from remaining expressions of good intention. Each commitment should be observable in the work. A forecast has an identified basis and date; an excluded test remains visible; a material concern receives a recorded response. The table should be agreed with the people who must use it, including operational representatives and suppliers whose evidence supports the decision. Agreement does not transfer the sponsor's authority to a committee or allow suppliers to define acceptance solely around their own deliverables. It establishes a common account of what will support a commitment and how disagreement will be handled. The charter should remain short enough to use and stable enough to mean something, while permitting authorised revision when experience shows that a requirement is ineffective or disproportionate. Any temporary exception to the charter should record its authorising owner, justification, compensating arrangements and review date, so that an expedient departure cannot silently become the organisation's normal practice.

Table 2.1 A Control Principles Charter for the fictional international rollout

Principle	Local commitment	Owner	Review consequence
Purpose	Release evidence must cover usable purchasing and payment integrity.	Operations and finance owners	Hold release where required operating conditions remain unmet.
Traceability	Show the US\$12 million approved budget and US\$14 million forecast separately; retain change history.	Controls lead	Refer any proposed baseline change for authorised review.
Calibration	Report routine and critical test results separately, with coverage and limitations.	Test lead	Obtain missing evidence; avoid unsupported confidence claims.
Challenge	Record material concerns and reasoned responses before release authorisation.	Sponsor with independent assurance	Escalate a disputed prerequisite before the commitment.
Accountability	Separate authority to prepare a correction proposal from authority to spend or release.	Sponsor and funding authority	Record the decision, conditions, owner and review date.
Adaptation	Pilot only after its prerequisites are met; retain a disposition for every failed test.	Delivery and operational owners	Expand, modify or hold through an authorised review.

The charter records local commitments for this illustrative decision. It is a proposed practice aid and does not replace professional judgment or the organisation's delegated authority.

In the example, the sponsor holds wider release and commissions a correction proposal within existing delegated planning resources. The delivery manager must submit the cost and schedule implications before committing the additional work. The finance owner checks payment integrity, the operational owner accepts the recovery demonstration, and an assurance reviewer examines whether the repeated tests cover the changed configuration. The revised US\$14 million forecast remains visible while the appropriate authority considers the funding decision; it does not become an approved budget by appearing in a report. A limited pilot remains an option only when its own defined acceptance conditions and containment arrangements have been satisfied. At the next review, the team considers the new test evidence, remaining defects, operating workload and updated estimate together. If correction has not produced the required evidence, the decision returns for further consideration. If it has, the sponsor can authorise the next bounded commitment with a stated basis. The method produces an inspectable chain from evidence to action and subsequent review.

Principles that travel without becoming rigid rules

An international handbook must distinguish principles that travel from arrangements that depend on context. The need to understand a forecast can apply across countries while reporting calendars, contracting practices and approval powers differ. A multinational programme may require common definitions for consolidation and local arrangements for confidentiality, language and operational decision making. Those differences should be documented rather than erased to make a global dashboard appear consistent. The cases in this chapter also draw on different kinds of evidence. Independent audits, organisational accounts, field studies, experiments and research reviews answer different questions. A government report can establish a documented weakness without identifying its sole cause; a laboratory experiment can illuminate a mechanism without estimating its prevalence in major programmes. International comparison is useful when these boundaries are retained. It becomes misleading when countries are ranked from a few selected projects or when a practice is transferred without examining the authority, capability and information that made it workable in its original setting.

The proposed framework should face its own discipline of credible control. A team can examine whether the method reveals material assumptions, makes disputed evidence traceable and connects concerns to timely decisions. It can assess whether actions are reviewed and whether the resulting learning changes future commitments. These observations would support an account of usability and practice; they would not, by themselves, prove that the framework causes better project outcomes. Stronger claims would require research that addresses alternative explanations and differences between projects. Until then, the six principles should guide professional inquiry rather than become another ceremonial compliance score. Their purpose is to protect the relationship between an investment promise and the decisions made in its name. The next chapter develops the governance arrangements needed to sustain that relationship: the authorities, information flows, approval points and escalation routes through which credible control becomes part of the organisation's ordinary work.

References

- Adler, P. S. and Borys, B. (1996). Two types of bureaucracy: Enabling and coercive. *Administrative Science Quarterly*, 41(1), 61–89. doi: [10.2307/2393986](https://doi.org/10.2307/2393986)
- Bevan, G. and Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538. doi: [10.1111/j.1467-9299.2006.00600.x](https://doi.org/10.1111/j.1467-9299.2006.00600.x)
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. doi: [10.2307/2666999](https://doi.org/10.2307/2666999)
- Edmondson, A. C. and Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational Behavior*, 10, 55–78. doi: [10.1146/annurev-orgpsych-120920-055217](https://doi.org/10.1146/annurev-orgpsych-120920-055217)
- Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European Planning Studies*, 16(1), 3–21. doi: [10.1080/09654310701747936](https://doi.org/10.1080/09654310701747936)
- Flyvbjerg, B. (2021). Top ten behavioral biases in project management: An overview. *Project Management Journal*, 52(6), 531–546. doi: [10.1177/87569728211049046](https://doi.org/10.1177/87569728211049046)

- Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A. and Vracheva, V. (2017). Psychological safety: A meta-analytic review and extension. *Personnel Psychology*, 70(1), 113–165. doi: [10.1111/peps.12183](https://doi.org/10.1111/peps.12183)
- Government Accountability Office (GAO) (2014). Healthcare.gov: Ineffective Planning and Oversight Practices Underscore the Need for Improved Contract Management. GAO-14-694, 30 July. www.gao.gov/products/gao-14-694
- Government Accountability Office (GAO) (2018). Health Insurance Exchanges: HHS Should Enhance Its Management of Open Enrollment Performance. GAO-18-565, 24 July; publicly released 23 August. www.gao.gov/products/gao-18-565
- Kutsch, E. and Hall, M. (2010). Deliberate ignorance in project risk management. *International Journal of Project Management*, 28(3), 245–255. doi: [10.1016/j.ijproman.2009.05.003](https://doi.org/10.1016/j.ijproman.2009.05.003)
- Lerner, J. S. and Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255–275. doi: [10.1037/0033-2909.125.2.255](https://doi.org/10.1037/0033-2909.125.2.255)
- Locke, E. A. and Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705–717. doi: [10.1037/0003-066X.57.9.705](https://doi.org/10.1037/0003-066X.57.9.705)
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. doi: [10.1287/orsc.2.1.71](https://doi.org/10.1287/orsc.2.1.71)
- Musawir, A. U., Serra, C. E. M., Zwikael, O. and Ali, I. (2017). Project governance, benefit management, and project success: Towards a framework for supporting organizational strategy implementation. *International Journal of Project Management*, 35(8), 1658–1672. doi: [10.1016/j.ijproman.2017.07.007](https://doi.org/10.1016/j.ijproman.2017.07.007)
- Office of the Auditor General of Canada (OAG Canada) (2018). Report 1 – Building and Implementing the Phoenix Pay System. Reports of the Auditor General of Canada to the Parliament of Canada, Spring 2018, tabled 29 May. publications.gc.ca/collections/collection_2018/bvg-oag/FA1-27-2018-1-1-eng.pdf
- Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25(9), 833–848. doi: [10.1287/mnsc.25.9.833](https://doi.org/10.1287/mnsc.25.9.833)
- PUB (2013). PUB Annual Report 2012/2013. Singapore, pp. 40–41. www.pub.gov.sg/-/media/PUB/Publications/Report/PDF/annualreport2013.pdf
- Samset, K. and Volden, G. H. (2016). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management*, 34(2), 297–313. doi: [10.1016/j.ijproman.2015.01.014](https://doi.org/10.1016/j.ijproman.2015.01.014)
- Singapore Government (2002). [Statement on acceptance of the Expert Panel’s recommendation for indirect potable use of NEWater]. 25 September. National Archives of Singapore, MSE_20020925001. Descriptive title supplied for the untitled archived statement. www.nas.gov.sg/archivesonline/data/pdffdoc/MSE_20020925001.pdf
- Sleesman, D. J., Conlon, D. E., McNamara, G. and Miles, J. E. (2012). Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal*, 55(3), 541–562. doi: [10.5465/amj.2010.0696](https://doi.org/10.5465/amj.2010.0696)
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16(1), 27–44. doi: [10.1016/0030-5073\(76\)90005-2](https://doi.org/10.1016/0030-5073(76)90005-2)
- Turner, N., Swart, J., Maylor, H. and Antonacopoulou, E. (2016). Making it happen: How managerial actions enable project-based ambidexterity. *Management Learning*, 47(2), 199–222. doi: [10.1177/1350507615610028](https://doi.org/10.1177/1350507615610028)
- Volden, G. H. and Samset, K. (2017a). Governance of major public investment projects: Principles and practices in six countries. *Project Management Journal*, 48(3), 90–108. doi: [10.1177/875697281704800306](https://doi.org/10.1177/875697281704800306)

- Volden, G. H. and Samset, K. (2017b). A close-up on public investment cases. Lessons from ex-post evaluations of 20 major Norwegian projects. Concept Report No. 52, English summary. Ex ante Academic Publisher, NTNU. www.ntnu.no/documents/1261860271/1262010703/CONCEPT_52_engelsk_web.pdf/2f03f5c0-7ddc-4631-aac0-86b2000819f4?version=1.0
- Zhang, Y. F. (2024). Too little or too much transparency? Influencing accountability in project management through the practices of using digital technologies. *International Journal of Project Management*, 42(6), 102636. doi: [10.1016/j.ijproman.2024.102636](https://doi.org/10.1016/j.ijproman.2024.102636)

CHAPTER 03

Connecting Governance to the Work



IN THIS CHAPTER

Making authority and information useful at the point of decision.

When authority and delivery become disconnected

Queensland's replacement health payroll system went live in March 2010, but the arrangements surrounding its delivery had left important questions unresolved. The Queensland Audit Office reported that overall accountability was unclear, that Queensland Health had not formally endorsed the project scope, and that more than 47 changes had contributed to delay and an increase in the IBM contract value from A\$6.19 million to A\$24 million. That contract movement was not a measure of the complete programme's lifetime cost. The audit also identified unmet acceptance criteria and inadequate arrangements for managing the business implications of the replacement (QAO, 2010, pp. 20–21). These findings raise a question that reaches beyond payroll. Who was responsible for joining the scope being commissioned to the operating conditions that the customer needed, and how should the consequences of a change have reached that authority? Separate approvals and progress reports cannot answer that question unless their relationships have been designed.

Chapter 1 introduced the Transformation Control Loop, and Chapter 2 established principles for making its evidence and decisions credible. This chapter explains how to connect that loop to ordinary organisational work. Governance determines who can commit resources, accept consequences and change direction. Project controls supplies and challenges the evidence through which those powers are exercised. The practical connection is a decision with an identified authority, an intelligible evidence requirement and a deadline derived from the work. A report without this destination can be accurate and still have little effect. An approval without a connection to delivery can express legitimate intent while leaving teams unsure what has changed. Figure 3.1 presents the proposed Governance Connection Map, an implementation aid within the existing framework. It links investment authority, sponsorship and delivery decisions to evidence flowing back from the work. Its purpose is to expose missing connections, rather than to prescribe the same committee structure for every transformation or introduce a universal measure of governance maturity.

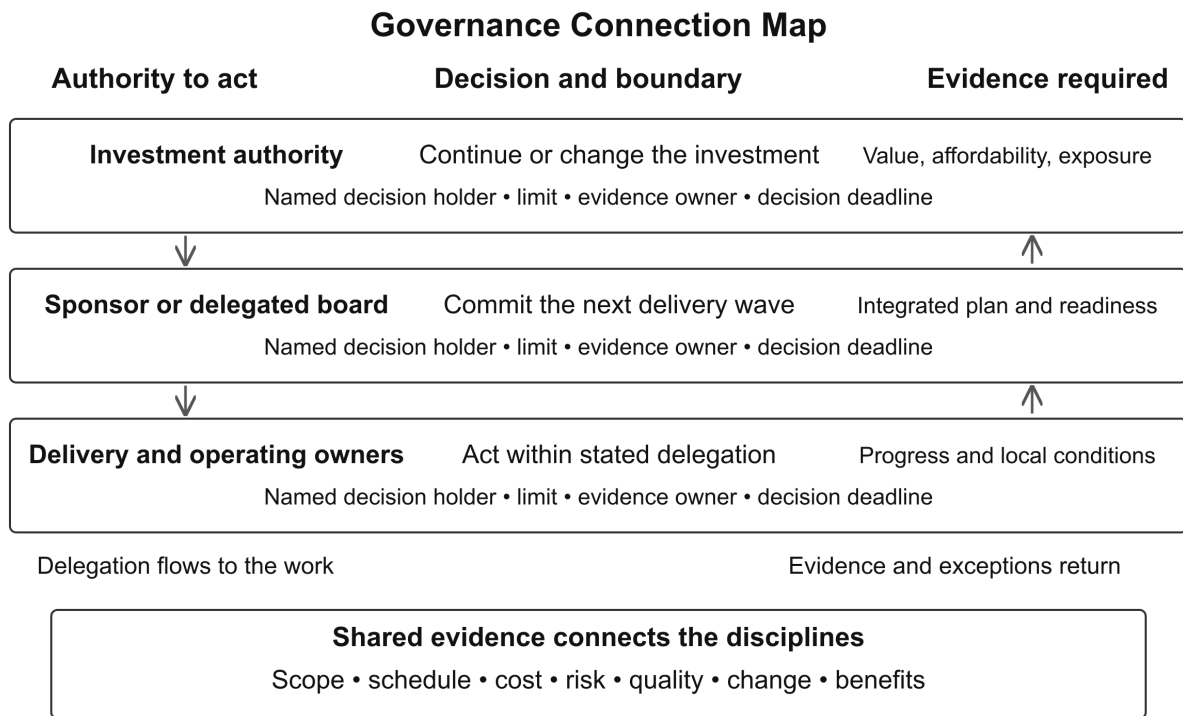


Figure 3.1 The Governance Connection Map links decisions at different levels to evidence, delegated limits and timing. Arrows show delegation towards the work and evidence returning to authority. Local roles may be combined; the functions and boundaries remain explicit.

Research helps distinguish these connections from an organisation chart. Too and Weaver's conceptual framework connects the governance of projects with portfolio selection, sponsorship, oversight and support, while separating governance from the management of individual projects (2014). Ahola and colleagues distinguish research concerned with governance inside projects from research examining arrangements outside them (2014). Biesenthal and Wilden's analysis of 97 papers further shows why the organisational level of a governance claim matters (2014). Together, these studies suggest a useful discipline for the controls professional: name the decision and its level before naming the forum. A project manager can rearrange authorised activities while lacking authority to reduce the service promised to customers. A sponsor may resolve priorities while lacking authority to increase the organisation's investment ceiling. A board may possess investment authority while needing another body to accept an operating responsibility. Describing all three as governance obscures the limits that determine what can actually happen. The map makes those limits visible.

Translate strategic intent into bounded authority

A strategic objective becomes governable when it is translated into commitments whose meaning can be tested. An ambition to improve payment reliability might justify investigating a shared platform, but it does not automatically authorise a global deployment. The investment authority first needs to establish the problem, the alternatives worth examining and the resources available for that investigation. Subsequent commitments should specify the scope being authorised, the resources that may be used, the conditions that must be met and the circumstances requiring reconsideration. This creates a chain of bounded decisions rather than an assumption that early enthusiasm authorises everything downstream. The chain also runs back towards strategy. If evidence shows that the proposed operating model will not achieve the intended improvement, the team needs a route for questioning the proposition itself. Corrective action cannot always remain inside the delivery plan. Sometimes the appropriate decision is to change the service design, revisit the investment rationale or stop an activity whose justification has weakened.

Delegation should therefore identify both the permitted action and its boundary. A delivery manager might be allowed to resequence work within an approved envelope while being prohibited from changing a customer commitment or consuming a centrally held contingency without further approval. The delegation needs to state whether a financial limit concerns an individual transaction, cumulative changes or forecast total expenditure. Otherwise, several individually small decisions can create an unauthorised aggregate commitment. It should also identify who acts when the usual decision holder is unavailable and how time sensitive exceptions are handled. These arrangements should be accessible to the people preparing requests, including suppliers. A responsibility matrix can identify participation, but a letter beside a role does not explain spending power, acceptance authority or the conditions for escalation. Those distinctions belong in the underlying decision record. Contractual permissions, technical acceptance and investment approval may involve different authorities; obtaining one should never be presented as obtaining all three. Stoppels and colleagues' study of six Dutch steering committees, drawing on eighteen interviews, found that responsibilities were often implicit and that members connected temporary projects with permanent organisations (2026). Its findings reinforce the need to distinguish attendance from authority and authority from responsibility for implementation. A representative must know which commitments they can make and which require a decision in their own organisation.

The appropriate balance between local freedom and central direction depends on the work. Chenhall's review of contingency research relates management control design to organisational circumstances rather than supporting one universally effective system (2003). Joslin and Müller's international survey analysed 246 responses after removing outliers from an initial usable set of 254 across 41 countries, examining associations between project governance and success; it does not establish that importing a particular arrangement causes better results (2016). The practical conclusion proposed here is to concentrate higher authority on consequential choices while giving teams enough room to solve bounded problems. A reversible configuration experiment can often remain local. A decision affecting multiple operating units, contractual obligations or an investment proposition usually requires wider consideration. Delegation must also be supported by competence and access to evidence. Giving someone responsibility without the

capacity to interpret the information simply relocates the problem. The controls function should make the boundary understandable and reveal when it has been approached, while the authorised manager remains responsible for the decision. Müller and Lecoivre's validated categorisation distinguishes shareholder–stakeholder orientation and behaviour or outcome control (2014). These dimensions help explain differences in governance emphasis; they do not identify a universally superior category or justify treating a classification as a quality score.

Decision rights become especially important where a transformation crosses organisations. The customer, supplier, regulator and future operator may each have legitimate powers that no programme board can absorb. A joint review can align their information without creating authority that its participants do not possess. Okhuysen and Bechky's integrative review explains coordination through accountability, predictability and common understanding, helping to show why meetings alone are insufficient (2009). A useful interface agreement should establish the evidence one party must provide, the condition another party must accept and the timing on which both depend. If a supplier completes a component while the operator lacks trained staff, each party's local report may be accurate even though the joint commitment is unachievable. The governance connection is the decision about that shared consequence. It requires a named route for resolving incompatible assumptions and an explicit account of what remains outside the programme's control. Escalation should clarify these boundaries, rather than pretend that seniority removes them.

Integrate the consequences before seeking approval

An integrated control system is defined by the relationships it preserves. Scope explains what the organisation is trying to deliver; schedule explains the sequence and timing; cost explains resource consequences; risk describes uncertain exposure. Quality establishes the conditions for acceptable work, change control governs alterations to commitments, and benefits management examines the improvement expected from use. These disciplines may employ different specialists and applications, but they must support a coherent account of the same proposed action. Malmi and Brown's discussion of management control systems as a package warns against examining individual controls without their wider relationships (2008). Their definition centres on directing behaviour; this handbook also includes decision support. That broader application makes the relationships among controls particularly important. Tight budget monitoring can encourage the removal of training that an adoption plan requires; an aggressive completion target can encourage acceptance of work that an operational measure would reject. Integration requires those consequences to be considered together before approval. A common dashboard is helpful only if the underlying definitions and dependencies allow the apparent agreement between its figures to be examined.

The initial technical task is to establish compatible control boundaries. A work package should be traceable to the requirement it satisfies, the activities that produce it, the responsible organisation and the estimate of resources required. Its acceptance evidence should explain what completion means. Not every system needs identical codes or the same level of detail, but mappings must preserve meaning when information is consolidated. A supplier invoice, for example, may span several work packages; allocating it entirely to the first available code can distort both progress

and remaining cost. An operational benefit may depend on several delivered capabilities and should not be counted in full against each one. GAO's Cost Estimating and Assessment Guide connects a technical baseline, work breakdown structure, assumptions, data, uncertainty analysis and updates using actual costs (2020). This guidance supports the discipline of a traceable estimate. It does not remove the need for judgment about where the transformation boundary lies or who accepts costs transferred into continuing operations.

Integration also requires attention to how controls interact. Tiwana's study of 120 outsourced systems development projects found different relationships between formal and informal controls: informal control complemented behaviour controls but substituted for outcome control in the studied relationships involving goal alignment and flexibility (2010). Here informal control concerned shared goals and values, rather than an absence of records or casual conversation. The lesson is more specific than saying that every additional control improves performance. A team may need agreed methods for a difficult interface and informal discussion to resolve ambiguity. For the controls professional, the practical task is to identify what each mechanism contributes. A weekly coordination conversation can reveal an emerging dependency, but it cannot silently amend an authorised baseline. A signed change request can authorise an adjustment, but it cannot create the shared technical understanding needed to implement it. The system needs both an appropriate decision record and working relationships through which its meaning is understood. Duplication should be removed where it adds effort without adding a distinct protection.

Consider a proposed correction to payment authorisation rules in the fictional manufacturer introduced in Chapter 2. The correction changes software scope and requires affected tests to be repeated. Testing needs specialist staff who are also preparing the next business unit, so the schedule consequence extends beyond the correction itself. Additional effort changes the estimate, while delayed introduction changes the timing of expected benefits. A temporary operating workaround may introduce further error exposure and support costs. Figure 3.2 follows these connected consequences. It is an original illustration of impact assessment, not a calculation claiming that all effects are known. The team should identify direct effects, dependencies and important uncertainties, then involve the owners capable of checking them. Some consequences will prove negligible; others may dominate the decision. The aim is a proportionate, dated account of the proposed change, including the consequences of declining it. This prevents the funding request from presenting only the supplier's price while leaving the organisation to discover its operating obligations later.

Follow a change through to its operating consequences

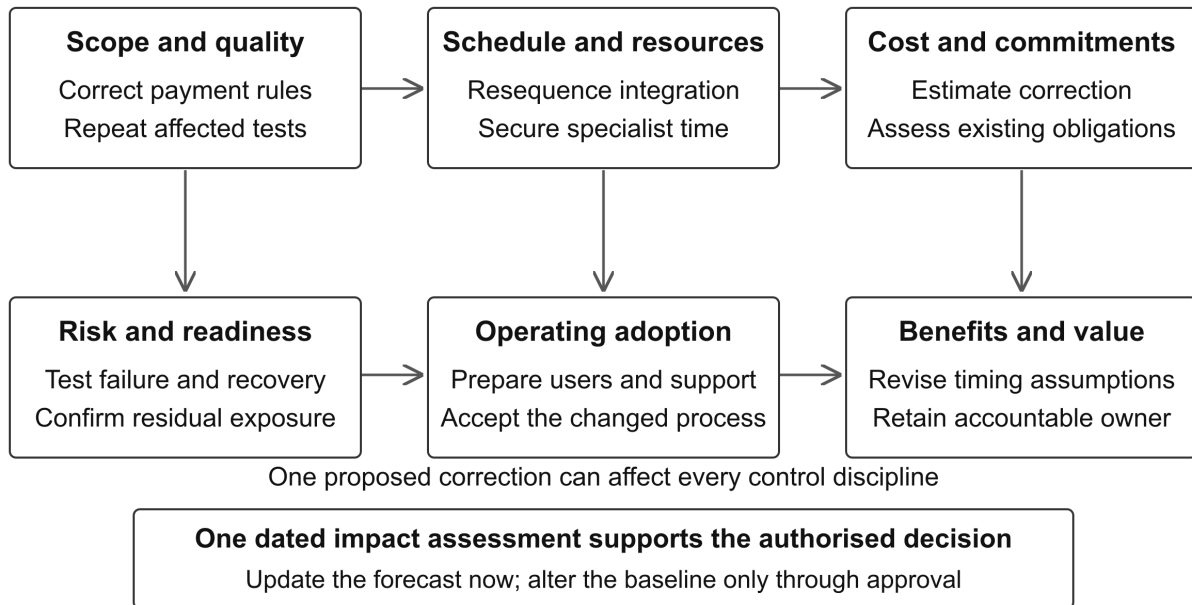


Figure 3.2 An integrated assessment follows a proposed payment correction through delivery and operating consequences. The links are illustrative, not an exhaustive dependency network. The forecast reflects current evidence; a baseline change requires approval.

Keep targets baselines forecasts and acceptance distinct

Four statements often appear beside one another in a report, yet they answer different questions. A target expresses an ambition, such as introducing a service by the end of June. An approved baseline records the authorised reference for scope, time and cost. A forecast describes the current expectation based on progress, remaining work and uncertainty. Acceptance criteria define the evidence or conditions required before the work or operating state can be accepted. A target can remain demanding while a forecast indicates delay; an approved budget can remain unchanged while the estimate at completion increases. Acceptance criteria should not weaken automatically because either date has moved. Keeping these statements distinct allows governance to deal with reality without losing accountability. The report should identify their owners, dates and approval status so that readers can see whether a number is an intention, a commitment, an expectation or a requirement. A traffic light alone cannot communicate these differences, especially when several statements happen to share the same numerical value.

Baseline change should consequently be a decision with a history, rather than a way of making variance disappear. The original baseline remains available, authorised changes explain the current baseline, and the forecast explains the latest expected outcome. When scope is added, the record should distinguish its approved resource implications from deterioration in the estimate for existing work. A revised schedule should similarly explain whether a milestone moved because its definition changed, a dependency changed or execution took longer than expected.

GAO's Schedule Assessment Guide emphasises an integrated, reliable schedule and the cost consequences of schedule slippage (2015). Applied to governance, this means that approving a new completion date without examining the funding and operating consequences leaves the decision incomplete. It also means that the project controls team must update its forecast when evidence changes, even while approval remains pending. The team should not preserve an implausible forecast to avoid appearing to request unauthorised funds. Forecasting the requirement and authorising the expenditure are separate responsibilities.

Uncertainty strengthens the need for that distinction. A forecast may contain a range of plausible outcomes, while an approved commitment still requires a defined decision basis. Acebes and colleagues propose an analytical approach to project control under uncertainty, using simulation and illustrative project networks to distinguish observed performance from expected variability (2014). Such methods can support interpretation, but their usefulness depends on credible assumptions and a model that represents the relevant work. A statistical threshold does not itself grant permission to spend or release a service. Nor is every movement inside a tolerance harmless. A small cost change may expose a critical dependency or breach an acceptance condition whose significance is poorly represented by its monetary size. Governance should define which types of consequence require referral regardless of their numerical magnitude. The resulting arrangement combines quantitative boundaries with explicit conditions, allowing control specialists to explain the evidence while preserving the authorised decision maker's responsibility for the response and its consequences.

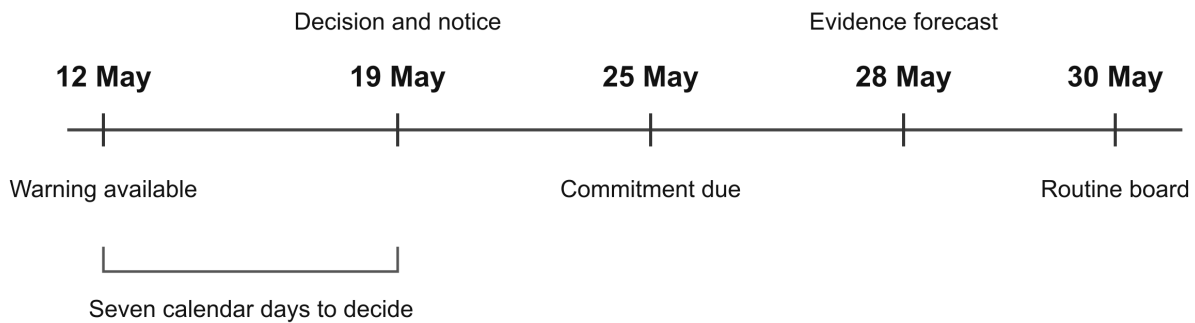
Set escalation timing from the work

A tolerance defines the room for action under a delegation, and an exception signals that the decision may need to move elsewhere. Useful tolerances identify their measure, boundary, period and response. A cost tolerance might concern forecast total expenditure against an approved ceiling; a schedule tolerance might concern a dependency needed by another programme. Neither should be interpreted as permission to conceal an adverse trend until the boundary has already been crossed. Escalation should occur when the current forecast or other credible evidence indicates that the delegated conditions are unlikely to hold, with time available to examine alternatives. Qualitative triggers may be equally important: an unresolved acceptance failure, an unsupported operating assumption or an unavailable specialist can warrant referral before a financial variance becomes large. The recipient needs to know the decision requested, the options available and the date after which those options narrow. Escalation is complete only when the concern receives a reasoned disposition or follows a defined further route if a response is overdue.

The timing of a governance meeting should therefore be tested against the commitment it serves. In a fictional example, a supplier commitment is due on 25 May and a local rule requires deferral notice to be received six calendar days beforehand, by 19 May. The sponsor sets the decision early enough that day for the commercial team to deliver the notice. On 12 May, the team forecasts that required evidence will not be available until 28 May. Its next routine board meets on 30 May. Waiting for that meeting would allow the useful decision window to close. Figure 3.3 makes the mismatch visible. The seven calendar days between the warning and the decision

deadline are available for clarifying options, obtaining evidence and convening the authorised response. These dates and the notice period are illustrative local rules, not universal scheduling guidance. The transferable practice is to work backwards from the consequence of delay. Where an approval requires consultation, commercial notice or operational preparation, those activities belong in the decision schedule as dependencies with owners and realistic durations.

A routine meeting can fall after the useful decision window



Escalate on 12 May because evidence is forecast after the commitment

Fictional dates • Positions are schematic • The six-day notice rule is local

Figure 3.3 Decision timing derived from a fictional commitment. A warning on 12 May must reach authority in time for a decision and receipt of any deferral notice by 19 May. Waiting for the 30 May board would miss the useful window. Spacing is schematic.

Research on temporality explains why a single reporting calendar cannot resolve every coordination problem. Söderlund develops a conceptual account of knowledge entrainment, illustrated with an in-depth transformation case, examining how differing knowledge processes and time orientations must be synchronised (2010). Dille, Söderlund and Clegg's processual study of an inter-institutional infrastructure project examines responses to competing temporal requirements across organisational boundaries (2018). These studies illuminate mechanisms rather than prescribe an optimal frequency of meetings. A multinational transformation may face quarterly investment approvals, monthly financial closure, weekly software releases and continuous operating obligations. The controls design should connect these rhythms where a shared decision requires it. Routine consolidation can follow a stable calendar while material exceptions travel through a faster authorised route. Information supplied between reporting cycles should state its date, maturity and reconciliation status. Otherwise, urgency can produce an unofficial account whose differences from the formal report remain unexplained and later undermine confidence in both.

The James Webb Space Telescope provides a documented example of changing governance connections during delivery. GAO's 2019 examination described NASA actions including management attendance at pre-test procedure reviews, a commissioning manager coordinating commissioning groups, and combined headquarters and centre status reviews. GAO also focused

on responsibility for sustaining the revised oversight arrangements (GAO, 2019, pp. 26–29). These were specific connections between technical work, coordination and management attention, rather than merely another statement of programme ambition. The report provides evidence that these arrangements were introduced; it does not isolate their causal contribution to the observatory's eventual performance. Their relevance here is the deliberate placement of oversight close to consequential work. A review of a test procedure can influence execution before an error is incurred, whereas a later cost report may only record its consequences. The controls professional should therefore ask which decisions need prospective examination and where routine retrospective reporting is sufficient, taking account of the reversibility and consequences of the next action.

Design information around the decision

The first question in designing management information should be what a named recipient needs to decide. Galbraith's information-processing account links uncertainty with the information demands placed on an organisation, and considers both reducing those demands and increasing the capacity to process them through vertical and lateral arrangements (1974). Applied to project controls, the implication is that adding reports can overload the very authority the system needs to inform. Local resolution within clear limits can reduce unnecessary referrals; direct coordination between dependent teams can resolve detail before senior attention is needed. Material choices still require a clear account of consequences and uncertainty. A decision paper should state the commitment under consideration, the relevant approved reference, the current expectation, the available options and the recommendation. Supporting detail should be traceable without forcing every reader to reconstruct the analysis. The appropriate level of compression depends on the decision, but the reasoning must remain recoverable, particularly where a favourable headline depends on an unresolved assumption.

Evidence quality requires more than a software connection between systems. Raymond and Bergeron's empirical study examines relationships between project management information systems, their use and benefits for managers and projects (2008). It supports attention to the usefulness of information, without establishing that purchasing an application creates success. Each important measure needs a definition, an owner, a source, a reporting date and an account of material limitations. Costs, progress and benefits may become available at different times. The organisation should distinguish a consistent reporting cut-off from an unrealistic claim that every source was updated simultaneously. Reconciliation should identify missing accruals, late progress evidence or an operating measure that covers an earlier period. A single authoritative record can preserve several legitimate perspectives, provided their boundaries are clear. The supplier's account of contractual completion and the operator's account of usable capability need not be forced into one percentage. The governance question is how their difference affects the next commitment and which evidence can resolve it.

Approvals also need to survive translation into action. The Dutch parliamentary inquiry into Fyra found that, following the loss of an underbody panel in January 2013, two of three safety measures were not implemented even though management assumed they had been. Communication procedures did not ensure that drivers knew and applied the instructions

(Parlementaire enquêtecommissie Fyra, 2015, p. 351). This finding concerns a specific breakdown in implementation; it should not be treated as a complete explanation of the service's failure. Its significance for controls is the gap between a decision recorded and an action verified. An action log should identify what observable change will demonstrate completion, who must provide that evidence and who accepts it. Sending an instruction, conducting a meeting or updating a status field may show activity without establishing the intended result. When the consequence is material, closure should depend on evidence from the work itself, with a route for reopening the action if subsequent observations contradict the original acceptance.

Benefits information must travel through the same governance connections while retaining its distinct meaning. Badewi's empirical research examines project management and benefits management practices together and proposes a project benefits governance framework (2016). Its relevance is the need to connect delivery responsibilities with the organisational arrangements through which benefits are pursued. In this chapter's application, the operational owner should review whether a delivery change alters the timing, feasibility or ownership of the expected improvement. A delayed release may postpone value; a changed process may shift workload; a reduced scope may remove the condition on which a benefit depended. These consequences belong beside cost and schedule effects in the investment decision. They should not be converted automatically into precise financial losses when attribution or valuation is uncertain. The controls team can present the changed assumption, its significance and the evidence needed for a stronger estimate. This preserves the connection to purpose without allowing an attractive benefit forecast to excuse an unsupported delivery commitment.

Apply the method to a real decision boundary

Return to the fictional manufacturer's shared purchasing and payment platform. The approved budget remains US\$12 million, while the current forecast is US\$14 million. The target date remains 30 June and the current forecast is 31 August. For this chapter's illustration, assume that 30 June is also the approved schedule baseline; the two retain distinct meanings even though their dates coincide. The critical release conditions introduced in Chapter 2 remain unmet. For this teaching example, the proposed supplier commitment on 25 May would begin additional correction work; existing delegated planning resources permit the team to develop options but do not authorise that contract change. The evidence forecast for 28 May concerns the integrated correction proposal, not completed release testing. This distinction prevents the example's dates from implying that a May decision accepts an August operating state. Table 3.1 sets out a Decision Connection Record that connects four different decisions to their evidence, authority and timing. The record applies Chapter 2's five-stage Credible Control Method. It adds governance detail to that method instead of creating a competing process that readers would have to reconcile.

Table 3.1 Decision Connection Record for the fictional manufacturer

Decision	Evidence and condition	Authority and owner	Deadline
Prepare options	Correction scope, reconciled estimate, dependencies and operating effects.	Delivery manager within existing planning delegation.	For the 19 May decision; identify missing evidence.
Protect deferral	Options, notice conditions and evidence that notice was received.	Sponsor under express delegation; commercial owner implements.	Decision early enough for receipt by 19 May.
Authorise funding	US\$14 million forecast, alternatives and proposed funding basis.	Investment authority decides; sponsor implements conditions.	Before additional commitment.
Accept and release	Required tests, recovery demonstration and operating readiness.	Release authority; finance and operations accept their conditions.	Before deployment.

The record separates preparation, protective action, funding and release. The forecast that proposal evidence will arrive on 28 May triggers the earlier deferral decision; it does not imply release readiness. All dates and delegations are illustrative.

Framing the decision means specifying what is being requested now. Here, the immediate question is whether to defer the supplier commitment, approve a bounded alternative or authorise the proposed correction when sufficient evidence is available. It is not a request for unrestricted continuation or an early declaration of operational readiness. The sponsor identifies the funding authority required for expenditure beyond the approved envelope, while the commercial owner confirms the local notice requirement and the consequences of each option. The controls lead records 19 May as the latest decision date and 12 May as the date on which the issue became visible. Operational acceptance remains a separate decision later in the sequence. The same person might hold more than one authority in a smaller organisation, but the decisions should retain their distinct evidence and conditions. The map thus describes functions that must be covered, rather than requiring additional job titles or a new committee for every function shown.

Establishing the position brings together the authorised reference, actual commitments and current forecast at an agreed cut-off. The delivery team defines the correction and its remaining uncertainties. The planner identifies its effect on integration and the availability of specialists; the estimator separates the correction effort from other causes of the US\$2 million forecast increase. The finance owner reconciles costs already incurred, outstanding commitments and the proposed additional expenditure. The operational owner explains effects on service continuity, staff preparation and the benefit assumptions. Each contribution should be supported at a level proportionate to the decision. A preliminary estimate can justify further investigation while being insufficient for a major commitment. The record should make that limitation visible rather than converting an early number into apparent certainty through repeated circulation. Where the account contains unresolved differences, the controls lead names their owners and explains

whether they could change the recommendation. Reconciliation is complete enough to decide when material differences are understood, not merely when every column adds correctly.

Challenging the account tests whether the proposed response follows from the evidence and whether another option deserves consideration. The reviewer checks the dependency logic, the basis of the estimate and the consequences of deferral. The operational owner tests whether the temporary workaround is feasible under ordinary workloads. A commercial specialist examines what can legitimately be changed within the existing arrangements. Where the evidence remains insufficient on 19 May, the authorised decision may be to defer commitment and commission defined further work within an available delegation. That response needs its own cost, duration and owner. If the authority instead approves an option, the record identifies the conditions and any remaining uncertainty it accepts. The controls team then updates authorised records while retaining the earlier baseline and decision history. An unresolved approval is never represented as completed because the recommended option appears likely. The sequence keeps an analysis, a recommendation and an authorisation distinguishable even when they are discussed in the same meeting.

Verification completes the connection. Following a deferral, the commercial owner supplies evidence that the required notice was issued and acknowledged where applicable. Following authorised correction work, the delivery owner demonstrates the changed configuration and the test evidence required for its acceptance. The operational owner retains responsibility for readiness and subsequent service performance. The next review examines whether the response changed the forecast and whether new dependencies require further action. A decision to spend can therefore be followed by a decision to accept work and, later, a decision to expand use, each with its own evidence. The record also keeps an overdue response visible. If funding remains undecided, the sponsor uses an expressly granted delegation to instruct protective deferral, with notice received by 19 May. Its permitted scope and financial limits are agreed in advance. The commercial owner verifies receipt; unresolved funding follows the further escalation route. Silence does not authorise commitment. This is a practical governance discipline: every material decision has a destination, every authorised action has an owner, and completion is examined against the consequence the action was intended to achieve.

Keep governance proportionate and test its connections

The map should remain usable in settings with different institutional arrangements. A public investment may require separate appropriation, procurement and operating approvals. A private organisation may delegate several decisions to one executive. An international partnership may need agreement from organisations whose powers overlap without forming a simple hierarchy. The common task is to identify the actual authorities and connect them to evidence and timing. The research and cases in this chapter support that inquiry, but they do not justify ranking national systems from a few examples or assuming that one structure travels unchanged. The independent audits, parliamentary inquiry and varied academic studies make different contributions. They can explain a mechanism, document a breakdown or identify an association without proving that the proposed Governance Connection Map will improve every transformation. Its practical value must be examined in use, with attention to local context and competing explanations.

An organisation can begin that examination with a small set of consequential decisions rather than a complete redesign of its reporting machinery. Trace a recent scope change, a forecast deterioration and an operating acceptance decision from the originating evidence to the resulting action. Establish whether the intended recipient had authority, whether the information arrived in time and whether the response was verified. Missing links suggest specific corrections: a clearer delegation, a shared dependency review, a revised reporting cut-off or a defined exception route. Measures of the process should avoid rewarding speed alone. A rapidly approved request may be poorly supported, while a justified pause may protect value. Review the timeliness of reasoned decisions, the visibility of unresolved obligations and whether authorised actions achieved their intended effects. These observations can inform improvement without turning governance into another aggregate score that hides decisive weaknesses. The objective is an inspectable connection between what the organisation knows and what it is prepared to commit.

Connecting governance to the work makes the first two chapters operational. The Transformation Control Loop provides the recurring logic; the principles of credible control protect the quality of its evidence and judgments; the Governance Connection Map identifies where those judgments become authorised action. None of these substitutes for professional competence or legitimate authority. Together, they direct attention to the points where an investment promise can become detached from delivery reality. A scope change must carry its consequences, a forecast must reach someone able to respond, and an approval must return to the work as an understood commitment whose effects can be checked. The next chapter examines the people and organisational relationships that sustain these connections. Sponsors, delivery managers, operational owners and controls specialists need distinct responsibilities, sufficient capacity and workable interfaces. The governance design developed here provides the questions their roles must answer before an organisation can reasonably expect its reporting system to influence transformation.

References

- Acebes, F., Pajares, J., Galán, J. M. and López-Paredes, A. (2014). A new approach for project control under uncertainty. Going back to the basics. *International Journal of Project Management*, 32(3), 423–434. doi: [10.1016/j.ijproman.2013.08.003](https://doi.org/10.1016/j.ijproman.2013.08.003)
- Ahola, T., Ruuska, I., Artto, K. and Kujala, J. (2014). What is project governance and what are its origins? *International Journal of Project Management*, 32(8), 1321–1332. doi: [10.1016/j.ijproman.2013.09.005](https://doi.org/10.1016/j.ijproman.2013.09.005)
- Badewi, A. (2016). The impact of project management (PM) and benefits management (BM) practices on project success: Towards developing a project benefits governance framework. *International Journal of Project Management*, 34(4), 761–778. doi: [10.1016/j.ijproman.2015.05.005](https://doi.org/10.1016/j.ijproman.2015.05.005)
- Biesenthal, C. and Wilden, R. (2014). Multi-level project governance: Trends and opportunities. *International Journal of Project Management*, 32(8), 1291–1308. doi: [10.1016/j.ijproman.2014.06.005](https://doi.org/10.1016/j.ijproman.2014.06.005)
- Chenhall, R. H. (2003). Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. *Accounting, Organizations and Society*, 28(2–3), 127–168. doi: [10.1016/S0361-3682\(01\)00027-7](https://doi.org/10.1016/S0361-3682(01)00027-7)
- Dille, T., Söderlund, J. and Clegg, S. (2018). Temporal conditioning and the dynamics of inter-institutional projects. *International Journal of Project Management*, 36(5), 673–686. doi: [10.1016/j.ijproman.2018.03.007](https://doi.org/10.1016/j.ijproman.2018.03.007)
- Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces*, 4(3), 28–36. doi: [10.1287/inte.4.3.28](https://doi.org/10.1287/inte.4.3.28)
- Government Accountability Office (GAO) (2015). GAO Schedule Assessment Guide: Best Practices for Project Schedules. GAO-16-89G, 22 December. www.gao.gov/products/gao-16-89g
- Government Accountability Office (GAO) (2019). James Webb Space Telescope: Opportunity Nears to Provide Additional Assurance That Project Can Meet New Cost and Schedule Commitments. GAO-19-189, 26 March. www.gao.gov/products/gao-19-189
- Government Accountability Office (GAO) (2020). Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Program Costs. GAO-20-195G, 12 March. www.gao.gov/products/gao-20-195g
- Joslin, R. and Müller, R. (2016). The relationship between project governance and project success. *International Journal of Project Management*, 34(4), 613–626. doi: [10.1016/j.ijproman.2016.01.008](https://doi.org/10.1016/j.ijproman.2016.01.008)
- Malmi, T. and Brown, D. A. (2008). Management control systems as a package—Opportunities, challenges and research directions. *Management Accounting Research*, 19(4), 287–300. doi: [10.1016/j.mar.2008.09.003](https://doi.org/10.1016/j.mar.2008.09.003)
- Müller, R. and Lecoivre, L. (2014). Operationalizing governance categories of projects. *International Journal of Project Management*, 32(8), 1346–1357. doi: [10.1016/j.ijproman.2014.04.005](https://doi.org/10.1016/j.ijproman.2014.04.005)
- Okhuysen, G. A. and Bechky, B. A. (2009). Coordination in organizations: An integrative perspective. *Academy of Management Annals*, 3(1), 463–502. doi: [10.5465/19416520903047533](https://doi.org/10.5465/19416520903047533)
- Parlementaire enquêtecommissie Fyra (2015). De reiziger in de kou. Tweede Kamer der Staten-Generaal, vergaderjaar 2015–2016, Kamerstuk 33 678, nr. 11. Presented 28 October. www.tweedekamer.nl/sites/default/files/atoms/files/rapport_dereizigerindekou_enquetcocommissiefyra_kst-33678-11.pdf
- Queensland Audit Office (QAO) (2010). Information systems governance and control, including the Queensland Health Implementation of Continuity Project. Auditor-General Report to Parliament No. 7 for 2010, June. www.qao.qld.gov.au/reports-resources/information-systems-governance-control-including-queensland-health-implementation
- Raymond, L. and Bergeron, F. (2008). Project management information systems: An empirical study of their impact on project managers and project success. *International Journal of Project Management*, 26(2), 213–220. doi: [10.1016/j.ijproman.2007.06.002](https://doi.org/10.1016/j.ijproman.2007.06.002)

- Söderlund, J. (2010). Knowledge entrainment and project management: The case of large-scale transformation projects. *International Journal of Project Management*, 28(2), 130–141. doi: [10.1016/j.ijproman.2009.11.010](https://doi.org/10.1016/j.ijproman.2009.11.010)
- Stoppels, J., Bosch-Rekvelde, M., Mooi, H. G. and Bakker, H. (2026). Roles and responsibilities in project steering committees. *Project Management Journal*, 57(4), 497–513. doi: [10.1177/87569728251406593](https://doi.org/10.1177/87569728251406593)
- Tiwana, A. (2010). Systems development ambidexterity: Explaining the complementary and substitutive roles of formal and informal controls. *Journal of Management Information Systems*, 27(2), 87–126. doi: [10.2753/MIS0742-1222270203](https://doi.org/10.2753/MIS0742-1222270203)
- Too, E. G. and Weaver, P. (2014). The management of project management: A conceptual framework for project governance. *International Journal of Project Management*, 32(8), 1382–1394. doi: [10.1016/j.ijproman.2013.07.006](https://doi.org/10.1016/j.ijproman.2013.07.006)

PART II

Designing the control system

People, design, economics, sequencing, information, assurance and uncertainty connected around decisions.



KENT BUSINESS COLLEGE
School of Project Control

CHAPTER 04

Authority Ownership and the Human System



IN THIS CHAPTER

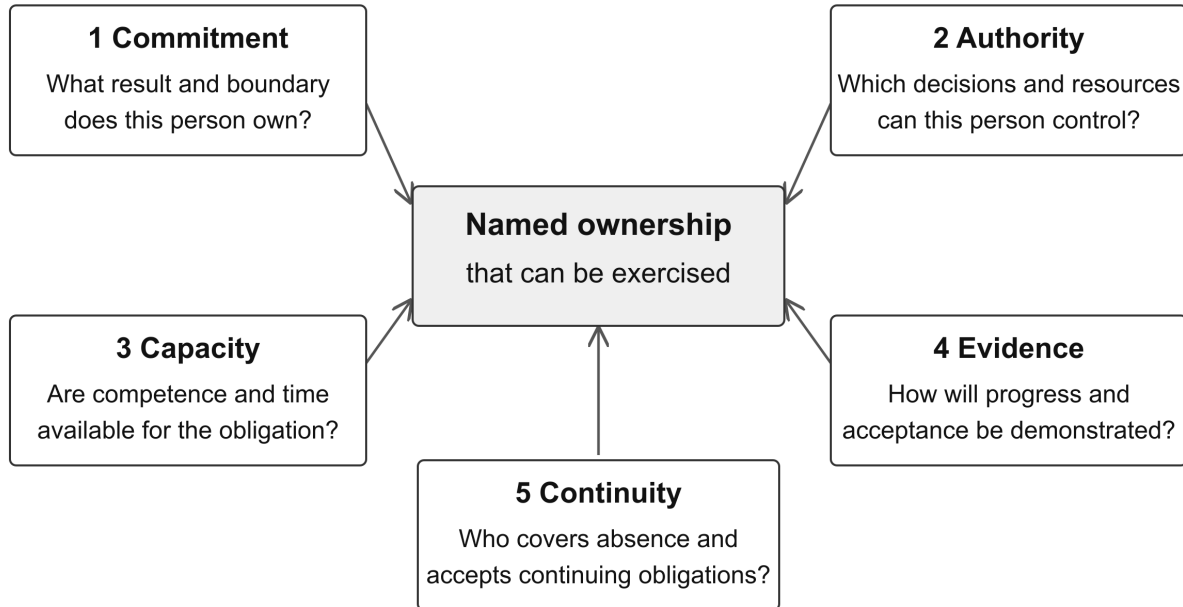
Giving people the authority, capacity and ownership to act.

An appointment must make action possible

The regional operations manager has been named as the owner of readiness for a new purchasing and payment service. Her name appears in the plan, the steering committee minutes and the handover checklist. Yet she still carries her ordinary service responsibilities, cannot release the staff needed for testing and has no agreed deputy. When readiness work falls behind, the report records an overdue action against her name. The record is accurate about the appointment but incomplete about the organisation that made it. This fictional situation extends the manufacturer introduced in the previous chapters. It illustrates a recurring design problem: assigning an obligation does not create the authority, competence, time or cooperation needed to fulfil it. Project controls should reveal that difference before a missed commitment becomes an argument about individual effort. The relevant question is whether the person can exercise the ownership being attributed to them, and what must change if they cannot.

Chapter 3 connected decisions to evidence, authority and deadlines. This chapter examines the human arrangements that make those connections workable. The proposed Ownership and Capacity Framework in Figure 4.1 tests five conditions: commitment, authority, capacity, evidence and continuity. Commitment defines the result and its boundary. Authority identifies the decisions and resources available to the owner. Capacity concerns competence and usable time, while evidence establishes how progress and acceptance will be demonstrated. Continuity covers absence, replacement and obligations that survive the project team. These conditions apply to sponsors, delivery managers, operational owners and controls specialists, although their actual obligations differ. The framework extends the book's existing control system rather than introducing a separate governance hierarchy. It is an original practice aid, not a validated scoring instrument. Its conditions should be examined together: additional expertise cannot compensate for missing authority, and an impressive mandate cannot supply hours that have already been committed elsewhere.

Ownership and Capacity Framework



Test every condition; a favourable average cannot compensate for a missing one

Figure 4.1 The proposed Ownership and Capacity Framework. Five conditions make an appointment examinable: commitment, authority, capacity, evidence and continuity. The framework extends the existing control system; it is not a numerical maturity or prediction model.

Clarity begins with the function being assigned. Responsibility concerns work to be performed; authority concerns permission to make decisions; accountability concerns answering for an obligation and explaining its fulfilment. Organisations use these words differently, so their intended meaning should be agreed rather than assumed. Zwikael and Meredith's conceptual analysis shows the inconsistency of project-role terminology and proposes distinctions between roles in funding and performing organisations (2018). The practical lesson is to define a role's purpose and boundaries before relying on its title. Here, an owner means an identified person answerable for a stated commitment, supported by explicit relationships with those who perform or authorise relevant work. That person need not execute every task or possess unilateral power over every dependency. Collective decisions may remain with a board, and partner organisations retain their own powers. A named lead makes coordination and explanation possible; the appointment should also make clear where their authority ends and who must resolve matters beyond it. A responsibility matrix can distinguish who performs work, who answers for it, who is consulted and who receives information. Its entries remain shorthand. The underlying agreement must identify the specific obligation, the person holding the role and the authority to resolve an exception. Otherwise, a completed matrix can conceal the same gaps as an organisation chart. The owner should know who will review these continuing obligations.

Separate the obligations behind familiar titles

The sponsor connects the transformation to the parent organisation's purpose and decision authority. This includes challenging the investment proposition, securing appropriate decisions and helping resolve priorities that exceed delivery management's mandate. Crawford and colleagues' qualitative research across 36 projects and programmes distinguishes governing a project from supporting it, with the balance shaped by context (2008). Helm and Remington's interviews with senior project managers also describe compensating behaviour when sponsorship is inadequate (2005). A capable manager can quietly find resources, negotiate access and protect a programme from interruptions, while leaving the executive deficiency invisible. That apparent resilience can become fragile when the manager leaves or a disagreement exceeds their influence. The sponsor's contribution should therefore be expressed through observable obligations: resolving a specified priority conflict, securing an operating commitment or presenting a continuation decision to the appropriate authority. Support does not mean defending every delivery proposal, and scrutiny does not mean personally directing all work. Both require informed attention at the points where executive involvement can change the result.

Ownership of benefits needs equally careful definition. Zwikael, Meredith and Smyrk's research develops project-owner responsibilities through executive interviews and a longitudinal organisational case, identifying an operations manager as a potentially suitable owner in operations-improvement settings (2019). Breese, Couch and Turner's study in one NHS hospital finds differing conceptions of sponsorship, including whether people understand it as a distinct responsibility with personal obligations for benefits (2020). These studies use role labels in particular ways; the handbook does not assume that their sponsor and owner are interchangeable in every organisation. The practical requirement is to name who will manage the operating changes on which each material benefit depends and who will examine whether improvement occurs. A finance specialist can validate a valuation method without controlling the service behaviour that produces the result. The benefit owner remains answerable for the agreed actions, assumptions and review process, while uncertainty about outcomes is retained. Appointing an owner cannot guarantee a benefit whose causes extend beyond the organisation.

Project and programme managers organise delivery within their mandates. A project manager coordinates the work needed for accepted outputs, including supplier interfaces, forecasts and corrective action. A programme manager addresses interdependencies among projects and operational activities where these require coordinated decisions. Neither role automatically absorbs the authority of the executive sponsor or the responsibility of a continuing service owner. Turner and Keegan's account of project-based organisations distinguishes relationship management at the client interface from stewardship of the resources needed for delivery (2001). This distinction helps explain why one title may conceal several demands. A manager can maintain a constructive customer relationship while lacking the power to release a specialist from another department. An integrated plan should show that dependency, and the relevant functional manager should confirm the resource commitment. Where a programme layer is unnecessary, these functions can be combined, provided their scope remains clear. Organisational complexity should follow the decisions and dependencies that need attention, rather than the prestige associated with adding another management level.

Project controls, finance and assurance contribute different forms of evidence and challenge. Controls specialists develop and reconcile the account of scope, time, resources, uncertainty and performance; delivery and operational managers remain responsible for the underlying work. Finance supplies accounting expertise and checks financial interpretations, while an assurance role examines claims with an appropriate degree of independence. The Institute of Internal Auditors' 2026 Three Lines Model distinguishes management responsibilities, specialist support and challenge, and independent internal audit, while emphasising coordination and safeguards against self-review (IIA, 2026). This professional guidance does not classify every project controller automatically as an auditor or a second-line function. The actual assignment matters. A controller who helps construct a forecast can exercise professional objectivity but should not present their own review as independent assurance of that same work. The organisation should specify evidence access, escalation rights and review arrangements so that useful collaboration does not erase the distinction between preparing a claim, accepting responsibility for it and independently examining its basis.

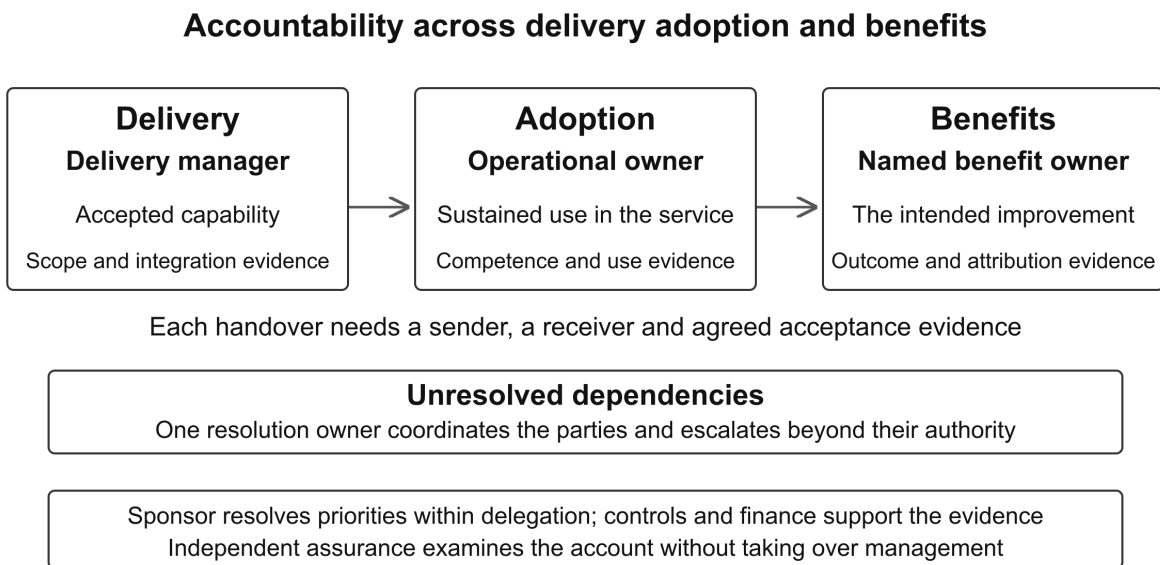


Figure 4.2 Accountability continues from accepted delivery through adoption to benefits. Arrows indicate dependencies, not an automatic transfer or discharge of responsibility. Unresolved interfaces require a resolution owner, while participating parties retain their own obligations.

Make the interfaces between people explicit

A matrix structure exposes people to more than one source of legitimate demand. A specialist may report to a functional leader for professional standards and staffing while receiving delivery priorities from a project manager. The difficulty is not simply that two names appear above them. It is whether those leaders have agreed how incompatible requests will be resolved. Ebbers and Wijnberg's qualitative study of dual leadership in film production examines role conflict, ambiguity and the space for negotiated coordination (2017). Its findings caution against treating all ambiguity as either harmful or desirable. A skilled individual may need discretion to connect neighbouring tasks, but discretion is different from uncertainty about who can set priorities. The controls response is to make competing commitments visible and locate the decision with the appropriate managers. A planner cannot solve a double booking by allocating the same specialist to two activities simultaneously. Nor should the specialist be expected to conceal the conflict through sustained unpaid availability or an unrecorded reduction in quality.

Interfaces also contain differences in knowledge and interests. Carlile distinguishes the transfer of information, the translation of meaning and the transformation of knowledge across boundaries where parties face consequential differences (2004). The paper develops a conceptual framework illustrated through an organisational case; it does not establish a compulsory sequence for every project. For transformation controls, the distinction helps explain why sending a specification may leave the receiving team unable to use it. A supplier may interpret completion as configuration delivered, while an operator understands it as a service that can handle exceptions during ordinary workloads. Agreement requires discussion of those consequences, not only common terminology. A useful handover describes what the sender supplies, what the receiver must be able to do, how acceptance will be established and what remains unresolved. The receiving owner should participate before the deliverable becomes difficult to alter. Otherwise, the organisation may discover at handover that an apparently complete output depends on operating work nobody had the time or authority to undertake.

New Zealand's Novopay school-payroll service went live on 20 August 2012. The June 2013 ministerial inquiry found that documented roles assigned project accountability to the Ministry's chief executive, delivery decisions to the Senior Responsible Officer and user-needs specification to the Business Owner. Yet participants disagreed about whether the Project Board advised or decided. The initial Business Owner's other responsibilities constrained attention to the project. Over four years, eight Ministry project managers, programme managers and programme directors held leadership roles. The inquiry also identified inadequate commercial experience for managing vendors and insufficient involvement of school users (Ministerial Inquiry into the Novopay Project, 2013, pp. 8–9, 15, 59–64 and 67). These findings concern the practical operation of roles, rather than an absence of formal structure. They do not attribute the entire failure to one individual. The controls implication is to test the customer's capacity to specify, challenge and accept the service, alongside the supplier's capacity to deliver it. A contract cannot supply missing operating knowledge by itself.

Formal descriptions establish expectations, but working relationships determine how those expectations are enacted. Bechky's ethnographic research across four film projects shows familiar

occupational roles providing continuity while participants negotiate their performance through everyday interaction (2006). The setting differs from a public service transformation, yet it illustrates why a responsibility chart cannot replace practical coordination. People need to know whom to approach, what another specialist considers adequate evidence and when a request changes the recipient's commitment. The proposed ownership agreement should therefore be discussed between neighbouring role holders rather than merely circulated for signatures. A supplier lead and an operational owner can examine an actual exception together; a planner and a finance partner can reconcile the same work package; a sponsor and a service manager can test a contested priority. These conversations reveal gaps that abstract descriptions often miss. They should result in recorded agreements where authority or commitments change, while leaving room for ordinary professional cooperation within those boundaries.

Treat stakeholder responses as information

Stakeholder influence shapes which concerns reach decision makers. Mitchell, Agle and Wood's conceptual framework explains managerial attention through perceptions of power, legitimacy and urgency (1997). It helps diagnose why a well connected supplier or an influential business unit may receive rapid attention while dispersed service users remain unheard. It should not be converted into an ethical rule that low-power groups matter less. A transformation can materially affect people who lack the time, confidence or organisational position to attend its meetings. The controls design should preserve channels through which their operating experience enters requirements, acceptance and performance review. Representation also needs a clear mandate: a person attending a workshop may provide valuable experience without being authorised to accept a change on behalf of every user. Record whose perspective has been obtained, which populations remain absent and how important differences will be investigated. The objective is a more complete account of consequences, rather than a stakeholder diagram that rewards proximity to the sponsor.

Objections deserve the same distinction between observation and interpretation. Piderit's conceptual analysis treats responses to change as potentially different across cognitive, emotional and intentional dimensions (2000). Someone can believe that a new service is useful, feel anxious about their competence and still intend to adopt it. Oreg, Vakola and Armenakis' review of 79 quantitative studies likewise places change reactions in the context of recipients, organisational conditions and the change process (2011). These sources support inquiry into the reasons behind a response. An objection may concern workload, professional identity, fairness, an unresolved defect or a perceived loss of control. The appropriate response depends on that meaning. Training will not resolve an authority dispute, and stronger advocacy will not repair an unusable workflow. The controls professional should connect a material concern to the condition it threatens and identify evidence that could clarify it. This approach allows legitimate challenge without assuming either that resistance is irrational or that every objection should prevent the organisation from deciding.

Victoria's Ultranet illustrates the work concealed by an adoption label. In December 2012, the Victorian Auditor-General reported that coaching had begun two years before deployment, when teachers could not apply learning in a live system. The department had underestimated

additional workload and had not provided sustained support. At the April 2012 operational transition, the IT Division assumed responsibility for the technical solution and the Student Learning Outcomes Division became the business owner, while some contracted functions remained undelivered. Business cases assumed use across all government schools, but departmental policy allowed schools to opt out (Victorian Auditor-General, 2012, pp. 26–27 and 30–31). The case therefore contained named owners alongside unresolved conditions for use. Its lesson for this chapter is to connect adoption responsibility with users' available time, practical support and local decision rights. A completed handover should preserve outstanding obligations rather than imply that their operating consequences have disappeared. The audit's findings describe those conditions without establishing a single explanation for the programme's difficulties.

Incentives can make these conversations more difficult. Eisenhardt's review of agency theory examines relationships in which parties have different interests, uneven information and varying attitudes towards risk (1989). The relevance to transformation is the possibility that locally reasonable behaviour produces an undesirable combined result. A supplier rewarded for delivery dates may favour rapid acceptance; an operating manager measured on uninterrupted service may avoid releasing staff for testing; a programme manager may defend scope because its removal threatens the original business case. The controls function should expose the interaction without assuming dishonest intent. Each party's obligation and incentive needs to be understood before a performance figure becomes an accusation. Shared objectives can help, but broad statements of partnership do not replace a decision about who carries a temporary cost or how a disputed result will be assessed. The sponsor should arrange a legitimate resolution of those trade-offs, while the evidence records both the agreed response and consequences that remain with another part of the organisation.

Test capacity before relying on commitment

A role can be clear and still be impossible to perform within the resources provided. Zika-Viktorsson, Sundström and Engwall's study of 392 people in nine Swedish companies examines overload in environments where individuals work across multiple projects (2006). Reported overload was associated with adverse experiences including strain and reduced opportunities for competence development, alongside schedule deviation. The study does not establish a universally safe number of projects per person. It does support examining the combined demand on individuals rather than treating each project's allocation as though it were their only obligation. The controls team should distinguish nominal allocation from usable capacity after existing service work, leave, required meetings and other commitments are considered. Skills also constrain substitution. Ten available hours from someone unfamiliar with payment exceptions may not replace ten hours from the specialist who understands them. Capacity assessment should therefore combine time, competence, timing and access to support. These are properties of the actual work arrangement, not qualities that can be inferred from a job title alone.

Readiness adds a collective dimension. Weiner's theory defines organisational readiness for change through shared commitment and shared belief in collective capability to implement the

change (2009). That belief should not be confused with an independently verified resource plan. People may be enthusiastic while underestimating the effort, or sceptical despite resources that appear sufficient on paper. The controls professional should bring these accounts together. Discussion can establish whether the change is considered worthwhile, while evidence about skills, staffing, dependencies and operating conditions tests whether the plan is feasible. A short readiness survey can help locate disagreement but should not become a substitute for observing the work. Differences between groups may be particularly informative: a central team may consider the service ready while local staff still cannot process an exception or obtain support. The relevant owner must investigate the discrepancy and identify the required response. Readiness becomes actionable when a concern is connected to a condition, an owner and a way of establishing whether that condition has improved.

The fictional manufacturer makes the capacity problem visible through a simple weekly account. Maya, the regional operational owner, has forty working hours available. Her proposed allocation contains twenty-four hours of continuing service work, ten hours of testing, eight hours of training and six hours of readiness preparation. The total is forty-eight hours, exceeding available time by eight hours. Figure 4.3 compares this allocation with a revised proposal. The arithmetic is simple so that the governance issue remains visible: naming Maya on every relevant action does not create the missing capacity. Nor would describing her as fifty per cent allocated to the transformation solve the contradiction without changing the activities. The figures are invented teaching assumptions, not measured task durations. In practice, the owner and line manager would examine how estimates were obtained, which duties can genuinely move and what interruptions are likely. The initial comparison identifies a gap that requires a resource or sequencing decision before the programme relies on the appointment.

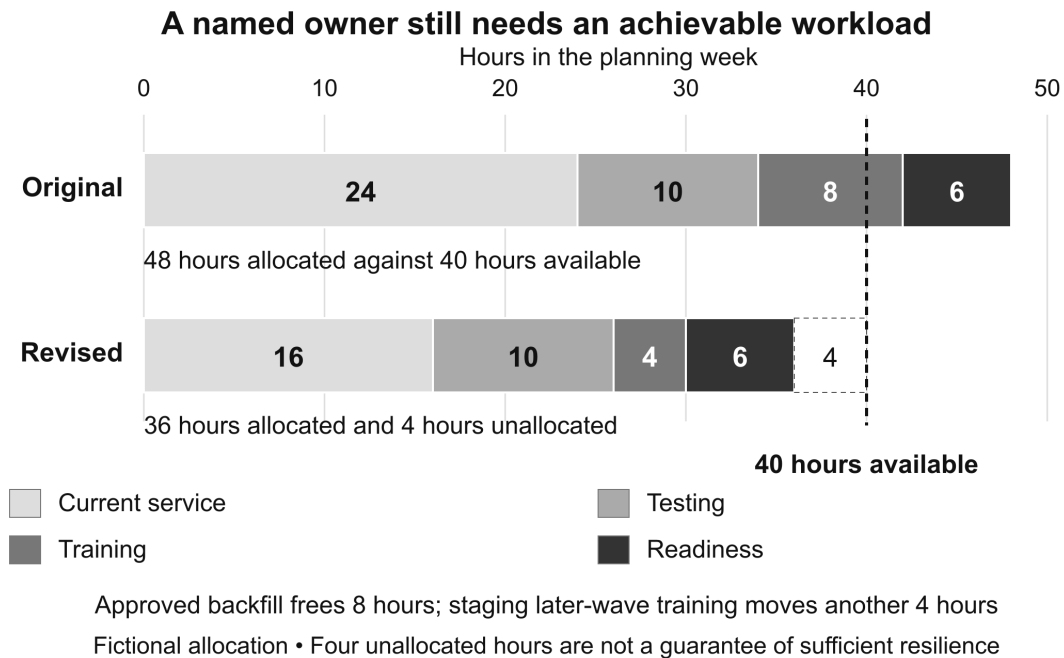


Figure 4.3 Fictional weekly capacity comparison. Original demand is 48 hours against 40 available. Approved backfill and explicitly deferred later-wave training reduce current demand to 36 hours. Four unallocated hours remain; the later training and backfill obligations still require management.

The revised proposal obtains approved backfill for eight hours of continuing service work and moves four hours of training for a later delivery wave to an explicitly scheduled later week. Maya's current allocation becomes sixteen hours of service work, ten of testing, four of training and six of readiness preparation: thirty-six hours, leaving four unallocated. The moved training remains an obligation, with an owner and a date before the later wave needs it. For this example, Maya reserves the deferred training in week three before the second wave in week four. That later allocation requires its own capacity check; the first wave's training record cannot demonstrate readiness for the second. The backfill also requires a competent person, an accepted handover and an authorised cost; it cannot be treated as free capacity. Any financial implication belongs in the current forecast and the relevant approval process. Four unallocated hours provide a visible allowance, but they do not prove that the plan can absorb every interruption. The revised arrangement must be checked against actual demand. Its value is that the sponsor and line manager can see what changed, which conditions support the new commitment and what evidence would require another adjustment.

Protect judgment while supporting delivery

Professional judgment is influenced by relationships and incentives as well as technical competence. Cain, Loewenstein and Moore's laboratory experiment examines conflicts between advice and incentives, finding that disclosure alone did not reliably protect the recipients of advice in the studied setting (2005). The study involved an artificial estimation task and should not be used to predict how practising project controllers will behave. Its useful warning is narrower: declaring a conflict does not necessarily remove its effect. A controls specialist involved in preparing a recovery estimate should be transparent about that role, while a consequential independent challenge may require another qualified person. A sponsor should similarly distinguish advocacy for a strategic objective from acceptance of a particular delivery claim. The organisation can support objectivity through access to underlying evidence, recorded assumptions and a route for unresolved professional disagreement. These arrangements make the quality of an account open to examination without demanding the unrealistic position that a knowledgeable participant must have no relationship with the work.

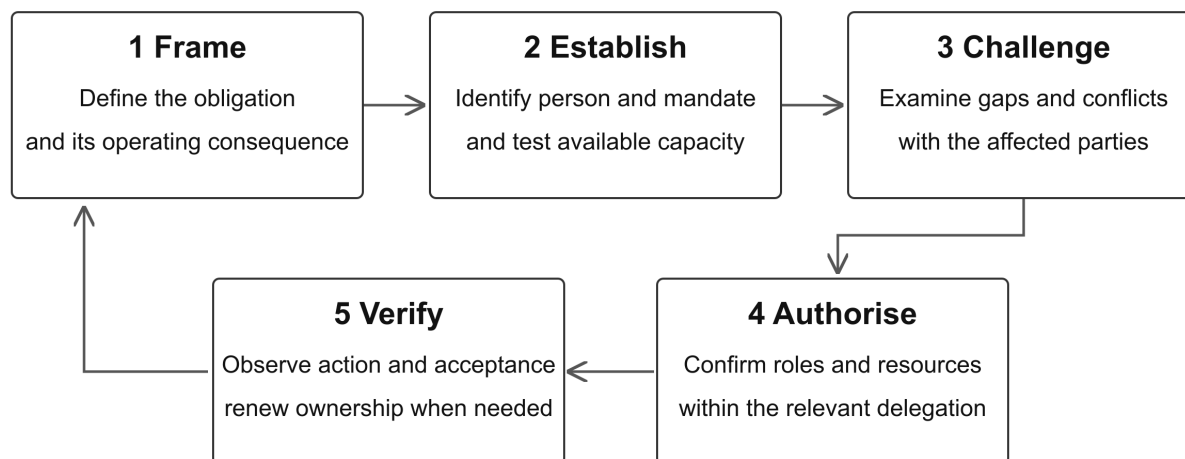
Constructive relationships are essential to that examination. The controls team needs enough proximity to understand the work and enough professional freedom to report what the evidence supports. Reporting to a delivery manager does not by itself establish either dependence or independence in every relevant sense; actual powers, access and escalation arrangements must be considered. A functional controls leader can help maintain methods and professional standards, while the project team provides operational context. Where these expectations conflict, the arrangement should state how the issue reaches the sponsor or another appropriate authority. Independence should not become an excuse for withholding useful analysis, and collaboration should not become a reason to alter an unwelcome forecast. The practical test is whether a material disagreement remains visible, receives an informed response and can be examined by someone with the required competence and authority. This builds on Chapter 2's principle of challenge while locating the people and working relationships through which it can occur.

Competence should be matched to the question rather than inferred from confidence or seniority. A capable scheduler may need support in evaluating adoption evidence; an experienced operations manager may need help distinguishing cash savings from released capacity. The organisation should identify the specialist knowledge required for consequential decisions and make it accessible before the decision is due. Development plans can combine formal learning, supervised application and review of actual work, with responsibilities increasing as competence is demonstrated. The controls lead should also examine whether the team's own reporting obligations leave time for analysis. If every available hour is consumed by assembling figures, the organisation has created a reporting service without protecting the capacity needed to interpret them. Simplifying duplicative reports or improving source quality can release useful effort, but promised efficiencies should be verified. A realistic staffing plan records both recurring production work and the analytical tasks necessary to understand emerging issues, challenge assumptions and support timely management action.

Mobilise ownership through the existing control method

The five-stage Credible Control Method from Chapter 2 can be used to establish these arrangements. Figure 4.4 applies its stages to ownership, retaining the same logic of framing, establishing, challenging, authorising and verifying. In the fictional manufacturer, Maya owns regional operational readiness and the agreed payment-reliability improvement; Alex manages delivery; Samira sponsors the transformation; and Lena leads project controls. These are invented people whose appointments make the example concrete. The first stage frames the obligation in terms of a result, its boundary and the decisions it requires. Maya's readiness obligation includes staff competence, operating procedures and support arrangements for the specified release. It does not make her responsible for personally correcting every supplier defect. Alex remains responsible for coordinating delivery and exposing dependencies, while Samira addresses priorities and commitments beyond their delegations. This allocation should be discussed with the role holders so that the same words describe the same obligations in the project plan and the continuing organisation.

Apply the Credible Control Method to ownership



Record the person, obligation, authority, capacity, evidence and continuity arrangements

Revisit the agreement when work, staffing, dependencies or operating conditions change

Figure 4.4 Applying the existing Credible Control Method to ownership. Frame the obligation, establish authority and capacity, challenge gaps, authorise the arrangement and verify its use. Revisit the agreement when its conditions change.

Establishing the position then tests whether the appointment can work. The team records the person's decision powers, relevant skills, available time, support arrangements and existing obligations. Maya's forty-eight-hour allocation reveals a capacity gap; the interface review identifies which supplier evidence she needs before accepting readiness. Lena traces these conditions into the integrated plan and current forecast without taking over Maya's operational judgment. The line manager confirms what work can be reassigned, while the appropriate financial authority considers any additional resource commitment. Where information is incomplete, the appointment can be bounded accordingly. Maya might lead investigation and preparation while a release decision remains conditional on further evidence. This prevents a provisional allocation from being presented as unconditional acceptance of the complete obligation. The same examination applies to Samira's sponsorship and Lena's controls role: their diaries, specialist support and authority must also be sufficient for the decisions and analysis on which the organisation intends to rely.

Challenge asks the neighbouring parties to test the proposed arrangement. Alex examines whether the operating tasks align with delivery dependencies; Maya checks whether supplier commitments cover realistic service conditions; finance reviews the resource implications; and assurance identifies where a separate examination would add value. The discussion should expose incompatible demands rather than distribute optimistic promises across several people. Authorisation then confirms the appointments and any supporting resource or sequencing changes through the relevant powers. An appointment letter or role agreement can record these decisions, but its length is less important than its clarity. It should identify the commitment, delegated limits, evidence obligations and route for resolving constraints. A person may

legitimately hold more than one role, provided the combined workload and potential self-review problems have been examined. Where a capacity or authority gap remains, the sponsor should amend the proposed commitment or seek a decision from the appropriate body. The agreement should not imply that willingness to help has removed a condition that remains unresolved.

Table 4.1 records the main ownership connections for the example. It separates delivery, adoption, benefits, unresolved dependencies and the control account. The distinction prevents the delivery manager's apparent completion from closing an operating obligation that belongs elsewhere. It also avoids assigning every uncertainty to the sponsor simply because the sponsor is senior. Samira owns the resolution route for the particular cross-organisational dependency shown, while the participating managers retain responsibility for their respective actions. The record identifies the evidence that permits review and the boundary beyond which referral is needed. In a live organisation, it would include contact details, deputies and links to the underlying delegations, kept in an appropriately controlled location. The published table remains compact so that the reader can examine the relationships rather than navigate an administrative form. Its purpose is to make these consequential obligations inspectable and usable in discussion, with the full role agreement carrying the detail required by the particular organisation.

Table 4.1 Ownership Agreement for the fictional manufacturer

Obligation	Named owner	Evidence for review	Authority boundary
Delivery	Alex—delivery manager	Accepted configuration and reconciled forecast of remaining work.	Act within delivery delegation; refer changes beyond it.
Adoption	Maya—operational owner	Staff competence, operating procedures, support and observed use.	Accept agreed service conditions; release authority remains separately defined.
Benefits	Maya—benefit owner	Payment reliability against the agreed baseline and explanatory factors.	Own operating actions and review; finance validates valuation.
Dependency	Samira—sponsor	Agreed supplier and service interface; unresolved conditions remain tracked.	Resolve within delegated powers; escalate beyond them.
Control account	Lena—controls lead	Reconciled delivery and operating evidence, assumptions and current forecast.	Maintain and challenge analysis; management authorises commitments.

Names and appointments are illustrative. The complete agreements retain deputies, current delegations and succession arrangements. Combining roles requires a review of capacity and potential conflicts; it does not merge their obligations.

Verification examines whether the arrangement is being exercised. Maya's participation in testing, staff competence evidence and accepted operating procedures provide a stronger account of readiness work than the presence of her name in a register. Lena compares actual demand on the team with the revised allocation and records conditions that require reconsideration. Alex

confirms that delivery changes reach the operating owner before their consequences are fixed. When a person leaves or becomes unavailable, continuity requires an authorised substitute and an accepted transfer of obligations. The successor needs the current position, unresolved assumptions, forthcoming decisions and access to the supporting evidence, including the reasons for material earlier choices. The transfer should identify its effective date and any limitations, while the relevant authority retains a route for matters awaiting acceptance. The organisation should not rely on an unattended inbox or an obsolete name to maintain accountability. Ownership persists through an explicit arrangement even when the individual carrying it changes. Where a successor has not accepted the role, the appointing authority must arrange interim coverage and state the limits of work that may continue. The gap should remain visible in the control account until the arrangement is confirmed.

Sustain ownership beyond the project team

Estonia's response to the 2017 ID-card vulnerability offers a bounded recovery example. The Information System Authority's account distinguishes digital functionality at that authority, identity documents at the Police and Border Guard Board, and certificates at SK ID Solutions. Following notification on 30 August 2017, public and private experts formed working groups and maintained a shared information portal. Remote updating began in October; high demand caused downtime, and additional service offices supported renewal when affected certificates were suspended on 3 November. The authority subsequently reported that, by 1 April 2018, 94 per cent of cards previously used electronically had been updated (RIA, 2018, PDF pp. 3–4). This is the responsible authority's retrospective account of an operating crisis, not an independent evaluation of a planned transformation. It illustrates coordinated technical, operational and communication work while institutions retained distinct responsibilities. It neither establishes uninterrupted access for every user nor proves that the role design caused recovery. Its relevance is the need to preserve operational coordination and user support after a digital service has been delivered.

The framework should itself be subject to review. A team can test a small set of consequential appointments by following an actual obligation through the five conditions. Was its meaning shared, could the owner make or obtain the necessary decisions, and did the promised time and competence exist? Did evidence reach the reviewer, and did unresolved work pass to an accepted successor? These observations can identify practical weaknesses without establishing that the framework causes better project outcomes. The underlying academic evidence includes conceptual arguments, interviews, ethnographic studies, surveys and a laboratory experiment; its settings and inferential limits differ. The international cases illuminate specific mechanisms rather than supply controlled comparisons of complete management systems. Readers should therefore use the framework to structure inquiry and improve arrangements, retaining the local institutional context. A successful allocation in one organisation may depend on relationships, powers or professional capabilities that another organisation must first establish. The common discipline is to test those conditions instead of assuming that the same titles carry the same capacity.

Authority and ownership make transformation a human system of commitments. People must be able to understand the result they are answerable for, obtain the decisions and support it requires, and demonstrate what happened. Clear records matter because they preserve that understanding, but they work through competent people and functioning relationships. Project controls contributes by making obligations, capacity gaps and unresolved interfaces visible, while preserving the distinction between analysis and the authority to act. The sponsor, delivery manager, operational owner and controls lead then have a shared basis for deciding what can responsibly be promised. That basis must survive staff changes and the eventual dissolution of the project team. The next chapter examines the object of these commitments: how to design a future operating state whose outputs, behaviours and benefits can be measured. Ownership becomes more useful when the claimed improvement is specific enough for its owner to explain what would demonstrate progress and what evidence would require a different response.

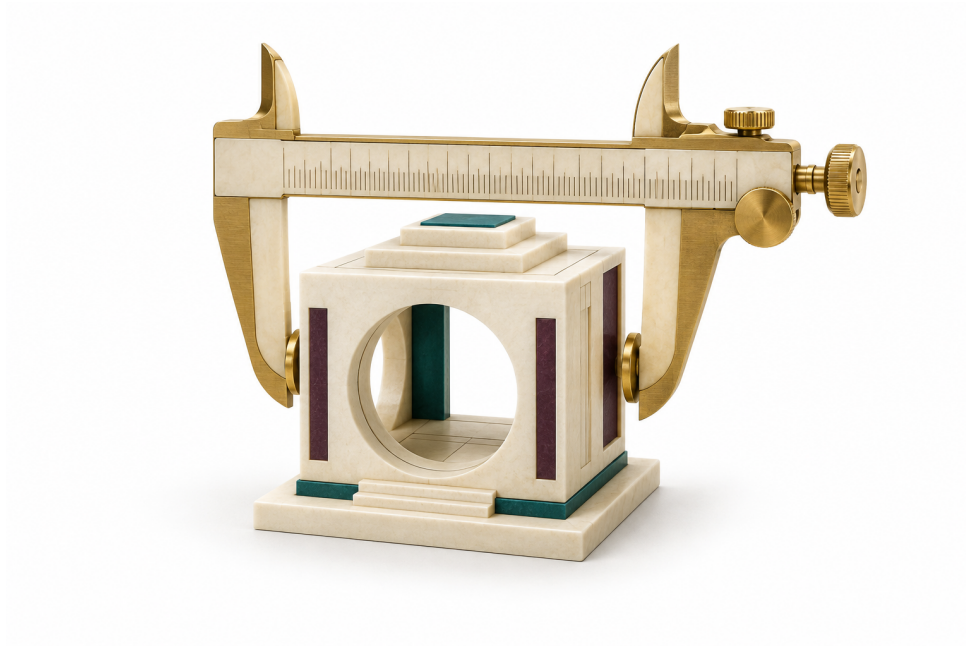
References

- Bechky, B. A. (2006). Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science*, 17(1), 3–21. doi: [10.1287/orsc.1050.0149](https://doi.org/10.1287/orsc.1050.0149)
- Breese, R., Couch, O. and Turner, D. (2020). The project sponsor role and benefits realisation: More than ‘just doing the day job’. *International Journal of Project Management*, 38(1), 17–26. doi: [10.1016/j.ijproman.2019.09.009](https://doi.org/10.1016/j.ijproman.2019.09.009)
- Cain, D. M., Loewenstein, G. and Moore, D. A. (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *The Journal of Legal Studies*, 34(1), 1–25. doi: [10.1086/426699](https://doi.org/10.1086/426699)
- Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, 15(5), 555–568. doi: [10.1287/orsc.1040.0094](https://doi.org/10.1287/orsc.1040.0094)
- Crawford, L., Cooke-Davies, T., Hobbs, B., Labuschagne, L., Remington, K. and Chen, P. (2008). Governance and support in the sponsoring of projects and programs. *Project Management Journal*, 39(1_suppl), S43–S55. doi: [10.1002/pmj.20059](https://doi.org/10.1002/pmj.20059)
- Ebbers, J. J. and Wijnberg, N. M. (2017). Betwixt and between: Role conflict, role ambiguity and role definition in project-based dual-leadership structures. *Human Relations*, 70(11), 1342–1365. doi: [10.1177/0018726717692852](https://doi.org/10.1177/0018726717692852)
- Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57–74. doi: [10.5465/amr.1989.4279003](https://doi.org/10.5465/amr.1989.4279003)
- Helm, J. and Remington, K. (2005). Effective project sponsorship: An evaluation of the role of the executive sponsor in complex infrastructure projects by senior project managers. *Project Management Journal*, 36(3), 51–61. doi: [10.1177/875697280503600306](https://doi.org/10.1177/875697280503600306)
- Information System Authority (RIA) (2018). ROCA Vulnerability and eID: Lessons Learned. Republic of Estonia. www.ria.ee/sites/default/files/documents/2022-11/Roca-vulnerability-and-eID-lessons-learned-2018.pdf
- Institute of Internal Auditors (IIA) (2026). Three Lines Model: Assurance and Advice in Support of Effective Governance. Statement of Position, 8 July. www.theiaa.org/globalassets/site/resources/statements-of-position/tl_m_assurance_advice_support_effective_gov_en.pdf
- Ministerial Inquiry into the Novopay Project (2013). Report of the Ministerial Inquiry into the Novopay Project. Wellington: New Zealand Government, June. img.scoop.co.nz/media/pdfs/1306/MIN130501InquiryReportweb.pdf

- Mitchell, R. K., Agle, B. R. and Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review*, 22(4), 853–886. doi: [10.5465/amr.1997.9711022105](https://doi.org/10.5465/amr.1997.9711022105)
- Oreg, S., Vakola, M. and Armenakis, A. (2011). Change recipients' reactions to organizational change: A 60-year review of quantitative studies. *The Journal of Applied Behavioral Science*, 47(4), 461–524. doi: [10.1177/0021886310396550](https://doi.org/10.1177/0021886310396550)
- Piderit, S. K. (2000). Rethinking resistance and recognizing ambivalence: A multidimensional view of attitudes toward an organizational change. *Academy of Management Review*, 25(4), 783–794. doi: [10.5465/amr.2000.3707722](https://doi.org/10.5465/amr.2000.3707722)
- Turner, J. R. and Keegan, A. (2001). Mechanisms of governance in the project-based organization: Roles of the broker and steward. *European Management Journal*, 19(3), 254–267. doi: [10.1016/S0263-2373\(01\)00022-6](https://doi.org/10.1016/S0263-2373(01)00022-6)
- Victorian Auditor-General (2012). Learning Technologies in Government Schools. Victorian Auditor-General's Office, PP No. 207, Session 2010–12, 12 December. www.audit.vic.gov.au/report/learning-technologies-government-schools
- Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science*, 4, article 67. doi: [10.1186/1748-5908-4-67](https://doi.org/10.1186/1748-5908-4-67)
- Zika-Viktorsson, A., Sundström, P. and Engwall, M. (2006). Project overload: An exploratory study of work and management in multi-project settings. *International Journal of Project Management*, 24(5), 385–394. doi: [10.1016/j.ijproman.2006.02.010](https://doi.org/10.1016/j.ijproman.2006.02.010)
- Zwikael, O. and Meredith, J. R. (2018). Who's who in the project zoo? The ten core project roles. *International Journal of Operations & Production Management*, 38(2), 474–492. doi: [10.1108/IJOPM-05-2017-0274](https://doi.org/10.1108/IJOPM-05-2017-0274)
- Zwikael, O., Meredith, J. R. and Smyrk, J. (2019). The responsibilities of the project owner in benefits realization. *International Journal of Operations & Production Management*, 39(4), 503–524. doi: [10.1108/IJOPM-02-2018-0086](https://doi.org/10.1108/IJOPM-02-2018-0086)

CHAPTER 05

Designing Change That Can Be Measured



IN THIS CHAPTER

Designing an operating change whose effects can be understood.

Turn the promise into an examinable claim

The purchasing platform has been installed, the demonstration is convincing, and the dashboard reports that most employees have signed in. The sponsor asks whether suppliers will now be paid more reliably. Nobody can yet answer because the project has measured deployment and access, while the promise concerns the performance of a payment service. This fictional situation continues the manufacturer introduced in earlier chapters. Its difficulty began before reporting: the design never specified how the delivered system, operating arrangements and changed behaviour would produce the promised improvement. A competent controls team can reconcile costs and dates without repairing that missing explanation. This chapter therefore treats measurable change as a design responsibility. The organisation must define what it expects to improve, for whom, through which operating changes, under what conditions, and with what evidence. Those questions make the promise examinable before expenditure turns an attractive assumption into an obligation that nobody knows how to test.

Chapter 4 established that ownership requires commitment, authority, capacity, evidence and continuity. Here the object of that commitment becomes more precise. The proposed Value and Evidence Map in Figure 5.1 connects outputs, usable capability, changed practice, operating outcomes and benefits. An output is something delivered; capability is an ability available under specified conditions; practice concerns what people actually do; an outcome is an observed change in performance or experience; and a benefit is an improvement valued against an agreed objective and reference position. The distinctions are working definitions for this handbook, not a claim that all professional vocabularies use these terms identically. A configured payment system is an output. The ability to resolve invoice exceptions before their due dates is a capability. Regularly using that process is changed practice. Fewer late payments are an outcome whose value must be examined for suppliers and the organisation. Every connection requires conditions that a diagram alone cannot satisfy.

The map extends Chapter 1's Transformation Control Loop and the principles developed in Chapter 2. It gives the loop's initial value claim a structure that can subsequently guide commitments, evidence and decisions. Each important arrow carries an explanation, a dependency and a question that evidence could answer. The map is an original practice aid informed by benefits management and evaluation research; its usefulness must be tested in application. It is not a validated prediction instrument or a formula for converting completed work into value. Readers should use it to expose gaps, competing explanations and decisions requiring investigation. The map can remain simple at the highest level while linking to detailed requirements, operating procedures and measurement records. Simplicity is helpful when it clarifies the argument; it becomes misleading when it removes the conditions on which that argument depends.

Benefits management research provides a foundation for this approach. Ward, Taylor and Bond (1996) examined practice across sixty UK organisations and developed a process perspective extending beyond the acquisition of information technology. Peppard, Ward and Daniel (2007) explained how benefits dependency networks connect investments to business changes and the benefits those changes are intended to enable. Their contribution directs attention to the work

between acquiring technology and obtaining value. For project controls, that work needs scope, resources, ownership and evidence just as technical delivery does. The practical implication is to begin a design discussion with the improvement and reason backwards towards the necessary operating changes, then examine the proposed sequence forwards for plausibility. This does not ensure success. It makes omissions discussable early enough for the organisation to change the proposal, establish a missing dependency or reduce a claim whose supporting conditions cannot credibly be supplied.

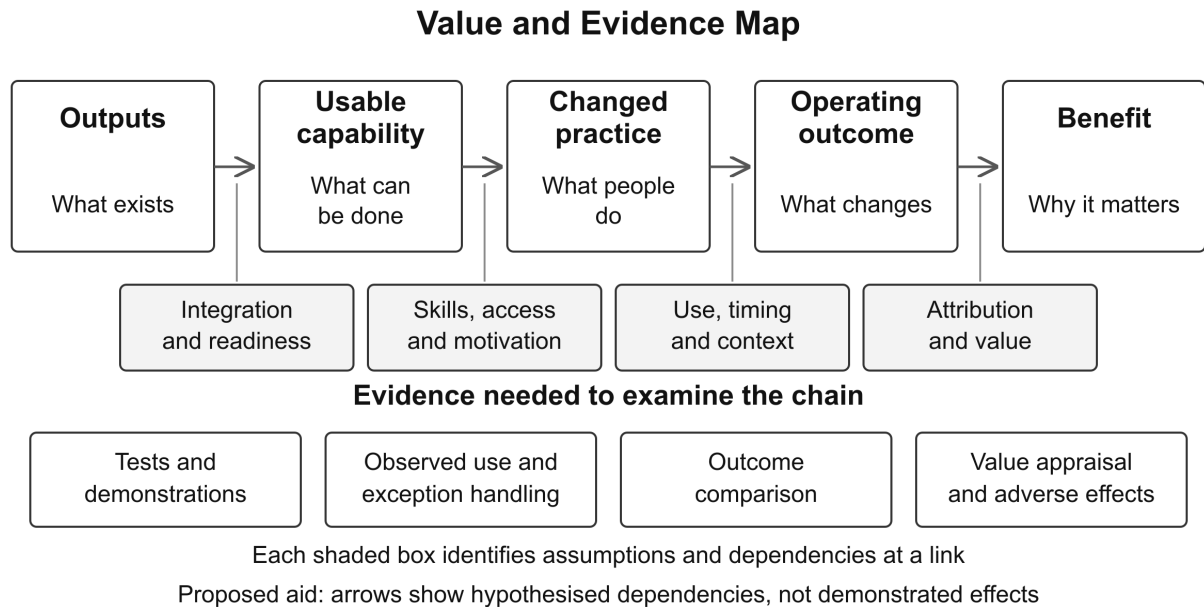


Figure 5.1 The proposed Value and Evidence Map connects delivered outputs to usable capability, changed practice, operating outcomes and benefits. Arrows express hypotheses and dependencies to examine, not proven causal effects. Evidence and assumptions accompany the pathway.

Strategic objectives determine which improvements deserve attention and how competing consequences will be judged. The same payment design could be justified through supplier resilience, reduced administration or stronger financial control, but these purposes are not interchangeable. A faster service may require additional staff; tighter checks may initially increase handling time. Before defining success, the sponsor should make the relevant priorities and constraints explicit, including who can resolve conflicts between them. The map can then connect several benefit claims to shared operating changes without counting the same improvement repeatedly. If two initiatives both claim the time released by removing one approval activity, the account needs a shared allocation or a single consolidated claim. The design should retain the disagreement where stakeholders value outcomes differently. A negotiated priority is an authorised choice whose reasons can be examined; it is not an objective fact created by assigning a weight in a spreadsheet.

Describe the operation that must become possible

A future operating model should describe how an organisation will perform the relevant work under ordinary and exceptional conditions. Its content includes people, processes, information, technology and organisational arrangements, with attention to the interactions between them. For the manufacturer, a diagram showing a supplier portal connected to an accounting system leaves important questions unanswered. Someone must recognise an incorrect invoice, have access to reliable purchasing data, obtain a decision from the appropriate manager and communicate with the supplier before the payment deadline. A system may route the case correctly while the designated resolver is unavailable or lacks authority. Figure 5.2 places the service flow alongside the conditions that make it workable. Its purpose is to help designers inspect a complete operating episode, including exceptions, instead of treating the organisation as a collection of independently satisfactory components.

Baxter and Sommerville's (2011) conceptual synthesis of sociotechnical systems engineering connects technical development with organisational change. It supports examining how work, people and technology interact, while also recognising difficulties in bringing these traditions together. Applied here, that perspective asks whether a design remains usable when it encounters the organisation's actual practices. A new approval rule may be technically enforceable but operationally unworkable when international time zones leave the approver unavailable. A process simplification may remove a visible administrative step while relocating its judgment to an already overloaded service team. The design review should therefore follow representative cases through the proposed operation, including the information, discretion and coordination each case requires. The lesson is a design question rather than a universal prescription: which interactions determine whether this particular service can produce the promised result, and how will the team investigate them before relying on the design?

Maya, the operational owner, can make the review concrete by walking through three fictional invoices: a routine transaction, an invoice with a disputed quantity and an urgent payment submitted during an approver's absence. Alex, the delivery manager, identifies what the platform and supplier must provide. Lena, the controls lead, records the dependencies and the evidence required before the team treats them as satisfied. The exercise should reveal more than user preferences. It may identify a missing delegation, inconsistent supplier identifiers, an unavailable banking interface or a support obligation absent from the estimate. These findings belong in the controlled design and integrated forecast. When a workaround is proposed, its staffing, duration and withdrawal conditions require the same attention. Calling it temporary does not remove the operating cost or the possibility that people will continue using it after the project team leaves.

An international example shows why the intervention boundary matters. Muralidharan, Niehaus and Sukhtankar (2016) studied a randomised rollout of biometric payment infrastructure in Andhra Pradesh, India. The intervention combined technology with changes to payment arrangements and local service delivery. Their evaluation reported that collecting an employment programme payment took about twenty-two minutes less in treatment areas, approximately twenty per cent of the comparison mean. The study also examined payment

access and retained attention to incomplete implementation and alternative payment routes. Its findings concern that combined intervention in its particular setting; they do not isolate the effect of a biometric device. For the controls professional, the design lesson is to specify what is actually being introduced. Equipment installation, payment agents, administrative rules and fallback arrangements may jointly determine the outcome. A schedule that tracks only the equipment can therefore omit activities central to the value claim, even where the evaluation of the complete intervention is strong.

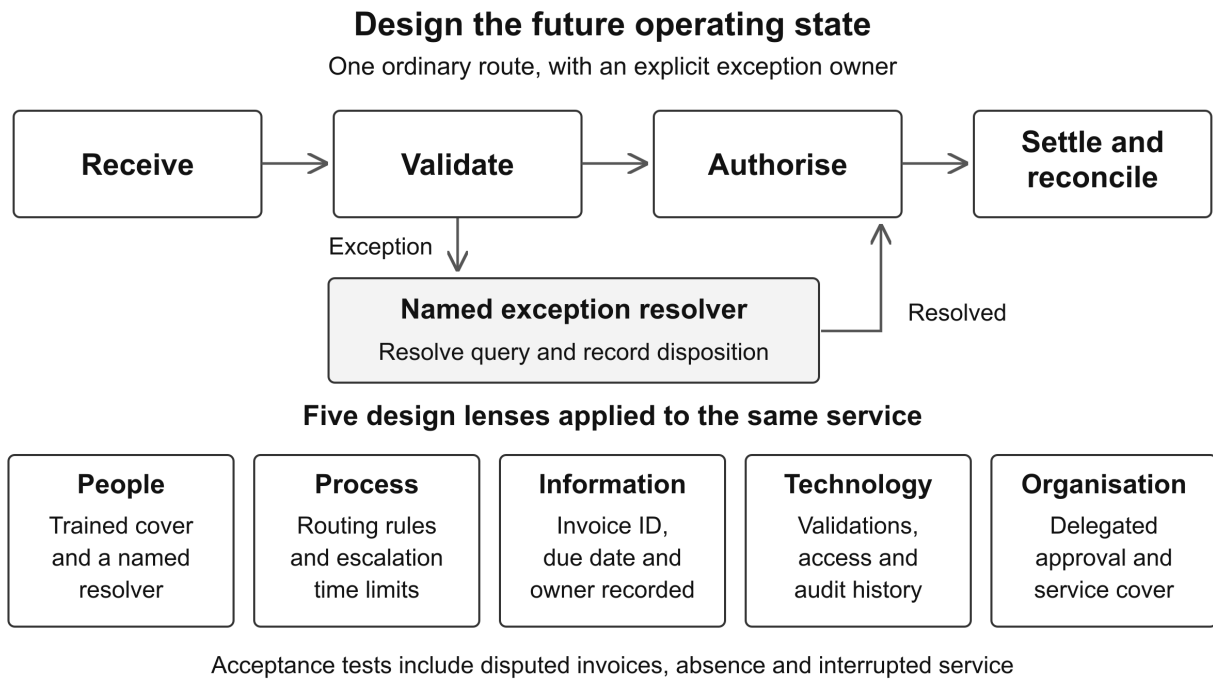


Figure 5.2 A fictional future payment operation. The routine flow and exception route depend on people, process, information, technology and organisational arrangements. The diagram supports an operating walkthrough; it does not replace detailed procedures or acceptance tests.

Explain why the proposed change should work

The connection between an operating design and an improvement is a theory that deserves examination. In the manufacturer, the proposed explanation might be that reliable invoice data and named exception owners make unresolved cases visible early enough for authorised staff to act, thereby reducing avoidable late payments. That sentence contains several assumptions: data arrive in time, the correct owner can be identified, staff can resolve the problem, and payment deadlines leave sufficient time for settlement. It also implies competing explanations. Late payment might principally reflect deliberate cash management, disputed commercial terms or unavailable funds, none of which a routing screen necessarily changes. A useful theory of change names these alternatives before selecting evidence. Otherwise the project may improve an activity that has little influence on the result, then interpret weak outcomes as employee resistance rather than reconsidering the design.

Markus and Robey (1988) examined the causal structures underlying explanations of technology and organisational change. Their conceptual argument helps distinguish accounts that treat technology as an independent cause from accounts concerned with human action and processes emerging over time. The relevance here is the need to explain the sequence and conditions through which a result is expected to arise. Rogers (2008) similarly examines how programme theory can represent interventions with multiple interacting components and emergent change. A single chain can communicate the central proposition, but it may need branches, feedback and contextual conditions to remain credible. The Value and Evidence Map should therefore be read as a navigational view of a larger argument. For example, improving invoice visibility might initially increase recorded exceptions because hidden cases become observable. That temporary increase could be consistent with better control rather than evidence that the service has deteriorated.

The theory becomes more useful when critical assumptions are converted into investigations. An assumption that supplier records are accurate can be examined through a representative reconciliation before migration. An assumption that managers can resolve disputes within the remaining payment window can be tested through timed operating scenarios and actual case histories. An assumption that a supplier will use the portal needs evidence from the affected supplier population, including organisations with limited administrative capacity. The team should state what an investigation can establish and what remains uncertain. A successful demonstration establishes that a particular scenario worked under its test conditions; it does not establish sustained performance at peak demand. The question for the next commitment is whether the evidence is sufficient for that commitment's scope and reversibility. This connects design to Chapter 3's distinction between investigating an option, authorising delivery and accepting a service into operation.

The UK's revised Magenta Book places evaluation within intervention design, including an explicit theory of change, assumptions, context and questions matched to the decisions the evidence must inform (HM Treasury, 2026, sections 2.2–2.3). It also explains why baseline data or a comparison arrangement may need to be established before implementation. This professional guidance supports treating evaluation as planned work. In the manufacturer's case, measurement access, data definitions and comparison arrangements therefore belong in the design brief, with owners and resources. They cannot safely be postponed until the sponsor asks whether the investment worked. The appropriate scale remains proportionate: a reversible local change may need a small investigation, while an expensive expansion with uncertain consequences deserves stronger evidence. The team should explain its choice of evaluation approach and the claims that approach can support.

Define measures that preserve the intended meaning

A measure is a rule for producing an observation, not simply a label on a dashboard. Payment reliability requires a defined population, a due date rule, a settlement event, a reporting period and a treatment of missing or disputed records. If one business unit measures the date it approves an invoice and another measures confirmed bank settlement, a combined percentage does not describe one consistent phenomenon. The definition should follow the intended improvement: suppliers need money when contractually due, so an internal approval timestamp is an intermediate event. The team must also decide whether the measure covers all valid invoices due in the period or only those processed through the new platform. Restricting the denominator to convenient transactions can remove exactly the difficult cases the transformation was meant to address. Definitions are therefore part of the design's substantive argument, not clerical detail added after the data arrive.

Cronbach and Meehl's (1955) discussion of construct validity concerns the relationship between a theoretical concept and the observations used to represent it. Its original setting is psychological testing, and applying its reasoning to controls requires care. It nevertheless provides a useful challenge: does the selected indicator capture the concept for which it is being used? DeLone and McLean's (2003) conceptual update of information systems success separates dimensions including system quality, information quality, service quality, use, satisfaction and net benefits. These distinctions explain why a login count cannot carry the whole claim of successful transformation. Frequent access may indicate productive adoption, repeated failed attempts or an inefficient process. The relevant observation depends on the question. For the payment service, evidence of correct exception resolution is more closely connected to the proposed mechanism than the number of hours employees spend inside the application.

Table 5.1 gives the manufacturer a compact Measurement Agreement. Its timeliness measure counts valid invoices due in the period that remain unpaid at the contractual deadline, using bank-confirmed settlement and retaining manual cases. Other records examine unauthorised payments, unresolved exceptions, staff effort and supplier burden. These are connected observations with different purposes. The first addresses the intended operating improvement; the others help explain the result and identify consequences that could make the apparent improvement unacceptable. Maya owns the operational interpretation, while finance checks payment and valuation definitions and Lena checks reconciliation and traceability. An analyst collecting the data should be able to reproduce the result from the stated rules. An affected supplier should be able to recognise whether the measure reflects its experience. Neither test requires every reader to inspect the database, but both require definitions capable of surviving challenge.

Table 5.1 Measurement Agreement for the fictional manufacturer

Measure	Operational definition	Evidence and owner	Review response
Payment timeliness	Late valid invoices divided by all valid invoices due; retain manual cases and flag disputes.	Contract dates and bank settlement; Maya, with finance reconciliation.	Compare periods consistently; investigate the causes of change.
Payment authority	Confirmed breaches of the agreed payment authorisation rules.	Control tests and transaction review; finance control owner.	Contain and investigate a serious breach regardless of the average result.
Exception age	Elapsed time for unresolved cases, including overdue and difficult cases.	Case timestamps and ownership history; Maya.	Investigate ageing and respond against agreed service thresholds.
Staff effort	Time spent on defined payment activities, with volumes and case mix retained.	Consistent observation or records; Maya with Lena.	Distinguish changed work, released capacity and actual expenditure.
Supplier burden	Time, access difficulties and unresolved issues across relevant supplier groups.	Structured supplier feedback and service records; Maya.	Examine whether improvement shifts work or excludes a group.

Illustrative summary. The complete agreement records population, period, inclusion rules, missing-data treatment, reporting cut-off, version, target and response thresholds. No numerical threshold here is offered as an industry standard.

Missing and delayed records deserve an explicit rule because they can systematically change the apparent result. An invoice with no reliable settlement record should not quietly disappear from the denominator or be assumed to have been paid on time. The report can identify unresolved records separately, explain their materiality and show how their eventual classification could affect the conclusion. Reconcile source totals with the reported population and investigate duplicate identifiers, reversals and corrections. Where a data migration alters historical fields, retain the source extract and the transformation rules needed to understand the resulting series. Observation timing matters as well: a reporting cut-off too close to a due date may precede confirmation from the bank. The team should agree the time allowed for reconciliation and distinguish provisional results from completed observations. These arrangements make uncertainty visible without turning every incomplete record into a presumed service failure.

Acceptance criteria, performance targets and response triggers should remain distinguishable. A critical payment authorisation control may require successful testing before release. A target to reduce late payments describes an intended operating result over a specified period. A response trigger, such as a serious unauthorised payment, calls for investigation or containment even when the average timeliness figure looks favourable. These statements serve different decisions

and need different evidence. They should be agreed early enough to avoid choosing a convenient interpretation after results become known. Where a definition changes for a valid reason, retain the earlier definition and show the effect of the change where feasible. A corrected measure may be better, but its discontinuity must remain visible. Otherwise an improvement created by excluding disputed invoices can be mistaken for an improvement in the service, and the control account begins to reward the redefinition of the problem.

Establish the comparison before interpreting improvement

An operating measurement baseline describes the starting position under specified observation rules. A counterfactual concerns what would have happened under an alternative condition, such as continuing the previous service. The distinction matters because the world continues changing while a project is delivered. Demand, staffing, regulation, supplier composition and other initiatives may alter the outcome even without the proposed transformation. A target is different again: it expresses an ambition, while a forecast expresses an expectation. Chapter 3 separated these concepts for delivery governance; the same discipline applies to operating performance. Retain the historical baseline, state the target and update the forecast as evidence develops. If an evaluation estimates a counterfactual, explain how it was constructed. Comparing an outcome with last year's value answers a useful descriptive question, but it does not automatically identify the project's effect. This remains true even when the difference is large and the chart looks persuasive.

Figure 5.3 provides an invented numerical example using the manufacturer. During a four-week baseline, the pilot unit has 120 late payments among 1,000 valid invoices due, a rate of twelve per cent. During a four-week pilot observation period, it has sixty among 1,000, or six per cent. A comparison unit changes from 100 late payments among 1,000 invoices to eighty among 1,000, moving from ten to eight per cent. The pilot unit's observed improvement is six percentage points; the comparison unit improves by two. Subtracting these changes gives a four-percentage-point contrast. Expressed another way, if the pilot unit would have experienced the comparison unit's two-point improvement without the intervention, its hypothetical rate would be ten per cent rather than the observed six. This is transparent arithmetic, not a demonstrated causal estimate. All counts, periods and rates are fictional, and the example supplies only one aggregate observation before and after in each unit.

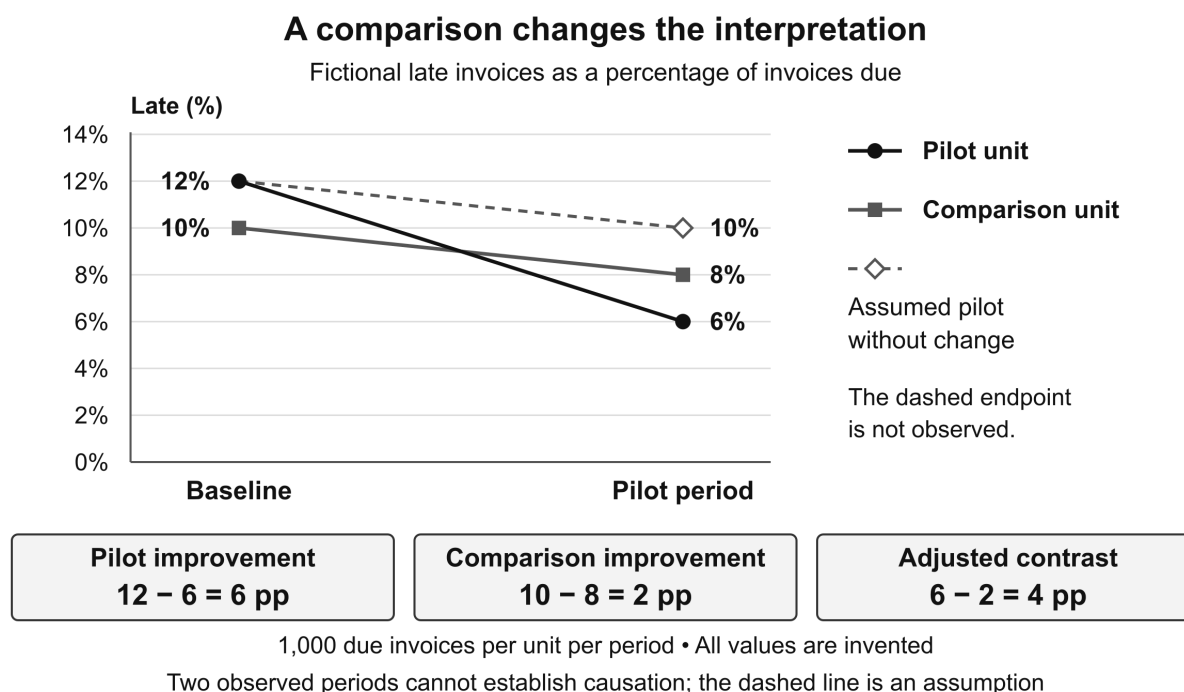


Figure 5.3 Fictional measurement comparison with 1,000 invoices due per unit in each period. The pilot improves by six percentage points and the comparison by two; the contrast is four. The dashed counterfactual is assumed, not observed. pp means percentage points. Two periods do not establish causation.

The example deliberately uses equal denominators to make the arithmetic transparent, but operational data rarely arrive so neatly. Reports should retain counts as well as rates, because six late payments out of one hundred and sixty out of one thousand have different statistical information. Transactions may also share suppliers, approvers or operating incidents, so they cannot automatically be treated as independent observations. A short favourable period may omit month-end workload or seasonal demand. Wing, Simon and Bello-Gomez (2018) review the design and inference requirements of difference-in-differences studies, including the choice of comparison and the assumptions needed for causal interpretation. Their methodological account reinforces the need for an analysis plan appropriate to the data. This chapter's contrast illustrates a question for that plan; it supplies neither a confidence interval nor a verified effect. Reporting such precision without the necessary design and analysis would conceal uncertainty behind a familiar calculation.

The assumed comparison requires more than similar labels on two business units. Lena would investigate whether their earlier trends, invoice mix, staffing, payment policies and exposure to other changes make the comparison credible. A pilot selected because its manager is unusually experienced may behave differently from later sites. Staff transferred from the comparison unit could also affect both units' results. Several earlier periods can help investigate trends, but cannot prove that unobserved influences would remain identical. Hernán and Robins (2016), writing about observational health research, make the design of a causal comparison explicit through eligibility, strategies, follow-up, outcomes and analysis. Applying that discipline here requires evaluation expertise appropriate to the setting; a two-line chart does not implement their method. The immediate controls contribution is to preserve who was included, when the change began,

which version was used and what else changed, so that the analysis does not acquire confidence by losing its context.

London's congestion charge illustrates the distinction in an actual operating system. TfL's sixth monitoring report found that vehicles with four or more wheels entering the original central zone during comparable charging hours were twenty-one per cent fewer in 2007 than in 2002. Yet measured congestion in 2007 had returned to the pre-charge level. TfL explained that effective road capacity had changed, including through road works and space allocated to other transport priorities, and distinguished the historical comparison from its assessment of conditions without charging (TfL, 2008, pp. 3–5, 40–41, 55–56). Traffic volume and experienced delay were different measures; the report's counterfactual was an assessment, not an observed uncharged version of the same London. The case supports retaining the starting record while investigating changing conditions. Returning to a historical level does not by itself establish that an intervention had no effect, just as reducing an intermediate measure does not establish that the intended outcome improved.

Investigate implementation as well as outcomes

When performance fails to improve, at least two broad explanations deserve examination: the intended change may not have been implemented, or its underlying explanation may be inadequate under the actual conditions. These possibilities require different responses. More training will not necessarily repair an intervention aimed at the wrong cause. Replacing the design may be premature if the relevant operating practice never became available. Moore and colleagues' (2015) guidance on process evaluation examines implementation, mechanisms and context together, using quantitative and qualitative evidence where appropriate. Its origins are in complex health interventions, but its questions can inform an organisational investigation. For the manufacturer, a result should be accompanied by evidence about whether exception owners received cases, had usable information and took the intended actions. Observing those activities strengthens understanding of the mechanism; it does not independently establish the attributable size of the benefit. Important questions should be planned in advance, with later exploratory explanations identified honestly.

Surgical checklist research offers a useful comparison of implementation and outcomes. Haynes and colleagues (2009) studied consecutive patients before and after a checklist programme in eight hospitals. Among patients aged at least sixteen undergoing noncardiac surgery, inpatient complications declined from 11.0 per cent of 3,733 baseline patients to 7.0 per cent of 3,955 subsequent patients. Recorded mortality fell from 1.5 to 0.8 per cent; follow-up ended at discharge or thirty days after surgery, whichever occurred first. The intervention included local implementation support and changes in team practice. These were prospective before-and-after associations, rather than results from randomised assignment. The useful design question is what staff actually did differently, including communication and checks at consequential moments. Treating the intervention as the distribution of a document would describe only a small part of the operating change that the study sought to introduce and observe.

Urbach and colleagues (2014) subsequently examined checklist adoption in Ontario, analysing 109,341 procedures before adoption and 106,370 afterwards across 101 hospitals. Adjusted operative mortality changed from 0.71 to 0.65 per cent, a difference that was not statistically significant; the change in complications was also not significant. Implementation was not standardised and formal team training was not required. The study used administrative data and three-month observation windows separated from adoption by three months. Its findings do not establish that checklists cannot work, and the earlier study does not guarantee the same result everywhere. Populations, implementation and measurement differed, so subtracting the studies' changes would not estimate the effect of implementation quality. Read together, they strengthen the case for specifying the intervention as enacted practice, investigating context and preserving the limits of each comparison. A shared intervention name is insufficient evidence that two organisations implemented the same operating arrangement.

Pilots should consequently be designed around consequential uncertainties rather than presented as miniature proof of an entire investment case. Skivington and colleagues (2021) emphasise programme theory, context, stakeholder perspectives, refinement and uncertainty throughout the development and evaluation of complex interventions. Their framework is methodological guidance, not a comparative demonstration that following it guarantees success. Applied to the manufacturer, an initial pilot might investigate whether exception resolution works during peak demand, while a later evaluation examines payment outcomes over an informative period. The team should record adaptations because evidence from one configuration may not apply unchanged to another. A pilot supported by unusually experienced staff also needs a credible account of how later sites will obtain sufficient competence and support. Expansion is a new commitment with its own evidence requirements. The decision should state what has been learned, what remains uncertain and which conditions must accompany the next operating wave.

Examine adverse effects and whose experience is represented

An improvement can be real and still conceal unacceptable consequences. Faster payments might depend on staff bypassing authorisation, transferring administrative work to small suppliers or delaying difficult invoices outside the measured population. These are design possibilities to investigate, not allegations about the fictional team. The measurement agreement should pair the main outcome with a small number of consequential adverse-effect measures and distributional views. For example, report payment reliability for suppliers using the portal and those requiring an alternative route, while retaining the combined result. Examine high-value transactions and recurrent disputed cases where consequences differ from the average invoice. The choice of breakdowns should reflect plausible mechanisms and affected groups, with adequate data and appropriate privacy protection. Collecting every available attribute creates cost and exposure without necessarily improving a decision. A proportionate design identifies what the aggregate could conceal and why that possibility matters.

Performance measurement itself can influence behaviour. Franco-Santos, Lucianetti and Bourne (2012) reviewed seventy-six empirical papers concerning contemporary performance measurement systems, identifying consequences for behaviour, organisational capabilities and performance, with substantial variation across settings. Their review excluded government and

not-for-profit organisations and was not a pooled estimate of a universal performance gain. Bevan and Hood (2006), examining English public healthcare targets, found evidence of gaming alongside reported improvement and could not determine how much improvement was genuine or offset elsewhere. Together these sources justify examining how a measure will be used, including its relationship with incentives and sanctions. They do not justify assuming every favourable result is manipulated. In the manufacturer, a timeliness target that ignores disputed cases could encourage exclusion instead of resolution. The design should make exclusions visible and provide a route for staff to report situations in which meeting the headline measure would undermine the service's purpose.

Adverse-effect evidence also needs interpretation. An increase in recorded exceptions after introduction may reflect better detection, a genuine deterioration or both. Staff interviews can help explain how the classification changed, while reconciled transaction records test its extent. Supplier complaints may increase because reporting becomes easier, so complaint counts should be read with access and resolution evidence. Conversely, an absence of complaints does not establish that every supplier can obtain help. The same reasoning applies to a zero count of unauthorised payments in a small pilot: it is a useful observation, but does not prove that the design eliminates the risk. Critical control tests and investigation of relevant failures remain necessary. The controls professional should resist converting every ambiguous signal into a favourable explanation. Record the alternatives, identify evidence that could distinguish them and connect the investigation to a responsible decision rather than leaving uncertainty as a permanent footnote.

Value must finally be examined from the perspective of those who bear costs and receive benefits. A reduction in internal processing time may be accompanied by additional supplier administration; a central service may improve the corporate average while making an isolated location less accessible. These effects belong in the same account, although they need not all be converted into money. Where staff time is released, record the activity reduced, the period and the credible alternative use. Multiplying minutes by a salary rate gives an indicative valuation under assumptions, not proof of reduced expenditure. Chapter 1 established this distinction, and Chapter 6 will examine the economics more fully. Here its design implication is immediate: measures must capture the operating change before valuation is attempted. An attractive monetary total cannot repair uncertain evidence about whether the underlying activity changed, who benefited or what additional work appeared elsewhere.

Trace design changes through to their consequences

Requirements traceability connects a need with the decisions, work and evidence intended to address it. Gotel and Finkelstein's (1994) investigation of requirements traceability, informed by more than one hundred practitioners, emphasised problems in tracing the origins and development of requirements as well as their downstream implementation. This matters when a seemingly minor design change removes a condition supporting a benefit. A requirement to identify an exception owner should retain the reason it exists: unresolved cases need a responsible person able to obtain action before payment deadlines. A link to a software component and its test is useful, but insufficient if the operating purpose has disappeared from view. Traceability should therefore preserve rationale, affected stakeholders and evidence alongside technical references. The amount of detail should match the consequence of losing the connection; a usable, maintained record is more valuable than a comprehensive register that nobody can interpret or keep current.

Figure 5.4 follows a fictional proposal to remove a mandatory exception-owner field because configuration is taking longer than expected. The immediate delivery effect may be less development and testing work. The operating effect could be cases remaining unassigned, creating longer resolution times and greater late-payment exposure. That consequence is conditional: an effective alternative allocation mechanism could preserve ownership without the field. The review must examine the mechanism, rather than defend a particular screen as inherently necessary. Maya tests the alternative procedure against realistic cases; Alex identifies delivery and supplier implications; Lena connects the changed scope to operating effort, acceptance evidence and the benefit forecast. Samira receives the integrated proposal at the appropriate decision boundary. This applies Chapter 3's governance connections and Chapter 4's ownership conditions to a design choice whose significance would be understated in a report limited to the supplier's saving.

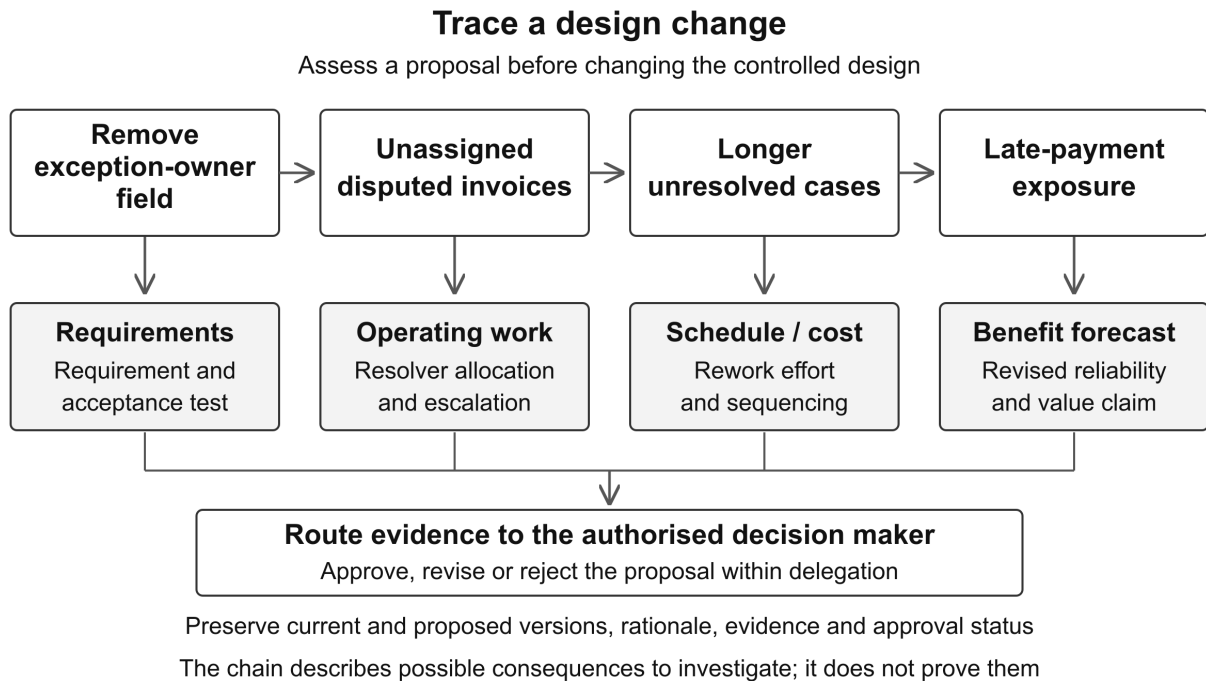


Figure 5.4 Tracing a proposed change through delivery and value consequences. Removing an exception-owner field could weaken case resolution unless an effective alternative preserves ownership. Review the mechanism, affected evidence and forecast before authorisation; retain the design and decision history.

Maintaining these links is continuing work with a cost and an owner. A useful review can sample a consequential requirement and follow it from the original need through its current specification, implemented version, acceptance evidence and expected operating effect. If the chain breaks, the team should repair the relevant connection and examine whether similar gaps affect other claims. This is different from requiring every minor design item to carry a lengthy justification. Concentrate detail where an error could change acceptance, investment or operating value. Automation can help identify references affected by a revision, but a person still needs to judge whether the relationship remains meaningful. A document link can remain technically valid after the underlying assumption has failed. The review should therefore examine the substance of the dependency and the continuing relevance of its evidence, rather than count the number of completed links as a proxy for control quality.

The record should distinguish the current approved design, the proposed change, the forecast consequences and the eventual authorisation. Evidence gathered against one configuration must remain identifiable when another is introduced. Otherwise a passed test can survive in the report after the condition it examined has been removed. Dependencies require similar care. A platform release may depend on a finance policy, a supplier data correction or a banking service that the project does not control. Each consequential dependency needs an owner, required condition and date connected to the operating sequence. Its uncertainty should influence the forecast without silently altering the approved commitment. When the design changes, revisit the affected benefit claims and measurement rules as well as the delivery plan. This is how a controls system retains purpose while adapting: it preserves the reasons for earlier decisions and makes the consequences of new choices visible before they are accepted.

Apply the existing method to a measurable commitment

The five-stage Credible Control Method provides a practical route from the map to an authorised commitment. Framing begins with Maya's proposed improvement in payment reliability, the population affected and the operating boundary. Establishing then assembles the current evidence, the future operating description, the measurement agreement and the dependencies. The four-week example supplies a historical observation, not a complete investment justification. The target of six per cent late payments should be identified as an illustrative ambition whose adequacy depends on the commercial context; it is not a recommended industry standard. Nor does reaching that number alone authorise expansion. Critical acceptance conditions, supplier experience and unresolved risks still matter. Lena should ensure that the decision pack separates delivered work, observed practice, operating outcomes and remaining uncertainty. Each statement can then be assessed on its own evidence instead of borrowing credibility from another favourable measure.

Challenge examines the mechanism and the proposed comparison before the results become a success narrative. Mayne's (2015) work on useful theory of change models makes assumptions between results explicit, including the connections among capacity, behaviour and benefits. The Value and Evidence Map draws on that reasoning while adding the handbook's links to control evidence and decisions. In practice, Alex and Maya should identify which link is least well supported and what investigation could change their judgment. Finance examines whether the proposed operating benefit could affect expenditure or other valued outcomes. Affected suppliers test whether the design improves their experience and preserves access. Evaluation specialists challenge the comparison and the claims it can sustain. The discussion should result in specific work or a bounded decision, not a general instruction to obtain more data. Additional measurement is worthwhile when it can resolve an uncertainty material to the next commitment.

Authorisation confirms the scope of the next step, the conditions on which it may proceed and the responsibility for examining its results. The sponsor may approve a bounded pilot while withholding approval for wider rollout, or require a revised operating design before further commitment. The authorised record should identify which assumptions remain open and how they will be tested. Verification then examines actual use, outcomes and adverse effects, retaining the applicable design version and measurement rules. Ashurst, Doherty and Peppard's (2008) research across twenty-five information systems projects found that benefits practices were not applied consistently and comprehensively. This supports attending to what the organisation actually does with its benefits arrangements, rather than relying on the existence of a template. The Measurement Agreement becomes useful when people use it to reconcile evidence, investigate a discrepancy and change a decision. Its completion is an enabling activity whose value depends on subsequent practice.

Badewi (2016) examined relationships among project management, benefits management, governance and success using survey data and structural equation modelling. The findings support considering delivery and benefits management together, while the design does not establish a universal causal gain from adopting a particular framework. That boundary fits this handbook's argument. Controls professionals should connect the disciplines sufficiently for a

delivery change to trigger reconsideration of operating value, and for weak benefit evidence to inform the next delivery commitment. The mechanism is organisational work: people exchange evidence, revise assumptions and exercise the relevant authority. A benefits register disconnected from scope and decisions cannot supply that integration through its existence alone. The practical test is whether an emerging observation, such as an unresolved exception burden, changes the current account and reaches someone able to authorise a response before further exposure is created.

Keep the explanation open to evidence

Some transformations cannot support a credible quantified counterfactual within the available setting and resources. That does not remove the need to investigate causation or permit every observed improvement to be attributed to the programme. Mayne's (2001) contribution analysis develops a reasoned account using programme logic, evidence about results, alternative explanations and further investigation where the account is weak. It aims to reduce uncertainty about contribution; it does not automatically assign a percentage of benefit to an intervention. For the manufacturer, such an investigation could combine transaction evidence, the timing of operating changes, case histories and interviews, while examining changes in staffing or payment policy. Contradictory evidence should remain in the account. The resulting conclusion may support a credible contribution with explicit limits, while a separate descriptive report states exactly what happened. This distinction allows the organisation to learn without converting a plausible explanation into a more precise claim than its evidence supports.

Measurable design is consequently an agreement about what the organisation intends to change and how that intention will remain open to examination. Its quality depends on connections: a benefit must link to an operating result, the result to a plausible mechanism, the mechanism to deliverable conditions, and each consequential claim to evidence and responsibility. The diagrams and records in this chapter make those connections visible within the control system already established. Their effectiveness should be assessed by whether they expose a missing condition, improve an investigation or inform an authorised decision. A reader can begin with one important value claim and follow it through an actual operating case before extending the approach. The next chapter examines whether the resulting proposal remains economically worthwhile and affordable. That judgment becomes more credible when the organisation first knows what improvement it is valuing, whose experience it represents and which conditions still have to be achieved.

References

- Ashurst, C., Doherty, N. F. and Peppard, J. (2008). Improving the impact of IT development projects: The benefits realization capability model. *European Journal of Information Systems*, 17(4), 352–370. doi: [10.1057/ejis.2008.33](https://doi.org/10.1057/ejis.2008.33)
- Badewi, A. (2016). The impact of project management (PM) and benefits management (BM) practices on project success: Towards developing a project benefits governance framework. *International Journal of Project Management*, 34(4), 761–778. doi: [10.1016/j.ijproman.2015.05.005](https://doi.org/10.1016/j.ijproman.2015.05.005)

- Baxter, G. and Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17. doi: [10.1016/j.intcom.2010.07.003](https://doi.org/10.1016/j.intcom.2010.07.003)
- Bevan, G. and Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538. doi: [10.1111/j.1467-9299.2006.00600.x](https://doi.org/10.1111/j.1467-9299.2006.00600.x)
- Cronbach, L. J. and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. doi: [10.1037/h0040957](https://doi.org/10.1037/h0040957)
- DeLone, W. H. and McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9–30. doi: [10.1080/07421222.2003.11045748](https://doi.org/10.1080/07421222.2003.11045748)
- Franco-Santos, M., Lucianetti, L. and Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research*, 23(2), 79–119. doi: [10.1016/j.mar.2012.04.001](https://doi.org/10.1016/j.mar.2012.04.001)
- Gotel, O. C. Z. and Finkelstein, A. C. W. (1994). An analysis of the requirements traceability problem. *Proceedings of the First International Conference on Requirements Engineering*, 94–101. IEEE Computer Society Press. doi: [10.1109/ICRE.1994.292398](https://doi.org/10.1109/ICRE.1994.292398)
- Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A.-H. S., Dellinger, E. P., Herbosa, T., Joseph, S., Kibatala, P. L., Lapitan, M. C. M., Merry, A. F., Moorthy, K., Reznick, R. K., Taylor, B. and Gawande, A. A., for the Safe Surgery Saves Lives Study Group (2009). A surgical safety checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine*, 360(5), 491–499. doi: [10.1056/NEJMsa0810119](https://doi.org/10.1056/NEJMsa0810119)
- Hernán, M. A. and Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, 183(8), 758–764. doi: [10.1093/aje/kwv254](https://doi.org/10.1093/aje/kwv254)
- HM Treasury (2026). Magenta Book: Central Government guidance on evaluation. Updated 15 May 2026. www.gov.uk/government/publications/the-magenta-book/magenta-book-central-government-guidance-on-evaluation-html
- Markus, M. L. and Robey, D. (1988). Information technology and organizational change: Causal structure in theory and research. *Management Science*, 34(5), 583–598. doi: [10.1287/mnsc.34.5.583](https://doi.org/10.1287/mnsc.34.5.583)
- Mayne, J. (2001). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16(1), 1–24. doi: [10.3138/cjpe.016.001](https://doi.org/10.3138/cjpe.016.001)
- Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, 30(2), 119–142. doi: [10.3138/cjpe.230](https://doi.org/10.3138/cjpe.230)
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D. and Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258. doi: [10.1136/bmj.h1258](https://doi.org/10.1136/bmj.h1258)
- Muralidharan, K., Niehaus, P. and Sukhtankar, S. (2016). Building state capacity: Evidence from biometric smartcards in India. *American Economic Review*, 106(10), 2895–2929. doi: [10.1257/aer.20141346](https://doi.org/10.1257/aer.20141346)
- Peppard, J., Ward, J. and Daniel, E. (2007). Managing the realization of business benefits from IT investments. *MIS Quarterly Executive*, 6(1), 1–11. aisel.aisnet.org/misqe/vol6/iss1/3/
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48. doi: [10.1177/1356389007084674](https://doi.org/10.1177/1356389007084674)
- Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M. and Moore, L. (2021). A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. *BMJ*, 374, n2061. doi: [10.1136/bmj.n2061](https://doi.org/10.1136/bmj.n2061)

- Transport for London (TfL) (2008). Central London Congestion Charging: Impacts Monitoring. Sixth Annual Report, July.
content.tfl.gov.uk/central-london-congestion-charging-impacts-monitoring-sixth-annual-report.pdf
- Urbach, D. R., Govindarajan, A., Saskin, R., Wilton, A. S. and Baxter, N. N. (2014). Introduction of surgical safety checklists in Ontario, Canada. *New England Journal of Medicine*, 370(11), 1029–1038. doi: [10.1056/NEJMsa1308261](https://doi.org/10.1056/NEJMsa1308261)
- Ward, J., Taylor, P. and Bond, P. (1996). Evaluation and realisation of IS/IT benefits: An empirical study of current practice. *European Journal of Information Systems*, 4(4), 214–225. doi: [10.1057/ejis.1996.3](https://doi.org/10.1057/ejis.1996.3)
- Wing, C., Simon, K. and Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual Review of Public Health*, 39, 453–469. doi: [10.1146/annurev-publhealth-040617-013507](https://doi.org/10.1146/annurev-publhealth-040617-013507)

CHAPTER 06

The Economics of Continuing



IN THIS CHAPTER

Testing the value, affordability and consequences of continuing.

Treat the next commitment as a new decision

The sponsor has explained why the organisation needs a new purchasing and payment service. Contracts have been signed, employees have invested time in preparation, and money has been spent. Now the remaining estimate has increased, the operating design has changed and the next payment requires approval. The original case for investment still matters as a record of what was promised, but it cannot answer the present question by itself. The decision concerns what the organisation should do with the resources it can still influence. This chapter develops that distinction through the fictional manufacturer used throughout the handbook. Its central argument is that a continuation decision needs an updated comparison of feasible futures, with explicit costs, benefits, obligations and uncertainty. Project controls makes that comparison credible by connecting the remaining work to the operating changes and evidence on which value depends.

Chapter 5 explained how an intended improvement becomes measurable. Economic appraisal asks what that improvement is worth relative to the resources required and the alternatives available. The answer can change even when the strategic purpose remains important. A payment service may still need improvement while a particular design no longer represents the best available response. Equally, an overrun does not establish that stopping would now be preferable. The relevant alternatives could include completing the present design, reducing its scope, changing the operating approach, pausing for information or ending the intervention while maintaining a necessary service. Each has consequences that should be examined from a common decision date. The practical discipline is to distinguish the accountability question about earlier commitments from the allocation question about resources still available. Both deserve an answer, and neither should be used to suppress the other.

Figure 6.1 introduces the proposed Continuation Review Map. It brings together value from the review date, affordability and funding, and operating feasibility, while retaining the historical account and nonfinancial conditions. The map extends the existing Transformation Control Loop rather than creating another lifecycle. It uses Chapter 2's Credible Control Method to move from a framed question through evidence and challenge to authorisation and verification. Its components are questions that must be answered together, not scores to be averaged. A favourable value estimate cannot supply cash on a date when the treasury has none, and available funding cannot make an unusable design worthwhile. The map is an original practice aid, informed by the research developed below. Its purpose is to expose the conditions supporting a recommendation and identify what evidence or authority is still missing before the next commitment can responsibly be made.

The economic review also needs a clear perspective. An organisation's incremental cash receipts differ from social benefits, accounting profit and changes in service quality. A payment penalty avoided by one organisation may be revenue lost by another; an environmental improvement may matter without appearing in either organisation's bank account. State whose costs and benefits are included, the geographical and organisational boundary, the comparison condition and the time horizon. Make the perspective visible in the title of the analysis and maintain it through the calculations. Combining a private cash saving with a wider social benefit can be

legitimate in an explicitly structured appraisal, but silently adding them produces an ambiguous total. The same discipline applies to transfers between business units: a central team cannot claim a corporate saving merely because it has moved an expense into another department's budget.

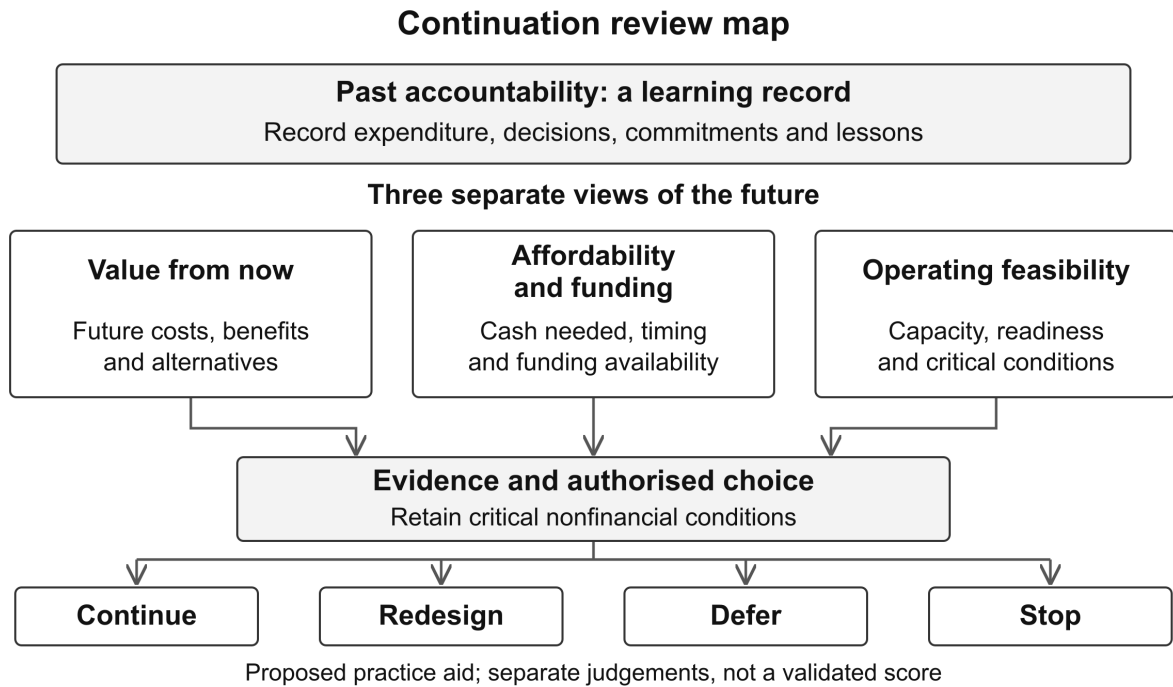


Figure 6.1 The proposed Continuation Review Map. Value, funding and operating feasibility inform a bounded choice, with the historical account and critical nonfinancial conditions retained. The map is a practice aid, not a scoring model or automatic decision rule.

Separate value affordability and available funding

Value concerns the consequences of using resources in one way rather than another. Affordability concerns whether the organisation can sustain the relevant expenditure and exposure within its financial constraints. Funding concerns the authorised sources and timing of the money required. These questions interact, but their answers can differ. A worthwhile service improvement may require a payment before funds become available; a funded project may have weak remaining benefits; an inexpensive pilot may create an unaffordable obligation to operate or expand. The review should therefore show the timing of gross expenditure and receipts as well as any net figure. A positive total over several years does not prevent a cash shortage next month. It also does not establish that an unapproved funding source will materialise. Controls should retain the difference between a proposed financing arrangement, an approved allocation and money available for use.

The Green Book illustrates why the appraisal perspective must remain explicit. HM Treasury's 2026 guidance concerns social costs and benefits in UK government decisions and distinguishes discounting from inflation. It excludes sunk costs from forward appraisal while retaining the opportunity cost of resources already owned, and treats economic transfers as offsetting social costs and benefits (HM Treasury, 2026). Those rules should not be confused with a company's

cash budget. A transfer can affect the organisation's ability to pay even when its social effects offset. The worked example in this chapter therefore uses an organisational cash-flow perspective and an expressly assumed discount rate; it is not a public-sector social appraisal. The professional task is to apply the relevant institutional rules consistently and explain material effects that fall outside the selected monetary account, rather than borrowing a rate or classification because it produces a convenient result.

Atkinson and Mourato's (2008) review of environmental cost-benefit analysis examines developments in valuation and the challenges of incorporating consequences that market transactions do not adequately represent. Its relevance to transformation is the need to examine the boundary of a monetary result. Some effects can be valued with a defensible method; others require explicit description, distributional analysis or separate criteria. A service transformation may affect accessibility, organisational resilience or trust in ways that a short cash-flow model does not capture. The review should distinguish an effect omitted for lack of evidence from one judged immaterial and one included elsewhere. These are different reasons with different consequences. A decision maker can then consider the monetary result alongside the remaining evidence without pretending that every important consequence has been converted into a precise financial amount.

Critical conditions should be distinguished from preferences that can be traded. A proposed saving obtained by removing a necessary payment authorisation control is not simply an attractive financial score to balance against a weak control score. The organisation must identify which conditions are required for an option to be feasible and who may accept any remaining exposure. Among feasible options, explicit nonfinancial criteria can illuminate differences in service, resilience and distribution. If a weighted comparison is used, define the scales, explain the weights and test whether plausible changes alter the ranking. Avoid counting a benefit once in money and again under a qualitative criterion without acknowledging the overlap. The aim is a reasoned comparison that preserves important judgments, not a composite number whose apparent precision hides disagreements about what the organisation values or what it is permitted to accept.

Build the remaining estimate from the work

An investment review needs an estimate whose scope matches the decision. The remaining delivery cost should include the work required to reach the specified acceptance condition, including integration, testing, data preparation, operational support and supplier interfaces where relevant. It must be reconciled with costs already incurred, accruals, outstanding commitments and proposed changes without counting the same obligation twice. A purchase order is not automatically an additional cost on top of the estimate if its work is already included. Conversely, an invoice not yet received does not make completed work free. The estimate should state its cut-off, technical and operating assumptions, exclusions, price basis and treatment of uncertainty. The US Government Accountability Office's estimating guide structures this work around scope, schedules, assumptions, data, methods, documentation and updates with actual costs (GAO, 2020). Its guidance supports a reproducible account rather than a number whose basis exists only in an estimator's memory.

Estimating methods should follow the maturity and character of the work. Analogy can help when comparable completed work exists, provided differences in scale and context are explained. Parametric methods relate cost or effort to relevant drivers and require data that support the relationship. Detailed estimating builds from activities, quantities and resource rates, but can create false confidence when the underlying scope remains unstable. Expert judgment remains necessary, especially where a transformation changes unfamiliar operating arrangements. Jørgensen's (2004) review of expert software effort estimation examines practices including relevant experience, historical evidence, independent approaches and explicit consideration of uncertainty. It does not establish that one method will always be most accurate. A proportionate review can compare independent estimates, investigate their differences and retain the assumptions that explain them. Agreement obtained by asking everyone to adjust to the sponsor's preferred figure is not independent corroboration.

Whole-life thinking extends the boundary beyond delivery. Licences, infrastructure, support, training, process ownership, upgrades, retirement and residual obligations may continue long after the project team leaves. Gluch and Baumann (2004) discuss the theoretical basis and practical limitations of life-cycle costing, including difficulties in representing environmental consequences and uncertain future conditions. Their analysis cautions against treating a complete-looking spreadsheet as a complete account of value. For the manufacturer, operating costs need owners who understand the service model and can explain how demand, defects or supplier changes would affect expenditure. The relevant horizon should reflect the decision and the useful life of the alternatives, with residual value or continuing liabilities considered consistently. If evidence is unavailable for a distant period, record the limitation and test its importance. An arbitrary end date can improve an apparent result by excluding expensive retirement work or the benefits of a longer-lived alternative.

The FBI's Sentinel programme provides a concrete cost-boundary example. In September 2012, the US Department of Justice Inspector General reported an FBI implementation estimate of US\$441 million against a US\$451 million budget, but noted that the comparison excluded two years of operations and maintenance that had been included in that budget. The report separately discussed estimated annual operations and maintenance costs of US\$30 million and noted that the US\$441 million development estimate excluded FBI employee costs; some legacy functions also continued (DOJ OIG, 2012, pp. 3–7). These were the estimates and boundaries examined at that date, not a complete final lifecycle account. The useful lesson is to reconcile what each number includes before interpreting an apparent saving. A favourable comparison can be misleading when one side ends at deployment and the other includes operation, even if both figures have been accurately copied from their source documents.

Preserve the historical account while valuing future choices

Figure 6.2 extends the manufacturer's fictional figures with an assumed review snapshot. The approved delivery budget remains US\$12 million and the current forecast for completing the existing design remains US\$14 million. Assume US\$6 million has been incurred and the remaining delivery forecast is US\$8 million. These quantities preserve the distinction between the approved commitment, actual expenditure and current expectation. For the forward comparison, the incurred US\$6 million cannot be recovered by selecting another option and is excluded from the choice calculation. It remains in the complete project account and in any assessment of performance against the original investment. Excluding it from the next decision does not erase the overrun or excuse weak earlier judgments. It prevents an irreversible past expenditure from being treated as an additional benefit of committing future resources to an unattractive course.

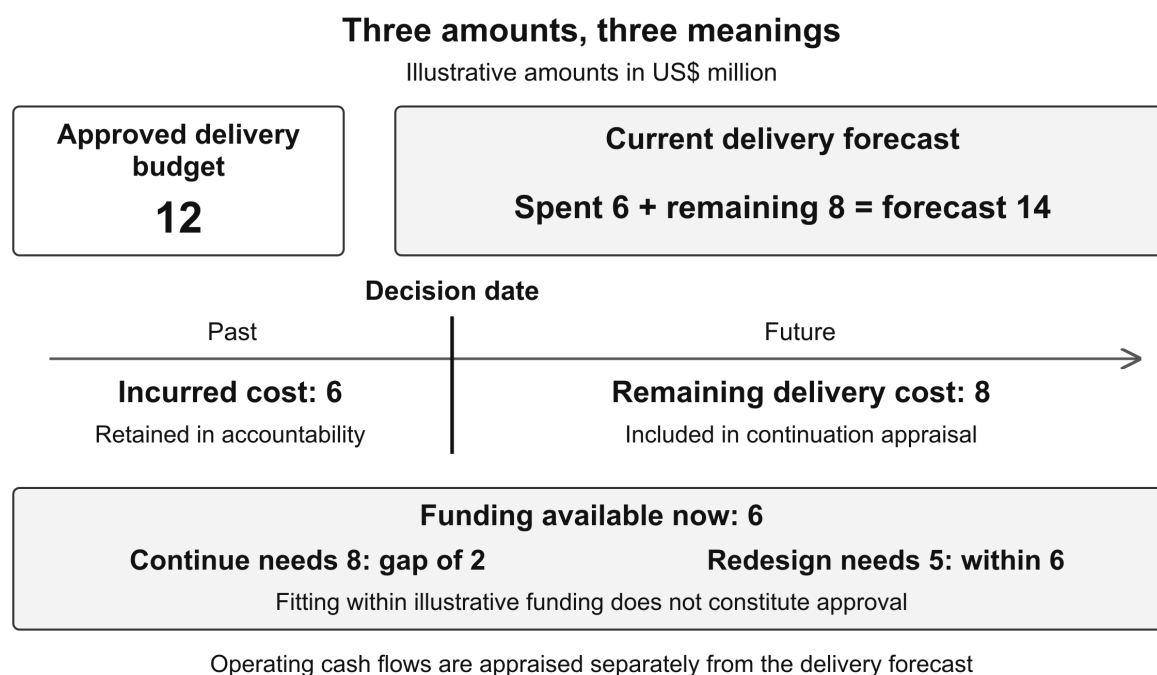


Figure 6.2 Fictional delivery and funding boundaries in US\$ millions. Approved budget 12, incurred cost 6 and remaining estimate 8 imply a current delivery forecast of 14. Immediate funding of 6 leaves a gap of 2 for continuation. Operating cash flows are assessed separately.

The word committed requires careful interpretation. Some future payments are avoidable through a change or termination; others remain payable under every feasible option. A signed agreement may permit cancellation with a charge, require payment for work already performed or create continuing support obligations. Commercial and financial specialists should establish the consequences rather than classify all contracted amounts as sunk. The comparison includes future amounts that differ between options, including exit costs and the opportunity cost of resources that could be sold or used elsewhere. Payments genuinely common to every option may be omitted from the difference calculation, but still belong in the affordability account. This distinction prevents a technically correct incremental analysis from understating the cash the

organisation actually needs. The review should reconcile its decision model with the complete expenditure forecast so that exclusions are understood rather than silently forgotten.

Discounted cash flow provides a way to compare amounts occurring at different times. Net present value is the sum of each period's incremental net cash flow divided by one plus the discount rate raised to that period number, with immediate cash flow included at its current value. The calculation requires consistency between cash flows and the rate. Constant-price amounts require a real rate; nominal amounts require a compatible nominal rate, with general inflation treated consistently. Relevant relative price changes still need consideration. Similarly, the financing and tax perspective must match the chosen rate and cash-flow definition. These are substantive assumptions, not formatting choices. Graham and Harvey's (2001) survey of corporate financial practice found variation in the use of appraisal techniques across firms. The existence of a familiar calculation does not ensure that organisations use it consistently or that its input estimates deserve the confidence attached to its output.

Benefits should enter the calculation through evidence-based operating changes. A reduction in processing time is not automatically a cash receipt. A cash saving requires a change in expenditure; released capacity may instead create separately justified service value without reducing cash outflows. A sales benefit normally requires the incremental contribution after associated costs, rather than gross revenue presented as free value. Changes in working capital also need their own cash-flow treatment instead of being repeatedly counted as annual savings. Chapter 5's payment-rate example cannot by itself establish any of these amounts. The financial assumptions used below are additional invented teaching inputs, which would require independent substantiation in practice. Each significant benefit should retain its owner, mechanism, observation rules, timing and valuation basis. Where several initiatives contribute to the same improvement, reconcile the claim across the portfolio before including it in more than one investment case.

Compare four feasible futures on the same basis

Table 6.1 sets out four fictional alternatives from a common review date. Amounts are US\$ millions in constant prices, assessed using an assumed eight per cent real discount rate over four years. The account contains organisational cash flows before financing and tax, relative to maintaining the existing service; it is not a social appraisal. The rate is a teaching assumption, not a recommended hurdle rate. Delivery and exit payments occur immediately at time zero unless stated otherwise, and later receipts occur at annual period ends. The four-year horizon and zero terminal value are common simplifying assumptions that require challenge in practice. Continuing the existing service is the accounting reference for recurring differences, not a claim that the organisation can abandon its transformation obligations without cost. The separate stop option includes the assumed cost of settling those obligations and maintaining the necessary service transition.

Continuing requires US\$8 million immediately and produces annual incremental net receipts of US\$3 million in years one to four. For illustration, each annual amount comprises US\$4 million of cash benefits less US\$1 million of incremental operating expenditure. Redesign requires US\$5

million immediately and produces annual net receipts of US\$2 million, comprising US\$2.8 million of benefits less US\$0.8 million of operating expenditure. The redesign preserves the essential service conditions but delivers a narrower benefit proposition; that feasibility assumption would need demonstration. Deferral costs US\$0.5 million immediately to preserve the option, followed by US\$8 million at the end of year one and US\$3 million in each of years two to four. Stopping costs US\$1 million immediately and produces no later incremental receipts. All relevant differences are assumed included in these figures, with no recoverable asset value; the incurred US\$6 million is common history, excluded from every forward option.

Table 6.1 Fictional incremental cash flows in US\$ millions

Option	Now	Year 1	Year 2	Year 3	Year 4	NPV
Continue	-8.00	3.00	3.00	3.00	3.00	1.94
Redesign	-5.00	2.00	2.00	2.00	2.00	1.62
Defer	-0.50	-8.00	3.00	3.00	3.00	-0.75
Stop	-1.00	0.00	0.00	0.00	0.00	-1.00

Negative amounts are outflows. Now is the common review date; later cash flows occur at annual period ends. NPV uses an assumed eight per cent real rate. Figures are before financing and tax, relative to maintaining the existing service, with no terminal value. The incurred US\$6 million is excluded.

At eight per cent, the present-value factor for four annual year-end receipts is approximately 3.3121. Continuing therefore gives US\$3 million multiplied by 3.3121, less US\$8 million: a prospective net present value of approximately US\$1.94 million. Redesign gives approximately US\$1.62 million, deferral minus US\$0.75 million, and stopping minus US\$1 million. Figure 6.3 compares these results. Continuing has the highest value under the stated monetary assumptions, exceeding redesign by about US\$0.31 million. This is not the whole-project return: the comparison deliberately starts at the review date. Nor does the negative value of stopping mean it is never appropriate. It means exit has an assumed future cost relative to the stated accounting reference. If continuing and redesigning failed critical conditions, stopping could remain the preferable feasible response despite that cost. An option's financial sign cannot settle a question that its boundaries exclude.

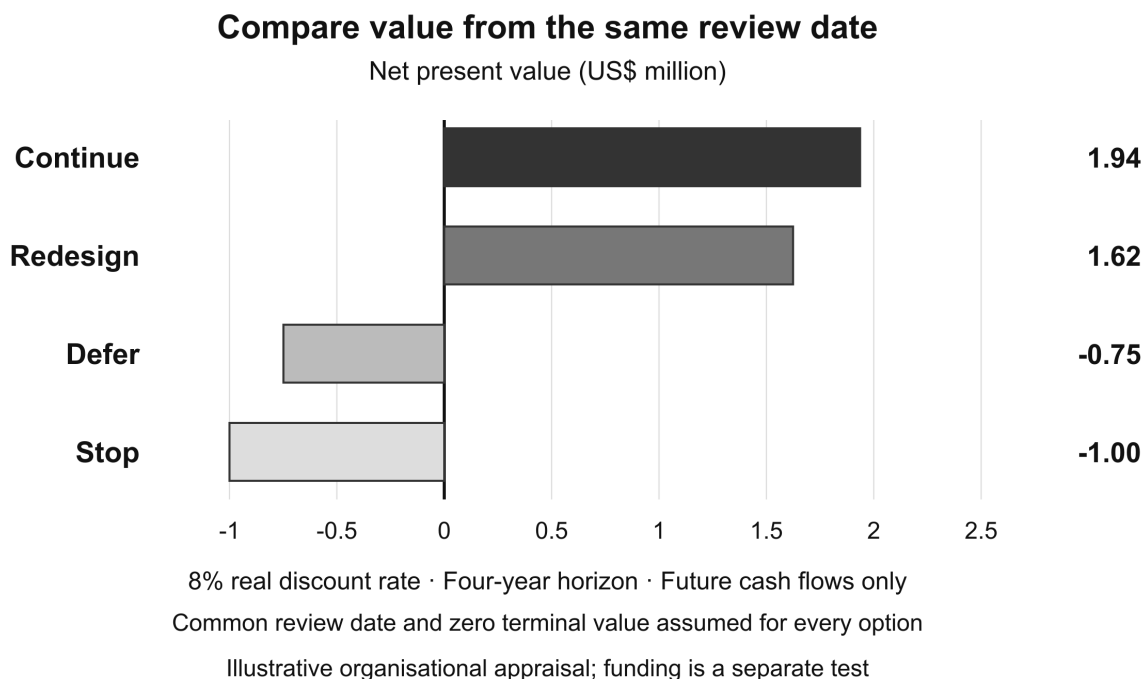


Figure 6.3 Fictional prospective net present values in US\$ millions. The common review date, four-year horizon, eight per cent real rate and zero terminal value apply to every option. The calculation excludes incurred expenditure and does not establish funding or operational acceptability.

Other summary measures can answer narrower questions, provided their limits remain visible. Under the example's annual receipt convention, both continuing and redesign recover their immediate outlay at the end of year three on an undiscounted basis. That payback observation does not distinguish their total value over the four-year horizon, and it says nothing about receipts after the recovery point. A simple percentage return can also obscure differences in scale and timing. Use these measures to explain a particular concern, such as exposure before recovery, rather than allow whichever indicator looks most favourable to determine the recommendation. Where options are mutually exclusive, examine the incremental consequences of selecting one over another. Continuing requires an additional US\$3 million compared with redesign and produces an additional US\$1 million annually under the central assumptions. This makes the reason for its marginal advantage transparent and identifies the extra commitment whose justification requires the strongest scrutiny.

Now assume that only US\$6 million is available for immediate expenditure and no additional funding has been approved. Continuing requires US\$8 million immediately and therefore has a US\$2 million funding gap. Redesign requires US\$5 million and fits within the assumed amount, but still requires proper authorisation and evidence of feasibility. The calculation should not quietly move the continuing payment into a later period to make the funding problem disappear. Deferral also retains a future funding requirement that must be addressed, while its lost first-year receipt and preservation cost are visible. The example's zero terminal value means it credits no receipts beyond year four, making the horizon particularly consequential for deferral. A real review would examine residual value, a longer common horizon or other consistent treatment of

different operating lives. The comparison is useful because it makes these choices explicit, not because its simplified timing is a complete investment model.

Test how uncertainty could change the recommendation

Figure 6.4 varies the scale of annual net receipts while holding timing, immediate costs and the discount rate fixed. At ninety per cent of the planned net receipts, continuing produces approximately US\$0.94 million of present value and redesign approximately US\$0.96 million. The ranking reverses even though both remain positive. The two lines cross at about 90.6 per cent of their planned receipts, because the extra US\$3 million required by continuation must be justified by its additional annual receipts. This is a conditional sensitivity calculation, not a probability that benefits will be achieved. The apparent advantage at ninety per cent is also very small relative to the uncertain inputs; reporting more decimal places would not make the decision robust. The useful question is which assumptions could change the preference and what evidence could reduce uncertainty about them before the commitment becomes difficult to reverse.

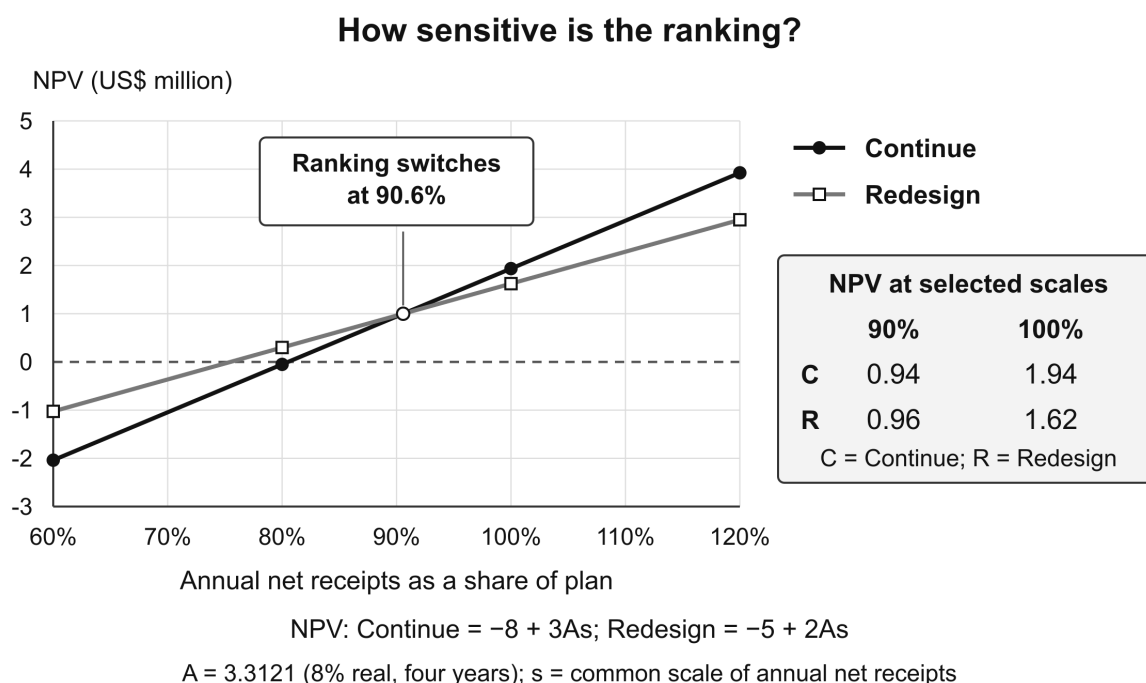


Figure 6.4 Conditional sensitivity of continuation and redesign to annual net receipts, with immediate costs, timing and rate held fixed. Their ranking switches at approximately 90.6 per cent of planned receipts. This comparison is not a probability distribution or a ranking of all feasible options.

Sensitivity changes one or a few inputs under a stated rule. Scenarios examine combinations that could plausibly occur together, such as delayed adoption accompanied by extended support costs and weaker early receipts. These effects should not be treated as independent merely because they occupy separate spreadsheet rows. Ward and Chapman (2003) argue for examining project uncertainty beyond a narrow list of adverse events, including uncertainty in assumptions and underlying variability. Applied here, the team should investigate the cost and benefit

mechanisms that connect the rows. Lower transaction demand might reduce both benefits and operating expenditure, whereas a difficult migration might reduce benefits while increasing support costs. A coherent scenario states those relationships and its rationale. Assign probabilities only where there is a defensible basis, and distinguish a scenario chosen to test resilience from an expected outcome. Quantitative risk methods receive fuller treatment in Chapter 10.

Contingency and other adjustments need the same discipline. The estimate should explain which uncertainties are already reflected in quantities, rates, scenario cash flows or explicit allowances, and which remain outside them. Adding an arbitrary percentage to every subtotal can double count uncertainty or conceal a missing activity. Conversely, removing an allowance to meet an approval limit does not remove the underlying exposure. Love and colleagues (2013) fitted overrun distributions using 276 completed Australian construction and engineering projects, measuring change from contract award. That reference point is central to interpretation: an allowance inferred from post-award outcomes does not automatically represent uncertainty at an earlier investment decision or the costs of operating a transformation. Use comparable definitions and examine the relevant population before borrowing an allowance. The funding decision should state the exposure it is intended to cover and who can authorise the use of any reserve.

An external comparison can challenge a persuasive internal story. Flyvbjerg (2006) proposes reference-class forecasting using actual outcomes from relevant completed undertakings. Its application requires a defensible class, compatible cost boundaries and attention to the stage at which the forecast is made. For the manufacturer, comparisons might concern migration effort, support duration or realised adoption timing rather than a generic percentage added to the total budget. The reviewer should investigate why the present forecast differs from that experience and retain the explanation. External evidence can itself be incomplete or selectively assembled, so an outside view deserves scrutiny as well. Neither an internal model nor an industry benchmark should acquire authority simply through its label. The two become useful when their disagreement identifies a material assumption, a missing cost or an operating condition that can be examined before further commitment.

Prevent past commitment from becoming future justification

Sunk-cost reasoning often enters discussion through ordinary language: too much has been spent to stop, cancellation would admit failure, or one more payment will protect the investment already made. Arkes and Blumer's (1985) experiments examined how prior expenditure could affect subsequent choices, while Staw's (1976) experiment investigated renewed resource commitment following negative feedback and personal responsibility for an earlier decision. These studies illuminate behavioural mechanisms; they do not establish that any particular sponsor is irrational or that continuation after an overrun is necessarily wrong. The controls response should make the future alternatives comparable and preserve the earlier account separately. An option may be worth pursuing because of its remaining benefits, existing usable assets or avoidable exit costs. Those are prospective reasons. The amount irrecoverably spent is not itself an additional future benefit, however understandable the wish to make that history end well.

Escalation can also arise through organisational conditions rather than a single person's attachment. Keil, Mann and Rai (2000) examined escalation in software projects through retrospective auditor assessments, connecting persistence with project and organisational circumstances. The practical implication is to examine reporting arrangements, perceived completion, external pressure and the ability to change course, rather than treating the problem as a personality test. Slesman and colleagues' (2012) meta-analysis draws together evidence across 166 independent samples and shows why determinants and study conditions need careful interpretation. In particular, sunk expenditure and closeness to completion can become entangled. For a review board, this means that a claim such as almost finished needs an operational definition and evidence of remaining work. A high percentage of expenditure or completed tasks does not establish that the unresolved integration, adoption or benefit conditions are small or inexpensive.

Forecast presentation can reinforce commitment before a formal decision occurs. Kahneman and Lovallo (1993) analyse how decisions framed narrowly around an individual undertaking can encourage overly optimistic forecasts and insufficient attention to broader evidence. Their conceptual argument supports asking what comparable experience and alternative uses of resources imply for the current proposal. Flyvbjerg (2009) examines the selection of major projects under optimism and strategic misrepresentation, warning that favourable initial representations can influence which proposals proceed. Neither account permits a reviewer to infer deliberate deception from an inaccurate forecast alone. The useful response is structural: retain the estimate basis, distinguish advocacy from analysis and make important contrary evidence visible to the authorising body. A proposal should survive informed comparison because its remaining case is credible, rather than because admitting uncertainty would make it less competitive for scarce organisational attention.

Challenge should make a different decision possible without making honest reporting personally ruinous. A delivery manager who identifies a weaker benefit mechanism should not have to defend an obsolete valuation to demonstrate commitment to the organisation. Equally, describing every disappointing forecast as new learning can obscure avoidable errors that require accountability. The review can ask both what should happen next and what the earlier process needs to learn. Preserve dated assumptions, approval records and the evidence available at the time, then distinguish genuinely new information from information previously ignored. This allows a sponsor to support redesign or stopping for defensible reasons while retaining a fair account of performance. The process should reward early clarification of a material weakness and require evidence for the proposed response. It should not reward either optimistic persistence or dramatic cancellation as a substitute for examining the feasible choices.

Opportunity cost can arise through scarce people and operational attention as well as money. If the same integration specialist is needed for a mandatory service correction and the proposed expansion, funding both does not make both achievable at the same time. The review should examine the value and consequences of the work displaced, including any delay imposed on another commitment. Chapter 4's capacity assessment and Chapter 7's sequencing methods provide the corresponding operating analysis. Do not automatically add a salary charge and the full value of forgone work if they represent overlapping measures of the same resource

consequence. State the basis and reconcile the treatment with the financial account. A portfolio authority may also judge priorities differently from an individual sponsor because it can see the alternatives competing for the resource. The local business case should expose that dependency rather than assume that a favourable project calculation establishes priority over every other use.

Make flexibility and exit genuine alternatives

Deferral can create value when information arrives before an irreversible commitment, but it can also destroy value through delay, preservation costs and lost opportunities. Dixit (1992) explains how uncertainty and costly reversal can make the ability to wait economically consequential. Trigeorgis (1993) shows through analytical modelling that multiple real options can interact, so their values cannot generally be added as independent bonuses. These theoretical results support examining genuine flexibility; they do not establish a standard premium to add to every staged transformation. The worked deferral option is a simple fixed plan to resume after one year. It does not value a contingent decision to stop or redesign after receiving new information. Such an analysis would need a credible description of the information, the decisions available afterwards and their conditional consequences. A delay with no learning purpose should not be presented as a valuable option simply because commitment has been postponed.

The team should specify what preserves a choice. A pilot may require interfaces or contractual rights that permit later expansion without repeating all the work. Deferral may require maintaining specialist capability, extending a licence or protecting an asset from deterioration. A break clause creates little practical freedom if no alternative service can be established before the existing arrangement ends. Include these conditions and their costs in the option description. Equally, identify choices lost after a commitment: a supplier mobilisation, data migration or site change may make reversal materially more difficult. The decision to investigate should be timed so its evidence arrives before that boundary. A useful review asks whether additional information could change the preferred action and whether obtaining it costs less than the improvement in decision quality it could reasonably support. The answer can be qualitative where the evidence cannot support a numerical information value.

Stopping also creates a programme of future work. The National Audit Office's July 2024 report examined the response to the October 2023 cancellation of HS2 Phase 2. At that historical snapshot, early works had left remediation obligations despite substantive construction not having begun. Remediation was estimated to cost up to £100 million in 2024 prices and expected to finish by summer 2027. Disposal of surplus land and property was a separate undertaking expected to take several years, with timing potentially affecting local markets (NAO, 2024, pp. 6, 20, 22–23). The report did not evaluate the cancellation decision itself, and the remediation estimate was not a total cancellation cost. Its relevance is the need to include exit obligations in a comparison. Cancellation should not become an artificially free alternative, and the existence of exit costs should not automatically justify continuing a different, more expensive course.

Renewal requires similar attention to the future proposition. Victoria's Auditor-General examined myki in June 2015 as Public Transport Victoria prepared to retender its continuing operation. The audit found no established evaluation process for expected benefits and objectives

that had not been updated after scope changes. It accepted a detailed procurement strategy instead of a separate new business case, but found proposed benefits vague and measures focused on outputs. Compressed procurement and transition timing also created cost exposure (VAGO, 2015, pp. 38–40, 47–49). These findings did not demonstrate that myki had produced no benefits or identify a superior alternative through a controlled comparison. They show why an existing asset and a continuing service need do not automatically justify a particular renewal. The next contract should be supported by achievable future value, defined obligations and a credible transition arrangement rather than by the fact that the previous investment was approved.

Use the control method to authorise a bounded choice

Table 6.2 translates the option comparison into an Investment Review Record. It retains the numerical account but adds the evidence and conditions needed for a decision. Applying the established method, framing states the precise commitment being requested, the review date and the alternatives available. Samira, the fictional sponsor, identifies the decision authority and the strategic conditions that the service must satisfy. Establishing brings together Alex's remaining delivery estimate, Maya's operating and benefit evidence, finance's cash-flow and funding analysis, and Lena's reconciled control account. Their names connect this chapter to the earlier ownership arrangements; they do not create authority beyond the agreed delegations. The record should distinguish approval to investigate an option from approval to incur its full expenditure. A promising calculation can justify further work while important acceptance or funding conditions remain unresolved.

Table 6.2 Investment Review Record for the fictional manufacturer

Option	Prospective account	Evidence needed	Decision condition
Continue	8 now; NPV 1.94 under the central assumptions.	Validate annual net receipts of 3 and essential service conditions.	Resolve the immediate funding gap of 2; obtain approval.
Redesign	5 now; NPV 1.62 for a narrower proposition.	Demonstrate retained critical requirements and annual net receipts of 2.	Fits assumed funding of 6; authorisation remains necessary.
Defer	0.5 now and 8 in year one; NPV –0.75.	Define preservation work, learning question, next decision date and future funding.	Fixed resumption is modelled; information value is not calculated.
Stop	1 now; NPV –1.00 with no later incremental receipts.	Verify exit terms, service transition, recoverable assets and residual ownership.	Zero salvage is assumed; confirm all remaining obligations.

All amounts are fictional US\$ millions. This summary accompanies the dated assumptions, estimate basis, cash profile, nonfinancial conditions, recommendation and approval record. It is not a substitute for those supporting records.

Challenge examines the boundaries, alternatives and assumptions before recommendation hardens into commitment. Finance checks the perspective, rates, timing and treatment of

common costs. Operations tests whether the redesign can preserve the essential service and whether forecast benefits have credible owners. Delivery examines remaining interfaces and the consequences of changing scope. An appropriately independent reviewer can examine consequential assumptions and evidence without taking over management's responsibilities. Marsh and colleagues' (2016) healthcare guidance on multiple criteria decision analysis distinguishes performance evidence from preference judgments and warns against overlapping criteria that duplicate weight. Applied cautiously here, it supports making the nonfinancial comparison explicit alongside the monetary one. A weighted score cannot establish that cash is available or compensate for an unmet critical condition. Where reviewers disagree, record the substance and whether it could change the preferred option, rather than concealing the disagreement in a consensus rating.

Authorisation identifies the chosen option, financial limits, conditions, owners and review triggers. In the fictional example, the existing funding limit prevents unconditional immediate continuation as modelled. The authority could commission a bounded redesign investigation, seek additional funding for a substantiated continuing case or approve another feasible option. It should not instruct Lena to report continuation as funded before that condition is satisfied. Keil and Robey's (1999) exploratory study of troubled software projects examined de-escalation through retrospective auditor accounts and included both redirection and termination. That distinction supports retaining redesign as a substantive choice, while the study does not prove that any single governance measure guarantees recovery. The decision record should explain why the selected course is preferable under the evidence now available and what would require reconsideration. Revising a course can be responsible management without becoming an excuse to obscure its original commitments.

Verification examines whether the decision changed the relevant obligations and whether its conditions continue to hold. A funding approval needs a usable allocation and timing, not merely supportive meeting minutes. A redesign approval needs an accepted scope and revised evidence requirements, with supplier and operating arrangements aligned. A deferral needs confirmation that preservation work occurred and that information will arrive before the next decision. A stop decision needs a controlled transition, settlements and responsibility for residual obligations. Lena updates the authorised records only after the appropriate decision, retaining the original baseline and the intervening history. The current forecast may change earlier when new evidence warrants it, but that does not itself authorise expenditure. This applies Chapter 3's governance distinction to investment economics: an analytical recommendation, a spending decision and demonstrated completion remain different events with different owners and evidence.

Maintain an investment case that can change honestly

An investment case should be maintained at the points where new knowledge could alter a material choice. That does not mean rewriting every calculation after every small event. Agree review triggers linked to consequential changes in scope, remaining cost, benefit timing, operating demand, funding or external conditions. Use routine control information to identify whether a trigger has been reached and whether the change could affect the recommendation. A stable reporting rhythm provides continuity, while an emerging issue may require an earlier decision. The review record should identify its data cut-off, design version, assumptions and next evidence date so that the reader can distinguish a current assessment from a once-valid approval paper. Its purpose is to support decisions while choices remain available. A documented appraisal completed after the relevant contract or operating transition has become irreversible can provide accountability while arriving too late to influence the commitment.

The whole-project account should continue alongside the prospective comparison. It records original expectations, approved changes, actual expenditure and the latest estimate of total consequences, including observed and forecast benefits kept distinct. This allows the organisation to examine the quality of earlier decisions and the eventual return on the investment without contaminating the next choice with irrecoverable expenditure. It also prevents a succession of individually attractive continuation decisions from being presented as proof that the original project was successful. Both statements can be true: the initial investment performed poorly, and a limited further commitment is preferable to the feasible alternatives now available. Explain that position directly. Equally, the original case may have been reasonable on the evidence available while changed conditions now favour stopping. Honest economic control preserves the chronology and the decision basis instead of requiring every new recommendation to validate the entire history.

Learning requires examining the sources of forecast movement. Separate changed quantities, prices, productivity, scope, timing and operating assumptions where their effects can be identified. Avoid attributing every difference to inflation, uncertainty or supplier behaviour without evidence. Compare actual operating demand and benefit mechanisms with what the earlier case assumed, then improve the relevant estimating data or decision process. Where a forecast was conditional, examine whether its stated conditions actually occurred. This is more informative than evaluating every estimate solely by whether its final number matched the outcome. Some uncertainty is irreducible, while other errors reflect missing work, inconsistent boundaries or unsupported optimism. The review should make those distinctions usable by the next team. Chapter 8 will develop the information and learning arrangements that allow an organisation to retain this evidence beyond the people who prepared the original case.

The economics of continuing is therefore a discipline of comparing remaining choices without losing responsibility for earlier ones. Its practical contribution is a coherent account of what each feasible option would require, what it could produce, when money is needed and which conditions remain uncertain. The four options in the example are alternatives to investigate, not a universal menu or an automatic stopping rule. Their ranking depends on evidence, boundaries, constraints and legitimate judgments that the calculations must expose. A controls professional

should leave the decision maker able to explain the choice, its conditions and the observation that would justify revisiting it. The next chapter turns to sequencing: even a worthwhile and funded proposal needs a delivery and operating plan that the organisation can absorb. Economic value becomes achievable only when the work, resources and timing required to obtain it can be coordinated in practice.

References

- Arkes, H. R. and Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140. doi: [10.1016/0749-5978\(85\)90049-4](https://doi.org/10.1016/0749-5978(85)90049-4)
- Atkinson, G. and Mourato, S. (2008). Environmental cost-benefit analysis. *Annual Review of Environment and Resources*, 33, 317–344. doi: [10.1146/annurev.enviro.33.020107.112927](https://doi.org/10.1146/annurev.enviro.33.020107.112927)
- Dixit, A. (1992). Investment and hysteresis. *Journal of Economic Perspectives*, 6(1), 107–132. doi: [10.1257/jep.6.1.107](https://doi.org/10.1257/jep.6.1.107)
- Flyvbjerg, B. (2006). From Nobel Prize to Project Management: Getting Risks Right. *Project Management Journal*, 37(3), 5–15. doi: [10.1177/875697280603700302](https://doi.org/10.1177/875697280603700302)
- Flyvbjerg, B. (2009). Survival of the unfittest: Why the worst infrastructure gets built—and what we can do about it. *Oxford Review of Economic Policy*, 25(3), 344–367. doi: [10.1093/oxrep/grp024](https://doi.org/10.1093/oxrep/grp024)
- Gluch, P. and Baumann, H. (2004). The life cycle costing (LCC) approach: A conceptual discussion of its usefulness for environmental decision-making. *Building and Environment*, 39(5), 571–580. doi: [10.1016/j.buildenv.2003.10.008](https://doi.org/10.1016/j.buildenv.2003.10.008)
- Graham, J. R. and Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, 60(2–3), 187–243. doi: [10.1016/S0304-405X\(01\)00044-7](https://doi.org/10.1016/S0304-405X(01)00044-7)
- HM Treasury (2026). The Green Book. Updated 5 February 2026. www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2026
- Jørgensen, M. (2004). A review of studies on expert estimation of software development effort. *Journal of Systems and Software*, 70(1–2), 37–60. doi: [10.1016/S0164-1212\(02\)00156-5](https://doi.org/10.1016/S0164-1212(02)00156-5)
- Kahneman, D. and Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1), 17–31. doi: [10.1287/mnsc.39.1.17](https://doi.org/10.1287/mnsc.39.1.17)
- Keil, M., Mann, J. and Rai, A. (2000). Why software projects escalate: An empirical analysis and test of four theoretical models. *MIS Quarterly*, 24(4), 631–664. doi: [10.2307/3250950](https://doi.org/10.2307/3250950)
- Keil, M. and Robey, D. (1999). Turning around troubled software projects: An exploratory study of the deescalation of commitment to failing courses of action. *Journal of Management Information Systems*, 15(4), 63–87. doi: [10.1080/07421222.1999.11518222](https://doi.org/10.1080/07421222.1999.11518222)
- Love, P. E. D., Wang, X., Sing, C. P. and Tiong, R. L. K. (2013). Determining the probability of project cost overruns. *Journal of Construction Engineering and Management*, 139(3), 321–330. doi: [10.1061/\(ASCE\)CO.1943-7862.0000575](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000575)
- Marsh, K., IJzerman, M., Thokala, P., Baltussen, R., Boysen, M., Kaló, Z., Lönnegren, T., Mussen, F., Peacock, S., Watkins, J. and Devlin, N. (2016). Multiple criteria decision analysis for health care decision making—Emerging good practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value in Health*, 19(2), 125–137. doi: [10.1016/j.jval.2015.12.016](https://doi.org/10.1016/j.jval.2015.12.016)
- National Audit Office (NAO) (2024). HS2: Update following cancellation of Phase 2. HC 128, Session 2024–25, 23 July. www.nao.org.uk/reports/hs2-update-following-cancellation-of-phase-2/

- Sleesman, D. J., Conlon, D. E., McNamara, G. and Miles, J. E. (2012). Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal*, 55(3), 541–562. doi: [10.5465/amj.2010.0696](https://doi.org/10.5465/amj.2010.0696)
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16(1), 27–44. doi: [10.1016/0030-5073\(76\)90005-2](https://doi.org/10.1016/0030-5073(76)90005-2)
- Trigeorgis, L. (1993). The nature of option interactions and the valuation of investments with multiple real options. *Journal of Financial and Quantitative Analysis*, 28(1), 1–20. doi: [10.2307/2331148](https://doi.org/10.2307/2331148)
- U.S. Department of Justice, Office of the Inspector General (DOJ OIG) (2012). Interim Report on the Federal Bureau of Investigation's Implementation of the Sentinel Project. Audit Report 12-38, 7 September. oig.justice.gov/archives/reports/2012/a1238.pdf
- U.S. Government Accountability Office (GAO) (2020). Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Program Costs. GAO-20-195G, 12 March. www.gao.gov/products/gao-20-195g
- Victorian Auditor-General (VAGO) (2015). Operational Effectiveness of the myki Ticketing System. Report 2014–15:33, PP No. 40, Session 2014–15, 10 June. www.audit.vic.gov.au/report/operational-effectiveness-myki-ticketing-system
- Ward, S. and Chapman, C. (2003). Transforming project risk management into project uncertainty management. *International Journal of Project Management*, 21(2), 97–105. doi: [10.1016/S0263-7863\(01\)00080-1](https://doi.org/10.1016/S0263-7863(01)00080-1)

CHAPTER 07

Sequencing Change at an Absorbable Pace



IN THIS CHAPTER

Sequencing delivery around dependencies and the capacity to absorb change.

Give the promise a workable sequence

The purchasing platform is nearly ready, the sponsor wants an earlier release, and the schedule shows several activities running in parallel. Lena, the project controller in our fictional manufacturer, notices that two of those activities require the same specialist. Maya, the operational owner, sees a different problem: the proposed release falls before her team can practise the new exception process. Alex, the delivery manager, can bring a demonstration forward, but cannot yet show when the service will operate reliably. All three observations concern time, although they describe different clocks. Technical completion, operational acceptance and benefit realisation have related but distinct sequences. This chapter explains how to connect them so that a promise about transformation becomes a practicable commitment. The objective is to help controllers identify the work that actually governs value, expose competing demands and make the next release sufficiently credible to authorise.

Chapter 5 connected delivered outputs to usable capability, changed practice, operating outcomes and benefits. Chapter 6 examined whether the remaining investment justified further commitment. Sequencing now asks when those conditions can be achieved together and what prevents earlier value. A benefits forecast needs more than a date attached to a completed asset. It needs the people, information, controls and operating arrangements that permit the intended improvement to occur, followed by an observation period capable of supporting the claim. These conditions should appear in the planning system even when they sit outside a supplier's contract. Otherwise the project may report progress against a complete delivery schedule while the organisation has no feasible route to the promised operating state. A controller's contribution is to connect these accounts without confusing their different owners, evidence requirements or levels of certainty.

The proposed Sequencing and Benefits Map in Figure 7.1 develops that connection. It follows technical work through a resource-feasible sequence into an accepted operating state and then into benefit observation. Ownership, evidence and authorised decisions connect the stages, while learning can require an earlier assumption to be reconsidered. The map is an original practice aid, not a validated prediction model. It extends the book's Transformation Control Loop and uses the existing Credible Control Method: Frame, Establish, Challenge, Authorise and Verify. Its practical question is whether a proposed date has a defensible chain of conditions behind it. A missing training commitment, an unresolved interface or an unstaffed support function is therefore a scheduling issue as well as an organisational issue. Making that dependency visible creates an opportunity to act before a deadline turns it into a crisis.

The US Government Accountability Office's Schedule Assessment Guide treats a reliable schedule as an integrated model of the work, with attention to logic, resources, critical paths, float and schedule risk analysis (GAO, 2015). This is professional guidance rather than experimental evidence that a particular scheduling method causes success. Its relevance here is the insistence that the schedule should explain the forecast. An attractive presentation cannot compensate for missing predecessors or impossible resource assignments. The chapter develops that principle through one small worked release, then extends it to uncertain durations, information feedback, congested operating teams and waves of benefit delivery. Every numerical

input in the worked release is fictional. The calculations teach how to examine a model; they do not provide benchmark durations or universal thresholds for another organisation.

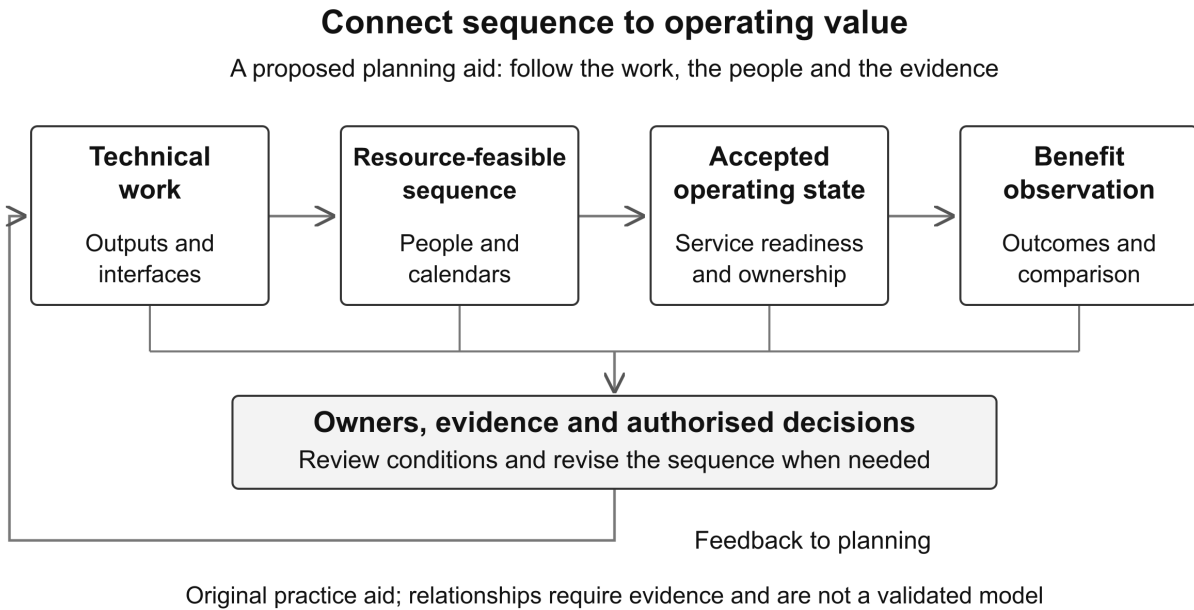


Figure 7.1 The proposed Sequencing and Benefits Map. A feasible sequence connects technical work, accepted operation and benefit evidence through named ownership and authorised decisions. This original practice aid is not a validated prediction model.

Organise the work around the operating state

A useful schedule begins with a defined boundary. Lena identifies the release as the capability to process an agreed category of supplier payments through the new service, including exception handling, authorisation and support. This description is more informative than naming a software version because it explains what operations must be able to do. The work breakdown structure, or WBS, divides the required work into manageable elements. The cost breakdown structure, or CBS, classifies the associated expenditure. The organisational breakdown structure, or OBS, identifies accountable units. These structures need compatible identifiers rather than identical shapes. A work package may incur several cost categories and require contributions from several teams, while retaining one accountable owner. Controls should make those relationships traceable so that scope, dates, resources, expenditure and acceptance evidence can be reconciled without relying on matching names in separate spreadsheets.

For the payment release, a work package could cover validated supplier data. Its scope includes cleansing, checks and acceptance, while its costs include specialist time and operating support. The accountable owner must be distinguished from the person performing each task and from the manager who releases that person's time. A control account can bring the relevant scope, schedule and budget together at an agreed management point, provided its boundary is explicit. Benefits should connect to that account through dependencies rather than being mechanically allocated across activities. Completing data validation enables reliable processing; it does not independently earn a percentage of the entire payment benefit. Several work packages may be

jointly necessary for one improvement, and one capability may support several improvements. The mapping must preserve these many-to-many relationships while preventing duplicate claims on the same saving.

Zwikael, Chih and Meredith (2018) developed and tested a scale for assessing the quality of target benefits through three studies. Their work supports careful formulation of the intended improvement, rather than treating a general aspiration as a sufficiently specified target. It does not validate a particular release sequence. In this chapter's application, the target benefit supplies a reason for asking what each wave must enable and what evidence will demonstrate progress. Lena therefore links the release boundary to the existing payment-performance definition, its denominator, its comparison condition and Maya's ownership. She does not invent a new success measure because the schedule needs a convenient milestone. If the benefit definition changes, that change belongs in the benefits account and governance record as well as the plan. A scheduling adjustment should not silently rewrite the investment's purpose.

Planning detail should concentrate where it can improve an actual decision. Near-term work needs sufficiently clear activities, dependencies, resources and completion evidence to support commitment. Later work may remain in planning packages with stated assumptions and dates for elaboration. This progressive detail is useful only when interfaces and long-lead obligations remain visible. A distant procurement decision can require attention now even when the installation sequence is still uncertain. Conversely, prescribing every task six months ahead can create detail that will soon become obsolete. The controller should retain the dated baseline, current forecast and reasons for change as separate records. Greater detail can improve a forecast without retrospectively changing what was approved. This preserves Chapter 2's principle of traceability while allowing the planning system to learn from work rather than defend an increasingly implausible original description.

Calculate what the dependency network actually says

Table 7.1 defines eight activities for a single release. Time is measured in relative working days from zero, using one calendar. Every relationship is finish-to-start, with no lag; tasks are non-preemptive, meaning that the exercise does not split them into interrupted fragments. These assumptions are deliberately simple. Activity A agrees the payment rules. B configures the platform and C validates the data. D tests their integration, while E provides staff practice after configuration. F is an operating trial, G the controlled release and H stabilisation and acceptance. The initial calculation assumes the required people are available whenever needed. It is therefore a precedence calculation, not yet a promise that the organisation can execute the plan. Keeping that distinction visible prevents mathematical correctness from being mistaken for operational feasibility.

Table 7.1 Fictional activities and precedence-only calculation

ID	Activity	Days	Depends on	ES	EF	Total float
A	Agree payment rules	2	None	0	2	0
B	Configure platform	4	A	2	6	0
C	Validate data	3	A	2	5	2
D	Integration test	2	B, C	6	8	1
E	Staff practice	3	B	6	9	0
F	Operating trial	2	D, E	9	11	0
G	Controlled release	1	F	11	12	0
H	Stabilise and accept	3	G	12	15	0

ES and EF mean earliest start and finish. Times are working-day boundaries from zero. One calendar, finish-to-start relationships, no lags and non-preemptive tasks apply. The shared B/C specialist restriction is introduced after this calculation.

The forward pass calculates each activity's earliest start as the latest earliest finish among its predecessors. Earliest finish then equals earliest start plus duration. A occupies days zero to two. B can occupy two to six, while C occupies two to five. D waits for both and therefore starts at six, finishing at eight. E also follows B, but takes three days and finishes at nine. F must wait for D and E, so it runs from nine to eleven; G finishes at twelve and H at fifteen. The important operation is the maximum at each merge. Adding every duration would incorrectly assume that all work is sequential, while considering only one branch could miss the predecessor that actually governs the merge. A schedule engine automates this arithmetic, but the controller should be able to explain its result without invoking the authority of the software.

The backward pass starts from the calculated finish of fifteen. Each latest finish is the earliest latest start among the activity's successors; latest start equals latest finish minus duration. H must start at twelve, G at eleven and F at nine. E must consequently start at six, while D can start at seven. Because B feeds both D and E, its latest finish is six and its latest start is two. C need only finish by seven and can start at four. Total float is latest start minus earliest start under these assumptions. C therefore has two days, D one, and the other activities zero. The critical path is A–B–E–F–G–H, with a duration of fifteen working days. Figure 7.2 makes the governing route visible alongside the branches that currently have room to move.

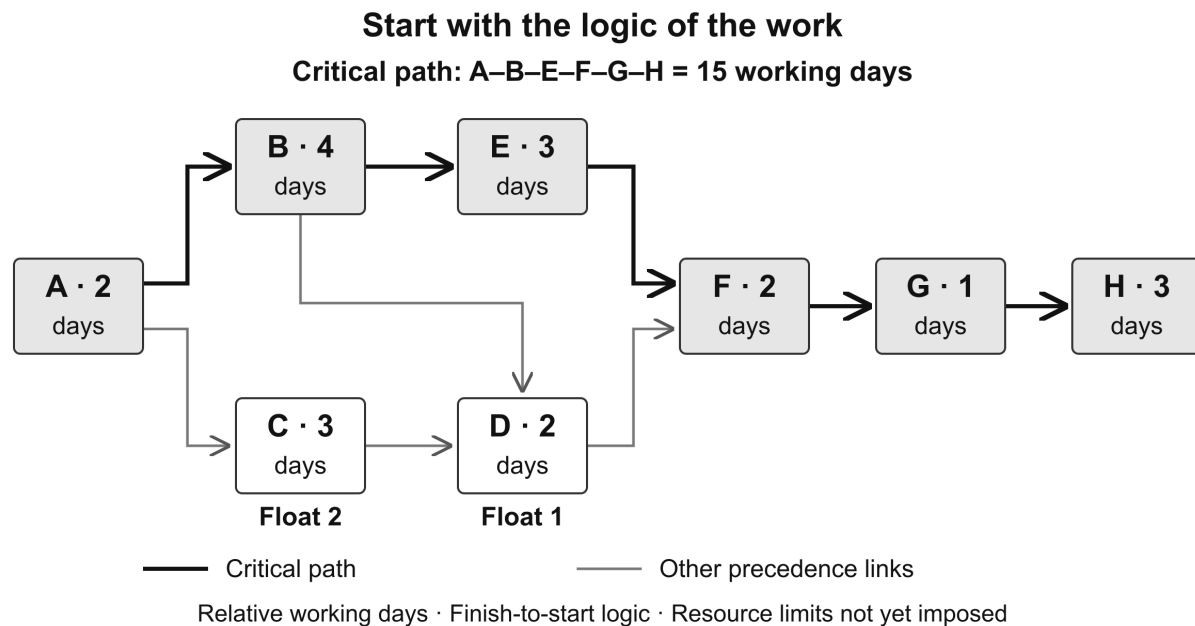


Figure 7.2 The fictional precedence network. The critical path A–B–E–F–G–H takes fifteen working days before resource constraints are introduced. C has two days of total float and D one under the stated calendar and finish condition.

Float needs interpretation before anyone spends it. C has two days of total float, but only one day of free float in this network: it can move from a finish at five to a finish at six without delaying D's earliest start. Its second day consumes room on the downstream branch. Total float is consequently not a set of independent allowances that every activity owner can use simultaneously. Nor is it automatically management contingency. It is a calculated property of the network, its calendars and its finish condition. Introduce an imposed deadline, a resource restriction or another successor and the values may change. Before treating float as available recovery time, Lena checks which commitments it protects and whether a proposed movement transfers pressure to another team. The number becomes useful when its consequences are understood.

This example also explains why progress percentages are insufficient. B could report ninety per cent complete while the remaining configuration includes the one rule necessary for E to practise safely. A duration forecast requires an assessment of remaining work, productivity, constraints and acceptance, not merely a subtraction from the original duration. Conversely, an activity can finish early without bringing the release forward if another predecessor governs the next merge. The controller should trace the consequence through the whole network. The practical exercise is to change one assumption, recalculate and explain which finish moved and why. Readers who can do this on eight activities can ask much better questions of a schedule containing thousands, because they understand the reasoning that the larger model is supposed to represent.

Make the sequence feasible for the people available

The first schedule fails a simple resource test: B and C require the same specialist, who cannot perform both at once. Kolisch and Hartmann (2006) investigated heuristics for resource-constrained project scheduling through computational comparisons on benchmark problems. Their work demonstrates the importance of schedule-generation choices under precedence and resource restrictions; it does not establish that one commercial tool will produce an optimal plan in every field setting. For Lena, the immediate lesson is to state the capacity assumption and examine the ordering decision. The model needs the specialist's actual availability, relevant skills and other commitments. A generic resource label cannot make two people interchangeable, and an eight-hour allocation cannot supply ten hours of uninterrupted work. Resource feasibility begins with credible commitments from the managers who control those people.

If B goes first, it occupies days two to six and C follows from six to nine. D then runs from nine to eleven, while E can run from six to nine using a separate operating team. F occupies eleven to thirteen, G finishes at fourteen and H at seventeen. The resource sequence B before C can be represented by an additional ordering link, provided the record identifies it as a resource decision rather than a technical necessity. If C goes first, it occupies two to five; B then occupies five to nine. D finishes at eleven, but E finishes at twelve, pushing F, G and H one day later. The resulting completion is eighteen. Figure 7.3 compares the two feasible orders. With these fixed durations and this one resource restriction, enumerating both alternatives establishes that B first is better; that small result does not make the same priority rule universally optimal.

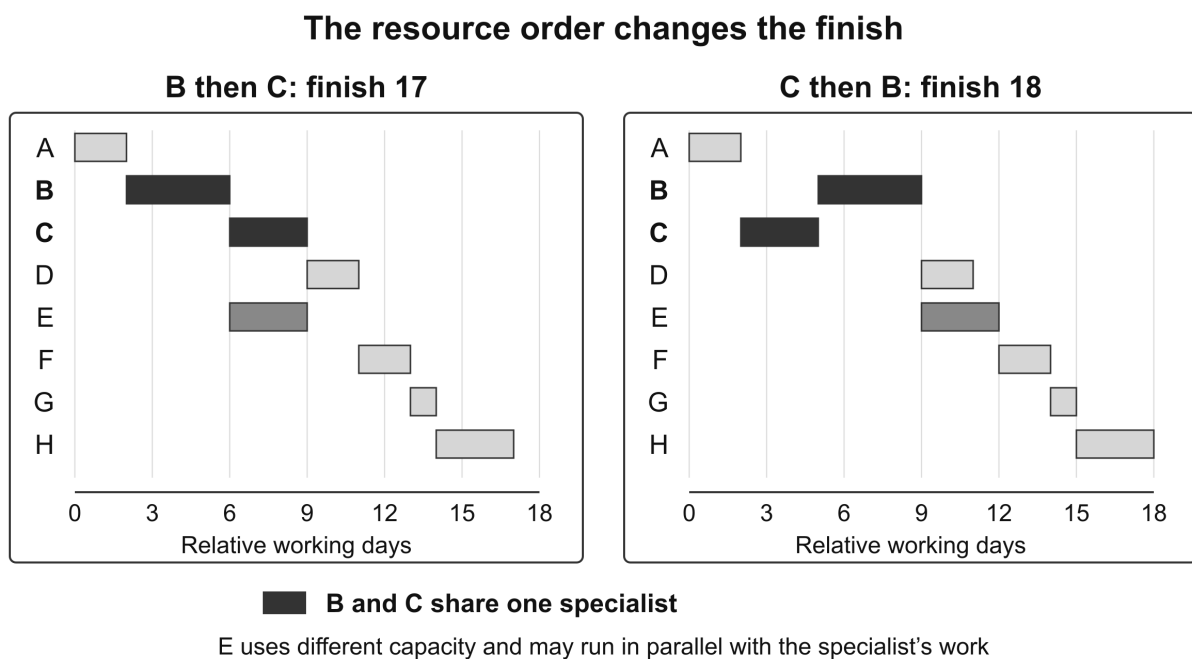


Figure 7.3 Alternative orders for one specialist shared by B and C. B before C gives acceptance at day seventeen; C before B gives day eighteen. Both use the same technical logic and deterministic durations. Times are relative working-day boundaries.

Recovery should target the constraint that actually controls completion. Under the seventeen-day sequence, shortening E from three days to two leaves the finish unchanged because D still governs F. Shortening D by one day moves completion to sixteen, assuming its acceptance requirements can still be met. Providing a second competent specialist permits B and C to overlap and restores the fifteen-day precedence result, if coordination and availability do not introduce additional work. Table 7.2 compares these interventions. The arithmetic prevents a common error: selecting the easiest activity to accelerate and then claiming the whole project will benefit. A recovery proposal should identify the changed dependency or capacity, the resulting forecast, the cost and the evidence that quality remains acceptable. Removing necessary testing is a change to the proposition being accepted, not simply a scheduling improvement.

Table 7.2 Test whether a recovery action changes completion

Intervention	Finish day	Change from 17	Condition to examine
Add a second specialist	15	2 days earlier	Competence, availability and coordination work must support parallel B/C.
Reduce E from 3 days to 2	17	No change	D still governs the start of F.
Reduce D from 2 days to 1	16	1 day earlier	The integration-test acceptance requirements must still be met.

All comparisons use the fictional B-before-C plan as the reference. A date improvement is not automatically an earlier benefit; operating and measurement conditions must also permit it.

Earlier completion is valuable only through the conditions it changes. Suppose, as an additional fictional assumption, that Chapter 6's annual net receipts of US\$3 million accrue uniformly across 250 productive working days and that the benefit horizon is fixed. This implies US\$12,000 per productive day, so bringing effective operation forward by two days could make US\$24,000 available before considering acceleration costs or discounting. The amount is a conditional illustration, not a new approved benefit forecast. If adoption, customer demand or a reporting cycle prevents earlier value, the gain may be zero. The controller should therefore trace the accelerated date into the operating and benefits model before comparing it with the cost of the extra specialist. A faster technical activity has financial significance only where a credible causal and temporal connection supports that significance.

Prioritising whole waves requires the same care. Suppose two independent, equally ready changes compete for one specialist, each needs two working days, and their assumed net benefits begin immediately after completion. If one produces US\$1,000 per day and the other US\$400, doing the first change before the second brings forward two days of the larger benefit while deferring two days of the smaller one. Over a common period starting before either change and ending after both are complete, the difference is US\$1,200 in favour of the first sequence. This original example assumes equal costs, no interaction, no discounting over the short comparison and adequate operating capacity. Unequal durations, shared prerequisites, contractual obligations or different adoption lags would require recalculation. A ranking based on benefit

divided by duration can offer a screening question, but should not replace the dependency model. The controller's task is to compare feasible sequences and explain the source of their value difference, including who experiences any deferred service or benefit.

Protect flow through scarce operating capacity

A feasible specialist schedule does not establish that Maya's team can absorb the release. The same employees may be handling ordinary invoices, learning new procedures, cleansing exceptions and supporting another initiative. Engwall and Jerbrant (2003), drawing on comparative case research in two multiproject organisations, described a resource-allocation syndrome involving recurring competition and reprioritisation. Their findings direct attention to the management of the wider project system rather than explaining every delay as a local planning failure. Lena therefore asks which commitments will be removed, postponed or covered when the new workload arrives. Capacity cannot be created by recording the same employee in several approved plans. The operational owner and resource manager need an agreed account of normal service, transition work, disruption and recovery, with an explicit decision when the combined demand exceeds what can credibly be supplied.

Zika-Viktorsson, Sundström and Engwall (2006) surveyed 392 project coworkers across nine Swedish companies. Perceived project overload was associated with stress reactions, reduced competence development and schedule deviations. The exploratory, self-reported evidence does not establish a universal maximum number of concurrent projects or prove a simple causal relationship. It does support looking beyond nominal hours when examining absorbable pace. Fragmentation, repeated handovers and interrupted concentration can make two apparently identical allocations very different in practice. In the manufacturer's release review, Lena asks how many distinct commitments Maya's team must switch between, which work demands continuity and what support is available after deployment. The answer should influence the sequence, rather than appear only as a wellbeing comment below an unchanged date. A credible plan makes room for the human work on which its technical assumptions depend.

Queueing offers a further explanation of why adding work can make an apparently busy organisation slower. Little's (1961) relation connects average work in the system, average throughput and average time in that system: $L = \lambda W$. Suppose an exception service has a stable average of twelve open cases, including those being processed, and completes three cases per working day. Average time in the defined system is then four working days. If average work in progress were nine while throughput remained three, the corresponding average time would be three days. This is a relation between matched long-run averages, not proof that imposing a limit of nine will preserve throughput or deliver every case within three days. A growing backlog, inconsistent boundaries or a snapshot count paired with an unrelated annual rate can invalidate the intended interpretation. The controller must first establish what enters, leaves and remains inside the measured service.

Kingman's (1961) heavy-traffic approximation makes the effect of variability and utilisation more explicit for a stable single-server queue. Expected waiting time is approximately q divided by one minus ρ , multiplied by half the sum of the squared coefficients of variation of arrival intervals

and service times, multiplied by mean service time. Here ρ denotes utilisation. Consider a separate fictional service with one server, a mean service time of one-third of a working day, and both squared coefficients equal to one. At eighty per cent utilisation, the implied arrival rate is 2.4 cases per day and approximate waiting time is 1.33 days. At ninety per cent, arrivals are 2.7 per day and waiting time is three days. Adding service gives approximately 1.67 and 3.33 days in the system. Figure 7.4 shows how a modest increase in loading can produce a much larger increase in waiting under the stated model.

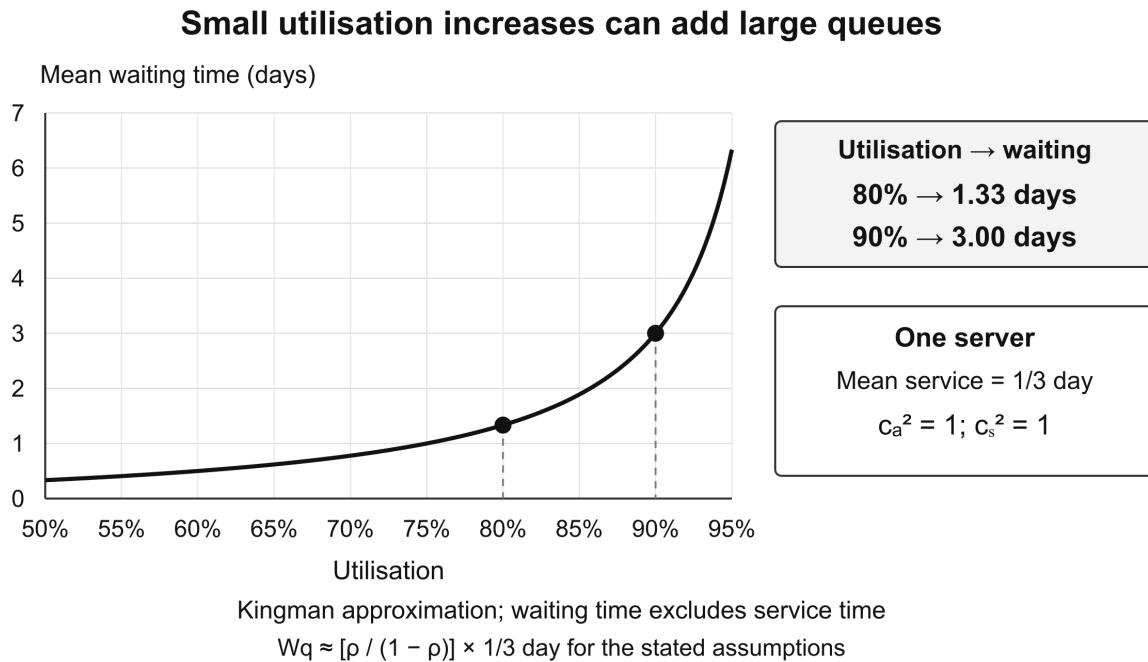


Figure 7.4 An original queueing illustration using Kingman's approximation. With mean service time one-third of a working day and both squared coefficients of variation equal to one, approximate waiting rises from 1.33 days at eighty per cent utilisation to three days at ninety per cent. The model does not prescribe a utilisation target.

The illustration assumes one first-come-first-served server with utilisation below one. Arrival intervals are independent draws from one distribution; service times are independent draws from another and independent of arrivals, with finite means and variances. It does not predict every department with outages, priorities or multiple specialists. Variability coefficients of one do not by themselves establish exponential distributions. Nevertheless, the example gives controllers a useful question: what happens to the queue when transition work consumes the capacity previously available for normal service? Adler and colleagues (1995) applied process-management and queueing perspectives to product development, showing why work release and congestion deserve managerial attention. The chapter's operational application is an adaptation of that reasoning. Lena might test a smaller release, protected exception-handling capacity or reduced variability at intake, then observe throughput and waiting together. She should avoid promising a particular improvement from a work-in-progress limit alone. A change to loading is a hypothesis about the operating system whose effects need measurement.

Expose information dependencies and the cost of overlap

Ordinary precedence diagrams work well when activities and their relationships can be represented as a directed sequence without feedback. Transformational work often contains information that develops through use. A test reveals that a payment rule is ambiguous; revising the rule requires configuration and data changes; those changes require another test. Eppinger and colleagues (1994) proposed a matrix-based method for examining technical task relationships in product development. A design structure matrix, or DSM, places the same activities on its rows and columns and marks where one requires information from another. Figure 7.5 adapts that idea to the payment release; its rules activity R corresponds to A in the schedule. Its convention is explicit: a marked cell means the row activity receives information from the column activity. The matrix helps the team distinguish a genuine information dependency from a sequence inherited merely because departments normally work in that order.

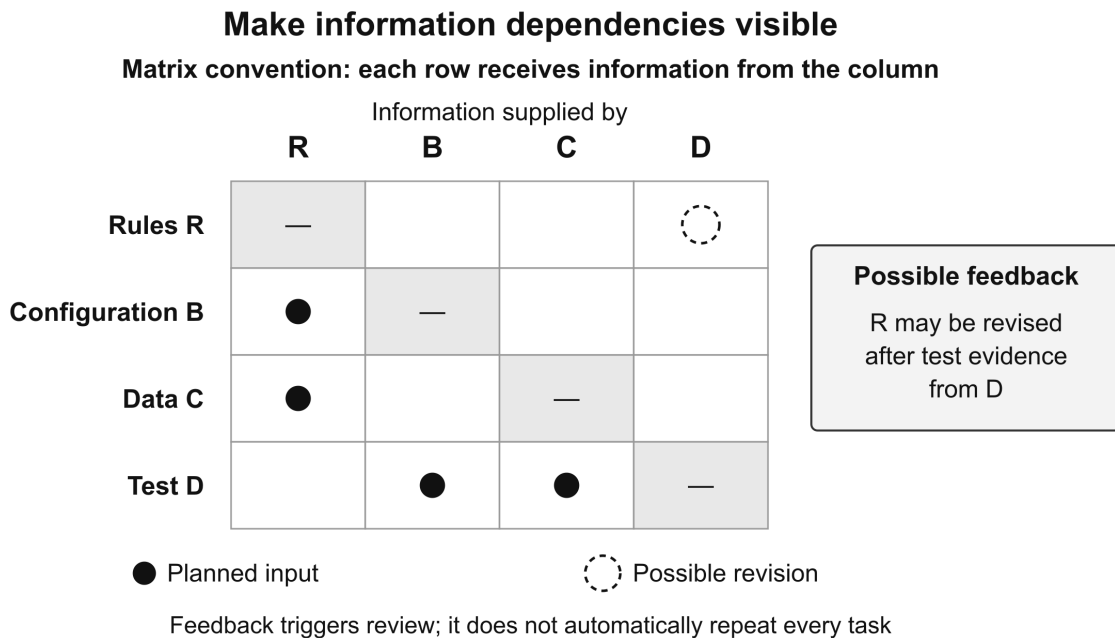


Figure 7.5 A small design structure matrix adapted for the fictional release. The row receives information from the column. Rules R corresponds to schedule activity A. Solid marks show planned inputs; the distinct feedback mark shows a possible test-driven revision requiring explicit treatment.

The matrix also reveals potential feedback from testing into rules. Smith and Eppinger (1997) used a work transformation matrix to examine controlling features of engineering design iteration, illustrated through automotive brake-system development. Their analytical treatment shows why coupled work may require repeated passes before convergence; it does not supply a universal rework percentage for transformation. In our example, Lena should ask whether the apparent eight-activity plan contains one test or an expected cycle of revision and retesting. If rework is credible, its scope, trigger, resource demand and acceptance condition need representation. A feedback arrow should not simply be inserted into the acyclic network used for the earlier arithmetic. The team can model a specified additional cycle, a defined risk event or a

more suitable iterative process, but must explain the treatment and avoid hiding material work inside a reassuring completion percentage.

Starting E before B is fully stable may appear to save time. Krishnan, Eppinger and Whitney (1997) developed a framework for overlapping product-development activities that examines the interaction between evolving upstream information and downstream sensitivity. Its value here is a disciplined question about what can safely proceed on provisional information. Maya may be able to teach general navigation early while postponing practice of the final authorisation rule. Treating those as separate learning activities is more credible than declaring the entire training package complete against an unstable design. An overlap proposal should identify the provisional input, the expected changes, the work exposed to revision, the confirmation date and the owner of correction. The schedule must include the resulting demand on people who may already be committed elsewhere. Otherwise an earlier start merely transfers unrecorded rework into operations.

A small calculation can help challenge the proposal without pretending to settle it. Suppose early practice advances a potentially governing activity by two days, but there is an assumed thirty per cent chance of three days of additional practice and correction. Expected additional activity duration is 0.9 working day under that simple scenario, but subtracting 0.9 from two does not establish a 1.1-day project acceleration. The network determines whether either the gain or the correction changes the finish, and the same trainers may be unavailable when correction is required. Both branches of the scenario need a feasible schedule. The percentage is fictional and should be replaced by evidence from comparable work or explicit expert judgement. The educational lesson is that expected activity duration, expected completion and the consequence of a delayed benefit are different quantities. A controller should preserve those distinctions when presenting an apparently economical overlap.

Forecast uncertainty across the whole sequence

The seventeen-day resource-feasible plan uses one duration for each activity. It says what follows from those inputs, not how likely the inputs are. Herroelen and Leus (2005) surveyed approaches to scheduling under uncertainty, distinguishing reactive, stochastic, proactive and other treatments. Their discussion helps separate a coordinating baseline from the policy used when actual work differs from that baseline. In the manufacturer's case, the forecast should examine uncertain durations and defined events while retaining the agreed reference for accountability. A baseline is not made more credible by repeatedly overwriting it with the latest result. Nor does protecting it justify ignoring new evidence. The controller needs a dated forecast, a clear execution policy and a record of what changed, so that stakeholders can understand both the current expectation and the performance against earlier commitments.

Table 7.3 supplies fictional triangular duration ranges, each expressed as minimum, most likely and maximum working days. These describe ordinary variability for the exercise. The ordinary duration draws are independent, and G remains fixed at one day. A separate shared data-quality event, independent of those duration draws, has an assumed twenty per cent probability and, if it occurs, adds two days to C and one day to D together. Those effects are expressly excluded from

the ordinary ranges to avoid counting the same exposure twice. For each simulation run, draw durations, draw the shared event once, apply its effects and recalculate the entire network. Repeat under the fixed B-before-C resource policy as well as the precedence-only policy. The model does not adaptively choose a different resource order in each run. This matters because a forecast conditional on a fixed policy should not be presented as the result of an optimisation that was never performed.

Table 7.3 Fictional duration inputs for the simulation

Activity	Minimum	Most likely	Maximum
A	1	2	3
B	3	4	6
C	2	3	5
D	1	2	4
E	2	3	4
F	1	2	3
G	1	1	1
H	2	3	5

Units are working days. A–F and H use independent triangular draws; G is fixed at one. An independent event draw with twenty per cent probability adds two days to C and one to D together. That event is excluded from the ordinary ranges. Resource runs retain B before C.

Van Dorp and Duffey (1999) examined statistical dependence in project-network risk analysis, developing a modelling approach and a shipbuilding illustration. Their work explains why durations influenced by common risks should not automatically be sampled as independent. The data-quality event here is an original teaching application of that concern, not a reproduction of their method. Logical dependency, resource competition and statistical dependence are different relationships. D waits for C because of logic; B and C compete because of a scarce specialist; C and D both lengthen because of a common event. A model can represent one and omit the others. Lena should ask what makes activities move together and which mechanisms the forecast actually includes. Adding more simulation runs cannot repair a missing resource constraint or an omitted source of shared delay.

Using 50,000 runs with the reproducible seed specified in the calculation record, the precedence-only model gives a median finish of 16.0 days and an eightieth percentile of 17.1. The B-before-C resource model gives 18.8 and 20.4 days respectively, rounded to one decimal place. Figure 7.6 compares the cumulative distributions. The median is the point at which half the modelled finishes occur at or before the indicated time; the eightieth percentile has eighty per cent at or before it. These are conditional results from invented inputs. They are not independently established probabilities that the real manufacturer will finish on those dates. The

deterministic seventeen-day plan is not the simulated mean or median because most-likely activity durations, merging paths and discrete events do not combine into a most-likely project duration through simple addition.

The same uncertainty produces different forecasts

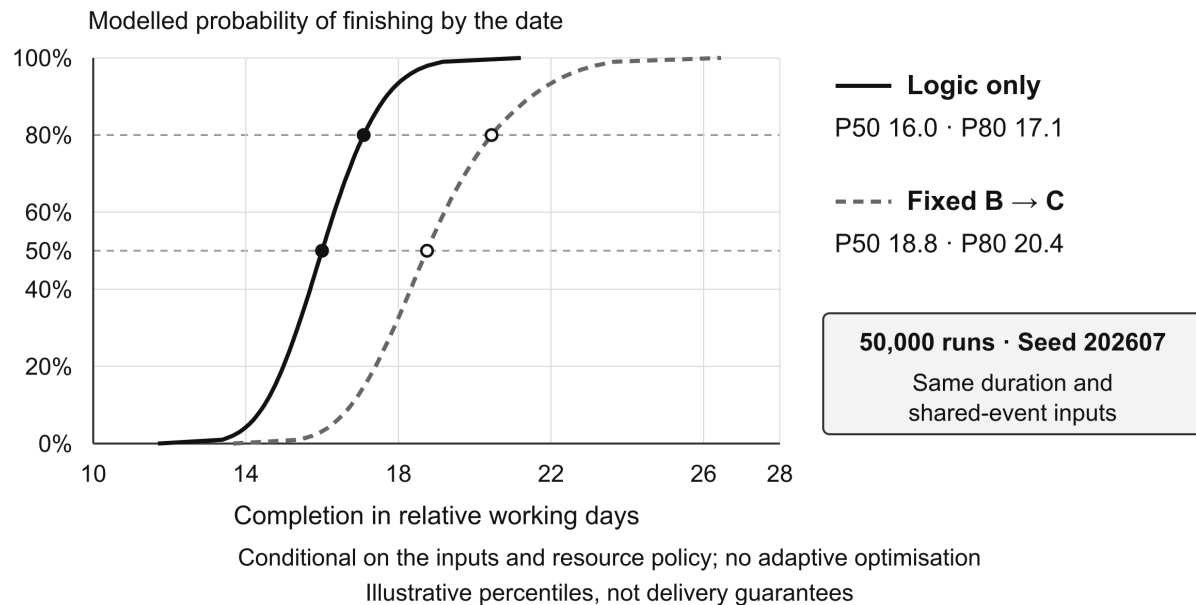


Figure 7.6 Conditional finish distributions from 50,000 fictional simulations. Precedence-only P50/P80 are 16.0/17.1 days; the fixed B-before-C resource policy gives 18.8/20.4. Both include the same shared data-quality event. These are model outputs, not independently established delivery probabilities.

The controller should test what materially changes the answer. Removing the specialist restriction, revising the common-event probability, adding a missing acceptance step or allowing a different resource policy may matter more than changing the number of runs. A criticality index can report how often an activity lies on a governing route in the chosen simulations, but that frequency is not the same as the causal value of accelerating it. Near-critical routes may become governing when another is shortened. Sensitivity analysis should therefore examine feasible interventions as well as statistical associations. The simulation record should preserve input sources, calendars, resource assumptions, exclusions, policy and update date. Where the consequences require whole working-day commitments, rounding and authorisation belong in the decision record. No universal requirement to promise the eightieth percentile follows from the calculation; the confidence level should reflect the decision and its consequences.

Place protection where it supports a commitment

Buffers make uncertainty visible when their purpose and ownership are clear. A project time buffer protects a delivery commitment, while a feeding buffer protects a merge into the sequence being protected. In critical-chain terminology, a resource buffer can function as a readiness warning rather than additional time. Herroelen and Leus (2001) examined merits and pitfalls of critical-chain scheduling, encouraging attention to its assumptions rather than automatic acceptance of a preferred label. The useful principle for this chapter is to connect protection to the resource-feasible sequence and the uncertainty it covers. It is not to remove every local allowance and then apply an arbitrary percentage elsewhere. Hidden padding, explicit event effects and an added reserve may cover the same exposure repeatedly unless their boundaries are examined. Protection should be explainable in the same language as the risks and commitments it serves.

Tukel, Rom and Eksioglu (2006) compared buffer-sizing techniques through simulation, including methods responsive to resource tightness and network complexity. Their results depend on the tested networks, duration assumptions and resource conditions; they do not establish a universal buffer percentage. In our example, the difference between the resource model's eightieth percentile and median is approximately 1.7 working days using the unrounded simulation results. That difference can describe the additional modelled time between two forecast positions, but is not automatically the correct project buffer. It depends on whether the organisation plans from the median, what protection already exists, which commitment is being made and which risks the model excludes. If the team instead retains seventeen as its internal reference, the distance to 20.4 is approximately 3.4 days. Calling both numbers the same reserve would conceal their different starting points.

Buffer monitoring should reveal decisions rather than provide another colour divorced from the work. Lena records the protected milestone, the reference forecast, the allowance, the owner and the evidence that triggers intervention. If C lengthens after poor data are discovered, the review examines the remaining sequence, resource availability and consequences for the operating trial. It does not simply divide consumed buffer by elapsed time and declare the project healthy or unhealthy. A late event near handover may have a different consequence from the same duration early in the work, especially where training or supplier windows cannot move easily. Agreed thresholds can help organise attention, but they are local decision rules to test and refine. Their value lies in prompting a timely, authorised response to a changing forecast, not in making uncertainty disappear behind an indicator.

Schedule robustness also concerns disruption to other commitments. A plan with the shortest expected finish may repeatedly move specialist bookings, training sessions and operational cover whenever an input changes. Another plan may finish slightly later while preserving more reliable handovers. The controller should make that trade-off visible when comparing alternatives. Recovery can include resequencing, obtaining capacity, reducing a wave's scope or changing a commitment through the proper authority. Each option needs its own consequences for cost, benefit timing and acceptance. A buffer does not give the project permission to consume somebody else's capacity without agreement. The original resource commitments and the current demand should remain reconciled, particularly when a recovery action protects one project by

destabilising another. This is where schedule control reconnects with the governance and ownership arrangements developed earlier in the book.

Release complete operating slices and let evidence mature

A transformation wave should leave behind an intelligible operating state. Delivering configuration in one wave, responsibility in another and support in a third may create activity completion without usable capability between them. The better question is what limited population, transaction type or service boundary can operate coherently, including its controls and exception arrangements. Ashurst, Doherty and Peppard (2008) developed a benefits-realisation capability perspective from exploratory research into information-system projects. Their work supports attention to organisational practices that continue beyond technical delivery. The chapter applies that perspective by making adoption, operation and evidence part of the wave definition. A small release can be valuable where it produces a usable service and answerable learning questions. Smallness alone is insufficient: an isolated fragment may require costly workarounds while providing little evidence about the eventual operating model.

Singapore's TradeNet provides a documented example of staged scope. An official APEC case records implementation on 1 January 1989, initially involving fifty users. The initial phase covered non-controlled and non-dutiable permits; controlled and dutiable goods followed, with later extensions including bank deductions and certificates of origin (APEC Committee on Trade and Investment, 2001, Singapore case, p. 2). The case illustrates sequencing by a defined transaction boundary within a connected service. It does not establish that staging alone caused the reported improvements, and the institutional setting differs from a manufacturer's payment operation. The transferable question is whether the first wave crosses enough organisational interfaces to function as a real service. A technically isolated pilot may hide precisely the coordination problem that expansion will expose. Controllers should ask what the chosen pilot represents and what remains untested before using its performance to justify scale.

Our worked sequence distinguishes release at day fourteen from stabilisation and acceptance at day seventeen. Suppose the first planned observation window then lasts five working days. Its earliest readout is day twenty-two, as Figure 7.7 shows. That readout may support a decision about the next wave, but it does not prove a sustained benefit or establish attribution. Payment measures may need a longer period, a sufficient number of transactions and a comparison that accounts for seasonality, invoice mix or other changes. The schedule should show when each type of evidence becomes available and what decision it can support. Evidence can mature at different rates: safe operation may be examinable before a financial saving appears in accounts. Waiting for every ultimate benefit before any expansion may be impractical, while treating an early operational signal as complete benefit proof is equally unsound.

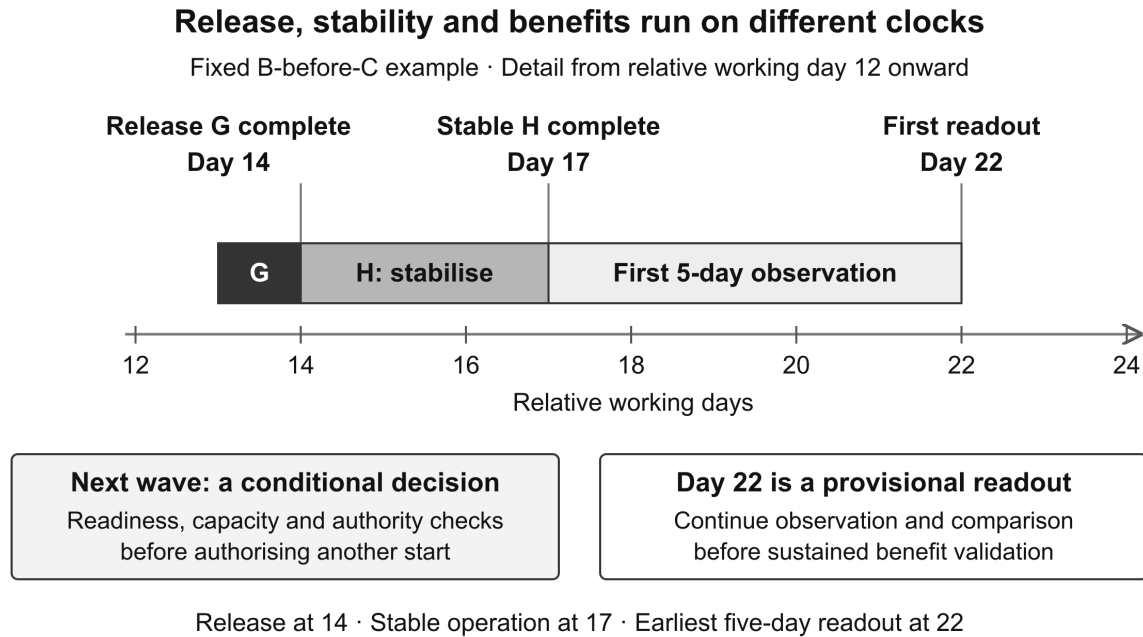


Figure 7.7 Distinct clocks in the deterministic B-before-C example. Release ends at day fourteen; stabilisation and acceptance end at seventeen; a five-day observation window first reports at twenty-two. Expansion requires an authorised decision, and sustained benefit assessment can require later evidence.

Serra and Kunc (2015) examined associations between benefits-realisation management and project success in Brazil, the United Kingdom and the United States. Their findings connect benefits practices with organisational value and success, while the survey design limits causal interpretation. For controllers, this reinforces the need to retain benefit responsibilities throughout delivery rather than attach a benefits report only at closure. Table 7.4 proposes a Wave Commitment Record linking the intended operating state to resources, acceptance, benefit evidence and the next authorised decision. It is a practical synthesis of the chapter, not an empirically validated instrument. The record should state which claims are observed, which remain forecasts and which depend on later action. Expansion then becomes a reasoned commitment based on specified evidence, with a continuing owner for benefits that will mature after the project team has moved elsewhere.

Table 7.4 Proposed Wave Commitment Record for the worked release

Control question	Proposed commitment	Evidence and next decision
Operating state	Process the agreed payment category with authorisation, exception handling and support.	Maya accepts demonstrated operation against defined criteria.
Capacity and sequence	One specialist; B before C. Operating team covers practice, trial and support.	Confirm bookings and ordinary-service cover before authorisation.
Milestones	Deterministic release 14; acceptance 17. Risk forecast assessed separately.	Retain reference, current forecast and any authorised commitment distinctly.
Benefit observation	First five-day window ends at 22; later assessment continues.	Retain denominator, comparison, cohort and owner; avoid treating the first readout as sustained proof.
Expansion	Samira authorises the next bounded wave when specified evidence supports it.	Record remaining assumptions, capacity and the next review; no automatic expansion at day 22.

All dates and conditions are fictional teaching assumptions. The record is an original practice aid connecting existing governance, ownership and benefits records; it is not a validated score or an approval already granted.

A cohort view helps preserve the timing of benefits as waves accumulate. Lena keeps separate records for the first business unit, the next unit and later adopters, with each cohort's release date, stable-operation date, eligible transaction population, observation period and benefit owner. An aggregate payment rate can improve merely because the mix of units changes, even when no individual unit improves. Comparing cohorts therefore requires attention to their starting conditions and transaction mix, rather than assuming that identical elapsed time means identical exposure. The benefit forecast can then be updated cohort by cohort, retaining the distinction between an expected improvement, an observed operating change and a verified financial consequence. This is especially useful where capacity is released gradually but a cash saving requires a later staffing or contract decision. The controller schedules that decision and its owner explicitly instead of converting every minute saved into immediate expenditure reduction. Benefits management becomes a continuing account of conditions and evidence, connected to the release sequence without being reduced to it.

Learn from the interfaces in international delivery

Heathrow Terminal 5 demonstrates the importance of the time between physical delivery and dependable operation. The UK House of Commons Transport Committee's 2008 inquiry examined the opening disruption; British Airways acknowledged that building delays had consumed testing and familiarisation time. Its account identifies shortcomings across preparation and operating interfaces, rather than establishing a single scheduling cause (House of Commons Transport Committee, 2008, pp. 3, 5–6). The lesson drawn here is that operational preparation has its own work content and cannot always be compressed to protect a public opening date. A building or system may be physically present while employees still need access, practice and integrated trials. In a transformation schedule, readiness activities should have real durations, resources and acceptance evidence. Treating them as expendable space at the end of construction or development transfers unresolved work into live service, where correction may become more disruptive.

The United States 2020 Census supplies a different view of the same interface. In its August 2018 report, GAO stated that, as of June 2018, thirty-six of forty-four systems supporting the 2018 End-to-End Test had completed development, while twenty had completed all system and integration testing (GAO, 2018, pp. 6–7). These were time-specific readiness measures for the test, not a verdict on the later census outcome. Their instructional value lies in the difference between developed components and tested integration. A controller reviewing a digital transformation should preserve both measures instead of reporting one overall completion percentage. A component can meet its own definition of done while the shared service remains unproven. The release sequence should identify the interfaces that must be tested together, the environments and people needed for that testing, and the evidence that permits progression to the next operating exposure.

Australia's migration to the National Broadband Network adds the perspective of continuity during transition. The Australian National Audit Office examined NBN Co's operations between 16 March 2018 and 28 March 2019 and judged its support for service continuity during migration largely effective, while identifying weaknesses in operational risk management and complaints (ANAO, 2019, summary paragraphs 10–14). Its account describes a typical eighteen-month migration window and dependencies involving retailers and end users. A network's availability did not by itself complete the customer's transition. The case helps distinguish infrastructure provision from the coordinated actions required to obtain an operating outcome. In another setting, the equivalent dependencies might involve suppliers registering payment details, customers adopting a service or local managers releasing staff. Those parties are not resources the project can simply schedule by command. Their commitments, incentives, information and support requirements need a place in the planning and escalation arrangements.

The three cases concern different technologies, institutions and measures, so they should not be combined into a numerical success formula. Together they help locate a recurring controls question: which interface remains unfinished when the headline milestone is reported complete? The answer may be physical access for practice, integration across systems or customer migration after infrastructure becomes available. The controller can translate that question into observable

acceptance conditions and named dependencies. Case evidence should retain its date, scope and source, especially where later performance changed. A documented problem is useful because it explains a mechanism worth examining, not because it permits every subsequent difficulty to be labelled the same failure. The handbook's international purpose is served by comparing the reasoning across contexts while preserving the institutional details that determine what a feasible response would actually require.

Join iterative delivery to shared commitments

Transformations often combine software discovery, fixed physical interfaces, procurement lead times and operational calendars. Collyer and colleagues (2010) examined delivery in dynamic environments and the importance of adapting planning and control to changing conditions. Their work supports questioning an assumption that every part of a project can be specified with equal certainty at the outset. It does not remove the need for commitments across organisational boundaries. In our manufacturer, an iterative configuration team may discover better ways to handle exceptions, while the operating team still needs notice to arrange cover and practice. The integrated plan should therefore state what the next release must enable, what remains provisional and when a decision must become stable enough for other teams to act. Detailed task planning can evolve within that structure without making every external dependency an unannounced moving target.

Serrador and Pinto (2015) analysed 1,002 projects and reported positive associations between agile use and project success. Gemino, Reich and Serrador (2021), studying 477 projects, found that hybrid and agile approaches could compare favourably on stakeholder success while traditional delivery dimensions showed similar outcomes across the categories examined. These observational studies do not establish that adopting a label causes benefits or that iteration automatically improves every project. The useful question is which planning practices fit the work and how their interfaces remain coherent. A software team can manage near-term tasks iteratively while sharing release assumptions, acceptance criteria, specialist needs and forecast changes with the programme. The controller needs evidence of integrated capability, not a conversion of story points into money or a declaration that all teams must use an identical method.

Consider a two-week development rhythm feeding a monthly operating release. Those cadences can coexist if the integration environment, acceptance work and operational preparation are planned explicitly. Completing an iteration does not guarantee a releasable operating slice; unfinished dependencies may span several iterations. Conversely, a release can contain changes from several teams without requiring them to estimate work in the same units. Lena maintains the shared milestones and dependency dates while allowing each team to manage suitable detail locally. She asks what evidence supports the current release forecast, which external commitment changed and what response is required. If a configuration discovery affects Maya's training material, the change should cross the team boundary immediately through the agreed ownership arrangement. Waiting until the next consolidated report would make administrative cadence more important than the real decision deadline.

Progressive elaboration should also retain a visible learning question. A pilot may test data quality, an operating assumption or the practicality of support, but its purpose should determine its scope and observation period. A pilot designed only to demonstrate favourable conditions cannot justify general expansion. The next wave needs an account of what was learned, which assumptions remain uncertain and how its population differs from the pilot. This connects Chapter 5's measurement discipline to the scheduling decision. Where evidence is weak, the appropriate response may be a narrower wave or a specific additional test rather than a blanket delay. Where the evidence contradicts the benefit mechanism, accelerating deployment can increase expenditure without improving value. The controller should ensure that learning has a route into scope, sequence and authorisation, so that iteration changes decisions as well as producing demonstrations.

Turn the method into a repeatable field practice

The Credible Control Method provides a practical way to use the chapter without introducing another lifecycle. Frame begins with the operating improvement and the decision requiring a date. Lena identifies the service boundary, benefit claim, proposed wave and consequence of delay. Establish connects the work packages, dependencies, resource commitments and acceptance conditions, then records the evidence behind durations and availability. The initial precedence calculation and resource test are different steps within that same examination. Challenge asks whether the network omits work, whether the people are actually available, whether uncertainty moves activities together and whether the operating team can absorb the proposed load. It also examines alternatives, including a smaller wave, changed resource order or additional capacity. These activities produce a recommendation that explains the date and its conditions, rather than asking the sponsor to approve a finish detached from the work.

Authorise concerns the commitment the organisation is actually making. Samira may approve a bounded wave, its funding and resource arrangement while retaining a separate decision before expansion. The approval should identify the acceptance authority, protected milestone, material assumptions, delegated tolerances and evidence required at the next review. It should also identify who can approve a change when the current sequence no longer fits. Verify then compares execution and operating evidence with the stated conditions. Completion evidence belongs with the activity; acceptance evidence belongs with the operating state; benefits evidence belongs with the improvement claim. The same report can connect all three, provided it preserves their meanings. The controller should avoid presenting authorisation as proof that the assumptions were correct. Approval gives authority to proceed under stated conditions; verification determines what subsequent evidence supports and what now needs reconsideration.

Routine schedule updates require a disciplined cut-off date. Actual starts and finishes should be supported by evidence, remaining durations should describe the work still required, and dependencies should reflect the current execution logic. Where work has started out of sequence, the controller should understand what happened before allowing the scheduling system to decide how the remaining work behaves. A partially completed test may still depend on an unavailable interface; simply marking it started does not remove that dependency. Retain the

baseline comparison, current forecast and explanation of material changes. Report the resulting effect on operating acceptance and benefit observation alongside the delivery milestone. This makes the update useful to Maya and Samira as well as Alex. The governing question is what changed in the conditions for the next decision, not merely which bars moved since the previous reporting cycle.

The field record should also permit another controller to reproduce a material conclusion. For the worked release, that means retaining the eight activities and their relationships, the shared specialist restriction, the chosen order, the duration ranges and the common-event definition. The calculation record uses 50,000 simulation runs, random seed 202607 and linear interpolation between ordered results for percentiles. These details support reproducibility, while input provenance determines whether the result deserves confidence. A review can then distinguish an arithmetic disagreement from a disagreement about resource availability or risk. Before issuing the next report, Lena reconciles the schedule's operating-acceptance date with the benefits register and asks Maya to confirm the observation window. If either has changed, the forecast and decision record change together. This modest reconciliation prevents two apparently authoritative reports from presenting incompatible versions of when value can begin. It also creates a usable record for later learning about which assumptions were reliable and which repeatedly failed in practice.

The final review brings the connected account back to the field. The unresourced network finished at fifteen; a feasible order finished at seventeen; a different order took eighteen; uncertainty shifted the conditional forecast; and benefit evidence arrived later than release. Each result answered a different question. Competent control preserves those distinctions while showing how they interact. An absorbable pace is the rate at which the organisation can deliver, adopt, operate and learn with credible commitments and acceptable consequences. It cannot be inferred from the number of activities completed or the percentage of employees allocated. The purpose of sequencing is to make the next valuable operating state achievable and examinable. Chapter 8 develops what happens to the evidence generated along that route: how information becomes organisational learning that can improve subsequent forecasts, commitments and decisions rather than disappear when the project team moves on.

References

- Adler, P. S., Mandelbaum, A., Nguyen, V., & Schwerer, E. (1995). From project to process management: An empirically-based framework for analyzing product development time. *Management Science*, 41(3), 458–484. doi: [10.1287/mnsc.41.3.458](https://doi.org/10.1287/mnsc.41.3.458)
- APEC Committee on Trade and Investment. (2001). Singapore's TradeNet system. In *APEC economies: Breaking down the barriers — Full case studies, 2001* (Singapore case, pp. 1–4). APEC Secretariat. www.apec.org/docs/default-source/publications/2001/12/apec-economies-breaking-down-the-barriers-full-case-studies-2001/01_cti_barriersfullstudy.pdf?sfvrsn=c21ca53_1
- Ashurst, C., Doherty, N. F., & Peppard, J. (2008). Improving the impact of IT development projects: The benefits realization capability model. *European Journal of Information Systems*, 17(4), 352–370. doi: [10.1057/ejis.2008.33](https://doi.org/10.1057/ejis.2008.33)

- Australian National Audit Office (ANAO). (2019, 28 November). National Broadband Network fixed line migration — Service continuity and complaints management. Auditor-General Report No. 15 of 2019–20. www.anao.gov.au/work/performance-audit/national-broadband-network-fixed-line-migration-service-continuity-and-complaints-management
- Collyer, S., Warren, C., Hemsley, B., & Stevens, C. (2010). Aim, fire, aim—Project planning styles in dynamic environments. *Project Management Journal*, 41(4), 108–121. doi: 10.1002/pmj.20199
- Engwall, M., & Jerbrant, A. (2003). The resource allocation syndrome: The prime challenge of multi-project management? *International Journal of Project Management*, 21(6), 403–409. doi: 10.1016/S0263-7863(02)00113-8
- Eppinger, S. D., Whitney, D. E., Smith, R. P., & Gebala, D. A. (1994). A model-based method for organizing tasks in product development. *Research in Engineering Design*, 6(1), 1–13. doi: 10.1007/BF01588087
- Gemino, A., Reich, B. H., & Serrador, P. M. (2021). Agile, traditional, and hybrid approaches to project success: Is hybrid a poor second choice? *Project Management Journal*, 52(2), 161–175. doi: 10.1177/8756972820973082
- Herroelen, W., & Leus, R. (2001). On the merits and pitfalls of critical chain scheduling. *Journal of Operations Management*, 19(5), 559–577. doi: 10.1016/S0272-6963(01)00054-7
- Herroelen, W., & Leus, R. (2005). Project scheduling under uncertainty: Survey and research potentials. *European Journal of Operational Research*, 165(2), 289–306. doi: 10.1016/j.ejor.2004.04.002
- House of Commons Transport Committee. (2008, 3 November). The opening of Heathrow Terminal 5. Twelfth Report of Session 2007–08, HC 543. The Stationery Office. publications.parliament.uk/pa/cm200708/cmselect/cmtran/543/543.pdf
- Kingman, J. F. C. (1961). The single server queue in heavy traffic. *Proceedings of the Cambridge Philosophical Society*, 57(4), 902–904. doi: 10.1017/S0305004100036094
- Kolisch, R., & Hartmann, S. (2006). Experimental investigation of heuristics for resource-constrained project scheduling: An update. *European Journal of Operational Research*, 174(1), 23–37. doi: 10.1016/j.ejor.2005.01.065
- Krishnan, V., Eppinger, S. D., & Whitney, D. E. (1997). A model-based framework to overlap product development activities. *Management Science*, 43(4), 437–451. doi: 10.1287/mnsc.43.4.437
- Little, J. D. C. (1961). A proof for the queuing formula: $L = \lambda W$. *Operations Research*, 9(3), 383–387. doi: 10.1287/opre.9.3.383
- 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. doi: 10.1016/j.ijproman.2014.03.011
- 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. doi: 10.1016/j.ijproman.2015.01.006
- Smith, R. P., & Eppinger, S. D. (1997). Identifying controlling features of engineering design iteration. *Management Science*, 43(3), 276–293. doi: 10.1287/mnsc.43.3.276
- Tukel, O. I., Rom, W. O., & Eksioglu, S. D. (2006). An investigation of buffer sizing techniques in critical chain scheduling. *European Journal of Operational Research*, 172(2), 401–416. doi: 10.1016/j.ejor.2004.10.019
- US Government Accountability Office (GAO). (2015, 22 December). Schedule assessment guide: Best practices for project schedules (GAO-16-89G). www.gao.gov/products/gao-16-89g
- US Government Accountability Office (GAO). (2018, 30 August). 2020 Census: Continued management attention needed to address challenges and risks with developing, testing, and securing IT systems (GAO-18-655). www.gao.gov/products/gao-18-655
- Van Dorp, J. R., & Duffey, M. R. (1999). Statistical dependence in risk analysis for project networks using Monte Carlo methods. *International Journal of Production Economics*, 58(1), 17–29. doi: 10.1016/S0925-5273(98)00081-4

Zika-Viktorsson, A., Sundström, P., & Engwall, M. (2006). Project overload: An exploratory study of work and management in multi-project settings. *International Journal of Project Management*, 24(5), 385–394. doi: [10.1016/j.ijproman.2006.02.010](https://doi.org/10.1016/j.ijproman.2006.02.010)

Zwikaël, O., Chih, Y.-Y., & Meredith, J. R. (2018). Project benefit management: Setting effective target benefits. *International Journal of Project Management*, 36(4), 650–658. doi: [10.1016/j.ijproman.2018.01.002](https://doi.org/10.1016/j.ijproman.2018.01.002)

CHAPTER 08

Turning Information into Organisational Learning



IN THIS CHAPTER

Turning reliable information into better decisions and changed practice.

Make the evidence change what happens next

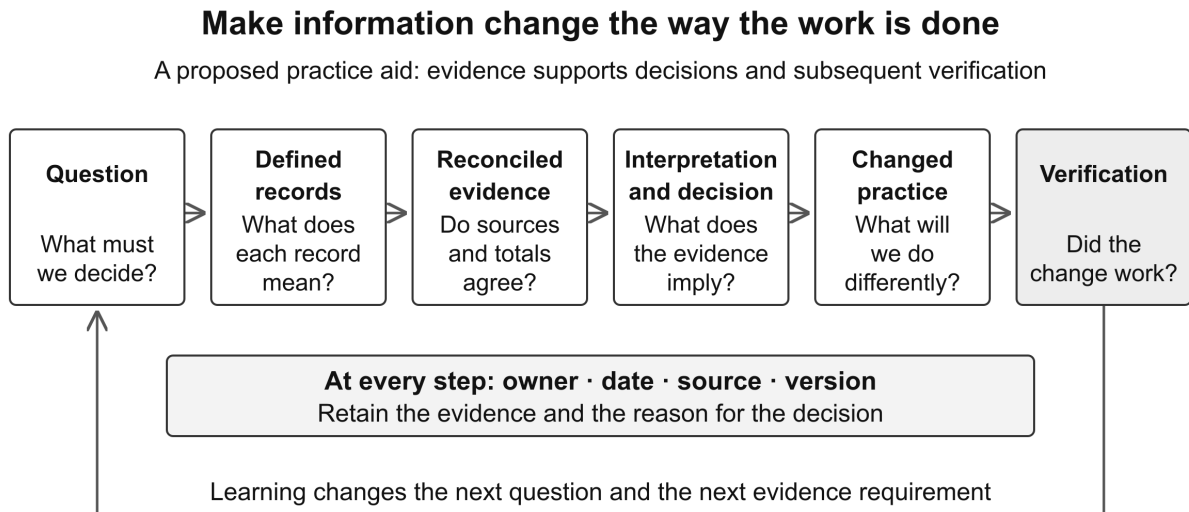
The manufacturer's monthly report has been corrected. A supplier's accepted work was missing from the cost account because its invoice had not arrived, and the benefits dashboard treated unresolved payment records as though they were on time. Lena, the project controller, has reconciled the figures and explained the difference. Samira, the sponsor, now asks whether the same problems will recur next month. A corrected report cannot answer that question. The organisation needs to understand how the omissions arose, decide what should change and examine whether the change works. This chapter develops that movement from information to learning. Its central argument is that project controls creates lasting value when evidence changes future estimates, operating practices and decisions, with a record that allows others to understand both the lesson and the conditions under which it applies.

Chapter 7 connected the delivery sequence to available resources, operational acceptance and benefit observation. Those connections generate information about what actually happens when an organisation attempts transformation. Some of that information confirms the plan; some exposes weak assumptions; some is too incomplete or inconsistent to support a conclusion. A learning organisation needs to distinguish these situations before converting experience into guidance. The controller therefore has two related responsibilities: establish a credible account of the event and help the organisation use that account. Recording a variance fulfils only part of the first responsibility. Circulating a presentation labelled lessons learned fulfils neither responsibility unless its explanation is defensible and its implications reach someone able to act. Learning should become visible in the work that follows, rather than be inferred from the volume of material stored.

Williams (2008) examined how projects learn lessons through a literature review and a survey with 522 usable responses. The respondent group was largely drawn from the Project Management Institute and subject to self-selection, so it should not be treated as a representative census of organisations. Respondents reported difficulty turning project experience into knowledge used beyond the team. This helps frame the chapter's proposed Evidence to Learning Map in Figure 8.1. The map connects a decision question, defined records, reconciled evidence, interpretation, authorised change and verification. It is an original practice aid within the Transformation Control Loop, not a validated maturity scale. Its test is concrete: what will someone do differently because of this evidence, who has authority to make that change and what later observation could show whether it helped?

The map uses Chapter 2's Credible Control Method rather than adding another lifecycle. Frame establishes the question; Establish assembles and checks the account; Challenge examines explanations and alternatives; Authorise gives effect to an agreed response; Verify determines what subsequent evidence supports. Information management runs through these activities. A reliable source cannot answer an undefined question, and a plausible explanation cannot implement itself. The method also allows proportionate action before every uncertainty has been resolved. A controller may correct a known coding error immediately while retaining a separate investigation into why it escaped review. The record should distinguish that justified correction from a causal conclusion still being tested. This preserves room for learning without allowing

provisional interpretations to harden unnoticed into organisational facts or mandatory procedures.



Original synthesis; the workflow is a proposed aid, not an empirically validated model

Figure 8.1 The proposed Evidence to Learning Map. Defined records support interpretation and authorised changes, followed by verification in later work. Ownership, sources and dates preserve the connection. This original practice aid operates within the existing Credible Control Method.

Agree what the records mean before combining them

The first task is semantic: establish what each important field represents. An approved budget, a purchase order, a posted cost, a payment and a forecast are different statements, even when each appears in a column labelled amount. A completion date may refer to installation, technical acceptance or operational acceptance. A benefit measure may describe all eligible transactions, only records with complete data or a selected pilot population. These differences should be explicit in a compact data definition linked to the relevant record. Lena specifies the identifier, scope, unit, reporting period, inclusion rule, source, owner and status of each material measure. The same definition should explain how a value is calculated and what would make it unsuitable for a particular decision. Agreement at this level is more useful than requiring every department to adopt the same software or report layout. The definition owner should approve each material revision and retain its effective date, so later comparisons can distinguish changed performance from a changed measurement basis.

Wang and Strong (1996) investigated data quality from the perspective of data consumers and showed why quality extends beyond accuracy alone. Their framework includes contextual, representational and accessibility considerations alongside intrinsic quality. For project controls, the implication is that a technically correct record may still be unusable because it concerns the wrong period, lacks an intelligible definition or cannot be obtained before the decision. Pipino, Lee and Wang (2002) similarly discuss combining subjective and objective assessment of data quality. A user's confidence and a measurable validation result can expose different weaknesses.

These contributions support asking who needs the information and for what purpose before constructing a quality score. They do not justify treating satisfaction with a dashboard as proof that the underlying transactions are complete, accurate or appropriately interpreted.

Lee and colleagues' (2002) AIMQ methodology links information-quality assessment to analysis and improvement. Its relevance is the movement from identifying a weakness to deciding which information process needs attention. The UK's Government Data Quality Framework distinguishes completeness, uniqueness, consistency, timeliness, validity and accuracy, and emphasises fitness for purpose (Government Data Quality Hub, 2020). These dimensions should remain distinguishable in a controls review. A payment date can be present and in the correct format while referring to the wrong transaction. Two reports can agree because both inherited the same error. An average quality score can conceal a decisive weakness in the one field needed for the current approval. Lena therefore reports the relevant defects and their consequences directly, rather than allowing several favourable indicators to compensate arithmetically for an unsupported claim.

Traceability needs stable identities as well as definitions. A supplier, invoice, work package and benefit cohort should have identifiers that permit their relationships to be checked. Joining two datasets by a familiar description can duplicate records where names are repeated or changed. A one-to-many relationship can also multiply an amount if the report sums a repeated invoice value after linking it to several acceptance records. Before interpreting a consolidated total, the controller tests the expected record count, uniqueness and relationship between the sources. The account should retain the source version, extraction time and transformation rules so that another competent person can reproduce it. Access should match responsibility and legitimate need; traceability does not require every user to see confidential supplier or employee information. A controlled route to evidence is sufficient where the decision maker can obtain the assurance needed.

Reconcile the accounts at one reporting cut off

Consider a separate fictional supplier work package, expressed in thousands of US dollars at one reporting cut-off. Its approved budget and purchase order are both 500. Posted cost is 200, supported by invoices for accepted work, while a further 80 of accepted service has not yet been invoiced or included in that posted amount. Under the exercise's agreed recognition basis, the 80 requires an accrual in the control account. Cash paid is 150. The accounting system's open-order report shows 300 because it reduces the original order only by the 200 invoiced. The remaining work is currently forecast to cost 250. These are deliberately specified system behaviours, not universal definitions used by every financial application. They concern one teaching package and do not replace the manufacturer's programme-wide budget or forecast established earlier in the book.

The reconciled incurred cost is 280: posted cost of 200 plus the unposted accrual of 80. The open order of 300 includes that same accepted but uninvoiced service, so only 220 relates to unperformed ordered work. Adding incurred cost of 280 to the whole open order would produce 580 by counting the 80 twice. It would also substitute the original order balance for a current

estimate of the remaining work. The current forecast is instead 280 incurred plus 250 remaining, giving 530. The remaining forecast exceeds the unperformed order value by 30, which needs an explanation and any required commercial or budget authority. Table 8.1 keeps the accounts separate, and Figure 8.2 shows their boundaries. A forecast of additional cost identifies an exposure; it does not itself authorise a larger order or an increase to the approved budget.

Table 8.1 Separate the cost and commitment accounts

Account	US\$000	Meaning at the fictional cut-off
Approved budget	500	Authorised package reference; not revised by the forecast.
Purchase order	500	Original authorised order for this package.
Posted cost	200	Invoices for accepted work recorded in the account.
Unposted accrual	80	Accepted service not yet invoiced or in the posted cost.
Incurred cost	280	200 posted plus 80 accrued.
Open order	300	500 less 200 invoiced; includes the accrued 80.
Unperformed order	220	300 open order less 80 already accepted.
Remaining forecast	250	Current estimate for unperformed work; includes 30 growth.
Forecast total	530	280 incurred plus 250 remaining.
Cash paid	150	Settlement of part of the posted 200; not added again.

Cumulative values at one reporting cut-off, excluding tax and currency changes. Source-system behaviour and recognition assumptions are specified in the text. The remaining forecast is not an additional commitment approval.

Cash adds a further distinction. Of the 200 posted cost, 150 has been paid and 50 remains unpaid in the exercise. The payment of 150 is a settlement of part of the recorded obligation, not another cost to add to 280. When the invoice for the accrued 80 arrives, the accounting treatment should replace or reverse the relevant accrual as appropriate under the organisation's policy, preventing the amount from appearing twice. The controller reconciles these movements with finance rather than maintaining an unexplained parallel ledger. The US GAO's Cost Estimating and Assessment Guide stresses updating estimates with actual costs, documenting changes and retaining reasons for variance so that estimates can improve (GAO, 2020). The teaching application here is to preserve the bridge between source records and the current forecast, including the recognition basis and unresolved differences.

The same work appears in different records

One fictional supplier package · Cumulative at one reporting cut-off · US\$000

Order 500	200	80	220
	Invoiced	Accepted unbilled	Unperformed
Current forecast 530	200	80	250
	Posted	Accrual	Remaining forecast

250 includes the 30 growth

Cash paid 150 is part of posted cost 200

Incurred 280 + remaining forecast 250 = forecast 530

The 30 growth is already included in the forecast; cash is a separate payment view

Figure 8.2 Fictional reconciliation in US\$ thousands. The order balance includes accepted but uninvoiced service. The current forecast uses incurred cost of 280 plus remaining work of 250, totalling 530. Cash paid is a settlement within posted cost, not an additional cost component.

Cut-off discipline makes that bridge reproducible. Lena records the period to which the service belongs, when it was accepted, when the invoice was received and when the report was extracted. A late-arriving invoice may clarify an earlier period without representing new work in the current one. The treatment must be explicit and consistent with the financial account. Where reports are restated, both the original issued version and the reason for the correction should remain accessible. A silent replacement prevents later reviewers from understanding what the decision maker actually knew. The same principle applies to schedule actuals, risk assessments and benefits data: the date of the event and the date on which it became known may differ. A useful learning record preserves both, because delayed information and delayed action can require different remedies.

An event timeline makes the distinction operational. Suppose service was accepted on the twenty-eighth, the monthly report was closed on the thirtieth, and the acceptance record reached finance on the second day of the following month. The service event preceded the cut-off, but the information did not reach the person preparing the account in time. The investigation should examine the handover and any alternative evidence available before closure, rather than assume that finance ignored a record it had received. Conversely, if finance had received the record on the twenty-ninth, the review would examine the processing and reconciliation arrangements. These are fictional alternatives for testing an explanation, not excuses established by the final invoice date. Lena retains the timestamps and ownership at each handover, identifies the latest useful arrival time and agrees what happens when a material record is missing. The reporting owner can then issue a qualified account, seek additional evidence or escalate the uncertainty according to its consequence, with the response visible in the decision history.

Investigate the uncertainty behind a performance claim

The benefits account needs equivalent care. In another explicitly fictional cohort, one hundred invoices are eligible for a payment-timeliness measure. Ninety-five have verified outcomes: eighty-nine were on time and six late. Five remain unresolved. The status field is therefore ninety-five per cent complete, but the late-payment proportion for the whole eligible cohort is not yet established. If every unresolved invoice was on time, the proportion would be six per cent; if all five were late, it would be eleven per cent. Figure 8.3 displays those possible completions of the missing information. This is not a statistical confidence interval or a claim that the values are equally likely. It is an arithmetic account of what remains possible given a fixed, complete denominator and the stated classifications. If eligible invoices were missing from the denominator itself, the same bounds would no longer be justified.

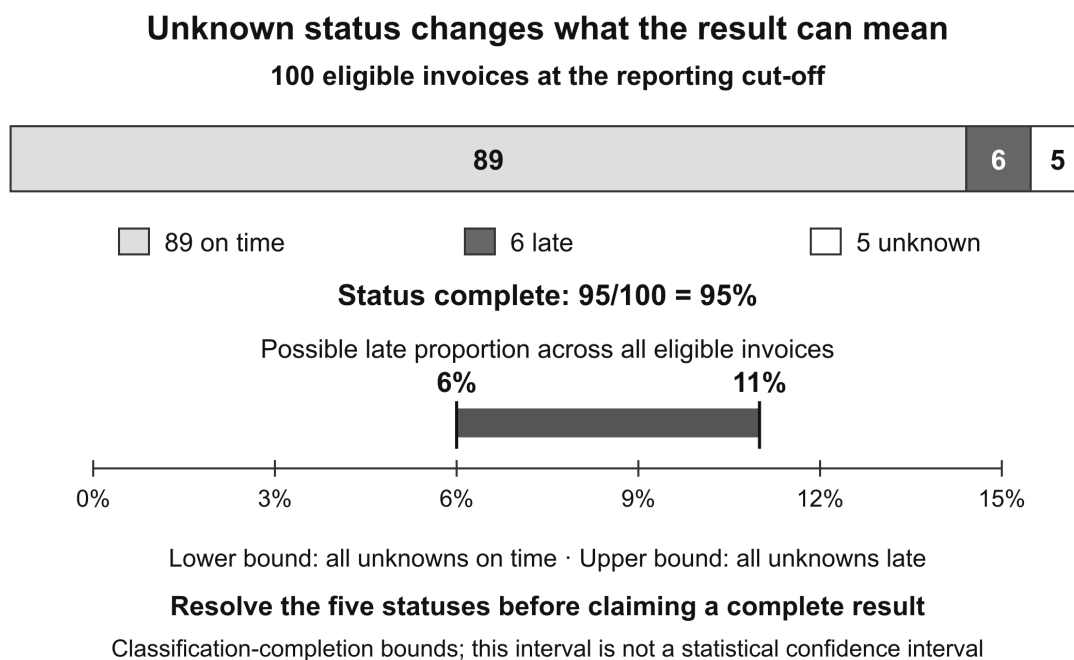


Figure 8.3 Missing classifications change what can be claimed. With one hundred eligible invoices, six known late and five unresolved, the final late proportion could range from six to eleven per cent. These are possible completions of the missing data, not a statistical confidence interval.

Reporting six divided by ninety-five, approximately 6.3 per cent, would answer a narrower question about records with known outcomes. It would not resolve the five unknown cases or establish that they resemble the observed cases. If difficult exceptions take longer to reconcile, missingness could be related to the outcome being measured. Lena should state the unresolved population, investigate its composition and explain whether the pending decision can tolerate that uncertainty. A claim that a six per cent target has been achieved is premature when the final proportion could be eleven. The next action may be to verify five records rather than commission a complex model. The broader lesson is that uncertainty sometimes arises from a specific information gap that can be closed, while at other times it remains part of the system being forecast. Effective controls identifies which situation it faces.

Interpretation begins when the team asks what the discrepancy means. Weick, Sutcliffe and Obstfeld (2005) describe organisational sensemaking as a process through which circumstances become intelligible enough to inform action. Their conceptual account directs attention to the social construction of explanations, rather than treating a report as though it speaks independently of its users. In our manufacturer, late records could reflect an unclear definition, a system interface, insufficient capacity or an incentive to preserve a favourable status. Those possibilities require different responses. Lena first reconstructs the sequence using dated records and the accounts of people involved, then distinguishes observations from explanations. Asking repeatedly why something happened can encourage inquiry, but stopping at the first plausible story can create a confident fiction. The explanation should identify evidence that would support it, evidence that would weaken it and alternative mechanisms still consistent with the facts.

The Columbia Accident Investigation Board provides a consequential example of the difference between possessing information systems and using organisational knowledge. Its 2003 report found that NASA's lessons-learned arrangements were not effectively informing senior decisions, despite existing mechanisms for recording and disseminating experience (CAIB, 2003, pp. 189, 193). This was an inquiry into a particular accident and organisational context; it does not establish that every unused lesson causes failure. Its relevance is that a repository cannot confer authority on an observation or ensure that it reaches a consequential decision. A controls review should therefore ask where a warning went, how it was interpreted and what response followed. Improving search or adding another dashboard may help access, but the investigation must also examine the organisational conditions that determine whether unwelcome evidence can change the course of the work.

Review experience while the work is still understood

A useful review occurs while participants can still explain the context, including the constraints and information available at the time. Schindler and Eppler (2003) examined methods for retaining project knowledge and the obstacles that impede their use. Their contribution supports treating learning as an organised part of project work rather than a document produced hurriedly at closure. For the payment release, Lena convenes a focused review after the disputed cut-off has been reconciled. Alex explains the delivery sequence, Maya describes service acceptance, and finance explains how the accrual and invoice entered the account. The review begins with the intended result and compares it with what occurred before assigning causes. It should preserve disagreement where the evidence cannot resolve it. A tidy consensus is not necessarily a better explanation than a record of competing hypotheses with a plan to examine them.

Tannenbaum and Cerasoli's (2013) meta-analysis found a positive average performance effect across forty-six debriefing samples, while the variety of settings and designs limits what can be promised for a particular transformation. The implication is to organise reflection around the work, rather than guarantee a percentage improvement from holding a meeting. A disciplined debrief examines what was intended, what occurred, why the difference may have arisen and what should change. It needs evidence, participation from relevant roles and an output that can influence subsequent action. In the manufacturer's review, the important result is not the

statement that communication should improve. It is an agreed change to how accepted service reaches the monthly cost account, with an owner, a cut-off requirement and a check at the next reporting cycle. Specificity makes the proposed learning examinable.

Successful work also deserves scrutiny. Ellis and Davidi (2005) compared reviews addressing successful and failed events with reviews focused on failures in a military navigation setting. The study supports the value of examining success as well as error, while its bounded population and company-level assignment limit generalisation. A successful release may have depended on an experienced employee's undocumented intervention, unusually simple transactions or spare capacity that will not exist in the next wave. Declaring the method proven can conceal those conditions. Lena therefore asks what made the satisfactory outcome possible and whether the organisation can deliberately reproduce it. This inquiry should also examine near misses and recoveries, because a favourable final result can follow a fragile process. The objective is to understand the mechanism sufficiently well to use it, rather than attach praise or blame to the outcome alone.

A review should produce a small number of consequential actions that people can complete and test. Each action needs an accountable owner with authority and capacity, as Chapter 4 established. It should identify the procedure, estimate, interface or decision rule that will change, the first occasion on which the new approach will be used and the evidence required for verification. Actions should remain distinct from recommendations still awaiting approval. Where the review identifies a problem beyond the project manager's delegation, the governance route must carry it to the appropriate authority. A recurring defect should not be assigned repeatedly to the person who reports it merely because that person is visible. The relevant question is who can change the conditions producing the defect. This protects the review from becoming another mechanism for recording obligations without making their fulfilment possible.

Question the rules that reproduce the problem

Some learning improves action within an existing rule; other learning examines the rule itself. Argyris (1976) developed the distinction between single-loop and double-loop learning in organisational decision making. Figure 8.4 applies that conceptual distinction to our fictional controls problem. Correcting the missing accrual brings the current report back into line with the agreed recognition basis. Asking why service acceptance did not reach finance before the cut-off examines the reporting arrangement that allowed the omission. A further question might challenge the performance rule that rewards prompt report submission while ignoring unresolved material inputs. These are different depths of inquiry, and the appropriate response depends on the evidence. Double-loop learning is not a requirement to reopen every policy after every error. It becomes useful when repeated corrections suggest that the governing assumptions or incentives need examination. Changing a procedure is not necessarily double-loop learning; that distinction depends on whether the inquiry changes the governing assumptions or objectives.

Correct the result and reconsider how it was produced

Same discrepancy, two responses · Accrual amounts in US\$000

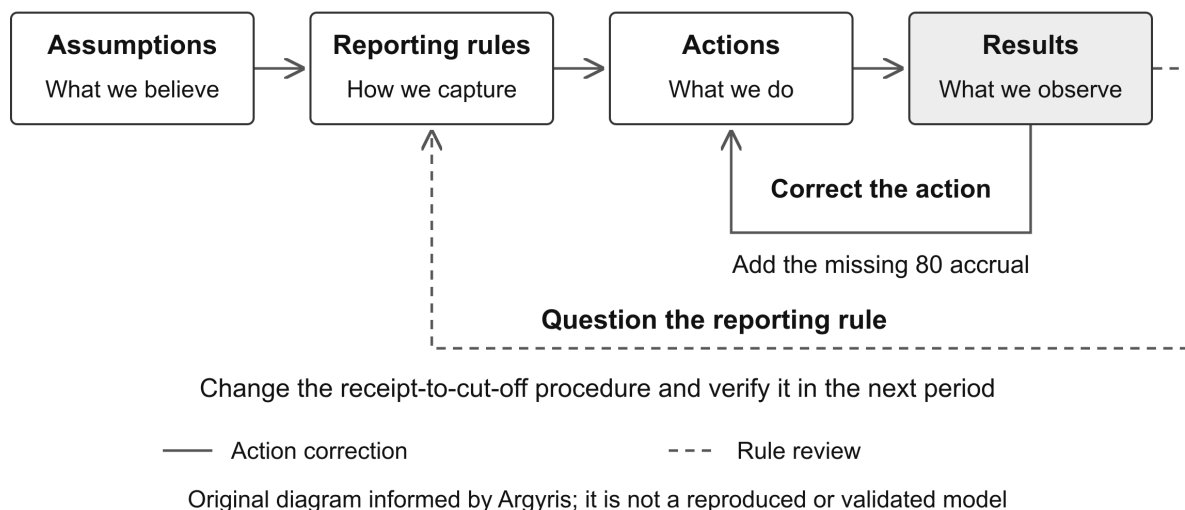


Figure 8.4 An original application informed by Argyris (1976). Correcting an action differs from reviewing the rule and its assumptions. A procedural adjustment is double-loop learning only where the inquiry changes governing assumptions or objectives; routine redesign alone does not establish that distinction.

Questioning a rule can create interpersonal risk. Edmondson's (1999) study of fifty-one teams in one manufacturing company found associations between team psychological safety and learning behaviour. The evidence does not establish that a friendly atmosphere automatically improves results or removes the need for accountability. Its practical relevance is whether people can disclose a mistake, ask for help or challenge an assumption without expecting humiliation. In the manufacturer's review, Samira can make this credible by asking what the organisation must change and by responding constructively when the evidence contradicts an earlier assurance. The controller still needs accurate records and clear responsibilities. A learning environment combines candour with examination of the work; it does not replace examination with reassurance. Where misconduct or deliberate concealment is suspected, the relevant process should remain distinct from a speculative search for someone to blame.

The IAEA's analysis of Fukushima illustrates the importance of deciding which experience deserves attention. Its 2015 technical account describes how TEPCO translated experience from the 2007 domestic earthquake into measures including emergency facilities and firefighting arrangements. It also describes how experience at India's Madras nuclear power station during the 2004 tsunami was screened out because the actual consequences were limited, despite the potential relevance of the initiating event (IAEA, 2015, Technical Volume 2, section 2.7.2.1). The comparison supports examining potential consequences and underlying mechanisms, not merely the severity of the observed outcome. It does not establish that one transferred lesson would have prevented every subsequent failure. For project controls, a near miss should prompt inquiry into what prevented harm and whether that protection will exist when the exposure recurs elsewhere.

The response must remain proportionate to the finding. If the accrual omission arose from one incorrectly mapped identifier, a corrected mapping and an effective reconciliation check may be

sufficient. If it recurs because nobody owns the acceptance-to-finance handover, a stronger validation rule alone leaves the organisational gap intact. If the rule deliberately excludes accepted but uninvoiced service from a performance claim, the meaning of the claim itself needs reconsideration. Lena records which level is being changed and who approved it. The resulting action should not erase the previous versions of the rule or report. Retaining that history allows a later reviewer to understand why the organisation changed its practice and whether the new arrangement addressed the original mechanism. Without this connection, process changes can accumulate until nobody knows which problem each control was intended to solve.

Transfer the reasoning as well as the lesson

A lesson is portable only to the extent that its meaning survives transfer. Argote and Ingram (2000) examine knowledge transfer in organisations through the interactions of people, tasks and tools. Their account helps explain why sending a document may not reproduce the practice it describes. The recipient needs to understand the problem, the conditions that made the response useful and the capabilities required to apply it. When the manufacturer prepares its next business unit, Lena should therefore connect the receiving controller with the people who performed the first reconciliation. A worked example, access to the relevant definitions and a supervised trial can reveal misunderstandings that a written instruction would leave hidden. The transfer should preserve the reasoning behind the rule while allowing appropriate adaptation to local systems and authority. Copying the wording is a weak substitute for demonstrating that the next team can use it.

Szulanski's (1996) investigation of internal practice transfer identified barriers associated with recipient capacity, causal ambiguity and relationships between the source and recipient. The study concerned transfers within a limited set of companies, not a universal estimate of how difficult every lesson will be to apply. Its relevance is that implementation effort belongs in the learning plan. A receiving team may lack time, access or experience even when it accepts the lesson's importance. The source team may understand a successful intervention without being able to explain which parts mattered. Lena consequently asks the receiving unit to describe its own version of the problem and demonstrate the proposed control on representative records. Differences become explicit before adoption. Where local conditions materially change the mechanism, the lesson should be treated as a hypothesis for testing rather than a standard that must be copied unchanged.

A practical transfer rehearsal can use three deliberately different records: accepted work awaiting an invoice, an invoice already matched to an accrual, and an order for work that has not started. The receiving controller explains how each enters the account and demonstrates the relevant source checks. The source team then introduces a changed identifier or a record arriving after the reporting cut-off. The purpose is to discover whether the receiving team understands the boundary, including when to stop and seek clarification. Passing the rehearsal is evidence of understanding under those conditions, not proof that every future transaction will be handled correctly. The next live cycle should therefore include an agreed review of actual use and unresolved exceptions. Lena records any adaptation made for the receiving system and obtains the relevant owner's agreement before calling the transfer complete. This approach turns a

general instruction to share lessons into a bounded piece of implementation work, with observable performance and a clear route for correcting misunderstandings before they affect a consequential report.

Brady and Davies (2004) developed an account of project capability building through longitudinal research in telecommunications organisations. Their work distinguishes learning within projects from the organisational development of capabilities that can support later work. Applied here, a successful reconciliation becomes more durable when it enters a maintained procedure, role expectation, training example and system check. Those elements need a continuing owner after the project closes. The organisational benefit is not that every future project must repeat the same process in full; it is that future teams can begin with better knowledge and identify where adaptation is necessary. A lessons repository can support this arrangement if entries are searchable by problem and conditions, linked to current practice and reviewed when assumptions change. Unmaintained guidance may become another source of error precisely because its formal appearance encourages later users to trust it.

The Australian Transport Safety Bureau's QF32 investigation provides a bounded example of a finding connected to changed practice. For a specific safety issue, Rolls-Royce reported revisions to failure-mode analysis and the component criticality classification, with associated process-control requirements. ATSB judged the action adequate; its safety-action entry dated 27 June 2013 lists the issue as closed (ATSB, 2013, safety issue AO-2010-089-SI-11). This does not establish the effectiveness of every action arising from the accident or quantify a general improvement in safety. It illustrates a traceable chain from an identified weakness to a named technical and procedural change and a review of that response. Controllers can seek the same clarity in less specialised settings: which approved document or control changed, how the change reached its users and what evidence supports closure of the particular finding?

Use retained forecasts to improve the next estimate

Forecasts become learning material when their original meaning remains recoverable. The controller should retain the forecast date, scope, maturity, assumptions, expected completion condition and information available at the time. Comparing a preliminary estimate with an actual outcome is useful, but it answers a different question from comparing a forecast made immediately before completion. Mixing those decision points can make forecasting appear more accurate merely because the organisation has begun measuring it later. Changes in scope, calendars, operating acceptance and external conditions also affect interpretation. Chapter 7's resource-aware forecast should not be judged against a technical-release date if it predicted stabilisation and acceptance. The target is a fair comparison between what was reasonably being forecast and what occurred, with material changes identified. That comparison protects learning from retrospective explanations that quietly alter the question after the result is known.

Table 8.2 contains five separate fictional wave observations, all measured in working days from an equivalent commitment point to operational acceptance. Forecasts are fifteen, eighteen, fifteen, sixteen and seventeen; actual durations are seventeen, seventeen, nineteen, sixteen and twenty. Define error as actual minus forecast, so the errors are two, minus one, four, zero and three. The

mean signed error is eight divided by five, or 1.6 days: actual completion was later on average in this small set. The mean absolute error is ten divided by five, or two days, which describes the average magnitude without allowing early and late outcomes to cancel. Figure 8.5 displays the individual errors because the average alone hides their pattern. These calculations describe the observations; they do not establish a reliable prediction interval or prove that every future estimate should receive a 1.6-day allowance.

Table 8.2 Retain the individual forecast comparisons

Wave	Forecast	Actual	Signed error	Absolute error
1	15	17	2	2
2	18	17	-1	1
3	15	19	4	4
4	16	16	0	0
5	17	20	3	3

All durations are fictional working days from an equivalent commitment point to operational acceptance. Signed error is actual minus forecast. Mean signed error is $8/5 = 1.6$; mean absolute error is $10/5 = 2.0$. Comparability still requires investigation.

The next question concerns causes and comparability. One wave may have missed a shared specialist, another may have required additional operating practice, and another may have changed scope after commitment. Lena should retain those distinctions before pooling the observations into an estimating rule. Where a missing work element recurs, the response may be to include that element explicitly rather than add an unexplained uplift. Where variation remains after relevant differences are understood, the organisation can use an appropriately supported range. Small samples and selected successes require particular caution. Excluding cancelled or troubled waves because their records are inconvenient can make the remaining history misleading. A forecast library should record why observations were included, excluded or adjusted, with a consistent comparison point. Its purpose is to improve the next decision, not to manufacture a flattering score for the people who produced earlier estimates.

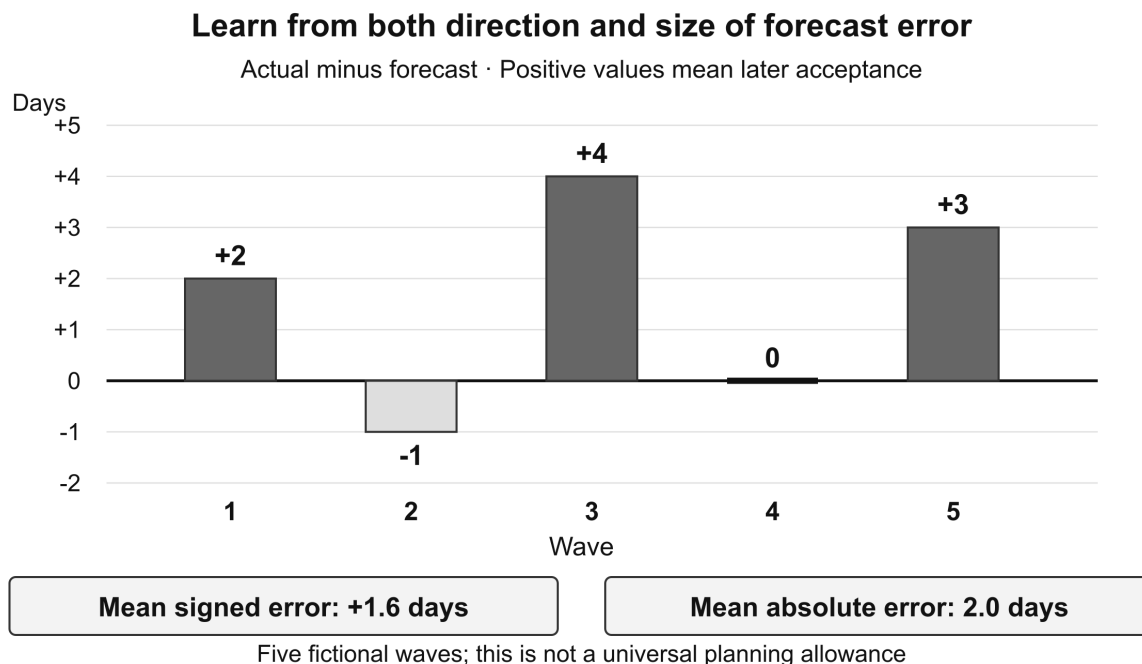


Figure 8.5 Errors in five fictional forecasts measured from equivalent commitment points to operational acceptance. Error equals actual minus forecast. Mean signed error is 1.6 working days; mean absolute error is two days. The observations do not establish a universal estimating allowance.

Verification should use later observations rather than declaring success immediately after changing the method. Suppose the team adds an explicit data-validation activity and a confirmed specialist booking to the next wave's estimate. It should record that change before execution and compare the resulting forecast with subsequent acceptance, while examining other differences. An improvement in one later wave does not isolate the effect of the new practice, but repeated, relevant observations can strengthen or weaken the explanation. Preserve the original procedure and the revised procedure long enough to understand which was used. The same approach applies to benefits forecasts: record the expected adoption path and benefit conditions, then compare them with observed use and outcomes. Learning becomes more useful when it explains which assumptions became more reliable, rather than merely reporting that a new spreadsheet produced a smaller error.

Automate the work without surrendering the evidence

Automation can reduce repetitive handling while changing where errors arise. Parasuraman, Sheridan and Wickens (2000) distinguish types and levels of automation across information acquisition, analysis, decision selection and action. Their conceptual model helps controllers avoid treating automation as one indivisible choice. A system might retrieve invoice records automatically while leaving classification, interpretation and approval with authorised people. Another system might calculate a defined reconciliation rule without being permitted to alter a ledger. The boundary should match the task and its consequences. Figure 8.6 proposes an original workflow for AI assistance in controls: approved source access, a traceable draft or extraction, numerical and source checks, named approval and a controlled published record. The output should remain visibly provisional until it passes the required checks. A fluent explanation does not acquire authority merely because it was generated quickly.

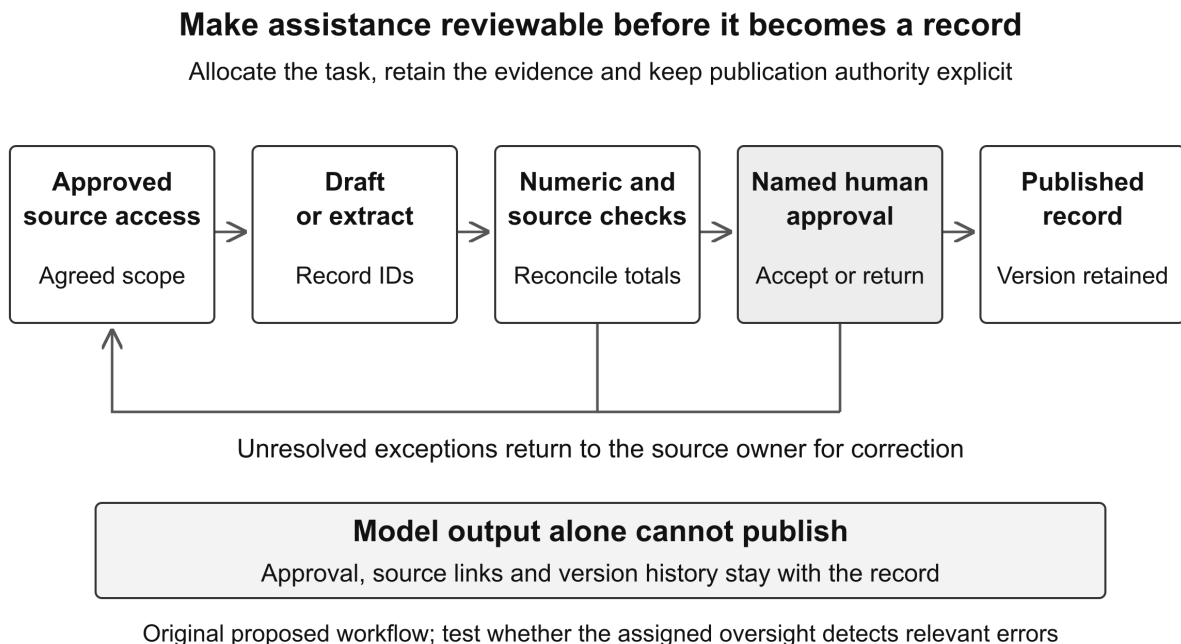


Figure 8.6 A proposed workflow for controlled AI assistance. Source access, traceable drafts, numerical and source checks, named approval and publication remain distinct. Unresolved exceptions return for examination. The arrangement requires evaluation; placing a person in the workflow does not guarantee reliable oversight.

Human involvement must itself be designed. Dzindolet and colleagues (2003) studied trust and reliance on an automated aid and showed that explanations of possible errors could increase reliance even when that reliance was unwarranted. The implication is more demanding than adding a reviewer to the workflow. The reviewer needs sufficient expertise, time, access to the source and a realistic ability to reject the output. Amershi and colleagues' (2019) guidelines for human–AI interaction likewise direct attention to communicating system capabilities, supporting correction and managing changing performance. In the manufacturer's process, an extracted accrual should carry the supporting acceptance record and the rule used to classify it. A reviewer should be able to inspect that evidence directly. Asking someone to approve a convincing

summary while withholding its basis creates the appearance of oversight without the practical means to exercise it.

Evaluation should use representative examples and consequential failure modes. Lena can assemble approved historical records containing duplicated invoices, late acceptance, changed identifiers and contradictory dates, then compare the assisted workflow with a verified reference. The test should distinguish omitted items, invented items, wrong amounts, wrong classifications and unsupported explanations. A single accuracy percentage can conceal a failure on the rare case that matters most. Numeric reconciliation should use deterministic calculations where the rule is known; language generation can help explain the result but should not replace the calculation. Retain the system version and relevant configuration so that a later change can be retested. Testing should also include whether users notice and correct errors, because the combined human and system process is what will operate in practice. A successful demonstration on convenient records is insufficient evidence for unrestricted use.

For a fictional evaluation set of one hundred items, suppose a verified reference identifies ten that require an accrual. An automated assistant flags twelve, of which eight are correct and four are false alarms; two necessary accruals are missed. Precision, the proportion of flagged items that are correct, is eight divided by twelve, approximately 66.7 per cent. Recall, the proportion of necessary accruals detected, is eight divided by ten, or eighty per cent. The measures answer different questions and should accompany the underlying counts. Neither establishes whether the missed amounts are financially material or whether reviewers can reliably resolve the false alarms. Lena examines the values and causes of the two omissions, the review burden created by the four false alarms and performance on records excluded from development. She then tests the complete working arrangement on later data. The example provides a way to describe performance; its invented results are neither an acceptable deployment threshold nor evidence about any actual AI product.

Raji and colleagues (2020) propose a framework for internal algorithmic auditing that connects documentation and examination across the development lifecycle. NIST's Generative Artificial Intelligence Profile similarly emphasises empirical evaluation, source verification and explicit human–AI responsibilities (NIST, 2024). These are frameworks and guidance, not proof that a particular product is reliable. Their practical application is to make evidence, boundaries and ownership part of the deployment decision. Source access should reflect existing permissions, and confidential project records should enter only approved processing arrangements. An assistant should not independently publish a revised forecast, change an approved baseline or mark a benefit achieved because it has generated a plausible narrative. Those acts belong to the relevant authority. The organisation should monitor errors and overrides after deployment and feed them back into the workflow, rather than treating initial approval as permanent evidence of suitability.

Verify that a lesson has become a different practice

The Evidence and Learning Record in Table 8.3 makes the chapter's argument usable in the field. Each entry connects an observation to its supporting evidence, an explanation with stated uncertainty, an authorised change, an owner and a verification point. The record can link to existing issue, change and decision logs rather than duplicate them. Its distinctive purpose is to preserve the connection between what was learned and what subsequently changed. In the accrual example, correcting the current report and changing the handover procedure are separate actions with different completion evidence. In the payment example, resolving the five unknown records differs from preventing the same loss of status information in the next cohort. A completed action is therefore not automatically a verified lesson. The review should identify what was implemented and what remains to be demonstrated through use.

Table 8.3 Illustrative entries in an Evidence and Learning Record

Observation and evidence	Proposed change and owner	Verification through use
Accepted service of 80 omitted from the account; acceptance record and extract retained.	Finance reporting owner and Maya agree a timely acceptance feed and reconciliation rule.	At the next cut-off, trace accepted service into the account and inspect unresolved items.
Five eligible invoices lack verified status; the whole-cohort claim remains uncertain.	Maya resolves those records and identifies the cause of the missing status before changing capture.	Check the next cohort and retain its denominator, completeness and exception evidence.
Five wave forecasts show signed error averaging 1.6 days; causes are not established.	Alex and Lena investigate comparable work and record any justified change to the estimating method.	Compare later retained forecasts with acceptance; examine scope and resource differences.

These are proposed teaching entries, not approvals or verified improvements. Link each to its source, explanation, authorisation, effective version, responsible person and dated review. Closing an action is distinct from establishing that a lesson works.

Applying the Credible Control Method starts with the decision that the information must support. Lena frames whether the next report and release can rely on the current evidence. She establishes the record through definitions, reconciliation and traceable sources, then challenges competing explanations with Alex, Maya and finance. Samira or the appropriate delegated authority authorises the response where it changes scope, resources, responsibilities or a governing rule. Verification occurs at a defined reporting cycle or operating event, using evidence agreed before the change is declared effective. This sequencing matters: an action described vaguely at the review is difficult to assess objectively later. The record should also state what happens if the intervention fails. Reopening an ineffective lesson is part of competent control, provided the organisation preserves the earlier reasoning and uses the new evidence to improve it.

Learning needs measures that preserve its purpose. Counting uploaded lessons rewards collection, while counting closed actions can reward administrative closure. More informative questions concern use and consequence: did the next estimator consult the relevant evidence, did

the receiving team apply the revised rule, did the previously omitted service reach the account, and did the new control detect the error it was designed to detect? These observations should be interpreted with attention to context and unintended effects. A reconciliation rule that removes omissions by delaying every report indefinitely may solve one problem while creating another. The organisation should review whether the benefit of the control justifies its continuing burden and whether a simpler arrangement could achieve the same purpose. Learning includes removing procedures whose original rationale no longer holds, with the reason documented rather than lost.

The chapter began with a corrected report and a question about recurrence. A credible answer now requires more than accurate numbers. It requires a defensible explanation, an authorised response, a recipient capable of using the learning and evidence from subsequent work. Information becomes organisational knowledge when people can understand and apply it; learning becomes visible when that application changes practice and remains open to examination. Project controls is well placed to support this movement because it already connects commitments, records, forecasts and decisions across organisational boundaries. Its contribution is strongest when it preserves meaning and uncertainty while making the next action clear. Chapter 9 develops the complementary question of confidence: how an organisation can subject these accounts and their supporting arrangements to proportionate challenge and assurance before consequential commitments depend on them.

References

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13. ACM. doi: [10.1145/3290605.3300233](https://doi.org/10.1145/3290605.3300233)
- Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, 82(1), 150–169. doi: [10.1006/obhd.2000.2893](https://doi.org/10.1006/obhd.2000.2893)
- Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly*, 21(3), 363–375. doi: [10.2307/2391848](https://doi.org/10.2307/2391848)
- Australian Transport Safety Bureau (ATSB). (2013, 27 June). Classification of the HP/IP bearing support assembly. Safety issue AO-2010-089-SI-11; safety action AO-2010-089-NSA-052. www.atsb.gov.au/safety-issues/AO-2010-089-SI-11
- Brady, T., & Davies, A. (2004). Building project capabilities: From exploratory to exploitative learning. *Organization Studies*, 25(9), 1601–1621. doi: [10.1177/0170840604048002](https://doi.org/10.1177/0170840604048002)
- Columbia Accident Investigation Board (CAIB). (2003, August). Report. Volume I. www.nasa.gov/wp-content/uploads/static/history/columbia/reports/CAIBreportv1.pdf
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, 58(6), 697–718. doi: [10.1016/S1071-5819\(03\)00038-7](https://doi.org/10.1016/S1071-5819(03)00038-7)
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. doi: [10.2307/2666999](https://doi.org/10.2307/2666999)
- Ellis, S., & Davidi, I. (2005). After-event reviews: Drawing lessons from successful and failed experience. *Journal of Applied Psychology*, 90(5), 857–871. doi: [10.1037/0021-9010.90.5.857](https://doi.org/10.1037/0021-9010.90.5.857)

- Government Data Quality Hub. (2020, 3 December). The Government Data Quality Framework. www.gov.uk/government/publications/the-government-data-quality-framework/the-government-data-quality-framework
- International Atomic Energy Agency (IAEA). (2015, August). The Fukushima Daiichi accident: Technical volume 2 — Safety assessment. STI/PUB/1710. IAEA. www-pub.iaea.org/MTCD/Publications/PDF/AdditionalVolumes/P1710/Pub1710-TV2-Web.pdf
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information & Management*, 40(2), 133–146. doi: 10.1016/S0378-7206(02)00043-5
- National Institute of Standards and Technology (NIST). (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile. NIST AI 600-1. doi: 10.6028/NIST.AI.600-1
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3), 286–297. doi: 10.1109/3468.844354
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211–218. doi: 10.1145/505248.506010
- Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. ACM. doi: 10.1145/3351095.3372873
- Schindler, M., & Eppler, M. J. (2003). Harvesting project knowledge: A review of project learning methods and success factors. *International Journal of Project Management*, 21(3), 219–228. doi: 10.1016/S0263-7863(02)00096-0
- Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17(S2), 27–43. doi: 10.1002/smj.4250171105
- Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors*, 55(1), 231–245. doi: 10.1177/0018720812448394
- US Government Accountability Office (GAO). (2020, 12 March). Cost estimating and assessment guide: Best practices for developing and managing program costs. GAO-20-195G. www.gao.gov/products/gao-20-195g
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5–33. doi: 10.1080/07421222.1996.11518099
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16(4), 409–421. doi: 10.1287/orsc.1050.0133
- Williams, T. (2008). How do organizations learn lessons from projects—and do they? *IEEE Transactions on Engineering Management*, 55(2), 248–266. doi: 10.1109/TEM.2007.912920

CHAPTER 09

Confidence That Survives Challenge



IN THIS CHAPTER

Building confidence through proportionate, independent challenge.

Ask what the confidence rests upon

Samira has received three favourable reports about the manufacturer's next payment-platform release. Alex reports that the delivery tasks are substantially complete. Maya confirms that the operating team has attended training. Lena has reconciled the latest cost and benefits records using the methods developed in Chapter 8. The sponsor is pleased, but her decision remains difficult: does this evidence justify moving another business unit onto the new service, and what does it establish about the benefits expected afterwards? Each report may be accurate within its own boundary. Their combination may still leave the most consequential question unanswered. The operating team could understand the normal transaction while being unable to resolve an exception; a credible delivery forecast could coexist with an unsupported savings claim. Assurance begins by making these gaps discussable before the organisation commits people, money and operational continuity to a conclusion that the evidence cannot yet support.

The chapter's central argument is that confidence becomes useful when its basis, limits and consequences can be examined. Project controls supplies much of the material: estimates, schedules, reconciliation, risk analysis, acceptance records and benefit measures. Assurance examines whether that material warrants a particular conclusion for a particular decision. It therefore extends the Transformation Control Loop rather than sitting outside it as a ceremonial inspection. Claims lead to commitments, commitments generate evidence, evidence informs decisions, and subsequent experience changes what the organisation knows. At each connection, the controller should be able to explain why the next step follows. The aim is neither universal certainty nor an endless search for defects. It is a proportionate, explicit judgement about what can responsibly proceed, what conditions must be met and what remains to be verified through use.

Bloomfield and Bishop (2010) review the history and practice of structured safety and assurance cases for software-based systems. Their discussion concerns assurance practice, including safety-related systems; it does not demonstrate that a well-presented argument predicts transformation success. Nevertheless, the distinction between possessing evidence and explaining its relevance is valuable here. A training register supports a claim about attendance. A witnessed transaction exercise supports a narrower claim about demonstrated performance under specified conditions. Neither automatically establishes that every shift can sustain the intended process at peak demand. The controller's practical task is to state the claim in terms that make this difference visible. "Ready for release" is too elastic unless the record identifies the service, users, operating conditions, acceptance criteria, remaining restrictions and person authorised to accept the residual exposure.

The proposed Claim and Challenge Map in Figure 9.1 provides a working structure for that task. It connects a bounded claim to criteria, the reasoning that links evidence to those criteria, the evidence itself and the strongest unresolved objections. A reviewer then gives a qualified conclusion that the authorised decision maker can use. The map is an original practice aid, not a validated confidence scale or a replacement for specialist assurance standards. It uses the existing Credible Control Method: Frame the decision, Establish the evidence, Challenge the inference, Authorise an appropriate commitment and Verify the conditions through subsequent

observation. This continuity matters because the organisation already has owners, decision routes and reporting cycles. Assurance should make these arrangements more dependable, rather than inventing another committee whose approval obscures who remains responsible for the work.

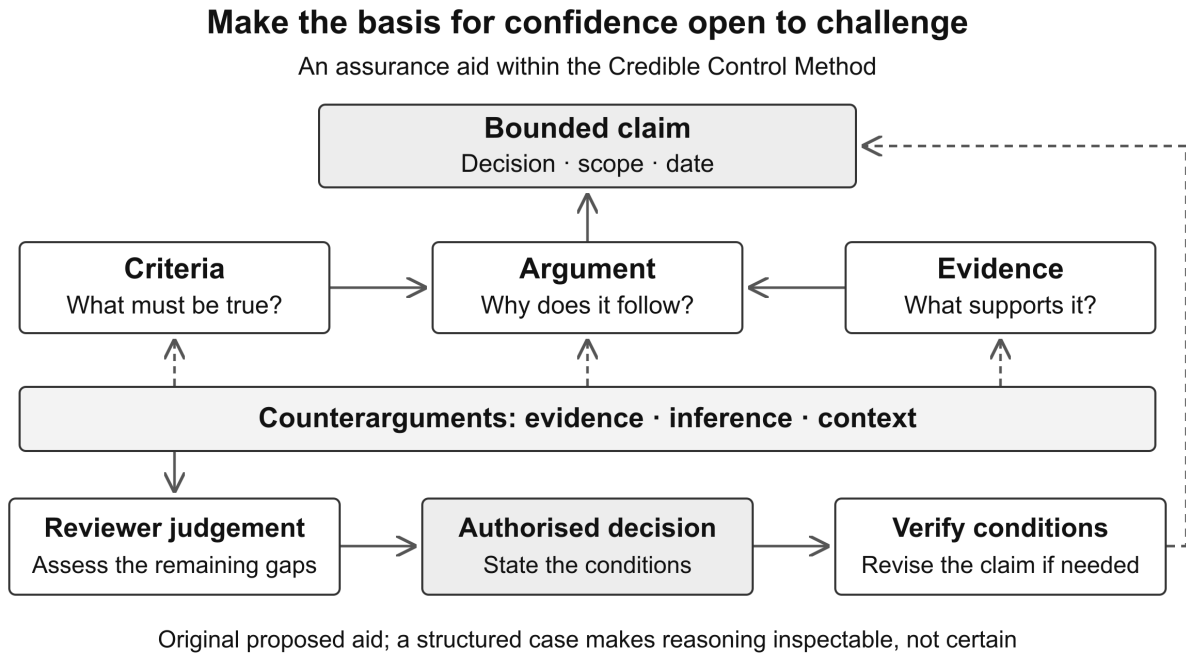


Figure 9.1 The proposed Claim and Challenge Map connects a bounded claim, criteria, reasoning and evidence to explicit objections and a qualified decision. It is an original practice aid within the Credible Control Method, informed by assurance-case literature; it is not a validated confidence scale.

The distinction also protects benefits management from premature declarations of success. Shenhar et al. (2001) examined project success as a multidimensional concept whose assessment changes with the dimension and time horizon considered. Their work helps explain why delivery efficiency and longer-term business value require different evidence; it does not supply a single universal weighting for every transformation. For Samira, this means that technical acceptance, operational adoption and financial benefit should remain separate conclusions. An early release decision can be justified even while the eventual benefit remains uncertain, provided that the uncertainty is explicit and the commitment is proportionate. Conversely, a completed delivery cannot be used to close a benefits claim simply because the original project team is leaving. Someone must retain responsibility for establishing whether the changed operation produces the value that justified the investment.

Give assurance a distinct job and a usable position

Monitoring, quality management, assurance, audit and independent review overlap in the evidence they use, but they should not be treated as interchangeable labels. Monitoring observes performance against a defined reference and helps management respond. Quality management establishes and operates arrangements intended to make outputs and processes fit their requirements. Assurance provides a reasoned assessment of whether evidence supports confidence in specified claims or controls. Audit is a structured examination against an explicit mandate and criteria, with conclusions whose scope depends on that mandate. An independent review is an arrangement for obtaining judgement with suitable separation from the work being judged; its value depends on competence, access and freedom to report. Table 9.1 uses the manufacturer's situation to make these distinctions practical. An activity may contribute to more than one function, but its report should identify which conclusion it actually supports.

Table 9.1 Give each examination an explicit purpose

Activity	Question in the manufacturer's review	Limit of the conclusion
Monitoring	Are recorded late payments moving against the agreed reference?	A trend alone does not explain causation.
Quality management	Are testing and acceptance processes designed and operated to meet requirements?	A compliant process does not establish every benefit.
Assurance	Does the available evidence justify this bounded release claim?	The conclusion retains scope, conditions and exclusions.
Audit	Do examined arrangements meet the mandate and stated criteria?	The opinion applies to the examination performed.
Independent review	Can suitably separate experts challenge the claim freely and competently?	External appointment alone does not prove independence or expertise.

These are working distinctions for this chapter. Functions can overlap; a report should state its actual mandate and evidential boundary.

The Institute of Internal Auditors' 2026 Three Lines Model distinguishes management responsibilities, specialist support and challenge, and the independent role of internal audit. It also treats independence as a combination of organisational positioning and safeguards, supported by access, authority and freedom from interference (IIA, 2026, pp. 8–16). Applied to the manufacturer, this means Lena can assemble an assurance plan and test routine information without becoming the independent reviewer of every method she designed. Where she constructed the cost model, another competent person should examine consequential assumptions and reproduce selected calculations. Where Maya owns the benefit, she supplies operational evidence and corrective action, while a reviewer challenges the adequacy of the claim. These arrangements preserve the value of knowledgeable internal staff while making self-review visible instead of concealing it behind a different job title.

DeAngelo (1981) conceptualised audit quality through the joint prospect that an auditor detects a breach and reports it, in a model concerned with auditor incentives and size. The transferable distinction is between the ability to discover a problem and the willingness or freedom to disclose it. It does not establish that a large assurance provider will always outperform a smaller specialist. For a transformation review, the appointment should therefore examine both dimensions. A reviewer may be independent of the delivery manager but unable to interpret a resource-constrained schedule. Another may understand the system exceptionally well but depend on the sponsor for future work and hesitate to challenge the release date. The review design needs relevant expertise, declared relationships, unrestricted evidence access and an escalation route that survives disagreement with the person whose claim is being examined.

Coordination should reduce unnecessary repetition while preserving scrutiny of important gaps. The controller can map the claims already examined by engineering, finance, security, operations and internal audit, recording their criteria, evidence dates and exclusions. Reliance on another review is then a considered decision: does its scope cover this question, was its work sufficiently thorough, and has the underlying situation changed? A security review of an earlier configuration may be excellent yet irrelevant to a later integration change. A finance review may validate the arithmetic of a benefit model while explicitly excluding the operational assumptions that generate the numbers. The assurance plan should show these boundaries rather than colouring the entire workstream green because several departments have signed something. Coordination becomes useful when it makes omissions visible and allows specialist effort to reach the claims that matter most.

Build a claim that can be tested and defeated

Consider the statement that the next unit can enter operation without undermining the intended payment benefit. Lena first separates its components: the configured service performs required functions; the operating team can use it under expected conditions; interfaces and exception handling work; the benefit mechanism remains plausible; and the organisation can fund and govern the remaining work. She then asks what would count as convincing evidence for each component. Technical testing may support the first, observed rehearsal the second and third, and a combination of process data and evaluation design the fourth. None should borrow credibility silently from the others. The resulting argument is not necessarily long. Its usefulness comes from making the inferential steps explicit enough that another competent person can identify where a conclusion goes beyond the evidence or depends on an assumption that has not been tested.

Hawkins et al. (2011) propose examining confidence in a safety argument through the assurance deficits that could undermine it. Their conceptual treatment is especially relevant where apparently separate parts of an argument share an underlying weakness. In a payment transformation, three reports might all depend on the same incomplete transaction extract. Reviewing the reports separately would not correct the missing records. Lena therefore adds evidence lineage to the claim record: who collected the observation, how it was transformed, which population it represents and which other conclusions depend on it. If the source is later found unreliable, the affected claims can be reopened together. The same principle applies to a

schedule that feeds both the cost forecast and the benefits start date. The repeated appearance of one assumption does not create additional support for that assumption.

Goodenough, Weinstock and Klein (2015) distinguish ways an assurance argument can be defeated, including challenges to the claim, the inference and the evidence. This technical framework can be taught through three different objections to Maya's readiness statement. A failed recovery exercise directly contradicts the claim that the team can recover the service. Successful normal transactions do not establish recovery capability, which challenges the inference. A test log from the wrong software version undermines the evidence's relevance. Each objection requires a different response: repair and retest the capability, strengthen or narrow the reasoning, or replace the evidence. Treating them all as requests for more documentation would waste effort and might leave the decisive weakness intact. The reviewer should identify what would resolve the objection and whether it prevents the proposed commitment or merely limits the claim that can accompany it.

Boston's tunnel ceiling collapse demonstrates the importance of matching a test to the mechanism being claimed. The NTSB's investigation, adopted in July 2007 after the July 2006 accident, found that the anchor adhesive could not sustain long-term loading because of creep. Following earlier displacement, failed anchors had been replaced and proof-test loads increased. The investigation found that stronger short-term tests could not establish long-term capacity, while replacement with the same epoxy did not prevent further displacement (NTSB, 2007, pp. 96, 107). The project-controls lesson is a specific one: a test must address the relevant operating condition and duration. More tests, stronger loads or more signatures cannot repair an argument about a mechanism they do not examine. The case concerns an engineering accident; it is not a measured general effect of assurance on transformation performance.

The manufacturer can apply that lesson without pretending that payment processing and structural engineering have identical risks. A demonstration with clean records and the supplier's experts present may establish that the configured service can perform a transaction. It may tell little about month-end congestion, absent staff, conflicting master data or a supplier disputing a rejected payment. The claim should therefore identify its operating envelope, and the review should deliberately examine the boundary. Lena asks the team to describe circumstances under which the favourable conclusion would no longer hold. This is often more revealing than asking for another confidence rating. A credible answer names conditions, warning signs and a response; an answer that says the service will be ready under all circumstances is a reason to inspect the argument more closely.

Put the review where it can change a decision

Assurance has its greatest opportunity to affect value while alternatives remain available. Reviewing an estimate before contract commitment can change scope, sequencing or commercial terms. Reviewing operational readiness before migration can change the wave size, support arrangements or release date. Reviewing benefits after sustained use can change the operating model or the remaining rollout. These are different questions at different moments, as Figure 9.2 shows. The controller should work backwards from the commitment, allowing time for evidence collection, challenge, response and authorised action. A review scheduled the day before an irreversible decision may create pressure to dilute its findings because there is no practical time to respond. Earlier scrutiny does not remove uncertainty, but it can reveal which uncertainty should be reduced before money or operating continuity is put at risk.

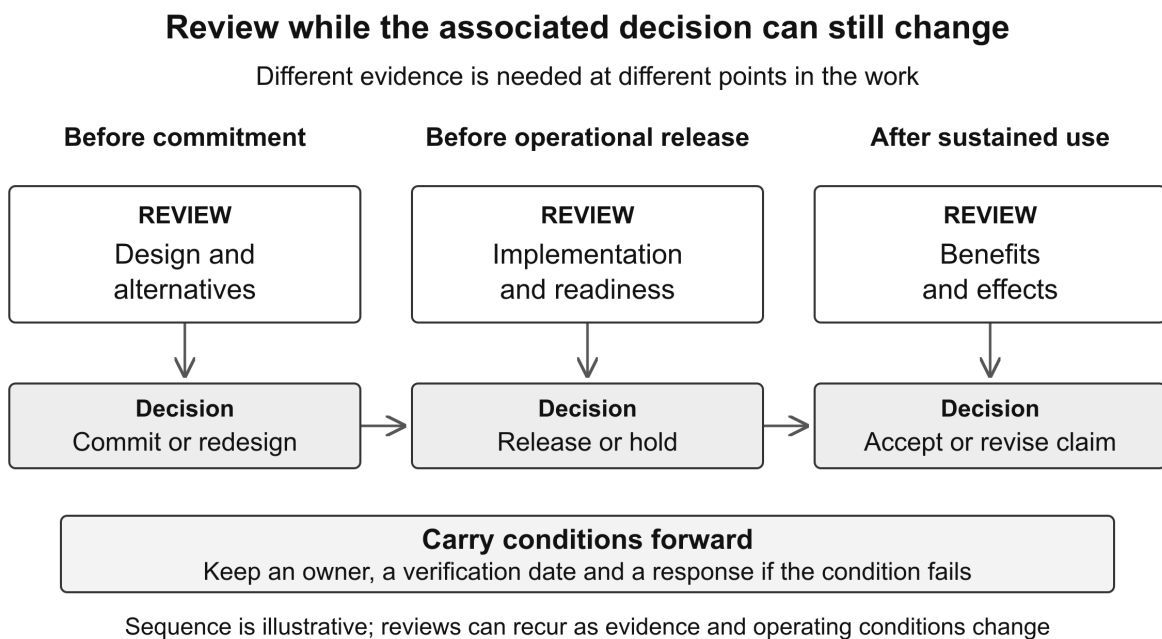


Figure 9.2 An original assurance-timing model. Review evidence while it can change the associated commitment, operational release or benefits decision. Conditions pass forward and may require a later review; technical release does not close the benefits argument.

Williams et al. (2012) studied early warning signs in complex projects and the difficulties of recognising and acting on them. Their work connects formal assessments with the organisational interpretation of emerging concerns, rather than offering a mechanical predictor of failure. The practical implication is to combine planned reviews with triggers that respond to changed conditions. A material supplier assumption may fail between gates; an operating team may lose the specialist on whom readiness depended; a benefit measure may change definition. Lena should record which changes reopen which claims, together with the person who decides whether a fresh review is needed. This makes assurance responsive to the transformation's actual development. A completed review remains evidence about a particular scope and time; it should not become a permanent certificate that follows the project regardless of subsequent change.

Proportionate planning begins with the consequences of being wrong and the uncertainty that matters to the decision. It also considers reversibility, exposure duration, the ability to detect deterioration and the availability of effective remedies. A low-cost, reversible pilot may justify proceeding with close observation and a narrow operating boundary. A migration that retires the only reliable fallback requires stronger evidence before commitment. These considerations should be discussed directly instead of hidden inside an unexplained product of ordinal likelihood and impact scores. A useful assurance plan states the decision, the claim that could mislead it, the proposed examination, the competent reviewer, the evidence deadline and the response route. Where resources are limited, the sponsor should see which important claims remain unexamined and what that gap means for the decision being requested.

New Zealand's Auditor-General reported anonymised examples of Gateway reviews affecting project decisions in September 2016. In one, reviewers helped a sponsor identify actions needed to secure stakeholder agreement. In another, advice supported a rewritten business case before Cabinet approval. These were examples of sponsor experience reported by the audit, rather than controlled comparisons. The same report noted that the Treasury had not formally assessed Gateway performance and that separating reviews' effects from other interventions would be difficult (Office of the Auditor-General, 2016, pp. 30–31). The constructive lesson is to trace assurance into what changed before commitment. Agreement, an improved business case or a positive sponsor response can show a useful contribution to the decision process, but none establishes eventual benefits or a quantified financial return from the review itself.

Review outputs should consequently distinguish a finding, a recommendation and a decision. A finding describes what the examination established against its criteria. A recommendation proposes a response. The decision belongs to the authorised owner, who may accept the recommendation, choose another justified action or explicitly retain the exposure where permitted. The record should preserve disagreement and its rationale. This does not mean every opinion is equally persuasive: an unaddressed critical finding should remain visible even when management wishes to proceed. It means the reviewer should not quietly become the delivery manager, and the sponsor should not treat the reviewer as the person who assumed the risk. The assurance plan is complete only when it connects evidence and judgement to a decision that somebody has both the authority and the obligation to make.

Reperform the forecast and challenge its construction

Forecast assurance needs two levels of examination. The first asks whether the model works as constructed: formulas, units, dates, logic, classifications and aggregation should be reproducible. The second asks whether that construction represents the decision adequately. A perfectly calculated schedule can omit a shared specialist; a reconciled cost account can exclude transition support; a benefit model can begin savings on the technical release date despite the time needed for adoption. Lena should preserve the forecast version submitted for review, reproduce important outputs and select assumptions according to their effect on the commitment. Where the team changes an assumption during challenge, both versions should remain identifiable. Otherwise the organisation may lose the ability to distinguish an error corrected by assurance from a genuine change in circumstances or an authorised change to the delivery strategy.

Chapter 7's fictional sequence illustrates why a reviewer should inspect constraints rather than merely inspect the final date. Its unconstrained network finished in fifteen working days, but the shared specialist made the selected sequence seventeen days. The modelled eightieth percentile of approximately 20.4 days applied only to the assumptions and resource treatment stated in that exercise. An assurance reviewer should reproduce the resource assignment, inspect whether risk dependencies and duration ranges are defensible, and test a plausible adverse alternative. The review should also ask whether operational acceptance and benefit observation are represented separately. Reproducing a percentile proves that the calculation can be repeated; it does not prove that the model includes every material uncertainty or that the organisation has an eighty per cent probability of achieving all its strategic benefits. The conclusion should retain the boundary of the calculation.

The same discipline applies to Chapter 8's separate supplier package. Its forecast of US\$530,000 comprised US\$280,000 incurred and US\$250,000 remaining. The incurred amount included an US\$80,000 accrual for accepted service; the open purchase order also included that service. Adding the entire open order would double count the accrual while missing the growth embedded in the remaining forecast. An independent re-performance should trace selected accepted services through the source record, accrual, commitment and forecast, then examine whether the remaining work is complete. A sample chosen only from posted invoices cannot establish that unrecorded accepted services are absent. The reviewer must inspect the route by which omissions could enter the account. This turns reconciliation into assurance of completeness and meaning, rather than a check that two totals happen to agree.

Sensitivity testing should focus on assumptions capable of changing the decision. Suppose the migration remains manageable with one week of additional support but becomes unaffordable or operationally unstable if the support period doubles. That threshold deserves explicit examination, including the mechanism that could produce the longer period and the evidence available before commitment. The controller can prepare a small set of coherent alternatives instead of changing every input independently and presenting a large collection of implausible combinations. Shared drivers should remain connected: low adoption might delay benefits while extending support costs, rather than allowing one to improve while the other deteriorates arbitrarily. The reviewer should record whether the recommended commitment survives these alternatives and which observations would require reconsideration. This is an assurance application of the earlier business-case and scheduling methods, not another forecast competing for organisational ownership.

Materiality in this setting extends beyond the largest financial variance. A modest untested interface can block the entire operating transition; a small group of vulnerable users can experience serious disbenefits; an apparently minor definition change can invalidate the benefit comparison. Lena therefore asks which defect could reverse the release decision or substantially change who receives value and who carries the burden. Examination follows those pathways as well as the large numbers. The resulting conclusion might say that the revised estimate is arithmetically sound and includes identified remaining work, while the start of cash savings still depends on an unapproved operating change. Such a conclusion is more useful than a single

confidence percentage because it tells Samira which part of the commitment has been supported and where a further decision or observation is still required.

Learn what a sample can and cannot establish

The next assurance technique addresses a deceptively reassuring statement: the reviewer tested sixty transactions and found no failures. Before interpreting that result, Lena needs to know what was eligible for selection, how cases were chosen, what counted as failure and whether the inspection could detect it. Sixty convenient records supplied by the team are different from sixty appropriately selected cases representing the process under review. A successful normal payment is also different from successful handling of a disputed invoice. The population and test should match the claim. Sampling can provide evidence about a recurring control, but a statistical result for routine transactions cannot establish that every critical failure mode has been examined. High-consequence exceptions may require deliberately targeted tests, whose findings should be reported separately from any probability-based inference about routine work.

The following original numerical example assumes a sample size fixed before testing, independent tests of a process with a constant but unknown failure probability, p , and effective identification of every failure in the tested cases. These are strong assumptions, stated so that the calculation can be interpreted correctly. If n tests produce zero failures, the probability of that result under a specified p is $(1 - p)^n$. Setting this probability to 0.05 and solving gives the one-sided 95 per cent upper confidence bound: $p_U = 1 - 0.05^{1/n}$. This is the zero-failure form of an exact binomial approach associated with Clopper and Pearson (1934). With sixty tests, the result is approximately 4.87 per cent. Zero observed failures therefore does not establish a zero failure probability, even when the tests and the sampling assumptions are sound.

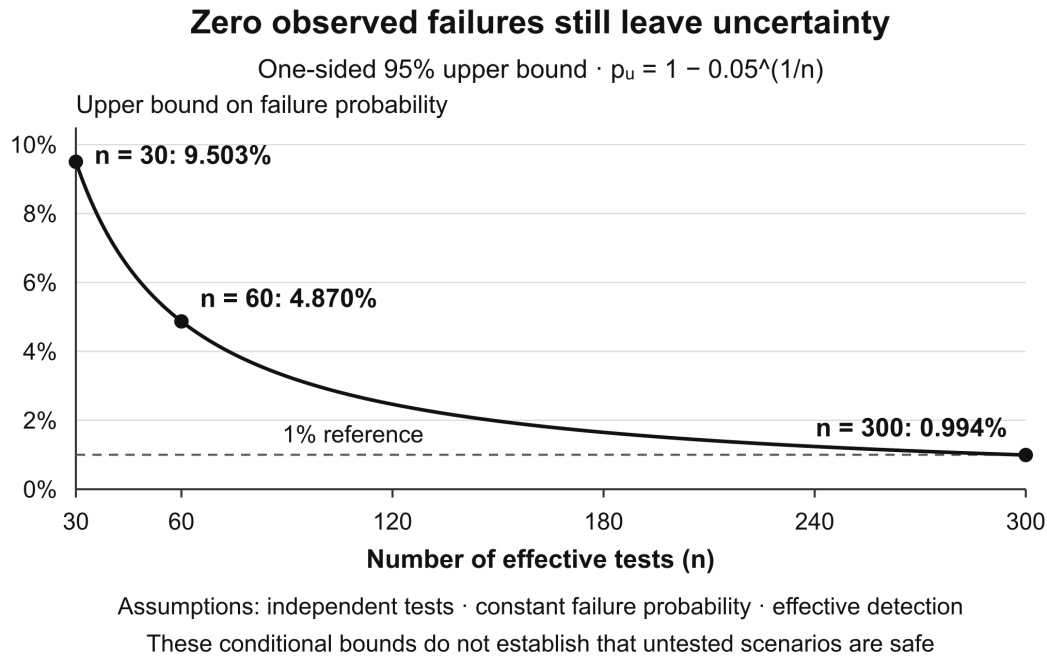


Figure 9.3 Original calculation of the one-sided 95 per cent exact binomial upper bound after zero failures. The marked bounds are 9.50, 4.87 and 0.99 per cent for 30, 60 and 300 tests. Interpretation requires a sample size fixed before testing, independent tests, a constant failure probability and effective inspection.

The word confidence has a technical meaning here. Under the stated model, the procedure for constructing the bound has at least the stated long-run coverage across repeated samples. It is not a claim that there is a 95 per cent probability that this fixed, unknown failure probability lies below the calculated bound. Lu (2020, pp. 15–17) explains the distinction between exact interval methods and Bayesian inference. It is also necessary to distinguish this one-sided 95 per cent bound from a two-sided 95 per cent interval: they allocate their tail probabilities differently. For the controller, the reporting rule is straightforward. State the method, population, number tested, failures observed, assumptions and resulting bound together. Removing the assumptions from the result makes the number easier to display but harder to use responsibly in an assurance conclusion.

Table 9.2 Zero failures still leave uncertainty

Tests	Failed	Upper bound	Interpretation
30	0	9.503%	A clean small sample leaves a comparatively high upper bound.
60	0	4.870%	The worked review does not demonstrate a one per cent threshold.
299	0	0.997%	Minimum test count for an upper bound at or below one per cent.
300	0	0.994%	A rounded extra test; the conclusion remains model-dependent.

One-sided 95 per cent exact binomial bounds, displayed to three decimal places. All rows assume sample sizes fixed before testing and independent effective tests with a constant failure probability. The threshold is a teaching choice, not a recommended risk tolerance.

Table 9.2 shows why an assurance sample should be designed from the question rather than from the time available. Thirty failure-free tests give an upper bound of approximately 9.50 per cent; three hundred give approximately 0.99 per cent. Under the same model, at least 299 tests with zero failures are needed for an upper bound no greater than one per cent. This follows by rounding $\log(0.05)/\log(0.99)$ upwards to the next integer. It does not mean that testing 299 cases guarantees acceptance: failures may occur, and the appropriate calculation then changes. Nor does it establish that one per cent is an acceptable operational risk. The threshold belongs to the decision criteria and consequences. The statistical design tells the reviewer what evidence could support that threshold under stated conditions, while the sponsor and relevant specialists determine whether the threshold is justified.

A complementary calculation asks how likely the review is to detect at least one failure if the true rate were five per cent. With sixty independent tests, the answer is $1 - 0.95$ raised to sixty, approximately 95.39 per cent. This is a detection probability under an assumed rate, not the posterior probability that the process is satisfactory after a clean sample. Keeping these questions separate prevents familiar percentages from becoming interchangeable forms of reassurance. The calculation can help Lena explain why a proposed sample is useful for one purpose but too weak for another. It can also support an explicit discussion of residual exposure: a review designed to detect a relatively common control failure may have little ability to discover a rare event with severe consequences unless the test design targets that mechanism directly.

The model must be reconsidered when transactions share a batch error, one operator or a common configuration, because sixty records may contain far less than sixty independent opportunities to reveal a problem. Changes during testing also challenge the constant-probability assumption. Sampling without replacement from a small finite population calls for an appropriate finite-population analysis rather than an unqualified binomial claim. Imperfect inspection introduces another layer of uncertainty: a reviewer who cannot detect the defect cannot support the same conclusion as a perfect test. In field practice, Lena should retain the selection method and rejected or unavailable records, describe the process coverage and involve a statistical specialist when the decision requires formal inference. The original graph and calculation teach the logic of evidence strength; they are not a universal audit sampling standard or permission to apply one sample size to every transformation.

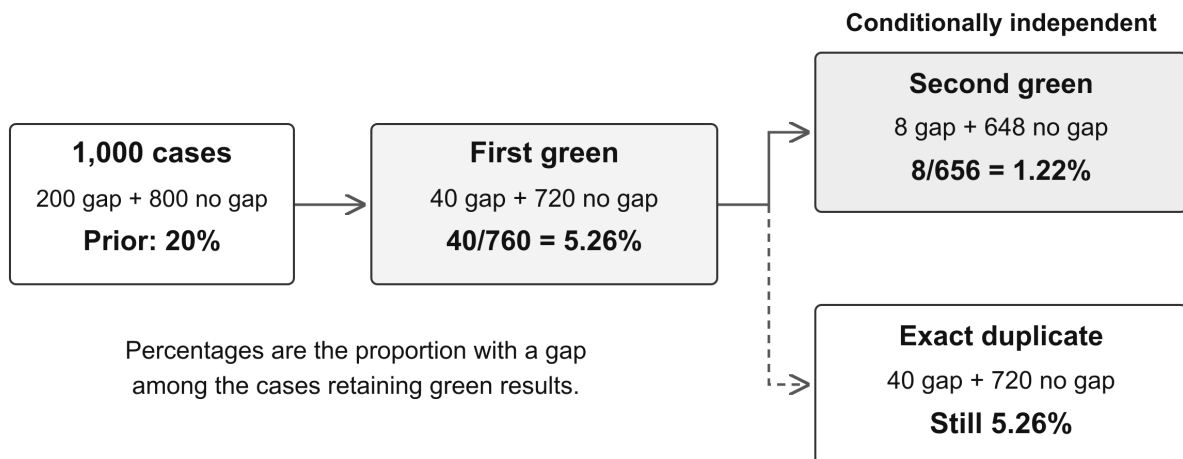
Distinguish another review from another source of evidence

Three reviewers can agree because they have reached the same conclusion through different examinations, or because each has copied the same mistaken premise. Organisational independence and evidential independence are therefore different properties. A reviewer outside Alex's reporting line may still rely entirely on Alex's test summary. Another reviewer within the organisation may independently observe an operating exercise that reveals information absent from the summary. Both relationships should be visible. The controller can draw a simple lineage map connecting conclusions to observations, source systems and assumptions. Shared roots identify where an apparently broad assurance portfolio is vulnerable to a common error. This does not imply that evidence should never be reused. It means the additional value of a review should be explained through the new examination or interpretation it contributes, rather than inferred from the number of signatures.

A fictional Bayesian example makes the consequence numerical. Suppose, solely for teaching, that a class of releases has a prior probability of a material readiness gap of twenty per cent. Assume a particular review raises a red flag in eighty per cent of releases with such a gap and in ten per cent without one. These rates are invented model inputs, not measured performance of project assurance. Among one thousand imagined releases, two hundred have a gap and eight hundred do not. The review gives a green result to forty of the gap cases and seven hundred and twenty of the no-gap cases. Among the seven hundred and sixty green results, forty still have the gap. The modelled probability of a gap after a green result is therefore $40/760$, or 5.26 per cent. A green result changes the judgement; it does not eliminate the possibility of a gap.

A second green result can repeat the same information

Fictional inputs: prior gap 20% · red given gap 80% · red given no gap 10%



Independence here is conditional on whether a gap exists; it is a model assumption
A second reviewer is not automatically a second independent source of evidence

Figure 9.4 Original fictional Bayesian example. The prior gap probability is 20 per cent; a red result occurs in 80 per cent of gap cases and 10 per cent of no-gap cases. A green gives 5.26 per cent; two conditionally independent greens give 1.22 per cent. An exact duplicate adds no evidence. These are model results, not measured assurance performance.

Now suppose a second review has the same detection characteristics and its result is conditionally independent of the first, both when a gap exists and when it does not. Among the two hundred gap cases, eight receive two greens; among the eight hundred no-gap cases, six hundred and forty-eight receive two greens. The probability of a gap after two greens is then $8/656$, approximately 1.22 per cent. If the second report is an exact duplicate of the first, however, it contains no new evidence and the probability remains 5.26 per cent. Conditional independence is the decisive assumption. Different reviewers, different report titles or separate meetings do not establish it. Shared data, common tests and a common misunderstood requirement can make the reviews strongly dependent, leaving much less additional information than the independent calculation suggests.

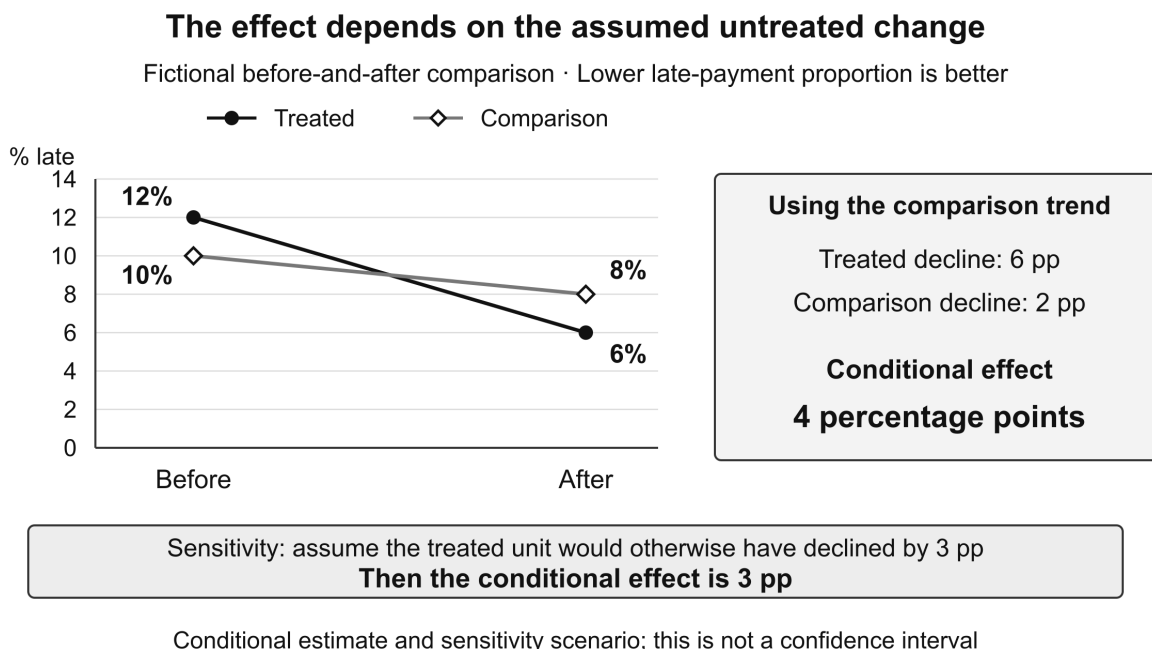
The same model also shows why an adverse result should not be interpreted mechanically. A red flag occurs in one hundred and sixty gap cases and eighty no-gap cases, giving a modelled gap probability of $160/240$, or 66.67 per cent. It is a reason to investigate, not proof of the gap in every flagged release. The prior also matters: changing the assumed prior gap probability from twenty to forty per cent raises the probability after one green to approximately 12.90 per cent. This sensitivity is useful precisely because it makes hidden judgement visible. In practice, the controller would need credible evidence for the prior, review performance and dependence structure, with uncertainty around those inputs. Without that basis, the calculation remains an educational scenario and should not be printed as the organisation's actual probability of readiness or benefit attainment.

Kelly and Weaver (2004) explain how structured safety arguments make relationships among goals, reasoning, context and evidence explicit. Their notation is not a statistical confidence calculator. The practical application here is to use a visible argument before attempting numerical combination. Lena should first identify what each review observes and how its conclusion bears on the release claim. Where several reports depend on the same transaction extract, the next useful activity may be to examine the extraction process or collect a genuinely different observation, such as a witnessed recovery exercise. This is more informative than commissioning another summary of the existing report. Quantification can clarify a defensible model, but it cannot compensate for an unspecified claim, unreliable inputs or an independence assumption that exists only because it makes the resulting percentage look reassuring.

Assure the benefit mechanism and the financial claim separately

Benefits assurance begins by identifying which proposition is under examination. A capability has been delivered; a changed practice is being used; an operating outcome has improved; the transformation contributed to that improvement; or a financial benefit has been realised. These propositions form a connected argument, but they require different observations. Badewi's (2016) survey analysis associated project-management and benefits-management practices with reported success and supports examining delivery and benefit governance together. The study was observational and based on organisational responses, so it does not establish a causal return from a particular assurance checklist. Lena's application is to retain the accountable benefit owner and the evidence required after technical completion. A review that confirms delivery but excludes changed practice and outcome measurement should say so, allowing Samira to understand which part of the value proposition still depends on future evidence.

Chapter 5's fictional payment example can now be examined as an assurance problem. Late payments fell from twelve to six per cent in the transformed unit and from ten to eight per cent in the comparison unit. The difference between the six-percentage-point reduction and the two-point comparison reduction is four percentage points. The calculation is correct. Its causal interpretation still depends on assumptions about what would have happened without the transformation. The reviewer should inspect earlier trends, population composition, measurement consistency, other policy changes and possible spillovers between units. If the transformed unit would have improved by three points without the intervention, the corresponding effect would be three points rather than four. Figure 9.5 shows this sensitivity without labelling it a confidence interval. It is a change in the assumed counterfactual, illustrating how a sound calculation can support different conclusions when the underlying causal assumption changes.



Conditional estimate and sensitivity scenario; this is not a confidence interval

Figure 9.5 The fictional Chapter 5 comparison revisited for assurance. The conditional difference in reductions is four percentage points. Assuming a three-point untreated reduction for the transformed unit gives a three-point effect. The alternative is a sensitivity scenario, not a statistical confidence interval or verified causal effect.

Hernán and Robins (2016) propose specifying a target trial when using observational data to investigate a causal question. Their methods paper comes from epidemiology, but its insistence on defining eligibility, strategies, follow-up, outcomes and the intended comparison is useful for benefits assurance. Lena can ask who was eligible for the rollout, when observation began and whether units classified as adopters had to survive an initial implementation period before entering the analysis. Selecting only successful adopters after seeing their performance can make the intervention appear stronger than an analysis of the original eligible group. The review should distinguish the effect of assignment to a rollout from the effect of actual use, because adherence is itself influenced by operational conditions. These questions improve the design; they do not turn a retrospective business comparison into a randomised experiment or remove unmeasured confounding automatically.

Where a long sequence of observations exists, an interrupted time series may be more informative than two isolated before-and-after values. Lopez Bernal, Cummins and Gasparrini (2017) teach segmented regression with attention to the expected impact pattern, seasonality, autocorrelation and concurrent events. An assurance reviewer should ask whether the intervention is expected to cause an immediate level change, a gradual change in trend or a delayed effect, and whether that expectation was specified before searching for a favourable result. In the manufacturer's setting, a short adoption period might make an immediate full benefit implausible. A change in invoice classification at the same date could create a false break in the series. The controller need not personally fit every advanced model, but should be able to commission the right analysis, preserve its assumptions and explain why the available design supports a narrower or stronger claim.

Australia's My Health Record provides a dated example of this distinction between implementation and benefit verification. The ANAO's November 2019 audit judged implementation under the opt-out model largely effective, while identifying a gap in the forward programme of benefits evaluation. The agency monitored intermediate outputs and had commissioned nine research studies, five completed at audit. It nevertheless lacked a structured forward plan specifying the timing and sequence of future research, data collection and evaluation checkpoints. The ANAO recommended such a plan and reporting of evaluation results, and the agency agreed (ANAO, 2019, pp. 7–8, 54–55). The finding does not mean that no benefits existed or that a benefits-realisation plan was absent. It demonstrates why confidence in implementation and confidence in the evidence for realised benefits should remain distinct, and why agreement with a recommendation is not its completed implementation.

Financial assurance adds another inference that requires explicit examination. In a separate fictional annual efficiency example, unrelated to the programme's overall business case, suppose 120,000 transactions are eligible for improvement. At seventy-five per cent adoption, four minutes saved per adopted transaction releases 6,000 hours. Valuing those hours at US\$30 produces US\$180,000 of annual capacity value. That figure is not automatically a cash saving. Assume that a feasible and authorised operating change can release forty per cent of that value as actual expenditure reduction, giving US\$72,000, and that incremental annual support costs US\$20,000. The resulting annual net cash benefit is US\$52,000 under these assumptions. Table 9.3 keeps the steps separate so that the reviewer can identify whether the disputed proposition concerns volume, adoption, time saved, financial conversion or additional cost. It is an annual operating illustration, not an investment appraisal or a revised programme forecast.

Table 9.3 Examine the conversion from capacity to cash

Step	Annual illustration	Evidence the reviewer would seek
Eligible and adopted volume	$120,000 \times 75\% = 90,000$	Defined eligible population and evidence of actual use.
Time released	$90,000 \times 4 \div 60 = 6,000$ hours	Comparable time observations; displaced work and service quality.
Capacity value	$6,000 \times \text{US\$}30 = \text{US\$}180,000$	A justified valuation rate and the use of released capacity.
Gross cash release	$\text{US\$}180,000 \times 40\% = \text{US\$}72,000$	A feasible, authorised expenditure-reduction mechanism.
Net cash benefit	$\text{US\$}72,000 - \text{US\$}20,000 = \text{US\$}52,000$	Complete incremental support cost and verified expenditure changes.

A separate fictional annual operating example, excluding investment cost, tax, discounting and transition timing. All inputs are assumptions; the result is a forecast, not a realised benefit. Capacity value and its cash-released share are not additive.

The practical examination follows the conversion mechanism. Released minutes scattered across employees may improve responsiveness without allowing a post, overtime payment or supplier

charge to be removed. That capacity can still be valuable, but recording the entire capacity value and the cash release as additive benefits would double count the same improvement. Lena should ask which budget or payment changes, when the reduction becomes effective, whether another team absorbs the work and whether service quality deteriorates. NIST's 2026 benefits-assurance guidance distinguishes financial release from efficiency savings and calls for continued ownership, measurement and review after delivery (NIST, 2026, pp. 17–20). The manufacturer's proposed record applies that distinction by connecting each financial claim to its operational mechanism and account. Non-financial improvements remain visible in their own measures rather than being forced into an unsupported monetary conversion.

Test readiness through the work people must perform

Operational readiness should be demonstrated through work that resembles the intended service closely enough to support the decision. Maya's team needs to perform a normal transaction, recognise and resolve an exception, and recover from a disruption without depending on support that will disappear after handover. Figure 9.6 proposes these as three complementary lines of inquiry, not an exhaustive acceptance standard. The reviewer selects scenarios from the actual operating risks and traces them through people, capacity, interfaces, authority and records. A payment exception may require a supplier response, a master-data correction and a finance approval; testing only the software screen leaves the organisational dependency unexamined. The exercise should record what happened, who intervened and which conditions differed from normal operation. The most useful finding may concern the work between teams rather than the component each team has already tested successfully.

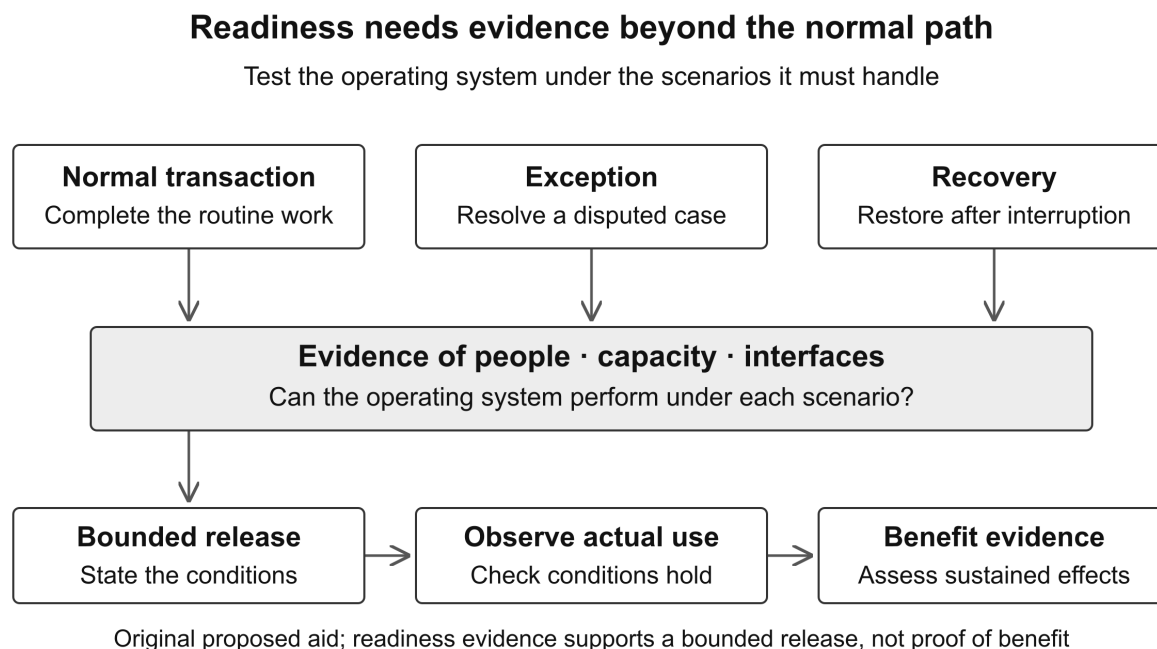


Figure 9.6 An original scenario-based readiness examination. Normal work, exceptions and recovery generate complementary evidence about people, capacity and interfaces. Release conditions connect this evidence to a bounded decision and later observation of use and benefits. The scenarios are not an exhaustive acceptance standard.

Shea et al. (2014) assessed a measure of organisational readiness that distinguishes commitment to change from beliefs about collective capability. Their psychometric studies supported aspects of the measure, while further predictive-validity work remained necessary. A readiness survey therefore supplies evidence about specified perceptions; it does not establish technical acceptance, prove staffing availability or provide a universal release threshold. Lena should inspect whose views are represented and whether disagreement is hidden by an average. A confident day shift cannot answer for a night shift that lacks an experienced supervisor. A small group facing the most difficult exceptions may matter more to the release decision than the overall mean rating suggests. Survey evidence can guide the review towards concerns and uneven preparation, while observed scenarios and resource evidence test whether the required work can actually be performed.

Implementation and benefits should also remain separate when performance disappoints. Durlak and DuPre's (2008) review connected implementation with outcomes across prevention and health-promotion research. Its setting differs from an industrial transformation, but the distinction is useful: an intervention may fail to produce the expected effect because the proposed practice was not enacted adequately, because the mechanism was weak, or because context altered its operation. Lena therefore examines actual reach, use, quality and adaptation before concluding that the platform itself cannot improve payment performance. At the same time, faithful use is not proof of worthwhile benefits. The review needs both an account of what changed in practice and an account of the outcome. Without the first, management may abandon a promising approach for the wrong reason; without the second, it may celebrate disciplined implementation of a process that does not create the intended value.

May and Finch's (2009) normalization process theory focuses on the work involved in embedding practices, including how people make sense of a change, participate, enact it together and appraise it. It is a conceptual explanation, not a sequence of mandatory gates. Applied here, it directs attention beyond the launch event. The controller can ask whether staff continue using the new routine when the project team is absent, whether exceptions return to unofficial spreadsheets and whether managers use the resulting information in ordinary decisions. These observations should inform the conditions for increasing the rollout. A bounded release may be justified with additional support and retained fallback, while wider deployment waits for evidence of sustained operation. The same observation period can reveal disbenefits such as displaced workload, slower service for particular suppliers or loss of useful local knowledge, which should accompany the benefit account.

Make challenge possible before the answer hardens

Assurance depends on people who can recognise an unwelcome implication and discuss it without turning the review into a contest of status. Kunda's (1990) review explains how motivation towards a preferred conclusion can affect the information and reasoning people use while remaining constrained by plausible justification. The issue is not confined to deliberate misrepresentation. Samira may sincerely want the release to succeed, Alex may identify with the solution and Lena may take pride in the model she built. A useful review design acknowledges these commitments and makes the reasoning examinable. Criteria should be recorded before the preferred answer is assessed, and reviewers should identify what evidence would change their view. These are reasoned safeguards informed by research; they are not a guarantee that motivated reasoning has been removed from the decision.

Moore et al. (2006) analyse how conflicts of interest can shape auditing judgements through individual and institutional mechanisms without requiring conscious corruption. Moore, Tanlu and Bazerman (2010) add three experiments in which assigned roles and affiliations affected judgements, including private valuations in later experiments despite accuracy incentives. Results varied by outcome, and the tasks do not estimate bias in a particular transformation review. Together, the studies support treating appointment, relationships and incentives as part of review design. Disclosure is useful, but it does not remove the relationship disclosed. A supplier that designed the benefit model can explain it and respond to challenge; its assurance of that same model should not be represented as equivalent to a competent examination with suitable separation and freedom to report an adverse conclusion.

Lerner and Tetlock (1999) show in their review that accountability has different effects depending on timing, audience and task. Requiring people to justify themselves can encourage careful thought under some conditions, but can also reinforce conformity or defence of an existing position. In practice, Samira should make it clear that the review may change the decision. Asking for independent challenge after announcing an immovable launch date sends a conflicting signal. The controller can arrange an initial examination before the team sees the sponsor's preferred recommendation, followed by a structured discussion of disagreements and missing evidence. This is not a rule that reviewers must work in isolation. It is a way of giving them room to form a judgement before social pressure narrows the range of acceptable conclusions.

A constructive challenge meeting should therefore concentrate on the claim, the reason for doubt and the decision consequence. The challenger identifies the evidence or assumption at issue and explains what further observation would resolve the concern. The claim owner can correct a factual misunderstanding, provide additional material, narrow the claim or accept a required action. Lena records the remaining difference without forcing artificial consensus. Where the disagreement is material, a competent decision maker or further specialist review is needed. Where it is immaterial to the bounded commitment, it can remain documented for later examination. This approach gives challenge a useful endpoint. The purpose is to improve the decision and its conditions, rather than reward the person who finds the largest number of faults or the person who can keep every dashboard green.

Close the finding through evidence of effectiveness

A recommendation accepted at a review is only the beginning of a control response. Vo et al. (2021) examined assurance in the UK government's major-projects portfolio using database analysis, a workshop, selected cases and interviews. Their study identified difficulties in recommendation framing and follow-through; observed changes in delivery-confidence ratings should not be interpreted as a controlled estimate of assurance's effect on realised benefits. The implication for the controller is to preserve the route from finding to response and subsequent evidence. A recommendation should be clear enough that the responsible owner understands the change required and the reviewer can later determine what was accomplished. "Improve benefits management" may describe a legitimate concern, but it is too broad to close reliably without specifying the missing arrangement, its decision consequence and the evidence that would demonstrate an adequate response.

Figure 9.7 distinguishes four stages: the action has been designed, the change has been implemented, the relevant operation has been independently reperformed where appropriate, and effectiveness has been observed over a suitable period. This is an original practice sequence for verification within the Credible Control Method, not a universal maturity scale. Some findings can be closed after a correction and re-performance; others require repeated operational evidence. The closure criteria should be proportionate to the finding. A corrected formula can be checked immediately, whereas a new reconciliation process needs evidence from its use. Where sustained observation is necessary, the action may be implemented while the assurance conclusion remains conditional. Keeping these states separate prevents an administrative action tracker from becoming an accidental statement that the underlying exposure has disappeared.

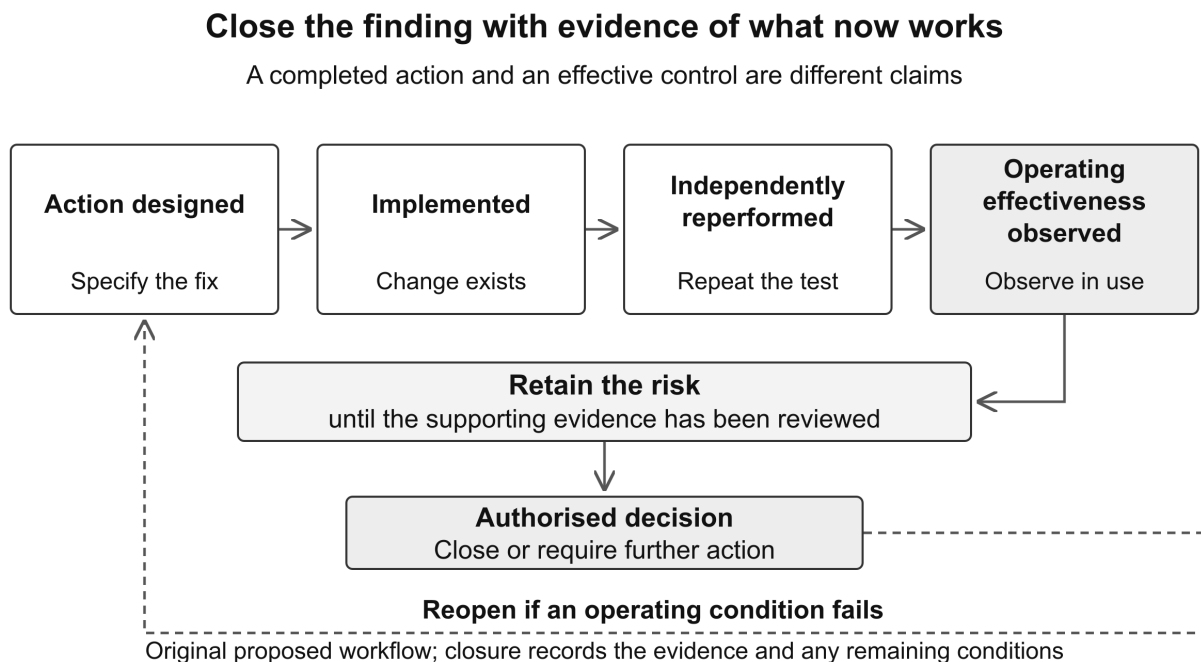


Figure 9.7 A proposed verification sequence distinguishes designing an action, implementing it, reperforming the relevant operation and observing effectiveness. The appropriate closure evidence depends on the finding; not every finding requires the same observation period. Reopen a conclusion when its supporting conditions fail.

Return to Chapter 8's missing accrual. The action owner might create an acceptance form and circulate guidance. That establishes a designed response and perhaps its introduction; it does not show that accepted service reaches the cost account on time. Lena should identify the next relevant cut-off, obtain the population of accepted services, trace selected items into the account and inspect exceptions. The reviewer should also check whether the change creates another problem, such as duplicate recognition when the invoice arrives. If the control works for one straightforward item but fails for disputed services, the conclusion should retain that limitation. The corrective action's effectiveness concerns the process and its operating conditions, not merely whether someone uploaded a completed document. This gives the organisation a defensible reason to close, extend or reopen the finding.

Denmark's IC4 experience shows why the conditions attached to expert advice must travel with its conclusion. Rigsrevisionen's October 2017 report examined DSB's work between late 2014 and late 2016. A commissioned assessment had judged completion for long-distance use in 2019 feasible subject to further activities and modifications. The audit found several critical activities had not been implemented; recommended maintenance staffing was reached only at the end of the period, while daily operations had been prioritised over modifications. DSB subsequently reduced intended deployment as reliability deteriorated and new faults emerged (Rigsrevisionen, 2017, pp. 1–3). The lesson is to retain the conditional conclusion together with owners and closure evidence. The audit did not establish that the original advice guaranteed success, nor should its later publication be presented as the cause of the earlier deployment decision.

A benefits finding needs the same discipline after the delivery team departs. If the review requires a credible baseline, a documented baseline is an output; a verified, consistently measured comparison is the relevant evidence. If it requires a cash-release mechanism, a benefit owner's signature is not the expenditure change. The assurance record should identify who will observe the result, when the evidence can reasonably exist and what happens if it contradicts the expected improvement. Responsibilities and funding for that work should survive project closure. A condition that nobody can monitor is a weak basis for proceeding, however reassuring its wording. Samira should see overdue conditions according to their consequence for the commitment, rather than only their age or the percentage of actions administratively marked complete.

Bring the evidence to one accountable decision

The practical instrument that brings the chapter together is the proposed Assurance and Decision Record. It builds on the earlier Decision Connection Record and Evidence and Learning Record rather than replacing them. For each consequential claim, it retains the decision and date, criteria, examination performed, reviewer and independence arrangements, evidence references and versions, unresolved objections, owner response and authorised conditions. Table 9.4 shows compact teaching entries. In a live programme, the record would link to detailed test results instead of reproducing them in every board paper. The reviewer states what the evidence supports and what it excludes; the decision maker records what may proceed within those limits. The record remains open to revision when a condition fails or material circumstances change. Its purpose is continuity between judgement and action, including benefits work that extends beyond the delivery team's formal existence.

Table 9.4 Illustrative entries in an Assurance and Decision Record

Claim and examination	Evidence boundary	Response and later verification
Routine control operates: examine selected transactions against defined criteria.	Sixty tests, zero failures; one-sided upper bound 4.87% under the stated model.	Retain the test scope and assumptions; do not extend the finding to all failure modes.
Team can recover service: witness a disruption and recovery scenario.	Recovery evidence is still missing in the final exercise.	Maya arranges the exercise; competent review precedes any release dependent on recovery.
Annual net cash benefit is achievable: examine adoption, time and expenditure conversion.	US\$52,000 is a conditional forecast; the expenditure mechanism awaits authorisation.	Benefit and budget owners secure the decision, then verify actual change and support cost.
Missing-accrual response works: trace accepted service at a later cut-off.	A new form demonstrates an action, not its operating effectiveness.	Owner implements; reviewer checks completeness and duplicate recognition before closure.

Original teaching entries, not approvals. A live record also identifies decision date, criteria, reviewer independence, evidence references and versions, owner, deadline, residual exposure, authorised conditions and reopening triggers.

Consider a final exercise. The next unit has completed normal-transaction tests, sixty independently selected routine control tests found no failures under the stated binomial assumptions, and the operating team has rehearsed common exceptions. Recovery from a lost interface has not been demonstrated. The financial model identifies US\$180,000 of capacity value, but only the assumed US\$72,000 gross expenditure release has a proposed mechanism, and that mechanism still awaits authorisation. Incremental annual support is forecast at US\$20,000. Two favourable review reports use the same underlying test extract. Samira is asked to authorise unrestricted rollout and declare the annual saving achieved. Before reading the proposed response, the learner should identify which claims are supported, which remain conditional, whether the proposed scope is reversible and what evidence could change the decision. The programme's overall approved budget and current forecast remain those established in earlier chapters.

The evidence does not support the decision as phrased. The clean sample supports a bounded statistical statement about the tested routine process, not proof of complete readiness. Reports sharing the same underlying test extract cannot automatically be treated as conditionally independent evidence. The untested recovery path requires examination before any release whose consequences depend on that capability. The US\$52,000 annual net cash figure is an assumption-based forecast, not an achieved benefit, and the US\$180,000 capacity value cannot be added to the cash release as if it represented unrelated value. A revised proposal might permit a limited, reversible release only after specified recovery evidence and support arrangements are accepted by the authorised owner. If no credible fallback or acceptable boundary exists, postponement may be necessary. The chapter does not prescribe release: it teaches how the evidence constrains the available justification.

Lena can now prepare a decision that is both concise and defensible. The paper states the exact release boundary, the evidence examined and its date, the unfulfilled conditions and the person responsible for verifying each one. Maya retains ownership of benefit observation and the proposed operating change; Alex retains delivery responsibilities; Samira authorises the commitment within her mandate. The assurance reviewer retains the freedom to report that a condition has not been met. Subsequent performance then returns to Chapter 8's learning process: an adverse observation may reopen the claim, improve the model or change the rollout. This preserves the connection between assurance and learning without allowing either activity to substitute for management. A review is useful when it helps the organisation make a better-supported commitment and recognise when that support no longer holds.

The chapter began with three favourable reports and a decision they did not fully answer. It ends with a more disciplined form of confidence: a claim bounded by purpose and conditions, evidence whose meaning can be examined, an argument exposed to fair challenge and an authorised response that remains connected to later verification. For project controllers, this is a practical extension of forecasting, risk analysis, information management and benefits ownership. It requires enough technical understanding to ask what a calculation establishes, enough operational curiosity to examine how work is performed and enough professional independence to preserve an unwelcome finding. Confidence that survives challenge remains open to correction. Its value lies in helping the organisation pursue transformation while retaining an honest account of what has been demonstrated, what remains assumed and what must happen next.

References

- Australian National Audit Office (ANAO). (2019, 25 November). Implementation of the My Health Record system (Auditor-General Report No. 13 of 2019–20).
www.anao.gov.au/work/performance-audit/implementation-the-my-health-record-system
- Badewi, A. (2016). The impact of project management (PM) and benefits management (BM) practices on project success: Towards developing a project benefits governance framework. *International Journal of Project Management*, 34(4), 761–778. doi: 10.1016/j.ijproman.2015.05.005

- Bloomfield, R., & Bishop, P. (2010). Safety and assurance cases: Past, present and possible future—an Adelard perspective. In C. Dale & T. Anderson (Eds.), *Making Systems Safer* (pp. 51–67). Springer. doi: [10.1007/978-1-84996-086-1_4](https://doi.org/10.1007/978-1-84996-086-1_4)
- Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4), 404–413. doi: [10.1093/biomet/26.4.404](https://doi.org/10.1093/biomet/26.4.404)
- DeAngelo, L. E. (1981). Auditor size and audit quality. *Journal of Accounting and Economics*, 3(3), 183–199. doi: [10.1016/0165-4101\(81\)90002-1](https://doi.org/10.1016/0165-4101(81)90002-1)
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350. doi: [10.1007/s10464-008-9165-0](https://doi.org/10.1007/s10464-008-9165-0)
- Goodenough, J. B., Weinstock, C. B., & Klein, A. Z. (2015). Elimination argumentation: A basis for arguing confidence in system properties (Technical Report CMU/SEI-2015-TR-005). Carnegie Mellon University, Software Engineering Institute. doi: [10.1184/R1/6573413.v1](https://doi.org/10.1184/R1/6573413.v1)
- Hawkins, R., Kelly, T., Knight, J., & Graydon, P. (2011). A new approach to creating clear safety arguments. In C. Dale & T. Anderson (Eds.), *Advances in Systems Safety* (pp. 3–23). Springer. doi: [10.1007/978-0-85729-133-2_1](https://doi.org/10.1007/978-0-85729-133-2_1)
- Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, 183(8), 758–764. doi: [10.1093/aje/kwv254](https://doi.org/10.1093/aje/kwv254)
- Institute of Internal Auditors (IIA). (2026). Three Lines Model: Assurance and Advice in Support of Effective Governance. Statement of Position. www.theiia.org/globalassets/site/resources/statements-of-position/tlm_assurance_advice_support_effective_gov_en.pdf
- Kelly, T., & Weaver, R. (2004). The goal structuring notation—a safety argument notation. Proceedings of the Dependable Systems and Networks 2004 Workshop on Assurance Cases. www.researchgate.net/publication/228990118_The_goal_structuring_notation-a_safety_argument_notation
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. doi: [10.1037/0033-2909.108.3.480](https://doi.org/10.1037/0033-2909.108.3.480)
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255–275. doi: [10.1037/0033-2909.125.2.255](https://doi.org/10.1037/0033-2909.125.2.255)
- Lopez Bernal, J., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International Journal of Epidemiology*, 46(1), 348–355. doi: [10.1093/ije/dyw098](https://doi.org/10.1093/ije/dyw098)
- Lu, J. (2020). Estimating Instrument Performance: with Confidence Intervals and Confidence Bounds (NIST Technical Note 2119). National Institute of Standards and Technology. doi: [10.6028/NIST.TN.2119](https://doi.org/10.6028/NIST.TN.2119)
- May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: An outline of normalization process theory. *Sociology*, 43(3), 535–554. doi: [10.1177/0038038509103208](https://doi.org/10.1177/0038038509103208)
- Moore, D. A., Tanlu, L., & Bazerman, M. H. (2010). Conflict of interest and the intrusion of bias. *Judgment and Decision Making*, 5(1), 37–53. doi: [10.1017/S1930297500002023](https://doi.org/10.1017/S1930297500002023)
- Moore, D. A., Tetlock, P. E., Tanlu, L., & Bazerman, M. H. (2006). Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review*, 31(1), 10–29. doi: [10.5465/amr.2006.19379621](https://doi.org/10.5465/amr.2006.19379621)
- National Infrastructure and Service Transformation Authority (NISTA). (2026). Guide on Assurance of Benefits Realisation in Major Projects (version 2.1, August). assets.publishing.service.gov.uk/media/6a832e841f783fe89daaa40f/Guide_on_Assurance_of_Benefits_Realisation_in_Major_Projects_V2.1_August_2026.pdf

- National Transportation Safety Board (NTSB). (2007, 10 July). Ceiling collapse in the Interstate 90 connector tunnel, Boston, Massachusetts, July 10, 2006 (Highway Accident Report NTSB/HAR-07/02). www.nts.gov/investigations/AccidentReports/Reports/HAR0702.pdf
- Office of the Auditor-General (New Zealand). (2016, September). Using Gateway reviews to support public sector projects. ao.parliament.nz/2016/gateway/docs/gateway.pdf@@download/file/gateway.pdf
- Rigsrevisionen. (2017, October). Extract from Rigsrevisionen's report on DSB's effort to complete the IC4 trains (1/2017). www.uk.rigsrevisionen.dk/Media/D/7/1-2017.pdf
- Shea, C. M., Jacobs, S. R., Esserman, D. A., Bruce, K., & Weiner, B. J. (2014). Organizational readiness for implementing change: A psychometric assessment of a new measure. *Implementation Science*, 9, Article 7. doi: 10.1186/1748-5908-9-7
- Shenhar, A. J., Dvir, D., Levy, O., & Maltz, A. C. (2001). Project success: A multidimensional strategic concept. *Long Range Planning*, 34(6), 699–725. doi: 10.1016/S0024-6301(01)00097-8
- Vo, H., Kirkham, R. J., Williams, T. M., Howells, A., Forster, R., & Cooke-Davies, T. (2021). An empirical study of assurance in the UK government major projects portfolio: From data to recommendations, to action or inaction. *International Journal of Managing Projects in Business*, 14(4), 865–897. doi: 10.1108/IJMPB-10-2019-0242
- Williams, T., Klakegg, O. J., Walker, D. H. T., Andersen, B., & Magnussen, O. M. (2012). Identifying and acting on early warning signs in complex projects. *Project Management Journal*, 43(2), 37–53. doi: 10.1002/pmj.21259

CHAPTER 10

Decisions under Uncertainty



IN THIS CHAPTER

Choosing and adapting when outcomes cannot be known with certainty.

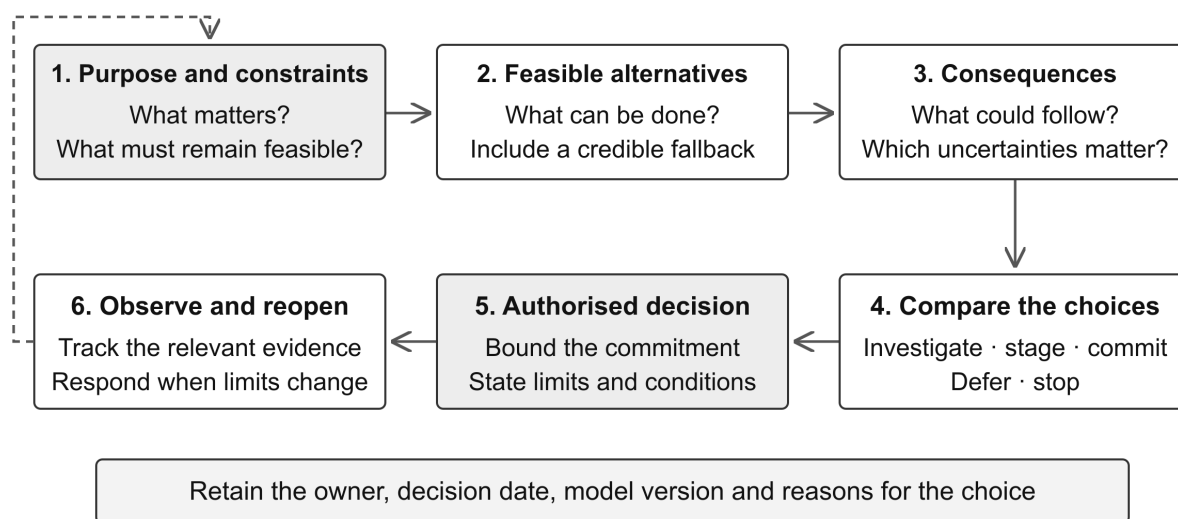
Turn uncertainty into a decision question

The manufacturer's assurance review has clarified what the evidence supports, but Samira still has to choose. Alex wants to commit the remaining units to the payment platform while supplier capacity is available. Maya prefers a smaller wave that gives operations time to learn. Lena can describe the exposure of each course, yet no further report will make the future completely observable. Waiting also has consequences: supplier options can expire, benefits can arrive later and the existing operation continues to consume resources. Chapter 9 established how to challenge confidence in a claim. This chapter develops the next responsibility: deciding what to do when the claim remains conditional. The controller's contribution is to connect the available alternatives with their consequences, the uncertainties that distinguish them and the authority needed to act before the choice disappears.

The central argument is that uncertainty should change the design of a commitment, not merely widen the range around a forecast. Sometimes the organisation should pay to learn more; sometimes it should proceed within a reversible boundary; sometimes it should protect a future option or stop an unattractive course. A sophisticated risk model becomes useful when it informs those choices. Its value is limited if the only result is a revised colour beside an unchanged plan. The proposed Decision and Uncertainty Map in Figure 10.1 connects purpose and constraints to feasible alternatives, uncertain consequences, comparison and an authorised response. It is an original practice aid within the existing Credible Control Method. Frame, Establish, Challenge, Authorise and Verify remain the organising activities, with the decision memorandum preserving the reasoning between them.

Make the decision and its uncertainty explicit

A practice aid within the existing Credible Control Method



Original proposed aid; calculations inform judgement within stated authority and constraints

Figure 10.1 The proposed Decision and Uncertainty Map connects purpose, feasible alternatives and uncertain consequences to an authorised response and later review. This original practice aid operates within the existing Credible Control Method; it is not a validated decision scale.

Several different problems are often hidden inside the word risk. A current issue has already occurred and requires a response, although its future consequences may remain uncertain. An estimate can vary because quantities, productivity or prices are imperfectly known. A discrete event may or may not occur. An assumption is a proposition used to make analysis possible and should have a basis, owner and review condition. Ambiguity may concern the mechanism itself or disagreement about what success means. These distinctions guide action. More observations may improve a productivity estimate, but they cannot settle a disagreement between faster processing and greater service accessibility without a decision about objectives. A risk register should preserve the relationship between these categories while preventing an existing failure from being presented as though it were merely a possibility still under surveillance.

Aven (2016) reviews risk foundations and explains why probability and expected loss can conceal weaknesses in the knowledge supporting an assessment. This is a conceptual argument for examining the basis of a judgement, not a reason to abandon probability whenever information is imperfect. Lena should therefore place the source of each important estimate beside its numerical result. A range derived from repeated comparable work has a different evidential basis from an expert's judgement about a novel operating change. Both can contribute to a decision if their meaning is clear. The sponsor needs to know which uncertainty the model represents and which plausible omissions or disputed assumptions sit outside it. Apparent precision should not remove the opportunity to ask whether the model describes the choice that the organisation is actually facing.

Pich, Loch and De Meyer (2002) develop a formal model linking project-management approaches to information adequacy and complexity. Their distinction between executing a sufficiently understood plan, learning, and trying alternative approaches in parallel before selecting helps explain why one control response cannot suit every unknown. If Maya knows the exception-handling process but is unsure about transaction volume, she may need capacity analysis. If nobody knows whether suppliers will adopt the new process, an investigation or bounded trial may be more informative than another detailed rollout schedule. This is a practical application of their theoretical reasoning, not a diagnostic instrument validated for every programme. The controller should identify whether the missing knowledge concerns a number inside an understood mechanism or whether the mechanism itself needs examination before a difficult-to-reverse commitment is justified.

Define the choice before calculating its value

A decision begins with an explicit question, a date and an accountable owner. “What is the risk position?” is a reporting question. “Should the next two units enter the service together, sequentially or after an additional trial?” is a decision question. Lena should specify which commitments each answer would authorise, what can still be changed and when an alternative becomes unavailable. She should also distinguish alternatives that are feasible now from those requiring a negotiation, design change or additional authority. Deferral must have an operating meaning: who continues the existing service, at what cost, with what consequences for benefits and supplier commitments? Stopping must include residual obligations and any value that can be retained. These definitions prevent the comparison from favouring the preferred option simply because its costs are described precisely while the alternatives remain vague.

Howard (1988) presents decision analysis as a discipline connecting a decision frame, alternatives, information, preferences and logical evaluation. The article is a methodological perspective rather than evidence that a particular software package improves programme outcomes. Its relevance to the manufacturer is the need to make the structure of the choice visible before calculating. Lena can begin with a short statement of purpose, such as improving payment reliability while maintaining service continuity and acceptable affordability. She then identifies which consequences are monetary, which require separate measures and which represent constraints that an option must satisfy. The resulting analysis should make a disagreement traceable. If Alex and Maya prefer different options because they value speed and operating stability differently, refining the same cost forecast will not resolve their underlying judgement without further discussion.

March and Shapira (1987) examine managerial perspectives on risk and the ways these can differ from conventional decision-theoretic representations. Their analysis helps explain why a sponsor may attend to a critical performance target or distinguish risk taking from gambling rather than respond only to a distribution's mean and variance. The controller should make those concerns explicit rather than assume they are irrational or already captured by expected value. An organisation that cannot finance a large downside may reasonably reject an option with an attractive average. Equally, a stated aversion to uncertainty should not become a convenient reason to ignore the risks of the current operation. The comparison needs a consistent account of consequences across alternatives, including the exposure retained by doing less or waiting. Preferences and constraints belong in the decision record where they can be challenged and authorised.

British Columbia's 2017 Site C inquiry illustrates how a recommendation can depend on assumptions that analysis cannot settle by itself. The Utilities Commission examined continuing, suspending or terminating the hydroelectric project, including alternative means of meeting future electricity needs. Its final report did not present a choice between completion and termination as independent of future demand and cost. The relative merits changed with assumptions about load growth, completion expenditure and the alternative portfolio (BCUC, 2017, p. 187). For the controller, the lesson is to expose the conditions under which the preferred course changes. An analysis can clarify those conditions without claiming that one forecast has

removed the decision maker's responsibility. The report is evidence about the inquiry's findings at that time, rather than a statement about the project's subsequent outcome or the present value of the alternatives.

Non-financial consequences should be handled with the same care. The organisation may compare payment accuracy, supplier accessibility, operational workload and service continuity alongside financial value. Where a consequence is a binding requirement, it should be treated as such before ranking feasible options. Where trade-offs are legitimate, the sponsor should understand what each option gains and sacrifices. A weighted score can help organise preferences only if scales, weights and permitted compensation are defensible; a high saving should not silently offset an unacceptable service failure through arithmetic. Lena can test whether plausible changes in preferences alter the recommendation and retain dissent where affected parties reasonably value outcomes differently. This keeps the technical comparison connected to governance rather than allowing the choice of units or scoring conventions to make the decision unnoticed.

Use expected value without losing the downside

Consider a separate fictional future investment choice, expressed in millions of US dollars of net present value. It does not replace the manufacturer's approved programme budget or current forecast. Full commitment produces a net value of twelve if adoption is strong and minus six if adoption is weak. A staged approach produces seven under strong adoption and two under weak adoption. These values already combine the delivery and operating consequences assigned to each option within the teaching model. Suppose the assessed probability of strong adoption is 0.60. Table 10.1 retains both states and their probabilities so that the comparison can be reconstructed. The two-state representation is deliberately simple: real application would require a justified account of adoption, benefits, cost, timing and the range of operating conditions rather than assuming that every outcome fits neatly into two categories.

Table 10.1 Compare conditional net values before averaging

Approach	Weak adoption 40%	Strong adoption 60%	Expected value
Full commitment	-6.00	12.00	4.80
Staged approach	2.00	7.00	5.00

US\$ millions of present value in a separate fictional choice. The option values exclude the later study fee and incremental delay cost. Maximising expected money is a risk-neutral criterion; feasibility and acceptable downside remain separate considerations.

The expected net value of full commitment is 0.60 multiplied by twelve plus 0.40 multiplied by minus six, giving 4.80. Staging gives 0.60 multiplied by seven plus 0.40 multiplied by two, giving 5.00. Under a risk-neutral criterion of maximising expected monetary value, staging is preferred by 0.20 million. The calculation does not predict that the organisation will receive five million; neither staged outcome equals that amount. It compares the probability-weighted consequences

of the available choices under the stated model. Expected values can be added without assuming independence, provided their boundaries are consistent and double counting is avoided. What generally cannot be added in the same way are percentiles, because the distribution of a combined outcome depends on how its components move together. That distinction becomes central when cost, schedule and benefits are assessed jointly.

The relatively small expected-value difference should encourage examination of the basis of the decision. Full commitment offers greater upside but exposes the organisation to a six-million loss in the weak-adoption state. Staging sacrifices some upside in return for a better outcome under weak adoption. A sponsor who faces a binding funding constraint may find the downside decisive even if a revised probability makes full commitment's expected value slightly higher. Conversely, another organisation may be able and willing to carry that exposure. The controller should report the full consequence pattern, not only the expected-value ranking. If preferences require a utility model, the shape of that model should be elicited and explained rather than chosen to manufacture a desired answer. The financial expectations, distributional consequences and authority to accept the downside should remain distinguishable.

The arithmetic also clarifies what a model can leave out. Suppose full commitment would remove the only reliable fallback before the new service has demonstrated recovery capability. The option may be infeasible under the organisation's service-continuity requirements regardless of its expected financial value. Alternatively, the staged option may consume scarce staff needed for another initiative, and its apparent protection may disappear once that opportunity cost is included. Lena should test the option definitions with the people who must implement them. The comparison is between actions with consequences, not labels such as ambitious and cautious. A staged strategy needs a genuine decision point, usable evidence and the practical ability to change the remaining commitment; splitting an unavoidable obligation into several payment instalments does not automatically create an option.

Before recommending further analysis, Lena asks what could change the preference. In this example, the two strategies are close at the assessed adoption probability, so better evidence about adoption could matter. If one option dominated the other under every credible state and met all relevant constraints, a study directed only at those states would have little decision value. This is a useful discipline for controllers facing a growing demand for reports. Analysis should have an identified route to a different action, a better-designed commitment or more effective protection. Information can still be required for compliance, accountability or learning beyond the immediate choice, but those purposes should be stated separately. They should not be confused with an unsupported claim that every additional study pays for itself by improving the current investment decision.

Decide whether further information is worth buying

The value of information is the improvement in a decision made possible by learning, before allowing for the cost of obtaining that knowledge. Howard (1966) connects information value to probabilities and economic consequences relevant to a decision. This is a decision-analytic concept rather than a property inherent in a report. In the fictional investment example, perfect information about adoption would allow full commitment in the strong state and staging in the weak state. The expected value would be 0.60 multiplied by twelve plus 0.40 multiplied by two, or eight million. Compared with the best current choice of five million, the expected value of perfect information is three million. Within this same decision model, that is an upper bound on the gross value of an imperfect study about the same uncertainty; it is not a proposed study budget or a guarantee that perfect prediction can be purchased.

A real study usually provides a signal rather than certainty. Suppose an adoption trial produces a positive result in eighty per cent of strong-adoption cases and twenty-five per cent of weak-adoption cases. These are fictional performance assumptions, introduced to teach the calculation. The trial supplies information without itself changing adoption states or payoffs, apart from the study and delay costs specified below. The overall probability of a positive signal is 0.60 multiplied by 0.80 plus 0.40 multiplied by 0.25, giving 0.58. Given that positive result, the revised probability of strong adoption is 0.48 divided by 0.58, or approximately 82.76 per cent. The negative result occurs with probability 0.42, and its revised strong-adoption probability is 0.12 divided by 0.42, or 28.57 per cent. Chapter 9 taught the logic of updating evidence. Here the additional step is to connect each possible result to the action that would then be selected.

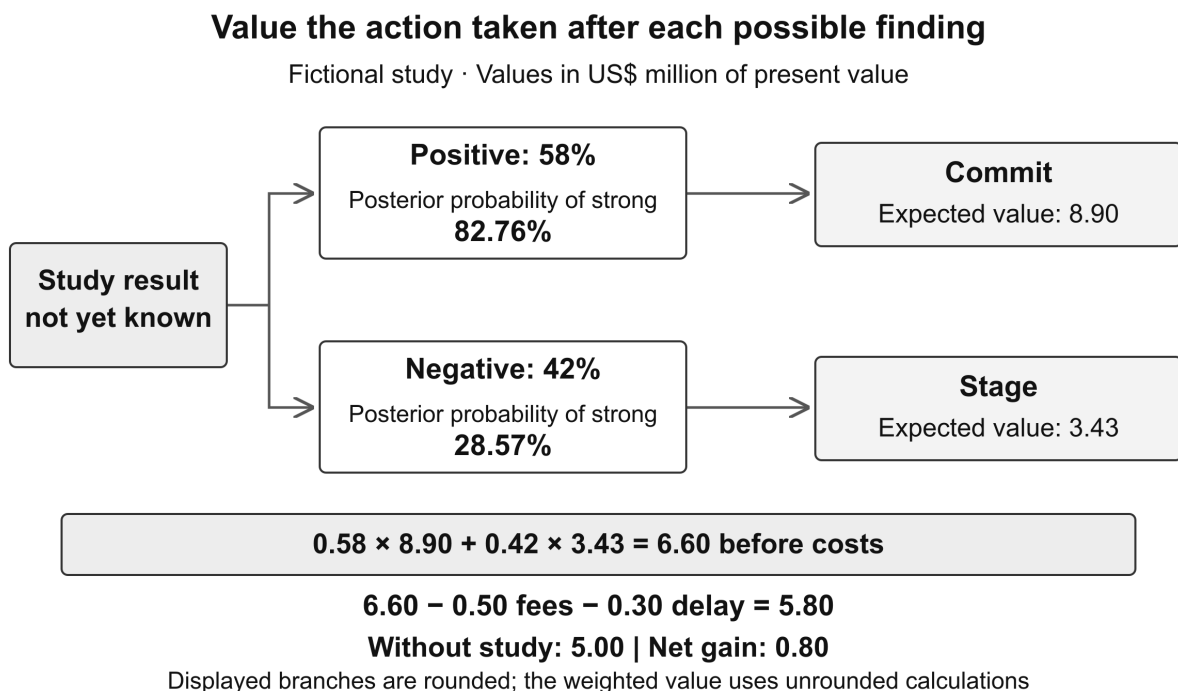


Figure 10.2 Original fictional information-value calculation in US\$ millions of present value. The best response after each signal produces 6.60 before information costs. Subtracting the 0.50 study fee and 0.30 delay cost gives 5.80, compared with 5.00 without the study. The signal informs the choice without itself changing adoption states or option payoffs.

After a positive signal, full commitment has expected value approximately 8.90 million, compared with 6.14 for staging. After a negative signal, staging gives approximately 3.43 million, compared with minus 0.86 for full commitment. The proposed policy is therefore to commit after a positive result and stage after a negative result, subject to the unchanged feasibility requirements. Weighting the selected values by the signal probabilities gives 0.58 multiplied by 8.8966 plus 0.42 multiplied by 3.4286, or 6.60 million. The expected value of sample information is consequently 1.60 million above the original best choice. The order of operations matters: evaluate the alternatives after each signal, choose the best feasible response and then average across signals. Averaging the information without allowing the decision to change would miss the source of its value.

Suppose the study costs 0.50 million and the incremental opportunity cost of waiting for it is 0.30 million. The original option values exclude those two items, so subtracting them here does not count them twice. The expected net value of the study-and-response policy becomes 5.80 million, an improvement of 0.80 over staging immediately. If study fees and additional delay consequences exceeded 1.60 million, this particular information would no longer justify its cost under the same assumptions. Delay may also change the available alternatives, the probability of adoption or the payoff itself; those changes require a revised model rather than an arbitrary deduction. The worked tree assumes the option definitions remain available while the signal is collected. That assumption should be tested before the controller recommends commissioning a trial whose result may arrive after the useful decision date.

Strong et al. (2015) develop computational methods for estimating expected value of sample information and demonstrate them in health-economic decision models. The paper supports treating information value as dependent on a specified study design, model and decision, rather than using a generic percentage allowance for research. Its setting and numerical methods differ from the manufacturer's simple tree. In practice, Lena should describe the population the trial would observe, its likely accuracy, possible results, time to completion and how each result would affect the commitment. A trial that tests convenient users while excluding the groups most likely to struggle may produce a signal with much less decision value than assumed. A study should be judged by the information it can credibly supply and the actions it can still influence, not by the precision of the report it promises to deliver.

Loch, Solt and Bailey (2008) examine an exploratory new-venture case in which breaking the problem into subproblems helped identify areas requiring search and learning. The authors' involvement and single-case setting limit generalisation, but the example is useful where the unknown cannot yet be represented by a stable signal model. Lena might first need to investigate why suppliers resist the new process before estimating the probability that a larger trial will predict adoption. That exploratory work can change the option set or reveal a mechanism absent from the original appraisal. Its expected value may be difficult to quantify credibly. The controller can still bound its cost, specify the question, identify who will use the findings and set a decision deadline. Buying information remains a commitment that needs purpose and governance, even when its value cannot responsibly be reduced to one number.

Find the assumptions that can reverse the choice

Sensitivity analysis should explain how the decision changes, not merely how individual outputs move. Return to the strong-adoption probability, now denoted p . Full commitment's expected value is twelve times p minus six times one minus p , which simplifies to $18p - 6$. Staging gives $7p + 2(1 - p)$, or $5p + 2$. Equating the two expressions gives $13p = 8$, so the switching probability is approximately 61.54 per cent. Below that value, staging has the higher expected monetary value; above it, full commitment does, subject to the same feasibility conditions. Figure 10.3 makes the crossing visible. The original judgement of sixty per cent lies close to the threshold, so the recommendation should not be described as insensitive to reasonable disagreement about adoption. The threshold is a property of the assumed payoffs and criterion, not a universal adoption target.

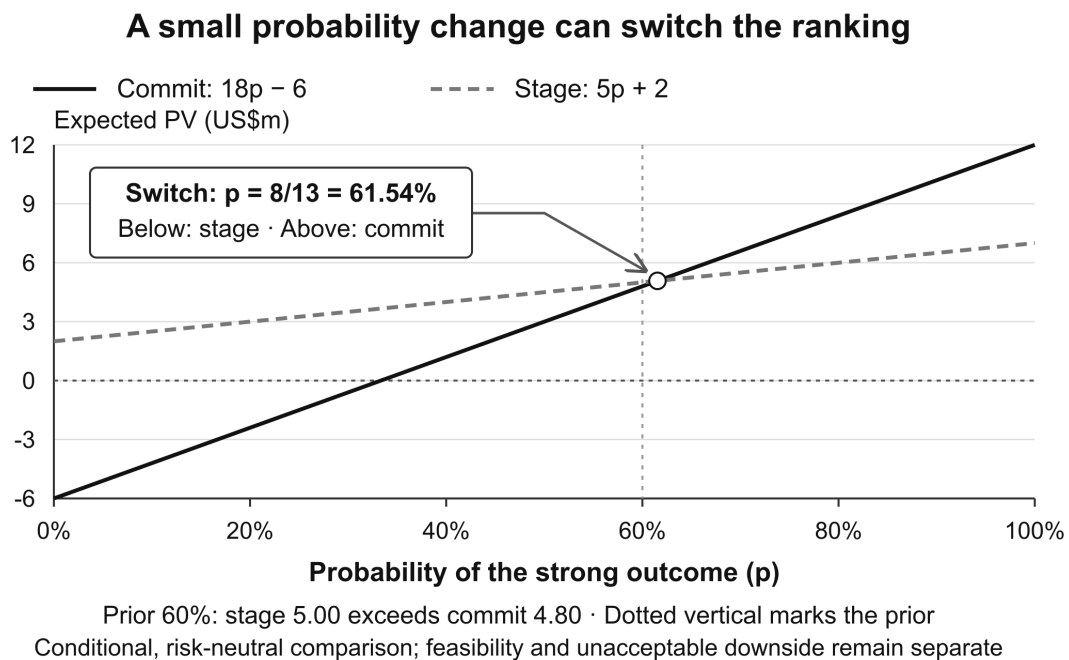


Figure 10.3 Expected-value preference switches at a strong-adoption probability of $8/13$, approximately 61.54 per cent. The original 60 per cent judgement is close to that threshold. The calculation holds payoffs and the risk-neutral criterion fixed; it does not establish the probability of crossing the threshold.

HM Treasury's Green Book (2026) asks practitioners to test uncertain assumptions and identify switching values at which an option ceases to be preferred or worthwhile. That appraisal guidance is useful here because it connects analysis to a change in choice. Lena can ask whether the evidence distinguishes a sixty per cent adoption probability from a value above the model's switching point. She can also identify the benefit or cost change that would make staging unattractive. This gives an investigation a specific purpose. A threshold does not say how likely it is to be crossed, and it should not be confused with a forecast. It tells the sponsor where the reasoning would need to change under the other assumptions held fixed. A different objective, constraint or payoff definition could move the threshold or remove the option altogether.

One-at-a-time tests are useful for understanding a model locally, but they can miss interactions and combinations. Saltelli and Annoni (2010) explain limitations of changing one input at a time when assessing sensitivity across a model's input space. Their methodological argument is relevant when adoption, benefit timing and support requirements can change together. Lena should distinguish a local switching calculation from a broader exploration of coherent scenarios. For example, a difficult migration could reduce early adoption, delay benefit onset and increase operating support simultaneously. Testing each at its favourable base value while changing only one at a time may understate the combined exposure. Conversely, combining every pessimistic extreme regardless of compatibility can create a scenario that no plausible operating state would produce. The analysis needs defensible relationships as well as ranges.

A practical method is to begin with a few transparent alternatives and then refine where the choice remains sensitive. Chapman and Ward (2000) propose a minimalist first pass for estimating and evaluating uncertainty, using an iterative approach to ranges and their implications. The contribution is methodological; it does not establish that any rough estimate is trustworthy. For the controller, the discipline is to record what each range means, why its endpoints are plausible and which unknowns deserve further effort. A triangular distribution, for example, requires a minimum, mode and maximum; calling three elicited numbers optimistic, likely and pessimistic does not establish that they are valid support limits. If the assessor actually means tenth, fiftieth and ninetieth percentiles, those values must be fitted or modelled accordingly. The distribution should express the judgement elicited, rather than silently changing its meaning.

The response to sensitivity should be operational. If adoption drives the preference, commission evidence that improves understanding of adoption or design a commitment that tolerates weaker use. If the decision depends on a narrow completion window, investigate the activities and approvals capable of shifting that window. If the preference reverses only under an implausible assumption far outside the evidence, further precision may add little value. Lena should retain the sensitive assumption, its source, the relevant threshold, the next evidence and the person who can act. This makes the analysis usable at a later review. A colourful tornado chart can help locate influential variables, but it should lead to a question and response rather than be treated as a finished explanation of causation or a ranking of which mitigation will deliver the greatest value.

Construct a joint model of delivery and value

The next teaching example examines a separate future wave rather than changing the programme's historical baseline. Its purpose is to show how one uncertainty can affect remaining cost, duration and benefits together. All monetary values are in millions of US dollars; durations are calendar days. The model uses an aggregate wave duration so that the relationship is easy to inspect. A live quantitative schedule-risk analysis would need the activity logic, resources, calendars and acceptance conditions developed in Chapter 7. The U.S. Government Accountability Office's schedule guide makes a credible schedule the foundation for risk analysis and examines the effects of risk drivers and dependence (GAO, 2015, Best Practice 8). Sampling a duration distribution cannot repair missing activities or an impossible resource plan. The simplified model therefore teaches integration, while leaving the detailed delivery network as a separate prerequisite for application.

Table 10.2 specifies the inputs. Base cost follows a triangular distribution with minimum four, mode five and maximum seven. Base duration uses eighty, one hundred and one hundred and forty days. Base adoption uses 0.45, 0.75 and 0.90 of the relevant operating population. These three base variables are assumed independent solely for the example, and all event draws are independent of them. A material interface problem has probability 0.30. If it occurs, it adds two million to cost, thirty-five days to duration and reduces the adoption fraction by 0.10. The base ranges exclude those particular event effects, so adding them does not count the event twice. The resulting adoption fraction remains positive within the stated support. Every input is fictional; the numerical precision of the simulation should not be mistaken for evidence that the manufacturer's real exposure has those distributions or probabilities.

Table 10.2 Specify the aggregate simulation before running it

Input	Assumption	Meaning and boundary
Base cost C0	Triangular 4, 5, 7	US\$ millions; excludes the specified event and mitigation fee.
Base duration D0	Triangular 80, 100, 140	Calendar days; aggregate duration, not a detailed network.
Base adoption A0	Triangular 0.45, 0.75, 0.90	Fraction of the relevant operating population.
Common event J	Probability 0.30	$C = C0 + 2J$; $D = D0 + 35J$; $A = A0 - 0.10J$.
Benefit stream value	12 at full adoption	US\$ millions as of benefit start, scaled by A.
Timing and discount	D + 30 days; 8% a year	Benefit value discounted back; all cost charged at time zero.
Proposed mitigation	Fee 0.35; event probability 0.10	Fee paid whether or not the event occurs; no added delay assumed.

Triangular parameters are minimum, mode and maximum. Base variables are independent; event draws are independent of them. The alternative comparison uses separate independent events for the three effects. All values are teaching assumptions. Mitigation leaves event severity and base distributions unchanged.

Ackermann et al. (2014) demonstrate stakeholder risk elicitation through causal mapping in an energy-project application. Their workshops exposed interacting concerns and feedback beyond an existing register, although participant responses and one application do not establish avoided overruns. Their approach helps explain why Lena should bring operational and delivery knowledge into the model. Maya may identify how an interface problem increases manual exceptions and suppresses use, while Alex explains its rework and testing consequences. The resulting links are propositions to investigate, not proven causal effects simply because they have been drawn as arrows. Figure 10.4 makes the common event explicit. This is preferable to entering three unrelated risks merely because cost, schedule and benefits occupy separate tabs in the reporting system. The modelling boundary should follow the mechanism under examination.

One event can move cost, time and adoption together

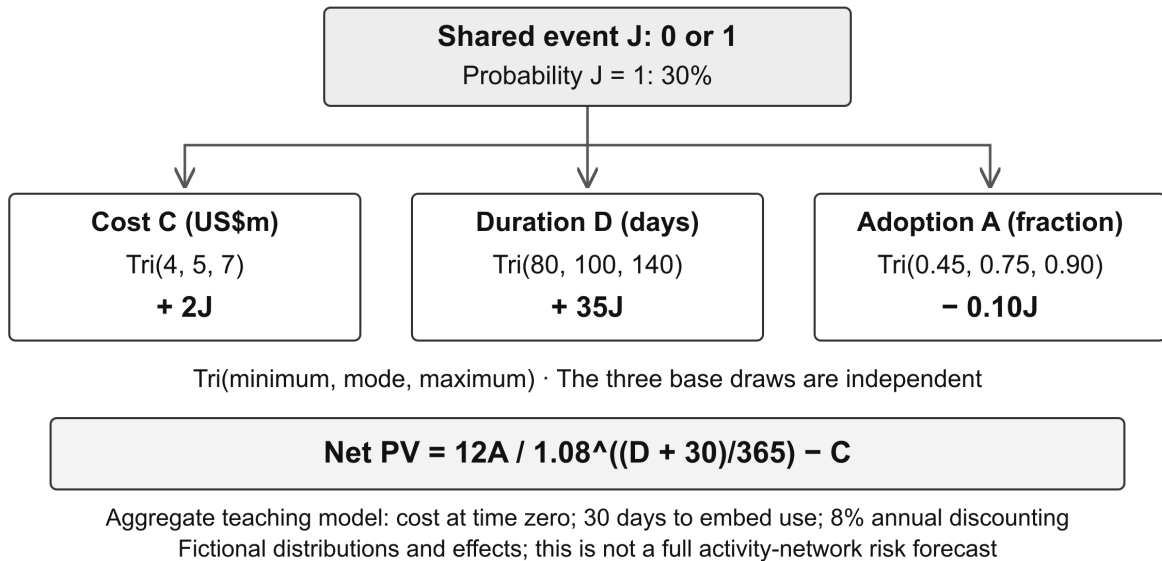


Figure 10.4 Original aggregate wave model. One event can increase cost and duration while reducing adoption. Base variables are independent and event draws are independent of them. All inputs are fictional, costs are charged at the decision date and the full-adoption benefit stream is valued at its start before discounting back. This is not a complete activity-network risk model.

For the benefit calculation, suppose a fully adopted stream of incremental operating benefits has a present value of twelve million at the date those benefits begin. Adoption scales that value, and benefits begin thirty days after delivery completion to allow embedding. An annual effective discount rate of eight per cent converts that value back to the decision date. If A is adoption and D is duration in calendar days, benefit value at the decision date is $12A$ divided by 1.08 raised to the power $(D + 30)/365$. Net value is this benefit value minus cost C . For teaching, all cost is charged at the decision date; a live appraisal would represent its actual timing and relevant cash-flow conventions. The benefit stream is assumed to be legitimate incremental value, already separated from capacity valuation and double counting using the earlier chapters' methods.

Each simulation trial samples the base variables and the event, calculates C , D and A , then evaluates the discounted benefit and net value. The worked results use 200,000 trials with a retained random seed of 202610, NumPy's default PCG64 generator and linear interpolation for simulation quantiles. The retained input definitions and calculation record allow reproduction. More trials reduce sampling noise in the simulation; they do not improve an unsupported adoption assumption, identify an omitted interface or make a triangular upper limit physically credible. Lena should inspect intermediate values, extreme combinations and selected individual trials before trusting a summary. A scenario should still describe an intelligible operating situation when read as a story. If it simultaneously assumes a prolonged interface failure and unaffected adoption despite the stated mechanism, the model's structure requires attention before its percentiles are used.

Understand dependence and joint confidence

The first comparison deliberately changes only the way the event effects occur. In the independent-events model, cost, delay and adoption effects each have probability 0.30 but occur independently. In the common-event model, one occurrence produces all three effects together. The marginal distributions of cost, duration and adoption are the same in theory across these two constructions, apart from simulation sampling variation in the reported estimates. Their joint behaviour is different. The question is not which dependence assumption produces the more convenient result, but which better represents the mechanism. If one interface problem drives all three effects, independent sampling creates too many mixed states in which the associated consequences separate. Other projects may contain partial dependence, several common causes or compensating effects; those structures require their own explanation rather than a default correlation chosen without reference to the work.

Table 10.3 compares the results. Mean net value is approximately 1.86 million in both unmitigated models, yet the probability of negative net value rises from approximately 13.7 per cent under independent events to 19.1 per cent under the common event. The tenth percentile of net value moves from about minus 0.35 million to minus 0.92 million. The marginal cost and duration percentiles remain similar, so a dashboard displaying those two figures alone would hide much of the change in the value downside. The means need not be exactly equal because adoption and delay interact in the discounted-benefit expression. The practical lesson is to examine the distribution of the decision-relevant consequence itself. Combining favourable-looking marginal summaries cannot establish how well the transformation protects value when several adverse consequences arrive together.

Table 10.3 Dependence and response change the value exposure

Measure	Independent events	Common event	Mitigated common event
Mean net value US\$m	1.86	1.86	2.15
Net value P10 US\$m	-0.35	-0.92	0.04
Negative net value	13.7%	19.1%	9.7%
Cost P80 US\$m	7.00	7.00	6.53
Duration P80 days	135.6	135.6	123.2
Both common P80 limits met	64.0%	73.0%	88.0%

200,000 trials, seed 202610. Joint results use the same unrounded common-model marginal P80 limits, shown here to four decimal places: cost US\$7.0011 million and duration 135.6330 days. They are not each model's own pair of limits. The mitigation fee is included in its cost. Results are rounded conditional simulation estimates.

The common-event model's marginal eightieth-percentile cost is approximately 7.00 million, and its marginal eightieth-percentile duration is approximately 135.6 days. Only about 73.0 per cent of trials satisfy both limits simultaneously. Under the independent-events construction,

approximately 64.0 per cent meet those same limits, close to the familiar product of two independent eighty per cent probabilities. Figure 10.5 plots a reproducible subset of the common-event trials and locates the two limits. Positive dependence has increased the chance of both favourable outcomes occurring together relative to independence, while also concentrating adverse outcomes. It would therefore be wrong to claim that dependence must reduce every joint success probability. The direction depends on the event being measured. Neither pair of marginal P80 statements establishes an eighty per cent probability of meeting both commitments, let alone achieving every intended benefit.

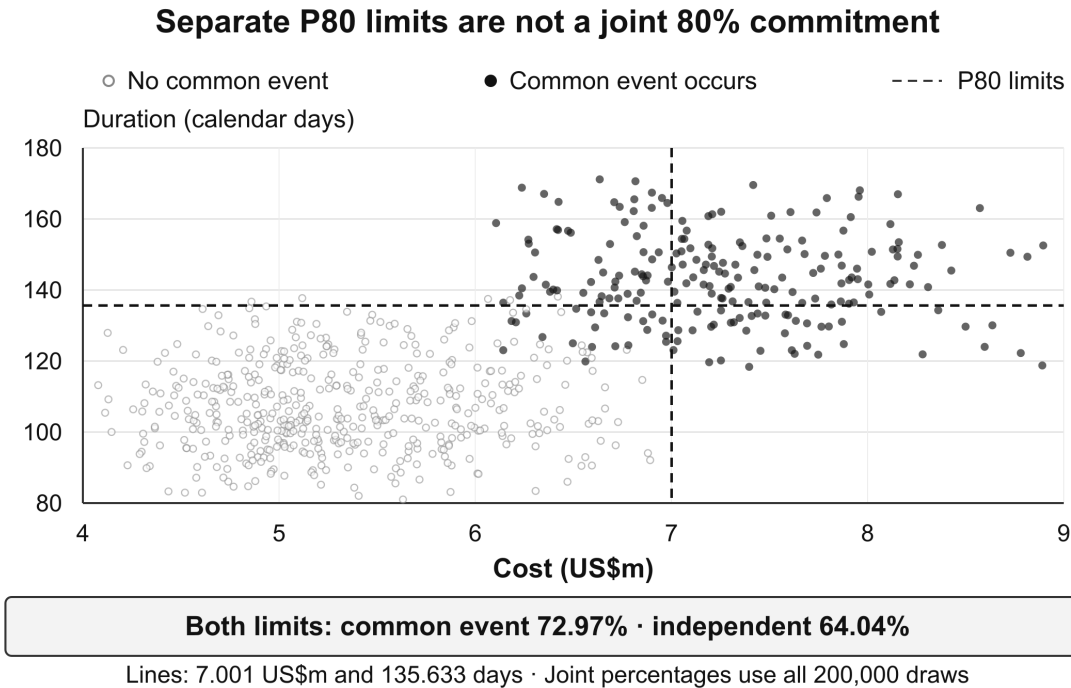


Figure 10.5 Seven hundred illustrative points from the 200,000-trial common-event model. The lines mark its marginal cost and duration P80 values. Approximately 73.0 per cent of all trials meet both limits; the independent-events construction gives about 64.0 per cent. Percentages use all trials, not just the plotted subset.

The United States' Orion review provides a dated warning about the meaning of a stated confidence level. GAO's July 2016 report examined a commitment baseline set at seventy per cent joint cost and schedule confidence. It judged the supporting estimates unreliable, finding that seven of twenty cost practices and one of eight applicable schedule practices were met or substantially met. The programme also pursued an earlier internal readiness goal, while reserves were concentrated later; technical difficulties had led to deferred or reordered work, whose aggregate effect on reserves had not been formally analysed since confirmation (GAO, 2016, pp. 21–22, 27–28, 33–35, 38). These findings concern the estimates and arrangements examined at that time. They do not convert the reported model percentage into a calibrated probability of mission success or imply that every management information source was unreliable.

The controller should consequently write probability statements as complete propositions. A result concerns a defined scope, model, information date, response policy and event, such as remaining cost below a stated amount and operational acceptance before a stated date. It should identify excluded uncertainties and the knowledge behind its inputs. The selected percentile also

needs a decision purpose: a forecasting estimate, a funding provision and a contractual commitment may require different consideration of consequences and authority. No arithmetic rule makes P80 universally appropriate. A sponsor may choose a more protective provision because failure would be severe, but the model's reliability still needs examination. Conversely, adopting a higher percentile from an incomplete model can create false comfort while leaving an unrepresented operational dependency untouched. The percentile is an output to interpret, not a substitute for the argument developed in Chapter 9.

Fund responses and retain a credible residual forecast

The risk model should distinguish current exposure from the result of an implementable response. Suppose an additional interface-validation activity costs 0.35 million at the decision date and is assumed to reduce the common-event probability from 0.30 to 0.10, without changing its conditional severity or the base distributions. Re-running the model with that response included increases mean net value from approximately 1.86 to 2.15 million and reduces the probability of negative net value from 19.1 to 9.7 per cent. Figure 10.6 shows the value distributions. These are conditional results, not evidence that the validation activity can deliver the assumed reduction. Lena should ask who will perform it, whether the required environment and skills exist, when it must finish and what observations would establish its effectiveness before the release depends on it.

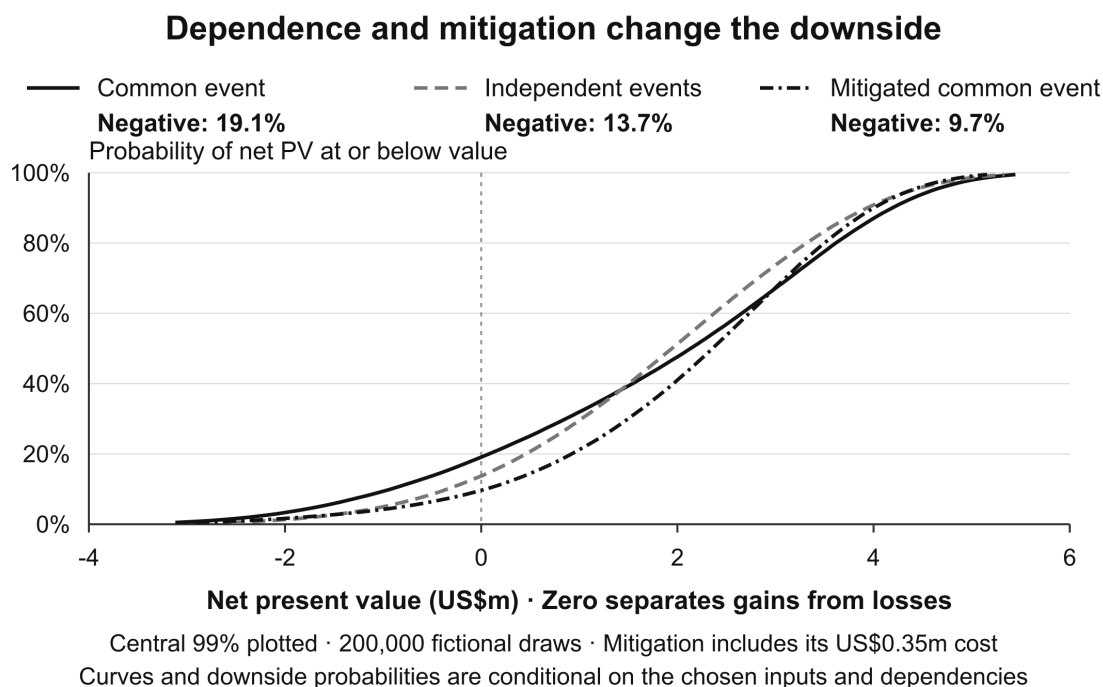


Figure 10.6 Original net-value cumulative distributions. The common event increases the modelled chance of negative value from 13.7 to 19.1 per cent; the proposed mitigation reduces it to 9.7 per cent under its effectiveness assumption, including its fee. Curves display the 0.5th–99.5th percentile range, not the complete support.

Williams (1999) argues for greater realism in network simulation by considering management responses and their sometimes imperfect or counterintuitive effects. His methodological

treatment supports including an explicit response policy rather than assuming that management will recover every adverse outcome. In the manufacturer's application, additional testing could consume a scarce specialist or delay the wave; those consequences would need inclusion if they were material. The present example assumes no extra delay and unchanged severity, which makes those assumptions visible for challenge. A useful refinement is to test the response's required effectiveness. The calculation record identifies the remaining event probability at which its expected benefit just covers the 0.35-million cost. If the team cannot defend a reduction beyond that threshold, the expected-value argument for the response weakens, even though other service or risk criteria might still justify it.

Contingency is a funding decision related to uncertainty, and it needs a clear reference. In the common-event model, the cost P80 of approximately seven million is about two million above a five-million point reference. That difference includes uncertainty already represented in the base cost distribution as well as the discrete event. It should not be labelled an additional allowance solely for the interface event and then added again to the risk-adjusted total. GAO's cost guide connects contingency to the difference between a specified estimate and a selected percentile, while emphasising dependence and the need to consider the estimate's basis (GAO, 2020, pp. 156–160). The controller should reconcile the base estimate, included risk allowances, response expenditure and funding provision. Moving money from a reserve into a work package changes the account; it does not by itself reduce the underlying exposure.

Mak and Picken (2000) compare contingency estimation in public-works projects using structured risk analysis with projects using other arrangements. Their observational comparison reported less unnecessary or exaggerated provision in the risk-analysis group, while also identifying scope for further refinement. It does not establish an optimal contingency percentage or prove that lower reserves always mean better control. The useful implication is to challenge both underfunding and unexplained overprovision. Excessive allowance can conceal weak scope definition, discourage investigation or crowd out other valuable work. Insufficient allowance can force reactive decisions and disrupt benefits. Lena should show the sponsor how the provision relates to exposure, when cash may be required and who can authorise its use. A reserve that arrives after the affected work needs it may offer little practical protection, even when the total funding appears adequate.

Australia's Inland Rail review connects this funding question to the maturity of the commitment. Kerry Schott's January 2023 report, released in April, found that approximately ninety per cent of the scope had low or moderate cost certainty. Estimates across sections used differing degrees of conservatism; the review recommended further independent cost and value-engineering work coordinated with scope definition. It also recommended examining staged completion, including an early operating section, while improving approval and cost certainty before later delivery (Schott, 2023, pp. 52, 56). The ninety per cent figure describes the share of scope assessed as having low or moderate cost certainty, not the probability of an overrun. These were recommendations at the review date, not verified implementation or realised benefits. The controls lesson is that some uncertainty calls for a better-defined or staged commitment, rather than a larger percentage added to an immature estimate.

Challenge the model with comparable experience

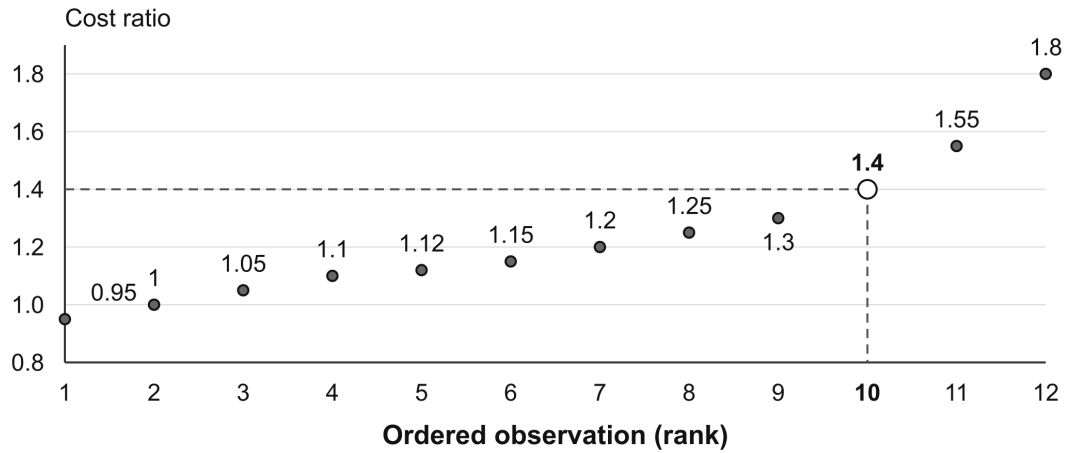
A detailed estimate can be internally coherent and still differ systematically from what comparable projects have achieved. Reference-class evidence offers an outside view, provided the comparison is meaningful. Lena should define the type of work, estimating stage, scope boundary, price basis and completion measure before selecting historical records. A forecast at initial approval cannot be compared casually with another made after most uncertainty has resolved. The same applies to benefits: first-year adoption, eventual steady-state use and discounted lifetime value are different outcomes. Missing, cancelled and unfinished projects also affect interpretation. A database containing only completed successes can make future performance appear more dependable than the relevant experience warrants. The comparison should expose these limitations rather than use a large count of records as a substitute for a defensible reference class.

Batselier and Vanhoucke (2016) compare forecasting approaches for a focal office-finishing project using nested reference classes. Reference-class forecasting performed well in their reported comparison, but the authors do not establish universal superiority. The broadest classes contained different numbers of schedule and cost histories, while the closest class comprised four earlier projects from the same company. The useful lesson is to examine comparability and the forecast's information date. Lena can ask whether a similar delivery method, operating environment or organisational capability makes a narrower class more informative, while recognising the greater uncertainty of a small sample. A broad class may reveal exposure excluded from the team's preferred analogues; a narrow class may better represent relevant conditions. The choice should be documented and sensitivity to alternative reasonable classes should accompany the resulting uplift or forecast.

Figure 10.7 uses twelve fictional ratios of actual final cost to an earlier estimate. They range from 0.95 to 1.80. Using the inverse empirical distribution, also called the nearest-rank convention, the eightieth-percentile observation is at rank equal to the ceiling of 0.80 multiplied by twelve: the tenth value, 1.40. Applied to an equivalent five-million estimate, this produces seven million, an uplift of two million. The convention matters because interpolation would produce a different number from such a small dataset. Ten of the twelve observations lie at or below 1.40, so this step in the empirical distribution covers approximately 83.3 per cent of the records rather than exactly eighty per cent. None of that establishes an eighty per cent chance of covering a new project's cost. The example teaches the calculation while leaving comparability and predictive calibration to be established with appropriate evidence.

A reference-class percentile needs a stated convention

Twelve sorted fictional actual-to-estimate cost ratios



$$\text{ceil}(0.80 \times 12) = 10 \rightarrow \text{ratio } 1.40 \rightarrow \text{US\$5m} \times 1.40 = \text{US\$7m}$$

Synthetic, uncalibrated cohort: this calculation does not establish coverage for a new project

Figure 10.7 Twelve fictional actual-to-estimate ratios. Under the nearest-rank convention, the empirical P80 is the tenth observation, 1.40. Scaling a comparable five-million estimate gives seven million. This small teaching cohort supplies no calibrated probability for a new project and does not justify adding another overlapping contingency.

The seven-million result happens to resemble the simulation's cost P80, but agreement between two fictional calculations is not validation. In practice, Lena should reconcile the mechanisms and boundaries behind an outside-view adjustment and the detailed risk estimate. Both may already include the consequences of scope development, productivity uncertainty and the same interface problem. Adding their allowances without examination could count the exposure twice. Conversely, the reference class might include market or governance failures absent from the detailed model. The right response is to investigate and explain the difference, not automatically select the smaller estimate or average two incompatible quantities. The forecast, funding provision and original commitment should remain separately visible. A changed outside view can justify revising the current forecast without rewriting the record of what was originally authorised.

Flyvbjerg and Bester (2021) analyse historical cost and benefit forecast errors across a large public-investment dataset. Their findings support scrutinising both sides of an appraisal, while the use of capital costs, mainly first-year benefit proxies and a much smaller set of paired observations limits what can be inferred about complete lifetime value. The controller should not apply a pooled uplift mechanically or infer a project-level relationship from separate averages. For the manufacturer, the outside view should ask whether similar transformations achieved their anticipated use, benefit timing and cash-release mechanisms as well as their delivery estimates. A corrected cost forecast paired with an untouched optimistic benefit forecast can still mislead the continuation decision. Relevant historical errors should inform assumptions and challenge, with the original measures and reasons for comparability retained so that future evidence can improve the reference class.

Choose a robust course when probabilities are contested

Some decisions involve more than uncertainty about parameters in an agreed model. Walker, Lempert and Kwakkel (2013) describe deep uncertainty as involving inadequate knowledge or agreement about models, probability distributions or the valuation of outcomes. This conceptual distinction is useful when, for example, the manufacturer faces competing explanations of future supplier behaviour or a regulatory change whose consequences are not adequately represented by the existing adoption model. Giving every scenario an equal probability because no better estimate is available does not make the uncertainty objectively quantified. Lena can instead retain competing models and plausible futures, showing how options perform under each. Probability-based analysis remains useful where its assumptions are defensible. The purpose is to match the representation to the knowledge available, rather than force every uncertainty into the same simulation merely because the software requires numbers.

One transparent comparison is regret: the value forgone by choosing an option instead of the best available option in a particular state. Using the two-state investment example without assigning probabilities, full commitment has regret of eight million under weak adoption, because staging would produce two rather than minus six. Under strong adoption its regret is zero. Staging has no regret under weak adoption and regret of five under strong adoption, because commitment would produce twelve rather than seven. Table 10.4 shows the comparison. A minimax-regret criterion chooses staging because its worst regret is five, below commitment's eight. This differs from maximising the worst absolute payoff, even though staging also wins that criterion here. The result is conditional on the alternatives, states and criterion considered. It is a clear decision rule, not a claim that the two states exhaust every future or that minimax regret suits every sponsor.

Table 10.4 Compare regret without assigning state probabilities

Approach	Weak-state regret	Strong-state regret	Maximum regret
Full commitment	8.00	0.00	8.00
Staged approach	0.00	5.00	5.00

US\$ millions. Regret is the best available value in the state minus the value of the chosen approach. Minimax regret selects staging in this limited example. This criterion does not establish the states' likelihoods or prove that the two-state set is complete.

Lempert et al. (2006) develop an iterative method for evaluating strategies across plausible futures, identifying vulnerabilities and constructing useful scenarios. Their computational demonstration is not a dataset of successful projects or evidence that robust methods always outperform probabilistic appraisal. Its value for controls is the reversal of emphasis: ask under what conditions a proposed strategy fails its requirements, then examine whether those conditions are plausible and whether a feasible response exists. Lena might test combinations of adoption, support capacity, supplier responsiveness and benefit value, retaining the combinations under which the staged rollout becomes unacceptable. The resulting account can guide further investigation or redesign. The proportion of tested scenarios meeting a criterion is a

description of that scenario set, not automatically the real-world probability of success. Its meaning depends on how the scenarios were selected and, where relevant, weighted.

Robustness also needs an explicit performance requirement. An option that merely avoids a financial loss may fail a service objective, while one that protects all service measures may exceed the funding available. Lena should expose these trade-offs rather than declare an option robust without saying to what. Stress testing can identify a region in which the preferred strategy performs adequately and the conditions outside that region requiring a change. The analysis should retain residual vulnerabilities and scenarios deliberately excluded, including consequences that cannot be reduced to the same monetary measure. Where stakeholders disagree about the criterion, the sponsor needs a governance conversation as well as technical work. The controller can show how the preferred option changes with reasonable criteria and identify the choices that remain defensible across several perspectives without pretending that this resolves every conflict.

The practical output is a conditional policy rather than a single prediction. The organisation might begin a bounded wave, expand only after observed capability and adoption reach specified conditions, and retain a prepared alternative if those conditions fail. Such a policy can be evaluated against the same futures as an immediate full commitment, including the costs and delays of monitoring and adaptation. It should not be credited with perfect foresight: decision makers observe noisy signals, recognise problems late and require time to act. A policy that performs well only when it switches instantly at the exact moment a forecast becomes wrong is unlikely to be useful in the field. Lena should model the information and response actually available, then show where the policy remains exposed. This connects robust comparison to the practical design of flexibility.

Make flexibility an option the organisation can exercise

De Neufville (2003) describes real options in systems planning and design as opportunities to change a system as uncertainty resolves. His conceptual engineering account helps distinguish practical flexibility from a general hope that the team will adapt. For the manufacturer, an option might require keeping an interface compatible with the existing service, preserving supplier capacity for a later wave or designing modules that can be expanded separately. Each provision has a price, an exercise cost, a useful life and conditions under which it can be used. These should enter the comparison. More uncertainty does not automatically make the entire project more valuable: preserving flexibility can help while the underlying investment becomes less attractive. The controller should identify the specific decision that the provision keeps available and the evidence that would justify exercising it.

Florice and Miller (2001), drawing on large engineering-project cases, distinguish arrangements for anticipated risks from the capacity to govern a response to disruptive change. Their case-based theorising does not establish that one governance structure guarantees success, but it highlights what a reserve alone cannot provide. A different operating model may need new authority, supplier agreement, staff capacity or revised benefit ownership. Lena should therefore test the organisational feasibility of each option alongside its technical design. A fallback

documented in a plan is weak protection if the supplier has withdrawn support or the staff who know the process have left. Preserving an option may require current expenditure and explicit agreements even when the future action is never taken. The decision record should show who maintains that capability and when its continued cost or usefulness will be reconsidered.

Haasnoot et al. (2013) develop dynamic adaptive policy pathways by connecting near-term actions, future alternatives, monitoring and the points at which an action ceases to meet objectives. Their water-management illustration is a model-based method demonstration rather than evidence of realised long-term performance. Its useful translation to transformation is to distinguish the operating limit from the earlier decision needed to avoid crossing it. The manufacturer can begin with one unit and define conditions for expansion, additional support, redesign or withdrawal. Figure 10.8 presents an original application. It does not reproduce the published pathway diagram or create a separate programme lifecycle. It adds an explicit view of possible future commitments to the existing control method, with each route dependent on evidence, preparation and the authority to change course while the alternative remains feasible.

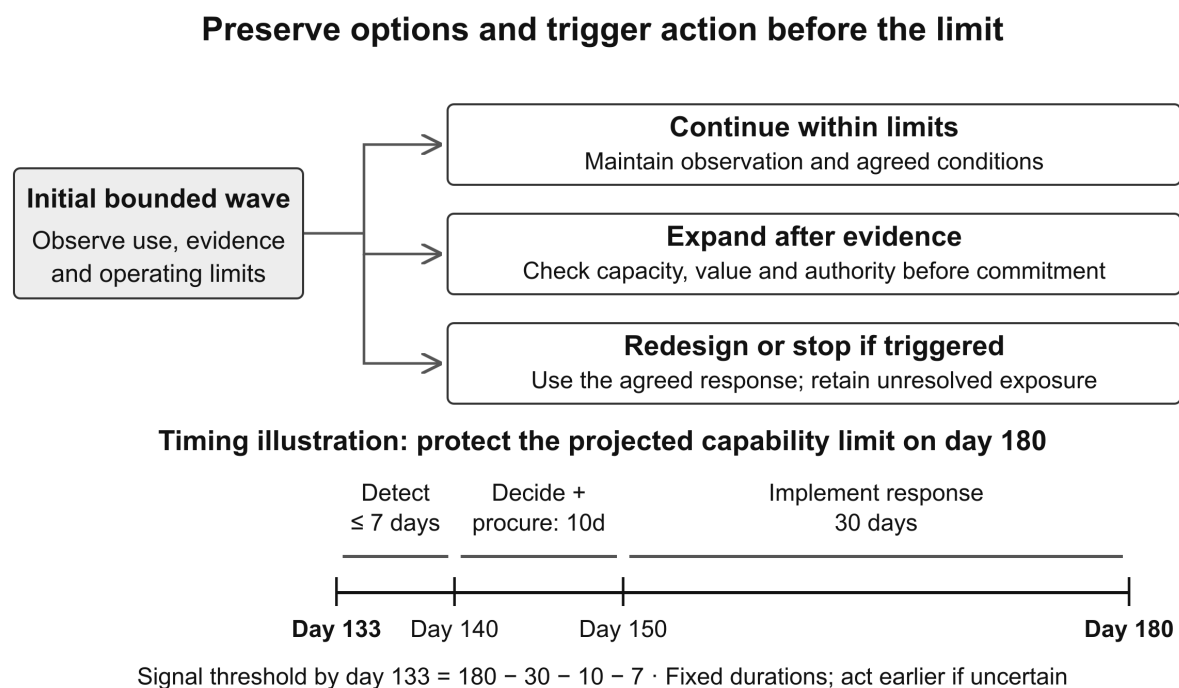


Figure 10.8 Original adaptive commitment and timing illustration. With a projected capability limit at day 180, up to seven days to detect a signal, ten for decision and procurement and thirty for implementation, the signal threshold must be reached by day 133. Fixed times and no further buffer are assumed; uncertainty may require earlier action.

The Environment Agency's April 2023 Thames Estuary 2100 update provides an observable example of adaptive planning. The published approach monitors indicators through five-year reviews and updates the plan every ten years. The update advanced the upstream defence-upgrade deadline from 2065 to 2050 and the deadline for choosing an end-of-century option from 2050 to 2040. An economic review found the options closer in costs and benefits, so all were retained for reassessment by 2030. Under the projections then used, the selected option was expected to operate by 2070, allowing time for planning, approval and construction (Defra and Environment Agency, 2023a; 2023b). These documents show a revised decision framework

that preserves alternatives. They do not establish that a final barrier option was selected in 2023 or that a future intervention's benefits have already been realised.

Kwakkel, Haasnoot and Walker (2016) compare robust decision-making and adaptive pathways in a model-based water-management case. Their analysis shows how examining vulnerabilities and trade-offs can complement attention to the sequence and timing of adaptation, without establishing universal superiority for either approach. Lena can combine those questions in a practical review: under which conditions does the preferred wave strategy fail, what could be observed before failure, and which response remains feasible at that time? This is more demanding than attaching a generic review date to the plan. It requires a monitorable sign, a decision criterion and a prepared action. Where the signal is too late or unreliable, the organisation may need an earlier commitment, a larger operating margin or a different design. Flexibility is valuable only when the pathway from observation to response can actually be completed.

Set triggers early enough to change the outcome

A trigger should identify a condition requiring attention or action before the relevant limit is breached. In a separate fictional timing example, suppose a capacity assessment indicates that the current operating arrangement will cease to meet requirements at day 180. Implementing the prepared response takes thirty days, and decision and procurement take ten more. Reviews occur every seven days, so a condition that develops just after a review might take up to seven days to be detected. The latest underlying trigger point is therefore day 180 minus thirty minus ten minus seven, or day 133. If the condition is reached then, detection by day 140 leaves forty days for decision and implementation. This is a deterministic illustration with no additional buffer; uncertainty in the forecast limit, decision latency or response duration would justify an earlier trigger or further protection.

The seven-day allowance is a detection lag, not time to subtract again after the condition has already been observed. If the review at day 140 confirms that action is necessary, the model has forty days left for the stated decision and implementation tasks. The record should distinguish the condition emerging, its observation, the authorised decision and the completed response. In real operation, a forecast of the capacity limit may move as demand changes, and the trigger needs recalculation or a margin that anticipates that movement. Lena should also consider false alarms and noisy measurements. A rule requiring several consecutive observations can reduce reactions to noise, but it consumes response time. The controller should test that trade-off explicitly rather than add confirmation steps after the trigger has been set and assume the original lead-time protection remains intact.

Escalation needs a recipient with authority and an action that can follow. A threshold saying that adoption below a target is red is incomplete if no one has agreed whether red means more support, a smaller wave, a revised benefits forecast or suspension of expansion. The trigger should identify the decision to reopen, the evidence required, the owner who assembles it and the person who can authorise the response. Some conditions require immediate protective action under delegated authority; others require a sponsor review because they change scope, funding

or intended benefits. The governance arrangements from Chapters 3 and 4 provide that distinction. Monitoring should retain the original condition and subsequent decision so that later reviewers can determine whether the organisation acted on the evidence or merely recorded repeated warnings without changing its commitment.

Verification then asks whether the selected response altered the exposure or outcome as expected. If the interface-validation activity was justified by reducing event probability, completing its tasks is evidence of implementation; performance in relevant tests and later use bears on effectiveness. If a smaller wave was chosen to preserve capacity, the review should examine workload and service performance rather than infer protection from wave size alone. Lena should update the current forecast when evidence changes while preserving the original decision model and assumptions. This allows the organisation to learn whether its thresholds were useful, whether response time was realistic and whether the expected benefits arrived. A trigger system should itself improve through experience. Its purpose is to make action timely and accountable, not to create an unchanging collection of limits whose continued relevance nobody examines.

Write the decision so that it can be reviewed

The practical output of this chapter is a Decision under Uncertainty Memorandum, linked to the earlier Decision Connection Record and assurance evidence. It states the choice, deadline and authority; describes feasible alternatives; identifies consequences and constraints; distinguishes estimates from assumptions and contested models; and records the reason for the recommendation. Table 10.5 gives illustrative entries for the manufacturer's separate teaching examples. Detailed calculations remain linked supporting material. The memorandum should make the response conditions visible enough that a future controller can understand why the commitment was reasonable on the evidence then available and what would require it to change. A favourable eventual outcome does not by itself prove the original reasoning was sound, just as an adverse outcome does not automatically make a properly considered decision irrational. Evaluation should examine both the decision process and what subsequently happened.

Table 10.5 Illustrative entries in a Decision under Uncertainty Memorandum

Choice and evidence	Conditional recommendation	Authority and review condition
Buy the adoption study: net strategy value 5.80 versus 5.00 without it.	Proceed only if signal quality, timing and available responses are credible.	Samira authorises the bounded study; the result informs the subsequent commitment.
Fund interface validation: simulated net gain approximately 0.30 after its fee.	Substantiate effectiveness; residual event probability below about 19.12% gives positive expected gain under these assumptions.	Alex supplies a feasible response; Maya confirms operating implications; Samira approves expenditure.
Choose funding protection: common-model cost P80 about 7.00 versus a point reference of 5.00.	Reconcile included uncertainty and existing allowances; do not add overlapping reserves.	Authorised funding owner agrees amount, timing and release rules; Lena retains the forecast basis.
Prepare a capacity response: projected limit day 180 and forty days for decision and action.	With up to seven days detection lag, the underlying trigger must occur by day 133.	Named operating owner monitors; decision authority acts in time; revise for uncertain lead times.

Original teaching entries, not approvals. A live memorandum also records scope, decision date, feasible alternatives, constraints, evidence versions, uncertainty, preferences, dissent, residual exposure, response owner, funding and later verification. The examples have separate financial boundaries.

Consider the sponsor's final exercise. The two-state choice still gives staging an expected value of five million and full commitment 4.80 at the assessed sixty per cent probability of strong adoption. The proposed study has the stated signal characteristics, costs 0.50 million and creates 0.30 million of additional delay consequences. Separately, the future-wave model identifies an interface response costing 0.35 million, but the team has not yet substantiated the claimed reduction in event probability to ten per cent. Samira is asked to fund both actions and authorise full rollout immediately. The learner should determine which parts of that proposal the current analysis supports, which assumptions require evidence and whether immediate rollout would remove the ability to use the study. The examples share a decision discipline, but their financial values concern different boundaries and must not be added into one invented programme benefit.

Under the stated model, buying the study and following its result has expected net value 5.80 million, above staging now. That supports considering the study if its accuracy, timing and the continued availability of the response options are credible. It does not support committing fully before the signal, which would remove the decision change that generates the study's value. The interface response is promising in the simulation, with an expected improvement of approximately 0.30 million after its cost, but that conclusion depends on effectiveness still requiring substantiation. Its financial break-even residual event probability is approximately 19.12 per cent under the other assumptions. Lena can recommend a bounded investigation of the response and retain the unmitigated exposure until a justified commitment and evidence support the revised forecast. Non-financial requirements and affordability remain part of the sponsor's decision alongside expected value.

The chapter completes the handbook's foundations and control-system disciplines by connecting uncertainty to purposeful action. Lena's task is to make the choice explicit, preserve the mechanisms linking delivery to benefits, examine what would change the recommendation and give Samira a commitment that can be governed after approval. A forecast describes possible consequences; a decision selects an action with responsibilities and conditions. Quantitative methods strengthen that selection when their assumptions are visible and their outputs answer the right question. Where knowledge is weaker, structured scenarios, bounded trials and prepared options can keep the organisation responsive without pretending to predict every disruption. The next chapters apply these disciplines through the transformation lifecycle, carrying forward the same obligation to connect evidence, authority, value and subsequent learning.

References

- Ackermann, F., Howick, S., Quigley, J., Walls, L., & Houghton, T. (2014). Systemic risk elicitation: Using causal maps to engage stakeholders and build a comprehensive view of risks. *European Journal of Operational Research*, 238(1), 290–299. doi: [10.1016/j.ejor.2014.03.035](https://doi.org/10.1016/j.ejor.2014.03.035)
- Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1), 1–13. doi: [10.1016/j.ejor.2015.12.023](https://doi.org/10.1016/j.ejor.2015.12.023)
- Batselier, J., & Vanhoucke, M. (2016). Practical application and empirical evaluation of reference class forecasting for project management. *Project Management Journal*, 47(5), 36–51. doi: [10.1177/875697281604700504](https://doi.org/10.1177/875697281604700504)
- British Columbia Utilities Commission (BCUC). (2017, 1 November). British Columbia Utilities Commission Inquiry Respecting Site C: Final Report to the Government of British Columbia. www.ordersdecisions.bcuc.com/bcuc/decisions/en/236681/1/document.do
- Chapman, C., & Ward, S. (2000). Estimation and evaluation of uncertainty: A minimalist first pass approach. *International Journal of Project Management*, 18(6), 369–383. doi: [10.1016/S0263-7863\(00\)00016-8](https://doi.org/10.1016/S0263-7863(00)00016-8)
- de Neufville, R. (2003). Real options: Dealing with uncertainty in systems planning and design. *Integrated Assessment*, 4(1), 26–34. doi: [10.1076/iaij.4.1.26.16461](https://doi.org/10.1076/iaij.4.1.26.16461)
- Department for Environment, Food & Rural Affairs (Defra), and Environment Agency. (2023a, 19 April). Major updates to Thames Estuary 2100 from 2012 to 2023. www.gov.uk/guidance/major-updates-to-thames-estuary-2100-from-2012-to-2023
- Department for Environment, Food & Rural Affairs (Defra), and Environment Agency. (2023b, 19 April). Taking an adaptive approach: Thames Estuary 2100. www.gov.uk/guidance/taking-an-adaptive-approach-thames-estuary-2100
- Florice, S., & Miller, R. (2001). Strategizing for anticipated risks and turbulence in large-scale engineering projects. *International Journal of Project Management*, 19(8), 445–455. doi: [10.1016/S0263-7863\(01\)00047-3](https://doi.org/10.1016/S0263-7863(01)00047-3)
- Flyvbjerg, B., & Bester, D. W. (2021). The cost-benefit fallacy: Why cost-benefit analysis is broken and how to fix it. *Journal of Benefit-Cost Analysis*, 12(3), 395–419. doi: [10.1017/bca.2021.9](https://doi.org/10.1017/bca.2021.9)
- Haasnoot, M., Kwakkel, J. H., Walker, W. E., & ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, 23(2), 485–498. doi: [10.1016/j.gloenvcha.2012.12.006](https://doi.org/10.1016/j.gloenvcha.2012.12.006)
- HM Treasury. (2026). The Green Book. www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2026

- Howard, R. A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1), 22–26. doi: [10.1109/TSSC.1966.300074](https://doi.org/10.1109/TSSC.1966.300074)
- Howard, R. A. (1988). Decision analysis: Practice and promise. *Management Science*, 34(6), 679–695. doi: [10.1287/mnsc.34.6.679](https://doi.org/10.1287/mnsc.34.6.679)
- Kwakkel, J. H., Haasnoot, M., & Walker, W. E. (2016). Comparing Robust Decision-Making and Dynamic Adaptive Policy Pathways for model-based decision support under deep uncertainty. *Environmental Modelling & Software*, 86, 168–183. doi: [10.1016/j.envsoft.2016.09.017](https://doi.org/10.1016/j.envsoft.2016.09.017)
- Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, 52(4), 514–528. doi: [10.1287/mnsc.1050.0472](https://doi.org/10.1287/mnsc.1050.0472)
- Loch, C. H., Solt, M. E., & Bailey, E. M. (2008). Diagnosing unforeseeable uncertainty in a new venture. *Journal of Product Innovation Management*, 25(1), 28–46. doi: [10.1111/j.1540-5885.2007.00281.x](https://doi.org/10.1111/j.1540-5885.2007.00281.x)
- Mak, S., & Picken, D. (2000). Using risk analysis to determine construction project contingencies. *Journal of Construction Engineering and Management*, 126(2), 130–136. doi: [10.1061/\(ASCE\)0733-9364\(2000\)126:2\(130\)](https://doi.org/10.1061/(ASCE)0733-9364(2000)126:2(130))
- March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science*, 33(11), 1404–1418. doi: [10.1287/mnsc.33.11.1404](https://doi.org/10.1287/mnsc.33.11.1404)
- Pich, M. T., Loch, C. H., & De Meyer, A. (2002). On uncertainty, ambiguity, and complexity in project management. *Management Science*, 48(8), 1008–1023. doi: [10.1287/mnsc.48.8.1008.163](https://doi.org/10.1287/mnsc.48.8.1008.163)
- Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12), 1508–1517. doi: [10.1016/j.envsoft.2010.04.012](https://doi.org/10.1016/j.envsoft.2010.04.012)
- Schott, K. (2023, January). The Delivery of Inland Rail: An Independent Review. Commonwealth of Australia, Department of Infrastructure, Transport, Regional Development, Communications and the Arts. www.infrastructure.gov.au/sites/default/files/documents/independent-review-of-inland-rail-report.pdf
- Strong, M., Oakley, J. E., Brennan, A., & Breeze, P. (2015). Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: A fast, nonparametric regression-based method. *Medical Decision Making*, 35(5), 570–583. doi: [10.1177/0272989X15575286](https://doi.org/10.1177/0272989X15575286)
- U.S. Government Accountability Office (GAO). (2015). Schedule Assessment Guide: Best Practices for Project Schedules (GAO-16-89G). www.gao.gov/products/gao-16-89g
- U.S. Government Accountability Office (GAO). (2016, 27 July). Orion Multi-Purpose Crew Vehicle: Action Needed to Improve Visibility into Cost, Schedule, and Capacity to Resolve Technical Challenges (GAO-16-620). www.gao.gov/products/gao-16-620
- U.S. Government Accountability Office (GAO). (2020). Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Program Costs (GAO-20-195G). www.gao.gov/products/gao-20-195g
- Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep uncertainty. In S. I. Gass & M. C. Fu (Eds.), *Encyclopedia of Operations Research and Management Science* (3rd ed., pp. 395–402). Springer. doi: [10.1007/978-1-4419-1153-7_1140](https://doi.org/10.1007/978-1-4419-1153-7_1140)
- Williams, T. (1999). Towards realism in network simulation. *Omega*, 27(3), 305–314. doi: [10.1016/S0305-0483\(98\)00062-0](https://doi.org/10.1016/S0305-0483(98)00062-0)

Applying controls through the transformation lifecycle

Applying the control disciplines to each consequential commitment through the transformation lifecycle.



KENT BUSINESS COLLEGE
School of Project Control

From control disciplines to transformation decisions

The manufacturer's next investment decision brings the earlier chapters together. Samira needs to know what she can responsibly authorise; Alex needs a deliverable commitment; Maya needs an operating change that her teams can absorb; and Lena needs evidence that connects those concerns. None can answer the whole question through a single cost report, benefits register or readiness assessment. The proposal must have a purpose, the intended operating state must be understandable, the next delivery wave must be feasible and the receiving organisation must be able to sustain its use. Part III follows these relationships through the transformation lifecycle. Its purpose is to show how the controls disciplines already developed in this handbook become a connected practice as an investment takes shape, produces capability, changes operations and eventually leaves the programme's temporary governance. The central question is what evidence and authority are needed for the next consequential commitment.

The Transformation Lifecycle Model in Figure III.1 provides the route through that practice. It is an original organising model for this handbook, rather than a prescribed organisational structure or an empirically validated maturity scale. Chapters 12 to 16 show the principal progression from an investigable proposal to a designed operating state, an authorised wave, integrated delivery and stable adoption. Chapter 11 establishes the control rhythm that supports the whole route. Chapter 17 supplies continuing evaluation and reforecasting, including the decision to revisit an earlier assumption or commitment. Chapter 18 explains closure and the transfer of responsibilities that must survive it. The arrows express relationships between decisions, not a claim that transformation always advances through cleanly separated phases. Different parts of a programme may occupy different positions at the same time, and evidence from operations may change a proposal while another wave remains in delivery.

The transformation lifecycle

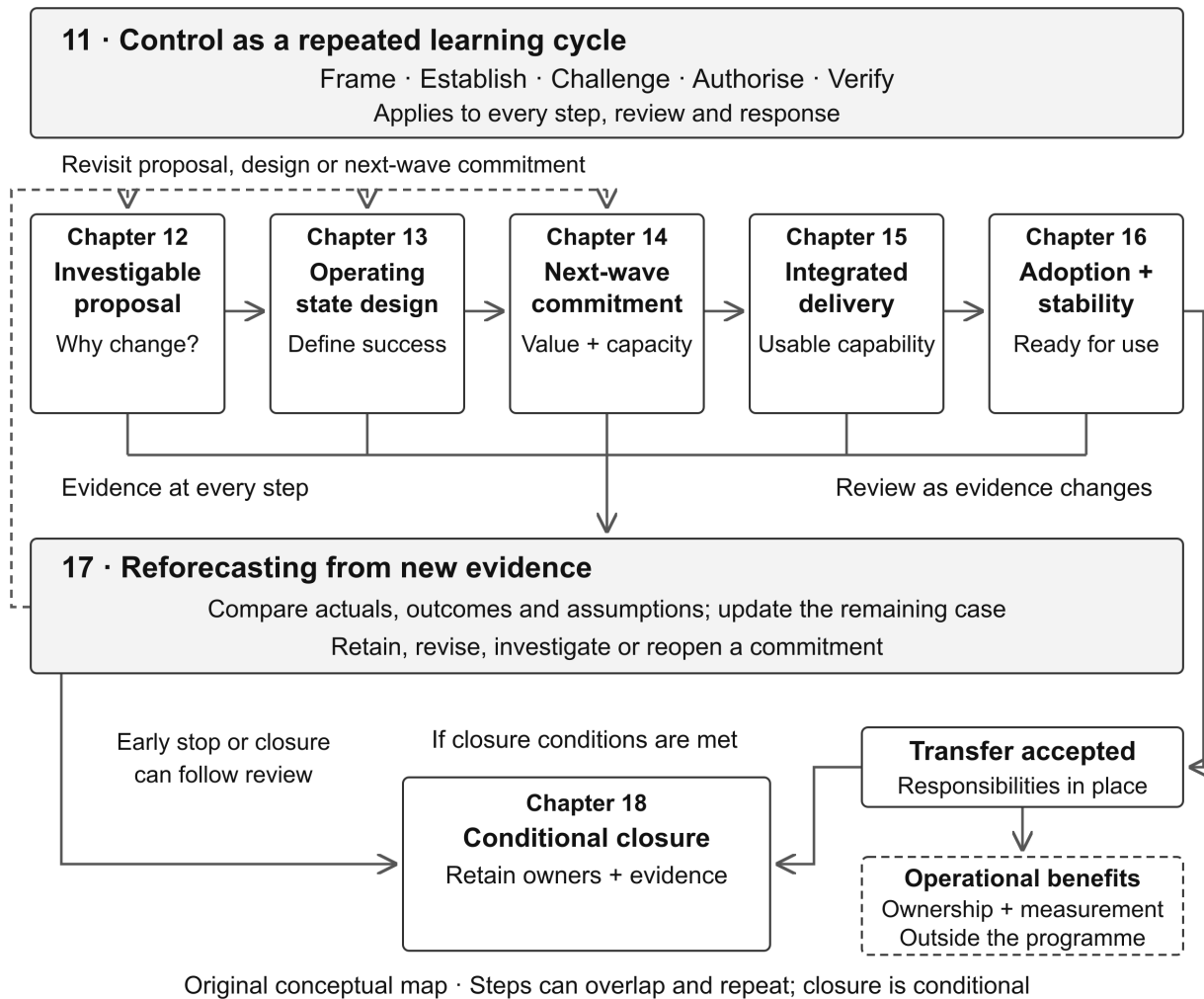


Figure III.1 The Transformation Lifecycle Model locates the decisions developed in Chapters 11–18. Control and evaluation operate throughout; evidence can reopen earlier choices or lead to closure. Activities may overlap across waves. Programme closure transfers continuing responsibilities rather than certifying that every forecast benefit has been achieved. This is an original organising model for the handbook.

The model works within the handbook's existing framework family. The Transformation Control Loop continues to connect value claims, commitments, evidence, decisions and learning. The Credible Control Method continues to organise the work through Frame, Establish, Challenge, Authorise and Verify. These activities recur inside every lifecycle area; they are not additional phases to complete before transformation can start. A controller might frame a readiness decision, establish the available operating evidence, challenge a recovery assumption, obtain a bounded authorisation and verify the resulting performance. That small cycle may take place within a larger investment wave. The lifecycle therefore locates the work, while the method explains how to conduct it. This distinction matters when knowledge is incomplete: Pich, Loch and De Meyer (2002) connect different management approaches to information adequacy and complexity, supporting a choice of response that reflects what remains unknown rather than automatic execution of an unchanged plan.

Reading the lifecycle through its decisions

Chapter 11, *Control as a Repeated Learning Cycle*, begins with the operating rhythm of control. It brings reporting dates, evidence preparation, challenge, decision and subsequent verification into a recurring process that people can actually run. The chapter will examine the difference between producing an update and completing a control cycle, including the point at which a recommendation becomes an authorised action. It will connect routine reviews with event-driven intervention, because a material change cannot always wait for the next scheduled meeting. The practical focus is a proportionate review cycle with clear inputs, accountable decision makers and a traceable route from findings to action. Readers will learn to ask whether the cycle changes commitments when evidence changes, and whether someone checks what happened afterwards. Its position above the lifecycle signifies a discipline applied throughout, including preparation for closure and the transfer of continuing monitoring into operations.

Chapter 12, *From Strategic Ambition to an Investigable Proposal*, concerns the first meaningful commitment: deciding whether an ambition deserves structured investigation. A strategic phrase such as improving supplier experience is insufficient as a control basis until the problem, affected population, intended value and plausible intervention are made explicit. This chapter will develop the connection between strategic intent, the current operating problem and the evidence needed to explore alternative responses. It will distinguish the authority to investigate from the authority to deliver, and it will examine the baseline information that can realistically be established before detailed design exists. The controller's task is to expose assumptions early enough for them to be questioned without pretending that uncertainty has already disappeared. The resulting proposal should identify an accountable sponsor, an investigable question, initial boundaries and the conditions under which further expenditure would or would not be justified.

Chapter 13, *Specifying the Future Operating State*, turns the proposal into a description of how the organisation is expected to work differently. Its concern reaches beyond the asset or system to the capabilities, practices, responsibilities and service conditions that must change around it. The chapter will use the existing Value and Evidence Map to connect outputs with capability, practice, outcomes and benefits, retaining the assumptions between those links. It will address measures, acceptance conditions, interfaces and the ownership of benefits that depend on several teams. For the manufacturer, a delivered payment platform is one element of the design; an operating team able to resolve exceptions and suppliers able to use the process are equally relevant. The control result is a testable account of the future state, with an agreed basis for recognising progress and questioning design choices when their expected contribution to value becomes doubtful.

Chapter 14, *Committing the Next Wave of Investment*, places that design within a bounded and affordable commitment. The full ambition may remain larger than the work that can responsibly be authorised now. This chapter will connect scope, schedule, cost, resources, dependencies, uncertainty and operating capacity within the next wave, using the sequencing and decision disciplines established earlier. It will examine the evidence needed to set credible control baselines, the distinction between a target and a forecast, and the conditions attached to releasing money or scarce people. It will also make future options visible, including what can be deferred

without damaging the purpose and what must be protected now to preserve later choices. The decision is whether the proposed wave has a coherent basis and a manageable exposure, with named owners and explicit boundaries for change. Later waves remain subject to the evidence available when their own commitments are considered.

Chapter 15, *Integrating Delivery without Losing the Purpose*, follows the authorised wave into execution. Its focus is the relationship between the work being performed and the capability that must become usable. The chapter will develop practical connections among integrated schedules, cost forecasts, progress evidence, interfaces and change decisions, while keeping the operating design and intended benefits in view. It will examine how apparently favourable reports can conceal incomplete integration or deferred work, and how a local recovery action can transfer difficulty to another team. The controller must help distinguish physical completion, accepted outputs and usable capability, because each answers a different question about progress. Delivery control should also preserve the reasons for an approved change, including its consequences for adoption and value. The resulting evidence supports corrective action and a credible decision about what can move into operational use, with limitations that the receiving organisation understands.

Chapter 16, *Adoption Readiness and Operational Stability*, addresses that movement into use. Readiness concerns the ability to operate under the conditions that matter, including workload, exception handling, support and recovery. It cannot be established solely by declaring that delivery is complete. This chapter will examine demonstrations of readiness, transition arrangements, adoption evidence, stabilisation measures and the ownership of remaining work. It will distinguish the installation of capability from changed practice and from a benefit that can be credibly observed. Badewi (2016) examines the relationship between project management, benefits management and project success; that research reinforces the need to consider their connection rather than treating delivery and benefits as separate professional concerns. The chapter's practical decision is whether a defined operating population can accept the change within understood limits, and what support, monitoring or fallback is required while performance becomes dependable. Benefits verification may continue well beyond this initial period.

Chapter 17, *Reforecasting from New Evidence*, makes learning consequential. Its position in the reading sequence provides a place to examine evaluation in depth, but its work runs across every lifecycle area. A changed demand assumption may alter the proposal; an integration result may alter the next wave; operational experience may alter both the benefits forecast and the design. The chapter will explain how to reconcile new evidence with previous commitments, update the remaining forecast and distinguish a changed expectation from a retrospective alteration of what was authorised. It will revisit continued justification, uncertainty and the choices to continue, modify, pause or stop. The controller's contribution is an intelligible account of what has changed, why it matters and which decision now requires attention. Evaluation closes a learning cycle only when its implications reach an accountable decision maker and the resulting action is subsequently checked. Producing a revised number alone leaves that responsibility unfinished.

Chapter 18, Ending the Programme while Protecting Its Value, examines the final programme commitment: agreeing that temporary governance can end under stated conditions. Closure requires more than a final delivery report or the exhaustion of a budget. The chapter will address acceptance, residual obligations, unresolved risks, records, lessons, ongoing benefits measurement and the receiving owners who must continue the work. It will distinguish a programme that has completed its intended contribution from one that is being stopped or materially curtailed, because those situations require different explanations and transfers. A benefit may still be developing when programme governance closes; its measure, owner, resources and review arrangements must remain usable. The closure decision should therefore state what has been achieved, what remains conditional and who has accepted responsibility for what follows. The lifecycle model places continuing operational value beyond closure because the organisation's obligation to understand and sustain the change does not end with the programme.

Returning to the model as evidence changes

Consider the manufacturer preparing its next supplier wave. The proposal may already be authorised and the payment platform largely designed, but a trial could reveal that exception handling requires more operating capacity than anticipated. Chapter 17's evaluation then connects evidence from Chapter 16's adoption work to Chapter 14's investment decision and, if necessary, Chapter 13's operating design. Samira may authorise a smaller wave, Alex may revise the integration sequence and Maya may secure additional support before expansion. Lena preserves the original commitment, explains the revised forecast and records the conditions attached to the new decision. This is a fictional application of the model, not a claim that every programme should follow the same response. It shows why returning to an earlier decision can be a disciplined consequence of learning. The important question is whether the change is supported, authorised and checked, rather than whether the diagram has been followed once without reversal.

Continuity across these decisions depends on evidence that can travel between teams without losing its meaning. The sponsor's value claim should remain recognisable in the operating design, the authorised wave, the delivery evidence and the eventual benefits review. This does not require one enormous document or a single system that contains everything. It requires stable definitions, identifiable owners, dated assumptions and references that allow a reader to trace a material claim to its supporting evidence. A handover should make clear which conditions have been demonstrated, which are forecast and which depend on further action. Where the programme closes, that distinction must survive the departure of its temporary team. Controllers therefore protect continuity of judgement as well as continuity of data. Each receiving owner needs enough context to use the information responsibly, challenge it when circumstances change and continue the obligations that have been accepted.

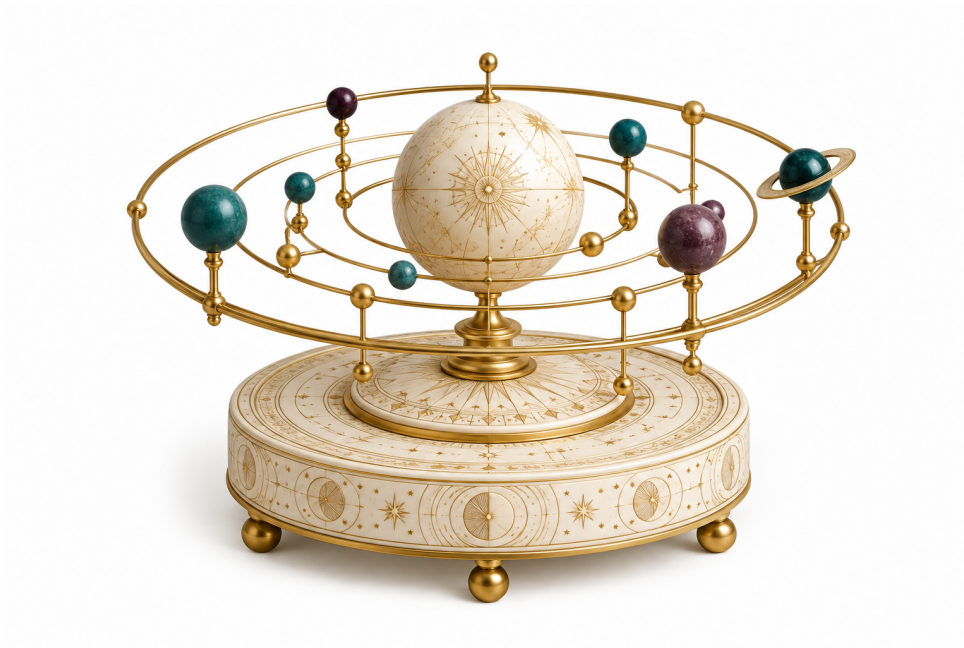
The following chapters use the same lifecycle drawing, with the relevant area highlighted while the surrounding relationships remain visible. Highlighting Chapter 11 or Chapter 17 identifies work that spans the route; highlighting another chapter locates its principal decision without isolating it from the whole transformation. Readers can therefore move through Part III in

sequence or return to the decision they currently face, while retaining a common orientation. Each chapter develops the evidence, responsibilities and practical methods needed at that point, building on the disciplines already established rather than introducing a replacement framework. The intended result is a usable connection between professional analysis and organisational action: a proposal that can be investigated, a commitment that can be governed, a capability that can be adopted and a transfer that leaves responsibility for value intact.

References

- Badewi, A. (2016). The impact of project management (PM) and benefits management (BM) practices on project success: Towards developing a project benefits governance framework. *International Journal of Project Management*, 34(4), 761–778. doi: [10.1016/j.ijproman.2015.05.005](https://doi.org/10.1016/j.ijproman.2015.05.005)
- Pich, M. T., Loch, C. H. and De Meyer, A. (2002). On uncertainty, ambiguity, and complexity in project management. *Management Science*, 48(8), 1008–1023. doi: [10.1287/mnsc.48.8.1008.163](https://doi.org/10.1287/mnsc.48.8.1008.163)

Control as a Repeated Learning Cycle



IN THIS CHAPTER

Using evidence and authorised action as a repeated learning cycle.

Complete the cycle beyond the report

Lena's report reaches the programme board on time, but a decision from the previous meeting remains unresolved. Alex has recorded the delivery action as complete because the revised procedure was issued. Maya's operating team still encounters the same payment exceptions, and nobody has examined whether the procedure changed the work. Samira is now being asked to approve another supplier wave using a report that describes activity without demonstrating its consequences. The difficulty is not simply missing data. The control process has allowed a recommendation to become an action, and an action to become a closed item, without preserving the question it was supposed to answer. Chapter 11 addresses that gap. A complete control cycle connects a defined decision with evidence, challenge, authorisation, implementation and verification, so that the next commitment reflects what actually happened rather than what the previous meeting intended.

Figure 11.1 locates this recurring work across the transformation lifecycle. The highlighted area applies when an ambition is first investigated, when a wave is authorised, during delivery and adoption, and when responsibilities transfer at closure. Its position above the route does not assign the programme board every operational decision. Different owners run control cycles at different scales, within their authority and with connections for escalation. Maya may correct a local operating problem, Alex may resolve a delivery interface and Samira may authorise a material change to the programme's commitment. Lena helps each receive usable evidence and makes the connections visible. The central design problem is to ensure that a signal reaches someone who can respond while there is still time to influence the outcome. Producing more frequent reports cannot solve a missing decision right or an action without resources.

The transformation lifecycle

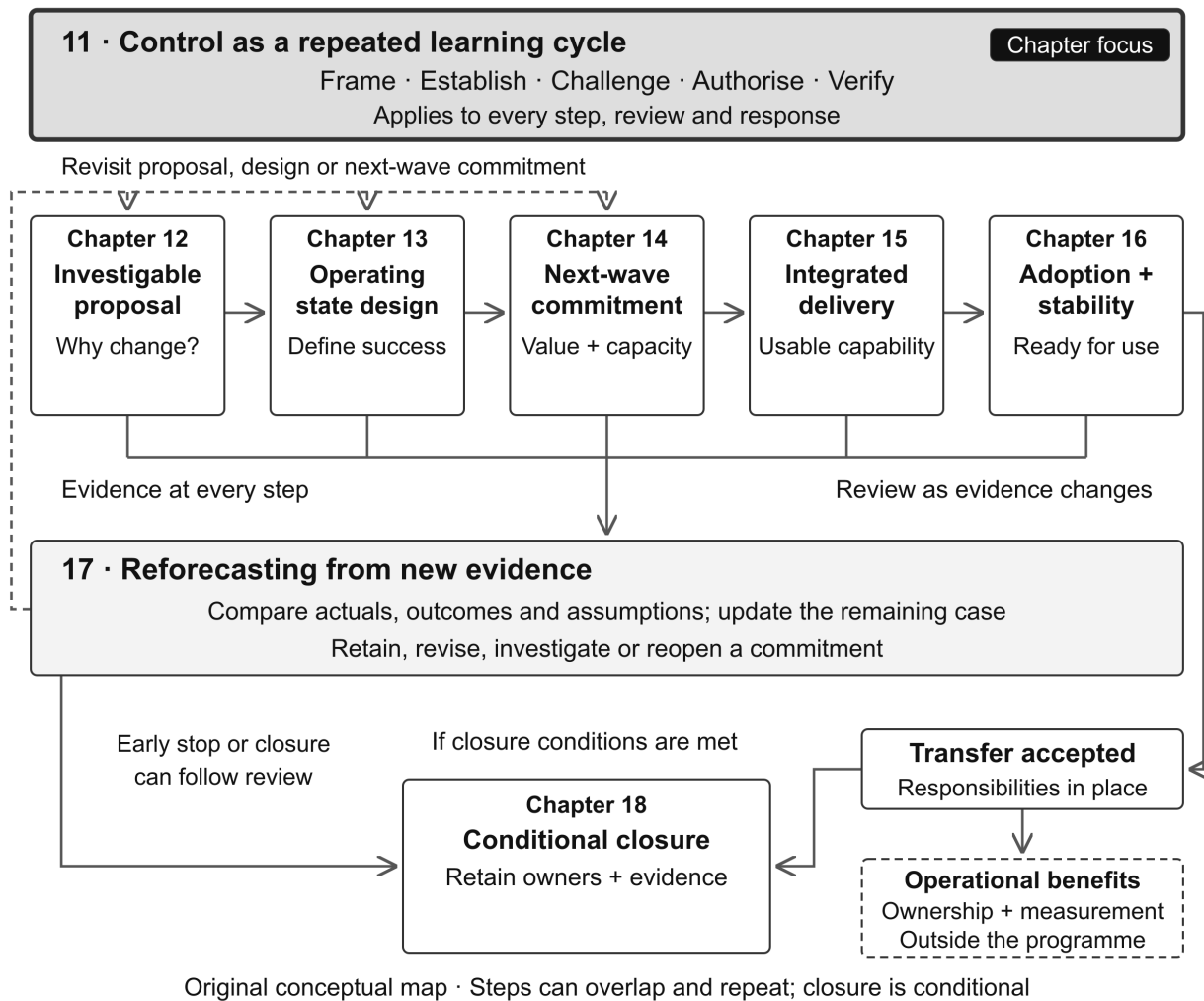


Figure 11.1 The shared transformation lifecycle highlights Chapter 11. The recurring control rhythm supports every decision and response, while Chapter 17 provides continuing evaluation. The unchanged model retains feedback, conditional closure and operational responsibilities.

The existing Credible Control Method supplies the organising sequence: Frame, Establish, Challenge, Authorise and Verify. Implementation is the work undertaken by the accountable owner between authorisation and verification, not a substitute for either. The method should leave a short but traceable account of the decision, the evidence available at the time, the response agreed and the observation that will test its effect. This chapter uses an illustrative Control Cycle Record to connect those items; it is a practical extension of the earlier decision and learning records and has not been empirically validated. The record may sit within existing systems. Its usefulness depends on whether people can follow the reasoning and act on it. A beautifully completed form that reaches the wrong authority after the decision has become irreversible represents a failed cycle even when every administrative field is populated.

Frame the decision and establish a usable position

Framing begins before the meeting. Lena identifies the decision that will be required, the person authorised to make it, its latest useful date and the evidence that could change the recommendation. The question might be whether to release the next supplier group, extend temporary operating support or retain an existing fallback. Each needs a different evidence set. A routine performance review can contain several such questions, but they should remain distinguishable from information supplied only for awareness. Preparation should also identify a feasible response if the preferred course cannot be supported. This prevents a meeting from discovering, after a long discussion, that the only person able to authorise a change is absent or that the recommendation depends on work nobody has costed. The agenda becomes a route through consequential choices rather than an inventory of departments reporting their activity.

Establishing the position requires an agreed information boundary. A schedule updated on Friday, financial actuals closed on Wednesday and an adoption measure extracted three weeks earlier do not describe one common moment merely because they appear on the same dashboard. Lena should retain each source's cut-off, population and known limitations, then explain material differences rather than disguising them through a shared reporting date. Where the decision depends on a more current position, she obtains the missing evidence or makes the uncertainty explicit. The approved baseline remains distinct from the latest forecast, and both remain distinct from actual performance. These distinctions preserve accountability while allowing the estimate of remaining work to improve. A revised date is evidence of a changed expectation; it becomes a changed commitment only through the relevant authority and a retained decision record.

Review frequency should follow the decisions and exposure, not a universal meeting calendar. Table 11.1 illustrates three connected rhythms. Operational exceptions may need attention each day, integrated delivery may need a weekly reconciliation and major investment choices may need a less frequent sponsor review. These are teaching examples rather than recommended frequencies for every organisation. Each rhythm needs an event route for information that cannot wait, and an explicit boundary beyond which its owner must escalate. A local decision should not be repeatedly reopened at every higher meeting unless a material condition has changed. Equally, a programme board should not assume that a favourable aggregate status removes a serious local constraint. Lena's role includes checking that issues can move between the rhythms without losing their original evidence, responsible owner or decision deadline.

Table 11.1 Connect review rhythms to decisions

Illustrative rhythm	Decision focus	Escalation condition
Operating review daily	Maya addresses workload, quality and recovery.	Exposure exceeds local authority or cannot await the next review.
Delivery review weekly	Alex reconciles interfaces, remaining work and resources.	A dependency threatens an authorised commitment or operating readiness.
Sponsor review at commitment dates	Samira authorises material changes, wave releases or continued investment.	A changed case or urgent exposure requires an earlier decision.

Frequencies are illustrative. Delegated authority, evidence ownership and event routes must suit the actual programme.

Haji-Kazemi, Andersen and Blakegg (2015) examine barriers between recognising early warnings and responding to them, using a survey of Norwegian project managers. The relevance is the gap between noticing a signal and making it consequential. Their findings do not establish a universal escalation interval or demonstrate that a particular meeting design prevents failure. For the manufacturer, a warning about unresolved exceptions is useful only if someone assesses its significance, identifies a response and takes it to the appropriate authority. Lena should therefore record when a consequential signal became available as well as when it entered the formal report. A long interval between those dates can reveal a problem in the control process even when the eventual analysis is technically sound. Reviewing that interval also helps distinguish late discovery from slow handling of information already known.

Challenge the account and authorise an executable response

Challenge should examine the explanation and the proposed response, not merely seek confirmation that figures have been checked. Maya may ask whether rising exceptions reflect a genuine operating deterioration or better identification of previously hidden cases. Alex may explain that a short-term correction consumes people needed to complete an interface. Lena should connect those consequences and distinguish what the evidence establishes from what remains an assumption. The review can then compare a bounded intervention, further investigation, a changed sequence or temporary containment. It should also consider what would happen if no additional action were taken. The purpose is not to reward disagreement or demand certainty before every decision. It is to prevent the preferred response from becoming inevitable before its operating consequences and alternatives have been examined by the people who understand them.

Edmondson's (1999) study of 51 teams in one manufacturing organisation associated psychological safety with learning behaviour. This observational evidence supports attention to whether people can disclose mistakes and uncertainty; it does not prove that a revised board agenda causes better transformation outcomes. In the review, Samira can make the distinction practical by asking for the strongest contrary evidence and responding to it without treating the disclosure itself as failure. That does not remove accountability for the quality of an estimate or

the completion of agreed work. It allows the team to separate an honestly revised forecast from an unsupported assurance. Lena should retain unresolved disagreement where it matters to the decision, including the reason an authorised course was chosen despite it. Silence, agreement under pressure and genuine convergence are not equivalent forms of evidence.

Feedback can also distract from the work it is intended to improve. Kluger and DeNisi's (1996) meta-analysis found that feedback interventions did not consistently improve performance; more than a third of the reported effects were negative. Their theoretical explanation considers where feedback directs attention, including towards the task or towards the self. The practical application here is proportionate: a controller should make a review specific enough to support correction rather than assuming that more visible rankings or sharper criticism will improve results. Comparing an overdue action with its required evidence and consequences is more useful than attaching a general label to the responsible team. The research does not determine the right wording for every meeting, but it gives a reason to test whether the review is producing clearer action or defensive reporting. The effect of feedback remains something to examine.

NASA's Mars Climate Orbiter investigation illustrates why a problem-reporting process must include effective follow-up. The March 2000 report identified weaknesses in reporting and tracking anomalies and in the team's use of the Incident, Surprise, Anomaly process (Mars Climate Orbiter Mishap Investigation Board, 2000, p. 18). This is evidence from a particular mission investigation, not a claim that paperwork alone could have prevented the loss. Its transferable question is whether a consequential concern remains visible until its technical and managerial implications have been resolved. For the manufacturer, authorisation should identify the action owner, resources, expected effect, due date, operating constraints and verification requirement. A recommendation to improve readiness is not yet executable. An agreement that Maya will demonstrate a defined recovery process with a specified team before the next release creates a commitment whose completion and adequacy can subsequently be examined.

Set the rhythm against the time available to respond

An apparently reasonable reporting interval can be too slow for the process being controlled. Consider a separate fictional queue of payment exceptions measured in comparable cases and calendar days. At the point a warning condition arises, the backlog is sixty cases. New cases arrive at 120 per day and the team clears 100, giving a net increase of twenty per day. An unacceptable operating limit is set at 200 cases for this illustration. With constant rates, the remaining margin is 140 cases and the time to reach the limit is 140 divided by twenty, or seven days. This simple stock-and-flow calculation describes a conditional trajectory, not a statistical forecast or a recommended limit for payment operations. Real cases can differ in effort, age and seriousness, and these differences would need to be included before using the model operationally.

A warning arising just after a weekly check may remain undetected for up to seven days. If validation and decision then take two days and implementation takes another three, the total delay before the response becomes effective is twelve days. Under the same constant-rate assumptions, the backlog reaches 300 before that response begins. Daily detection reduces the

corresponding total to six days, with a backlog of 180 and one day remaining before the limit would be reached without action. Figure 11.2 compares these timings. The calculation assumes that implementation immediately changes capacity once complete and that the stages occur consecutively. If an approved contingency allows some activities to run in parallel, their actual dependency structure should replace the simple sum. A controller should establish those timings from the feasible response, rather than shorten them on paper to make a preferred review calendar appear adequate.

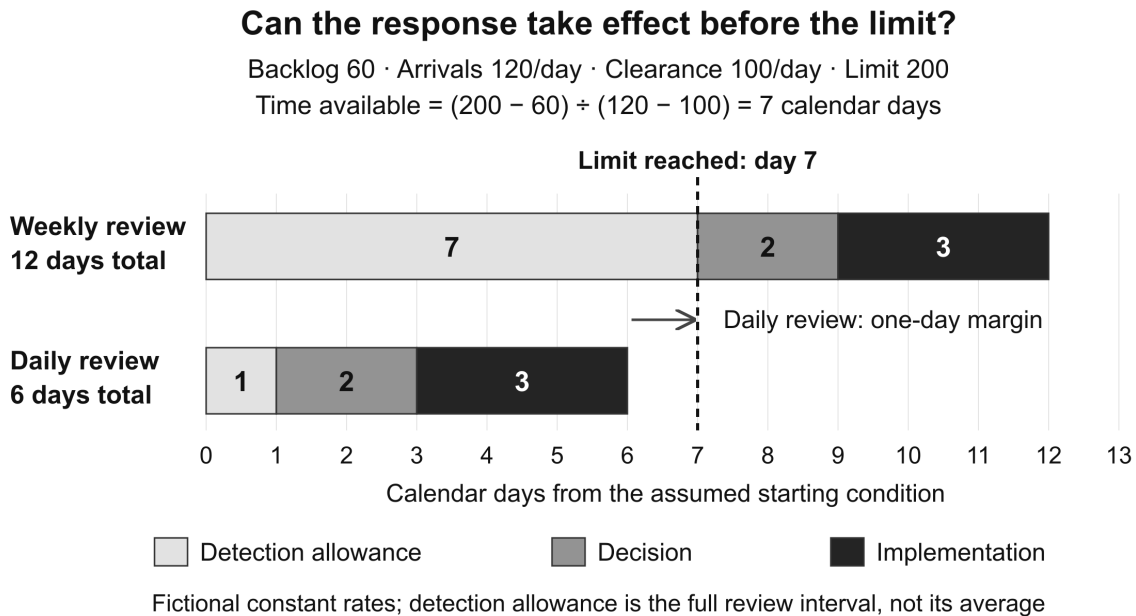


Figure 11.2 Fictional response timing. The backlog reaches its limit after seven days at constant net growth. Sequential detection, decision and implementation can require twelve days with weekly checks or six with daily checks. The one-day margin is conditional, not a confidence bound.

The one-day margin is a warning about sensitivity, not reassurance that daily review solves the problem. If net growth is thirty cases per day instead of twenty, the time to the same limit falls to approximately 4.67 days, while the six-day response produces a backlog of 240. The organisation would then need earlier detection, faster authorised action, temporary containment or another feasible response. This is the same anticipatory principle developed in Chapter 10, now applied to the design of a recurring process. Before setting a reporting cadence, Lena should compare the time available with the time needed to recognise, decide and act, allowing for uncertainty and operating consequences. Some conditions also require immediate intervention under established procedures. A reporting calendar must not become a reason to defer an action that the responsible owner is already authorised and required to take.

Increasing observation frequency has costs and can create repeated alarms without better decisions. The practical aim is to make a consequential condition visible early enough to respond, using reliable measures and suitable authority. A daily automated count may be appropriate while a full management meeting remains weekly; an exception can trigger a short decision conversation outside that meeting. Thresholds should be connected to consequences and feasible responses, with a rule for handling recurrence, missing data and material changes in exposure.

Where a measure fluctuates, the team should investigate its normal behaviour before treating every movement as a new problem. Conversely, averaging can conceal a serious local condition. These choices require judgement about the operating process. The controller should document why the chosen signal and response arrangement provide useful time, and revisit that reasoning when workload or capacity changes.

Verify implementation and examine the operating effect

Suppose Maya's approved response raises clearance capacity from 100 to 140 cases per day once implemented. Under the daily-detection scenario, the response begins with 180 cases outstanding. If arrivals remain at 120, the backlog then falls by twenty cases per day and reaches eighty after five further days. This gives the team a testable expectation and a review date; it does not establish that an increase in staffing will deliver the assumed capacity without training, equipment or other support. Alex must confirm the dependencies that make the response feasible, and the authorising owner must accept its costs and effects on other work. Table 11.2 preserves the fictional decision and its verification basis. The illustration remains separate from the manufacturer's established programme budget and forecast, which are not recalculated by this operating example.

Table 11.2 Worked entries in the Control Cycle Record

Record item	Fictional entry
Decision	Authorise a resourced capacity response before the backlog reaches 200.
Responsibilities	Maya implements; Alex confirms dependencies; Samira authorises resources; Lena retains evidence.
Response timing	Up to 1 day detection, 2 days decision and 3 days implementation; effect begins at day 6 in this scenario.
Implementation evidence	Confirm capacity for 140 comparable cases per day is available and the agreed constraints are met.
Effect review	Five full days later, test the expectation of 80 cases remaining; inspect arrivals, clearance, quality and case mix.
If unsupported	Reopen the explanation or response; retain residual exposure and escalate before the next consequential commitment.

Original teaching record. Constant arrivals of 120 cases per day and immediate full effect after implementation are assumptions to verify, not operational guarantees.

The first verification question is whether the agreed action actually occurred. An approved resource request, an issued procedure and a completed training session are different observations from staff working effectively in the intended process. Maya should supply evidence appropriate to the commitment, including when the additional capacity became available and any conditions that limited it. Lena then updates the status of implementation without automatically closing the original concern. Figure 11.3 separates implementation evidence from evidence about the effect.

If the intervention is late or incomplete, the expected operating trajectory must be reconsidered immediately. If it is implemented as planned but the backlog does not improve, the team must examine arrivals, actual clearance, case mix, rework and the original explanation. Repeating the same action more emphatically is not a substitute for investigating why its expected mechanism did not operate.

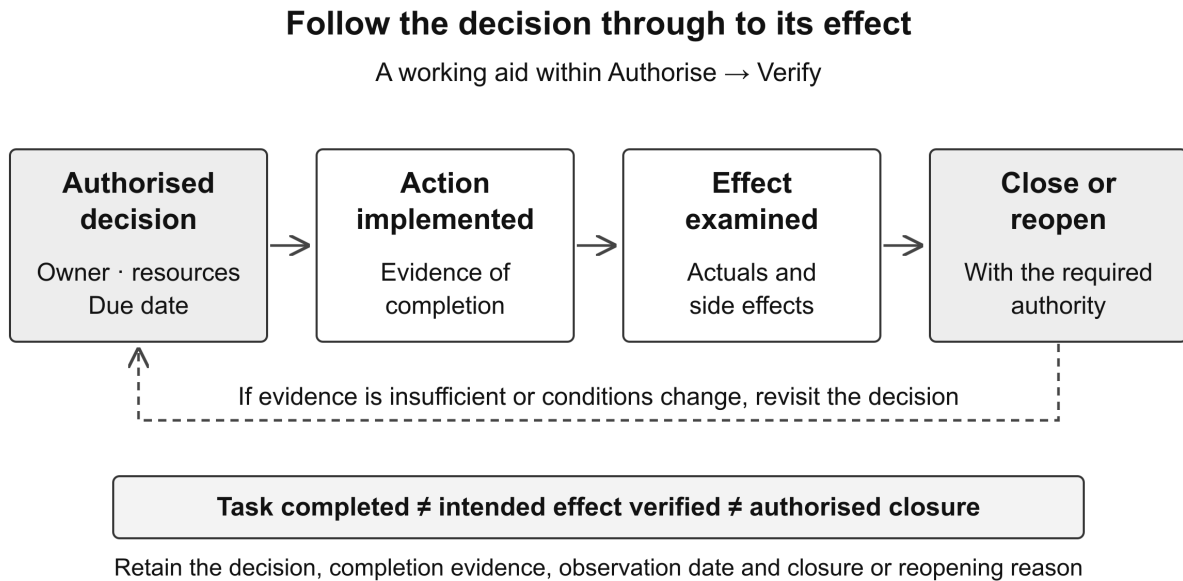


Figure 11.3 Original practice aid connecting Authorise to Verify. Evidence that an action was implemented remains distinct from evidence about its effect and from authority to close the concern. Reopening returns the decision to its responsible owner.

The second verification question concerns performance and unintended effects. A reduced backlog may reflect genuine clearance, a fall in demand, transfer of work elsewhere or a change in counting rules. Faster processing may also increase errors that return later. Lena should compare the observations with the stated expectation, reconcile the measurement boundary and examine credible alternative explanations before attributing an improvement to the intervention. A short before-and-after comparison can support a cautious operating judgement without proving causality. Closure should therefore state what was demonstrated and what still needs monitoring. A completed action may leave the outcome under observation; a favourable outcome may occur before the proposed action is complete. These are legitimate states of the record, provided they are visible. The authority closing the concern should be satisfied that the remaining exposure has an owner and an appropriate route for review.

New Zealand's Auditor-General reported in March 2022 on follow-up to earlier recommendations concerning the Covid-19 vaccination programme. The Ministry reported routine monitoring of vaccine supply and use and monthly reviews of onboarding processes. The Auditor-General assessed progress while retaining unresolved equity concerns (Office of the Auditor-General, 2022). The follow-up used selected evidence and limited stakeholder discussions; it was not another full performance audit. This provides a bounded illustration of recommendations returning for examination after implementation activity, rather than proof that a particular review cycle caused vaccination outcomes. The useful distinction is between an

organisation reporting that it has acted and a reviewer establishing what that action supports. For project controllers, a follow-up should retain the original concern, the response claimed, the evidence inspected and the matters still open. An overall positive assessment need not erase specific obligations that require continuing ownership.

Let experience change the routine

A control process needs enough consistency to make evidence comparable and responsibilities predictable, but consistency does not mean freezing every practice. Feldman and Pentland (2003) develop a theoretical account of organisational routines that distinguishes an abstract understanding of the routine from its particular performances. Their argument helps explain how repetition can also generate change. It does not establish that every local variation is beneficial or that documented procedures are unnecessary. In this chapter's application, Lena compares how the review is supposed to operate with what participants actually do. Perhaps operational evidence repeatedly arrives after the recommendation is settled, or verification dates are routinely postponed. Those observations concern the functioning of the routine itself. The team can then revise the preparation, participation or timing through an explicit decision, preserving useful consistency while correcting a process that no longer serves its purpose.

Structured reflection can help identify that correction. Tannenbaum and Cerasoli's (2013) meta-analysis of debriefs supports their potential to improve subsequent performance across the studied settings, but it does not provide a guaranteed improvement for a programme board or validate a particular template. A short debrief after an important control episode should reconstruct what was expected, what was observed, how the decision was made and what should change in the next cycle. It should examine successful responses as carefully as disappointing ones, because favourable results can depend on temporary conditions that will not recur. The resulting lesson needs an intended use and an owner. A general instruction to communicate better supplies neither. A decision to obtain the operating extract before the recommendation is drafted, with a named source owner and trial period, creates a change whose use can be checked.

Learning may concern the response, the decision process or the objective being pursued. If the backlog remains high because the original design creates avoidable exceptions, temporary clearance capacity may only contain the consequence. The appropriate next cycle can return the question to the operating-state design or the next investment wave, using the routes shown in the lifecycle model. That does not authorise a local team to replace an agreed objective or quietly lower an acceptance standard. Material changes still need the relevant authority and a preserved account of their consequences. Chapter 8 established how evidence becomes organisational learning; here the emphasis is on the recurring opportunity to act on it. The test is whether the lesson changes a subsequent preparation, commitment or operating practice and whether the receiving owner can demonstrate that change under relevant conditions.

Keep the record short and the responsibility continuous

The Control Cycle Record should be as small as the decision permits while retaining the connections needed for accountability. It identifies the decision and its owner, evidence cut-off, original commitment, current expectation, material uncertainty, authorised response and verification requirement. Source material can remain in the systems that maintain it, provided references are stable and accessible. A record should not duplicate every schedule, risk entry or benefits calculation. Its purpose is to allow another competent person to understand why the action was chosen and what evidence would justify closing or reopening the concern. Where several actions support one outcome, their dependencies should be visible. Where one action affects several outcomes, its consequences should remain connected. Giving every document a new identifier without preserving these relationships increases administrative volume while making the underlying judgement harder to reconstruct.

Status language should reflect that judgement. A recommendation awaiting authority, an authorised action awaiting implementation, an implemented action awaiting effect evidence and a concern closed under continuing monitoring are meaningfully different conditions. Collapsing them into one completion percentage conceals work that the next owner still needs to perform. Lena can review overdue verification separately from overdue implementation, and examine whether the same concerns recur after apparently successful closure. These observations help diagnose the control process, but they should not become targets that encourage premature closure or superficial verification. A shorter action list is not necessarily a healthier programme. The sponsor needs confidence that important concerns are being resolved or deliberately retained within understood limits. Reported closure should therefore be supported by the decision, evidence and residual responsibility, rather than by a desire to improve the colour of the dashboard.

For example, suppose twelve actions have reached their implementation deadlines and eight have completion evidence. The implementation measure is eight out of twelve, or 66.7 per cent. Only six actions have reached their separately agreed effect-review dates, and four of those reviews have been completed: verification coverage is four out of six, also 66.7 per cent. The equal percentages conceal different populations and responsibilities. Neither establishes that two-thirds of the intended benefits have been achieved. The completed effect reviews must still distinguish supported improvement, no demonstrated change and an adverse result. Recording these states avoids treating an effect review whose agreed date has not yet arrived as overdue, while exposing reviews that should already have occurred. Original dates and authorised changes should remain traceable so that repeated rescheduling cannot manufacture a favourable score. These measures help examine follow-through; they do not replace judgement about the seriousness of each unresolved concern.

Reopen the next commitment with better evidence

At the next manufacturer review, Lena does not simply report that Maya's response is complete. She distinguishes when capacity became available, what happened to the backlog, whether quality and operating conditions remained acceptable and what uncertainty still affects expansion. Alex explains the implications for delivery resources. Maya confirms the support needed to sustain the operating position. Samira can then authorise the next supplier wave, change its scope or retain a bounded pause, with conditions that another review can test. If the evidence remains insufficient, the record says what is missing and when a useful decision must be made. This completes the cycle without pretending that the programme has reached certainty. The practical achievement is that the next commitment is informed by the consequences of the previous one, with responsibility attached to both the decision and its follow-through.

Control as a repeated learning cycle is therefore an operating discipline throughout transformation. It begins with a question that matters, preserves an intelligible evidence boundary, makes challenge useful and turns authority into feasible action. It then returns to the work to examine implementation and effects before closing the concern or changing course. The process should be proportionate to the decision, fast enough for the exposure and open to correction when its own assumptions fail. The remaining lifecycle chapters apply that same discipline to more specific commitments. Chapter 12 begins at the front of the route, where strategic ambition must become a proposal that can be investigated responsibly. The same obligation already applies there: establish what the organisation is deciding, identify what could change that decision and arrange the next opportunity to learn before a larger commitment becomes difficult to reverse.

References

- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383. doi: [10.2307/2666999](https://doi.org/10.2307/2666999)
- Feldman, M. S. and Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1), 94–118. doi: [10.2307/3556620](https://doi.org/10.2307/3556620)
- Haji-Kazemi, S., Andersen, B. and Blakegg, O. J. (2015). Barriers against effective responses to early warning signs in projects. *International Journal of Project Management*, 33(5), 1068–1083. doi: [10.1016/j.ijproman.2015.01.002](https://doi.org/10.1016/j.ijproman.2015.01.002)
- Kluger, A. N. and DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. doi: [10.1037/0033-2909.119.2.254](https://doi.org/10.1037/0033-2909.119.2.254)
- Mars Climate Orbiter Mishap Investigation Board. (2000, 13 March). Report on project management in NASA. National Aeronautics and Space Administration. discovery.larc.nasa.gov/discovery/PDF_FILES/mars_climate_orbiter_phaseII.pdf
- Office of the Auditor-General (New Zealand). (2022, 10 March). Implementation of recommendations – Preparations for the nationwide roll-out of the Covid-19 vaccine. Letter to the Chairperson of the Health Committee and Attachment 2. ao.parliament.nz/2022/vaccines-follow-up
- Tannenbaum, S. I. and Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors*, 55(1), 231–245. doi: [10.1177/0018720812448394](https://doi.org/10.1177/0018720812448394)

CHAPTER 12

From Strategic Ambition to an Investigable Proposal



IN THIS CHAPTER

Making the strategic question precise enough to investigate.

Keep the first commitment open to investigation

To examine the first commitment, return to the period before the manufacturer's transformation programme was approved. Samira wants suppliers to experience reliable payment and less administrative effort. A presentation has already translated that ambition into a proposed global payment platform, although the organisation has not established why payments are late or which changes would address the problem. Alex is asked for a delivery date, Maya is asked to prepare operations and Lena is asked to create a control baseline. Answering those requests too quickly would make an untested solution appear to be an authorised programme. The immediate decision is smaller and more consequential: whether the need warrants a structured investigation, what that investigation should compare and what authority it should receive. The fictional examples in this chapter concern that earlier decision and do not replace the programme budget or forecast established in previous chapters.

Figure 12.1 highlights this position in the transformation lifecycle. An investigable proposal connects a strategic intention with an observed need, plausible responses, material assumptions and a bounded request for evidence. It does not require the final operating design, detailed delivery schedule or complete benefits case before inquiry can begin. It does require enough clarity to prevent investigation from becoming an expensive exercise in defending a favoured answer. The controller helps distinguish the authority to examine a possibility from the authority to implement it. That distinction preserves choice while making the initial expenditure accountable. Samira needs to understand what the study could establish, what it will leave uncertain and which decision follows it. Maya and Alex need to know what work they are committing to, while Lena needs a traceable basis for reporting whether the inquiry remains useful and within its limits.

The transformation lifecycle

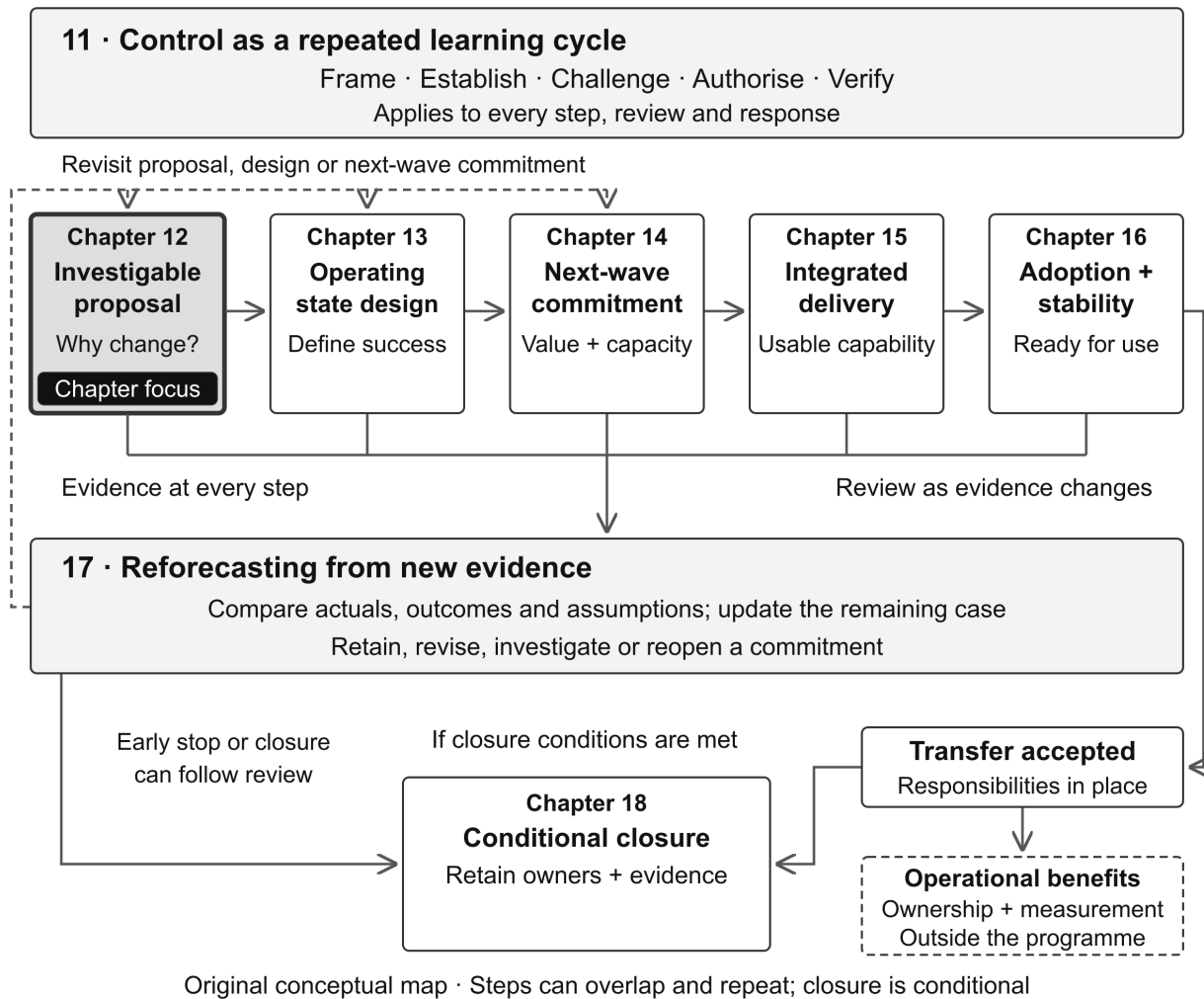


Figure 12.1 The shared lifecycle highlights the investigable proposal. Evidence may reopen the need, design or next commitment. Initial authorisation supports bounded inquiry; later operating design and investment require their own decisions.

The Credible Control Method applies from this point. Frame identifies the decision and the need behind it; Establish assembles a credible description of the present position; Challenge examines explanations, alternatives and assumptions. Authorise bounds the investigation and its responsibilities, while Verify checks what the evidence supports when the team returns. The proposed Proposal and Investigation Record connects those activities without adding another lifecycle or claiming empirical validation. Figure 12.2 shows the reasoning that the record should preserve. A strategic aspiration such as dependable supplier relationships is a starting point, but it becomes investigable through questions about a defined population, an operating condition and the consequences that matter. The record should allow a reviewer to understand why further work is warranted without having to accept the preferred technology, supplier or delivery model as an established requirement.

Turn ambition into a question worth investigating

An inquiry aid within the Credible Control Method

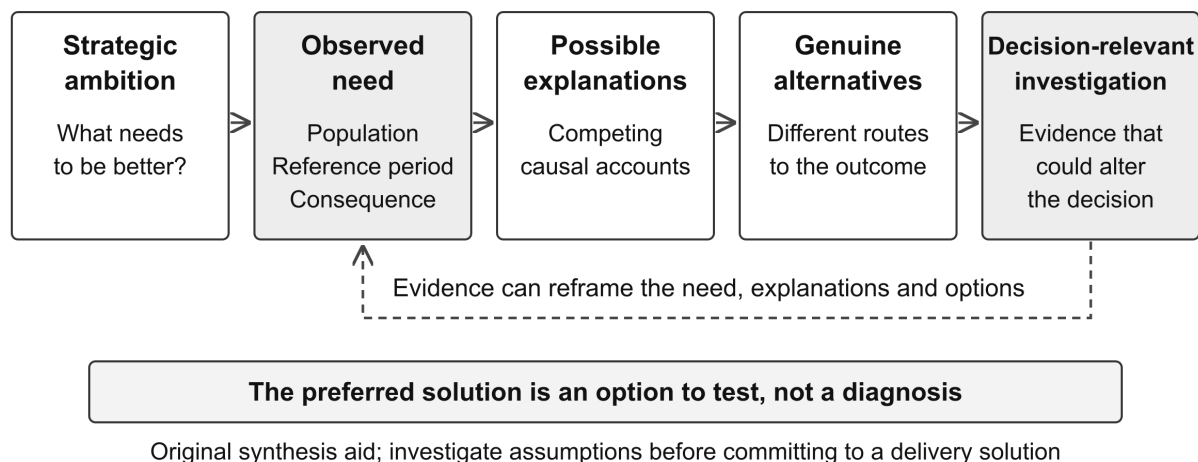


Figure 12.2 Original inquiry aid within the Credible Control Method. Strategic ambition becomes an investigable question through a defined need, competing explanations and genuine alternatives. Findings can revise that formulation. A preferred solution remains an option to examine.

Establish the need before estimating its remedy

Samset and Volden (2016) examine paradoxes in front-end project definition, including the tendency to concentrate on delivery detail while more fundamental choices remain weakly explored. Their paper synthesises research and cases; it does not establish that one standard investigation process causes successful investments. Its practical relevance is the distinction between improving the estimate of a chosen solution and improving the choice itself. Lena can request a more precise platform schedule, but that will not reveal whether the platform addresses the principal causes of late payment. Before demanding precision, she should identify which uncertainty matters to the next decision. Some questions concern the scale of an observed problem, others its causes, and others the organisation's ability to change them. Treating these as separate questions helps the team avoid spending its evidence budget on detail that cannot change the recommendation.

An initial problem statement should identify the affected population, reference period, observed condition, consequences and relevant constraints. For the manufacturer, the population might be invoices contractually due within a stated period, with a consistent rule for identifying those paid late or still unpaid after their due date. The source and extraction date matter because a payment system can change status after the reporting period ends. Missing due dates or disputed eligibility should remain visible rather than being silently treated as satisfactory performance. Supplier experience may also differ by size, location, payment method or access to support. The initial account need not resolve every difference, but it should identify those capable of changing the decision. An assertion that payments are unreliable becomes more useful when the team can say for whom, under which conditions and with what practical consequences.

Table 12.1 gives a fictional, consistently defined cohort of four thousand eligible invoices across three business units. Unit A has two hundred late invoices among one thousand, or twenty per cent. Unit B has one hundred among two thousand, or five per cent. Unit C has one hundred among one thousand, or ten per cent. The combined rate is four hundred divided by four thousand, or ten per cent. Averaging the three unit percentages without their denominators gives approximately 11.67 per cent, which answers a different question: the average unit rate with each unit given equal weight. Neither statistic should be substituted for the other without explanation. The combined rate describes this invoice cohort; it does not establish the share of suppliers affected, the duration of delay or the value of payments outstanding. Those require their own definitions and evidence.

Table 12.1 Keep the population attached to the rate

Business unit	Eligible invoices	Late invoices	Late rate
A	1,000	200	20.0%
B	2,000	100	5.0%
C	1,000	100	10.0%
Combined	4,000	400	10.0%

Fictional mature cohort with one consistent eligibility and lateness definition. The combined rate weights invoices; the equally weighted mean of unit rates is 11.67 per cent. Neither measures supplier prevalence or delay duration.

Suppose an initial review assigns each of the four hundred late cases to one primary recorded reason: 120 involve missing order information, one hundred master data, eighty approver availability and one hundred other or unresolved matters. The first three categories contain three hundred cases, or seventy-five per cent of the late cohort. That is a reason to investigate the relevant processes, not a forecast that seventy-five per cent of late payments can be prevented. A recorded reason may describe where a case became visible rather than its underlying cause; multiple interacting causes may be hidden by the primary-category rule. Lena should retain the classification method and examine representative case histories with operating staff. The immediate value of the analysis is to locate questions that could discriminate between interventions. It cannot yet justify attributing a financial benefit to the proposed platform.

Form the problem with those who understand the work

Baer, Dirks and Nickerson (2013) develop a conceptual account of strategic problem formulation, examining how a useful representation can emerge when relevant knowledge is distributed and participants bring different perspectives. Their framework is theoretical rather than a validated workshop intervention. For the controller, the implication is to make the formulation process visible enough to challenge. Maya may describe missing approvals, procurement may identify incomplete purchase orders and suppliers may describe inaccessible dispute handling. Those accounts can concern different parts of the same mechanism rather than competing truths from which the sponsor must choose one. Lena should record the observations, proposed explanations and unresolved differences separately. A facilitated discussion can then examine where the accounts connect and what evidence would distinguish them. Reaching agreement on a preferred solution before those questions are explored may create apparent alignment while leaving the underlying problem poorly specified.

Matinheikki and colleagues (2016) study inter-organisational value creation in the front end of the Rehapolis project. Their retrospective qualitative case examines activities through which participating organisations developed relationships and a shared basis for the project; it does not evaluate whether those activities ultimately produced successful value outcomes. The useful connection here concerns the organisations whose cooperation makes the proposal feasible. A manufacturer's finance function cannot redesign supplier behaviour, purchasing practice and operational approval responsibilities entirely through its own mandate. The investigation should identify those dependencies and secure proportionate participation from the relevant owners. An expression of support is different from an agreement to provide people, information or changed working arrangements. Lena should make that distinction visible before estimating an achievable programme, while avoiding the assumption that every affected party needs the same decision right or the same degree of involvement.

The emerging proposal should also state whose improvement is sought and which constraints must be respected. Faster payment processing may be valuable to suppliers, but it must preserve appropriate checks and a workable route for exceptions. Reducing internal handling effort may release capacity without reducing cash expenditure, and moving work to suppliers may improve an internal metric while worsening their experience. Maya should therefore help identify a small set of outcomes that would make the change worthwhile, together with consequences that would make an option unacceptable. At this stage these are provisional decision requirements, not a completed benefits baseline. Each needs an owner who can explain its meaning and a plan for strengthening the evidence. Chapter 13 will develop the operating state in greater detail; the present task is to establish a credible reason to undertake that design work.

Compare genuine alternatives before preference hardens

Williams and Samset (2010) discuss front-end decision issues, including the importance of considering genuinely different concepts and the zero option. Their paper is a research-based conceptual discussion, not a prescribed sequence that every programme must follow. In the manufacturer's investigation, meaningful alternatives could include improving existing data and approval practices, adding targeted workflow support or replacing a larger part of the payment capability. These differ in their mechanism and operating demands, not merely in the brand of software supplied. The reference position also needs a realistic account of continued operations, including necessary maintenance or unavoidable change. Calling it do nothing should not imply that it has no cost, risk or consequence. Lena should compare options against the same need, population and period, retaining the assumptions that make their consequences different.

Nutt's (1998) study of 352 strategic decisions examines how their framing relates to reported success. It identifies associations between performance-based framing and more successful decisions, while prematurely specifying an action can narrow the search. This observational evidence does not prove that a particular phrasing will improve a programme's outcome. Its application is to keep the purpose of the investigation distinct from the solution preferred by its advocates. Investigating how to reduce avoidable payment delays permits several responses; investigating how to introduce the chosen platform begins after much of that choice has already been made. Where an external requirement genuinely fixes part of the solution, the proposal should identify that constraint and the choices still available. The controller should not invent alternatives that cannot meet a binding obligation, but neither should convenience or sponsorship be presented as such an obligation without examination.

Victoria's 2026 Commonwealth Games planning illustrates the consequences of a narrow early comparison. The Victorian Auditor-General's March 2024 audit found that the business case considered hosting and not hosting without analysing alternative delivery models, and that agencies did not recommend delaying commitment until a fit-for-purpose case was available (Victorian Auditor-General's Office, 2024). The case concerns the adequacy of advice and decision material at that time, rather than a universal conclusion about hosting major events. For controllers, the question is whether a binary choice conceals feasible ways to meet the underlying objective at different cost and exposure. A study that compares a fully specified preferred option with a weakly described reference position can create an appearance of appraisal without preserving meaningful choice. The proposal should state which alternatives will be examined and justify exclusions before detailed estimation makes one option difficult to dislodge.

Ask for evidence that can change the next decision

An investigation becomes manageable when each important assumption is connected to an evidence question and a possible consequence for the recommendation. Suppose the preferred explanation is that approver availability causes delays that workflow automation would remove. The team can trace selected cases across their timestamps, examine who could authorise payment and test whether a different allocation of responsibility could resolve the constraint. If the delay occurs before a valid order exists, automating an approval reminder may have limited effect. The question is therefore broader than whether users like a demonstration of the proposed system. Lena should ask what result would lead the team to prefer another response, limit the proposed scope or abandon further investment. If no plausible finding could change the recommendation, the work may be solution development rather than the investigation for which authority is being requested.

Investigation priorities should follow the consequences of being wrong and the likelihood that evidence could change the decision. If all plausible responses depend on data that the organisation cannot legally or practically access, that question may need resolution before detailed option costing. If an uncertain operating assumption would make only one option infeasible, a small targeted examination may be more useful than refining every estimate equally. The team should record the decision affected, evidence sought, responsible investigator and consequence of each possible finding. This applies Chapter 10's information-value reasoning without requiring a monetary calculation for every question. An issue can be important because it concerns a binding requirement or a potentially irreversible commitment. A numerical priority score should not silently trade away such a requirement. Lena's role is to make the ordering of inquiries defensible and revisable as the initial findings arrive.

The evidence plan should match the type of uncertainty. A complete extract may establish volumes and recorded status but leave causes unresolved. Interviews can explain practice and expose missing categories while remaining vulnerable to recall and selective participation. Observing a small trial can test whether a mechanism is workable under defined conditions without proving that it will transfer across all business units. These sources should complement one another rather than being treated as interchangeable counts of evidence. Lena should specify how cases will be selected, what comparison is feasible and which populations or conditions remain outside the study. Uncertainty that cannot be resolved within the investigation should be carried forward explicitly. The purpose is to obtain enough justified understanding for the next decision, with a clear account of its limits, rather than to promise that an initial study will remove every future delivery risk.

Potential benefits belong in the proposal as claims whose mechanisms require investigation. For each important claim, the team should connect an operating change with a valued outcome and identify who would own it. A reduction in handling time requires evidence about case volumes, changed effort, adoption and the use of released capacity. A cash-saving claim additionally requires a feasible change in expenditure and the relevant financial authority. Two teams should not independently claim the same improvement because one reports fewer exceptions and the other reports less rework. The initial benefits account can remain a range or a structured

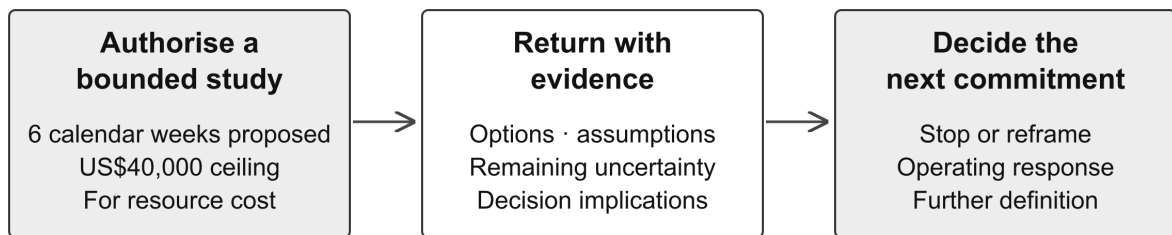
hypothesis while these boundaries are clarified. That is more useful than a precise total assembled from incompatible assumptions. The resulting evidence can support the later business case without allowing an attractive headline to become an unquestioned commitment before its basis has been examined.

Authorise a bounded investigation

Figure 12.3 places a clear boundary around the initial authorisation. The sponsor approves specified inquiry, with resources, a proposed completion date, evidence requirements and a return decision. Approval does not automatically authorise procurement of the preferred delivery solution, deployment into live operations or the full programme. Any limited trial needs its own operating conditions, responsible owner and authority. This boundary should be understandable to suppliers and internal teams so that study activity does not create unintended commitments through announcements, contracts or mobilisation. Lena should also distinguish investigation outputs from its purpose. A report and presentation may be required outputs, but their completion matters because they support a defined choice. If changed circumstances remove that choice, the sponsor should reconsider the study rather than continue simply because its original timetable and budget remain available.

Authorise the investigation, then decide again

Fictional study: authority ends at the investigation boundary



US\$33,500 base resource cost + US\$6,500 controlled allowance = US\$40,000

No automatic delivery approval

53 person-days does not establish a six-week elapsed plan.

Demonstrate capacity, sequencing and availability; distinguish resource use from new cash.

Figure 12.3 Fictional investigation boundary. A US\$40,000 resource-cost ceiling comprises US\$33,500 of planned work and US\$6,500 of controlled allowance. Six calendar weeks is a proposed duration requiring a feasible plan. Returning with evidence leads to another decision, without automatic delivery approval.

Consider a fictional investigation proposed over six calendar weeks. Table 12.2 values thirty analyst days at US\$600 each, fifteen operating-review days at US\$500 and eight specialist-assessment days at US\$1,000. The resulting base resource cost is US\$33,500. A controlled allowance of US\$6,500 gives a maximum resource-cost authorisation of US\$40,000 under the stated scope. These figures are teaching assumptions, not evidence about the actual cost of the manufacturer's programme. The day rates should have a consistent cost boundary,

and internal resource valuation should remain distinct from additional cash expenditure. Existing payroll does not make staff time free, but allocating its value to the study does not necessarily require an equal new cash payment. The approving authority needs both the resource commitment and the cash implications before accepting the proposed investigation.

Table 12.2 Bound the resources committed to investigation

Resource	Days	Day rate US\$	Value US\$
Analyst	30	600	18,000
Operating reviewers	15	500	7,500
Specialist assessment	8	1,000	8,000
Base resource cost	53	—	33,500
Controlled allowance	—	—	6,500
Authorised ceiling	—	—	40,000

Fictional consistently costed day rates. Internal resource values are not automatically additional cash expenditure. Allowance release needs authority, and elapsed feasibility requires a separate assessment.

The fifty-three person-days do not establish that the study can be completed in six weeks. Some work can occur concurrently, while access to evidence, operating availability and specialist dependencies can determine the elapsed time. Maya must agree when staff can participate without undermining current operations, and Alex must identify any technical work required before options can be examined credibly. Lena should build a short dependency-based plan for the study itself, preserving the difference between effort, capacity and duration. The allowance also needs a release rule and owner; it should not become automatic permission to extend scope. If data access fails or a material new question emerges, the team should explain the implications and request an explicit decision to narrow, extend or stop the work. A small investigation still requires the control discipline taught in Chapter 11.

Proportionate stopping and continuation conditions make that discipline practical. The sponsor might require evidence that the observed problem is material, that at least one feasible intervention addresses a credible mechanism and that the operating owners can support further definition. Those conditions should be tailored before results are known, with room to explain evidence that challenges the original framing. An inability to substantiate the preferred option is a legitimate finding, not necessarily a failed study. Equally, identifying a promising concept does not remove the need to examine affordability, dependencies and uncertain benefits in the next stage. Lena should distinguish findings that support immediate operating correction from those that justify further programme development. This helps prevent every useful observation from becoming a reason to establish a larger temporary organisation, while preserving a route for change that genuinely crosses organisational boundaries.

Return with a decision rather than a larger assumption

The Forth Replacement Crossing provides a contrasting example of early investigation. Audit Scotland's 2018 report describes examination of repair and replacement and a peer-reviewed appraisal that began with 65 options (Audit Scotland, 2018, pp. 13–16). The chronology matters: an initial business case supported the December 2007 decision, while the final business case followed in March 2011 after procurement for the principal contract. The case therefore illustrates structured examination of alternatives without implying that every detail was settled before the first commitment. Its lesson for the manufacturer is to make the maturity of the decision material visible. A proposal can support further definition under stated conditions even when later work remains necessary. The controller should explain what has been examined, which choices remain open and what commitment the present evidence can sustain, rather than borrowing confidence from a later, more complete version of the case.

At the return review, the Proposal and Investigation Record should bring the original question back into view. It identifies the evidence obtained, changes to the problem definition, the alternatives examined, the reasons for exclusions and the remaining assumptions. It also records resources consumed and any material limitations in the study. The decision can be to stop, retain an operating response, reframe the question or authorise further definition of a promising transformation. Where conditions attach to continuation, they need owners and dates rather than a general request for more detail. Lena should retain the difference between an agreed finding and the sponsor's judgement about what to do with it. Independent challenge can strengthen that judgement, but it does not transfer responsibility for the commitment away from the person authorised to make it.

The record should also protect the history against gradual rewriting. An ambition may change as evidence develops, and a new intervention may become more credible than the original suggestion. That is a reason to update the proposal with traceable authority, not to present the latest understanding as though it had always been agreed. The initial baseline should remain identifiable alongside corrected data, revised definitions and new forecasts. This allows later reviewers to distinguish learning from avoidable misrepresentation and to understand why the programme took its eventual form. It also supports a proportionate account of study performance. A well-controlled investigation can end without a programme, while a poorly examined proposal can obtain approval. The relevant test is whether the authorised work produced a defensible basis for the decision it was intended to support.

Carry a credible proposition into operating design

When Samira returns to the manufacturer's proposal, Lena can now show the difference between a strategic aspiration, a measured problem and a preferred response. The ten per cent late-payment rate describes a defined invoice cohort. The primary-reason analysis directs investigation without promising that all associated cases can be prevented. Maya can explain the operating mechanisms and ownership questions that need further work, while Alex can compare technical interventions with changes to existing practice. Samira's authorisation can therefore fund a bounded investigation or further definition, with a clear account of the alternatives and the evidence required at the next review. The result is useful even if the original global-platform proposal changes substantially. The control contribution is to make the commitment match the available knowledge and to preserve the organisation's ability to learn before the consequences of its choice become difficult to reverse.

Chapter 12 establishes the first usable connection between strategic intent and governed work. The controller helps define the need, expose the population and evidence behind it, keep alternatives meaningful and connect uncertainty to a bounded inquiry. Resources and authority are attached to the study, while later investment remains subject to a further decision. This approach provides a credible starting point for transformation without pretending that its final design is already known. Chapter 13 develops the next part of the lifecycle: specifying the future operating state. It will connect the emerging proposition with the capabilities, practices, outcomes and responsibilities needed to make the intended change measurable. The investigation carried forward should make that design work better founded, with its assumptions visible and its purpose still connected to the need that justified further commitment.

References

- Audit Scotland. (2018, 2 August). Forth Replacement Crossing. Prepared for the Auditor General for Scotland. audit.scot/uploads/docs/report/2018/nr_180802_forth_crossing.pdf
- Baer, M., Dirks, K. T. and Nickerson, J. A. (2013). Microfoundations of strategic problem formulation. *Strategic Management Journal*, 34(2), 197–214. doi: 10.1002/smj.2004
- Matinheikki, J., Artto, K., Peltokorpi, A. and Rajala, R. (2016). Managing inter-organizational networks for value creation in the front-end of projects. *International Journal of Project Management*, 34(7), 1226–1241. doi: 10.1016/j.ijproman.2016.06.003
- Nutt, P. C. (1998). Framing strategic decisions. *Organization Science*, 9(2), 195–216. doi: 10.1287/orsc.9.2.195
- Samset, K. and Volden, G. H. (2016). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management*, 34(2), 297–313. doi: 10.1016/j.ijproman.2015.01.014
- Victorian Auditor-General's Office. (2024, 20 March). Withdrawal from 2026 Commonwealth Games. www.audit.vic.gov.au/report/withdrawal-2026-commonwealth-games
- Williams, T. and Samset, K. (2010). Issues in front-end decision making on projects. *Project Management Journal*, 41(2), 38–49. doi: 10.1002/pmj.20160

CHAPTER 13

Specifying the Future Operating State



IN THIS CHAPTER

Specifying the capability, behaviour and evidence the future requires.

Design the operating change that the investment must support

The manufacturer's proposal can now support a serious design discussion. Chapter 12 established how to investigate the need without treating the preferred platform as an inevitable answer. We continue that early planning story, before the later programme commitments described elsewhere in this handbook. Samira wants more dependable payments and a defensible investment; Alex needs to understand what must be delivered; Maya must explain how the receiving organisation will work; and Lena must connect their claims with evidence. A future operating state gives those responsibilities a common object. It describes the service the organisation intends to provide, the conditions under which it must work and the arrangements needed to sustain it. The controller's contribution is to make that description testable before it becomes a delivery commitment. This chapter develops a practical specification that can support design challenge, investment appraisal and the later assessment of readiness.

The design begins with the intended operating difference. A configured workflow might make an invoice visible, yet dependable payment also requires accurate information, an available approver, an authorised exception route and a reliable settlement interface. The future state should explain how these arrangements work together for the defined population. It should also identify what remains unchanged, because existing policies and services may provide essential conditions for success. A programme cannot claim the whole payment result while leaving these dependencies outside its reasoning. Equally, it need not rebuild every contributing service. Maya can negotiate an explicit service obligation with an existing team, while Alex includes the relevant interface and Lena records the evidence needed to rely on it. The resulting boundary is organisational as well as technical: it follows the work far enough to explain the promised improvement and to identify where responsibility passes between owners.

The transformation lifecycle

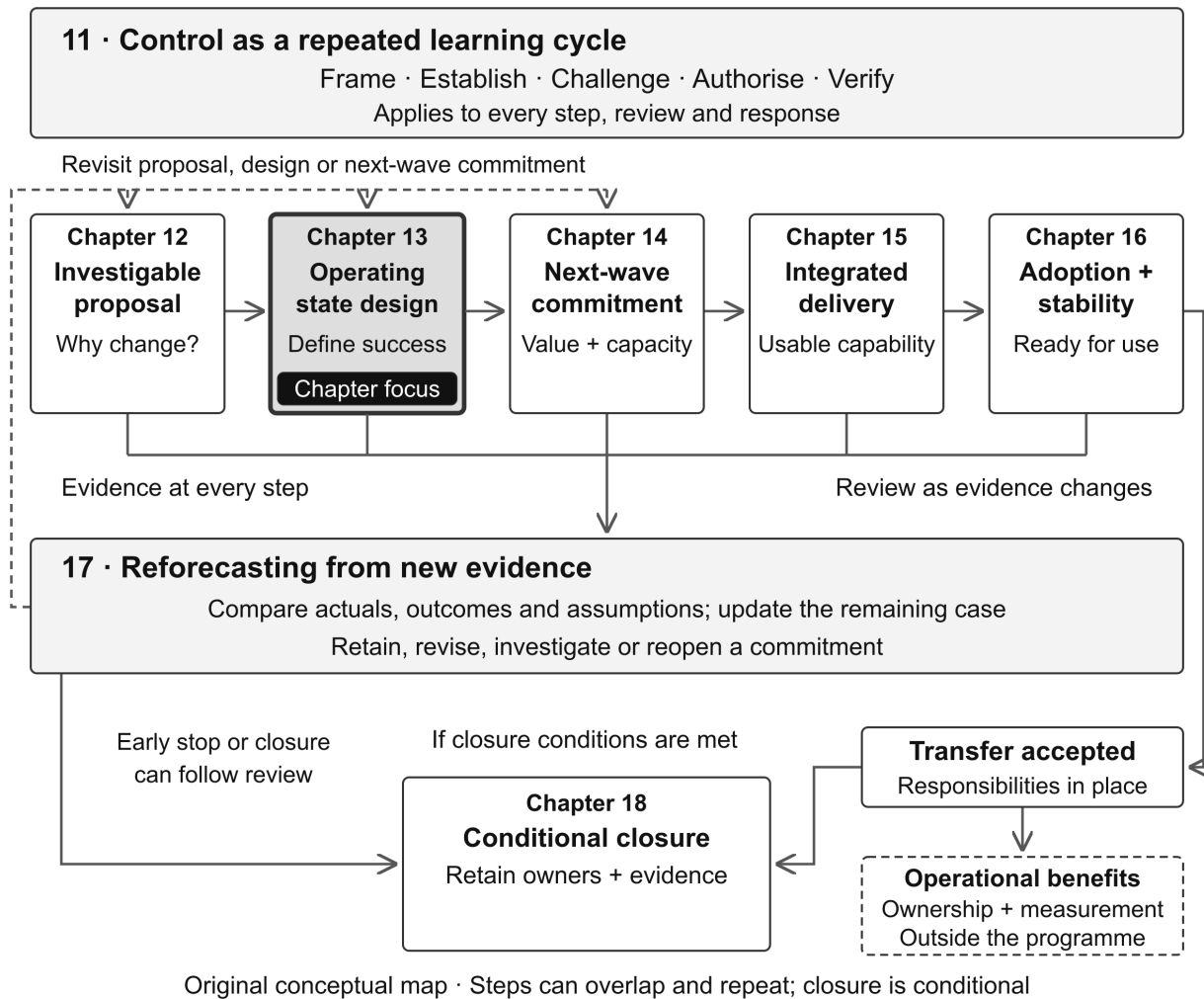


Figure 13.1 The shared lifecycle highlights the future operating state. The investigated proposition becomes a testable design. Evidence can reopen the proposal, design or next commitment, while readiness and continuing benefit evaluation remain connected.

The existing Value and Evidence Map provides the organising logic. Outputs enable usable capability; people employ that capability in changed practice; operating outcomes may improve; and those improvements acquire value against the agreed objectives. Each connection retains its assumptions. For Chapter 13, we propose a Future State Evidence Specification as a practical extension of that map, not an empirically validated instrument. It records the relevant population, operating conditions, required behaviour, accountable owner, evidence and unresolved dependency for each material claim. The Credible Control Method gives the work its sequence: Frame the intended difference, Establish the design and reference conditions, Challenge its feasibility, Authorise a bounded design position and Verify it through suitable evidence. These activities repeat as knowledge improves. A specification accepted for investment planning may still contain conditions to be demonstrated before operational use, and its approval must preserve that distinction.

Describe the work under realistic operating conditions

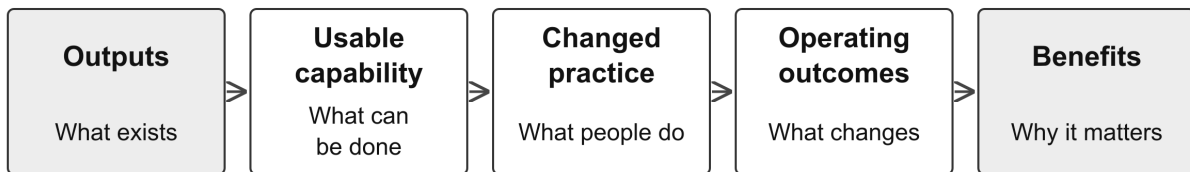
Clegg (2000) treats social and technical arrangements as interdependent aspects of work-system design. His synthesis of sociotechnical principles draws on earlier literature and experience, offering guidance for design rather than evidence from a controlled implementation trial. The relevance to the manufacturer lies in examining the interactions among people, processes, information, technology and organisational arrangements already introduced in Chapter 5. A routing rule can be technically correct while assigning work to someone without the authority, information or time to resolve it. That example applies the principles to our fictional service; it is not a finding reported by Clegg. The practical task is to design the whole operating episode with the people who understand it. A sociotechnical label cannot guarantee that the resulting arrangement will work, and its assumptions still require examination through appropriate evidence.

Maya starts by walking through a complete payment journey with the people who perform it. The population includes invoices arriving through supported channels, including permitted manual routes, and the operating boundary extends to confirmed settlement and reconciliation. The team identifies ordinary work, exceptions and recovery from interruption. For an invoice lacking an order reference, the future design states who investigates, what information that person can access, how the supplier is contacted and what happens when the contractual deadline approaches. Those choices affect job design, information permissions, staffing and service performance. Lena records them at a level that allows the next decision to be made; she does not attempt to document every keystroke. An observable operating condition is more useful than a vague instruction to improve collaboration. It tells the delivery team what capability must exist and gives the operating owner something concrete to examine.

Interfaces deserve particular attention because local completion can conceal incomplete service. Suppose procurement classifies an exception as resolved when it supplies an order reference, while finance uses the same status to mean that payment has been reconciled. A shared dashboard can then report completion while money has not moved. The specification should retain distinct events, identifiers and timestamps, together with the owner who confirms each transition. An interface agreement can state what is transferred, how receipt is acknowledged, how duplicates are detected, and who acts when acknowledgement does not arrive. It should also preserve the case through handoffs so that a supplier does not disappear from the operating population merely because responsibility moved. Lena can test this agreement with a worked transaction rather than relying only on signatures. The exercise reveals whether the teams share an understanding of completion and whether their records support that understanding.

Specify the operating state and how it will be tested

Practical aid within the Value and Evidence Map



Apply a common evidence basis across the map



Arrows represent hypotheses: outputs do not automatically cause benefits.

Retain the assumptions and the evidence needed to support each connection.

Figure 13.2 Original specification aid within the Value and Evidence Map. Every proposed connection needs defined conditions, ownership and evidence. A clear chain makes the argument examinable; it does not establish that the intended benefit will occur.

Preserve the connection between need design and evidence

Nuseibeh and Easterbrook (2000) connect requirements with the problems they are intended to address, their rationale and the consequences of change. Their conceptual roadmap distinguishes checking a description's coherence from examining its correspondence with stakeholders' real-world needs. This helps explain why a consistent catalogue can still specify an inadequate service. In engineering terms, verification against stated requirements and validation for intended use answer different questions; the handbook's broader Verify activity includes both and subsequent outcome evidence. Scenarios and prototypes can expose misunderstandings before a requirement becomes expensive to change. Lena should therefore preserve the origin and rationale of a requirement as well as its connections to delivery and testing. Trace links make the argument examinable, but their existence does not establish that the original need was understood correctly or that the proposed response remains suitable.

A useful specification makes that connection visible without becoming an oversized requirements catalogue. Table 13.1 illustrates four linked conditions for the fictional service. The entries are proposed design commitments, not test results. Each identifies an operating need, the circumstances in which it matters, the evidence to be obtained and the person responsible for interpreting it. Detailed scripts, data definitions and decision records can remain in their working systems, connected by stable identifiers. The table should distinguish evidence needed before use from evidence that can only arise during operation. A configured authorisation control can be tested before release; dependable performance across a mature invoice cohort requires subsequent observation. Lena should challenge any entry whose owner lacks authority over the relevant work or whose evidence date comes after an irreversible commitment that depends on it. The remedy may be to change the design, obtain earlier evidence or narrow that commitment.

Table 13.1 Proposed entries in the Future State Evidence Specification

Operating need	Service condition	Evidence and timing	Owner
R1 Payment identity	Lost acknowledgement and retry	Before use: one instruction reconciled to one settlement, with an audit trail	Alex
R2 Work allocation	Usual approver absent	Before use: authorised deputy resolves the case without bypassing control	Maya
R3 Supported access	Permitted manual channel	Before use: case identity and deadline preserved across handoffs	Maya
R4 Payment outcome	Mature eligible invoice cohort	After use: reconciled lateness and adverse effects under the Measurement Agreement	Maya

Illustrative design entries, not completed tests. Alex owns the technical demonstration; Maya accepts its operating implications. Detailed definitions, scripts, records and any finance review remain linked to each entry.

Burton-Jones and Grange (2013) develop a theoretical account that distinguishes using an information system from using it effectively in relation to its purpose. Their contribution does not establish a universal acceptance threshold, but it helps explain why access statistics alone are insufficient evidence for the intended operating state. Staff may log in frequently while continuing to resolve difficult invoices outside the supported process. In the manufacturer's design, a demonstration should therefore examine whether people can complete relevant work with the information and authority available, including exceptions that matter to suppliers. That is an application of the theoretical distinction, not a test prescribed by the paper. Successful performance in a rehearsal supplies evidence about the demonstrated tasks and conditions. It does not establish sustained adoption, attributable improvement or realised financial benefits, all of which require further observation and judgement.

Test whether the service can carry its workload

Capacity makes the operating description numerically examinable. Consider a separate fictional design assumption for an exception service: 48 cases arrive per working day, each requires twelve minutes of handling, and two trained people each provide six productive hours daily. These assumptions are not derived from Chapter 12's late-invoice cohort; exceptions and late invoices are different populations. Daily handling demand is 576 minutes, or 9.6 hours, against twelve productive staff-hours. Nominal capacity is therefore sixty cases daily, with an average load equal to eighty per cent of that capacity. Table 13.2 retains the units and the staffing assumptions. The calculation establishes internal arithmetic consistency; it does not establish acceptable waiting times. Productive hours already exclude other duties, and any further allowance must avoid counting those deductions twice. Maya needs evidence that the assumed handling time, skills and availability represent the work the team will actually receive.

Table 13.2 Check staffing assumptions against a consistent daily workload

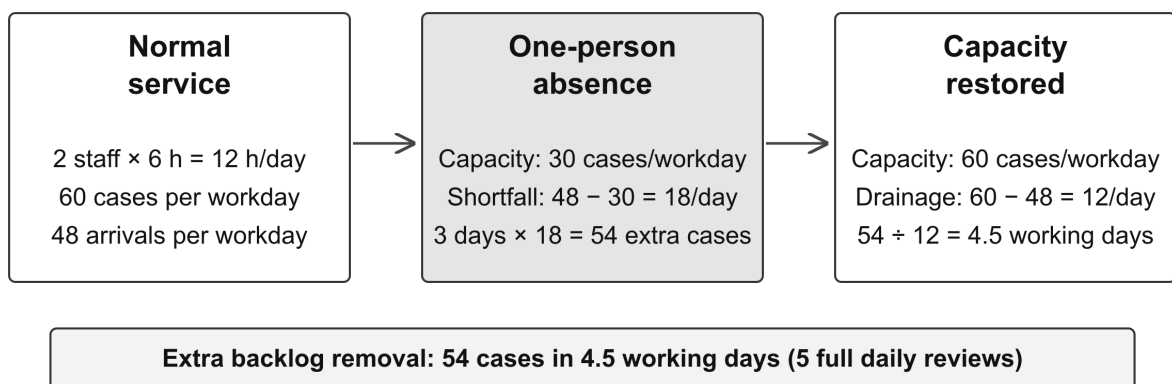
State	Productive staff hours per day	Capacity cases per day	Arrivals less capacity cases per day
Normal staffing	12	60	-12
One person absent	6	30	+18
Staffing restored	12	60	-12

Fictional arrivals of 48 cases per working day and twelve minutes of handling per case. Negative balances drain existing backlog only; they do not create negative queues. Constant rates exclude variability, additional rework and further lost time.

An absence scenario exposes a dependency that the average conceals. With one person unavailable for three working days, capacity falls to thirty cases daily. At the assumed arrival rate, eighteen additional cases accumulate each day, producing 54 extra cases by the end of the absence. Restoring both people provides twelve cases of spare daily capacity after new arrivals, so removing that additional backlog requires 4.5 working days under ideal continuous arithmetic, or completion during the fifth full daily review interval. Figure 13.3 makes the sequence visible. This is a deliberately simplified calculation with constant arrivals and handling time, no additional rework and no loss of productive hours. The design response could involve trained cover, a changed service boundary or another means of managing demand. Lena should ask for the consequences of each response, including displaced work elsewhere, before treating recovery capacity as available.

Test capacity when assumed staffing changes

New fictional exception service · 48 arrivals/workday · 12 minutes/case



Idealised constant-rate arithmetic, not a demonstrated forecast.

Assumes productive hours, constant case effort and immediate restoration; no extra rework.

Figure 13.3 Fictional capacity stress test. Three days with one person absent add 54 cases. Restored staffing removes that extra backlog in an ideal 4.5 working days, with clearance first visible at the fifth daily review. The calculation restores the prior position rather than proving zero total open work.

Little (1961) provides a separate consistency check by relating the mean number of items in a system to their mean arrival rate and mean time within it, under specified stationary conditions. If the same stable service has a long-run mean of 72 open cases and throughput of 48 cases per working day, mean residence time is 1.5 working days. Open cases here include those waiting and being handled. These are matched averages, not a snapshot backlog divided by nominal capacity. A one-day mean at unchanged throughput would correspond to 48 cases on average, but imposing a limit of 48 does not prove that throughput will be maintained. The relationship cannot establish a percentile deadline or predict the temporary absence scenario. It helps Lena expose inconsistent assumptions while preserving the need to examine the actual behaviour of the proposed service.

Service promises require evidence beyond those means. A mean residence time of 1.5 working days is compatible with very different experiences for individual cases, particularly where a small group requires specialist approval. A proposed condition that ninety-five per cent should finish within two working days needs a defined clock, eligible population and treatment of weekends, reopened cases and missing records. It also needs evidence about variability, priorities and the queues through which cases pass. The team can rehearse representative demand patterns, use historical arrivals where suitable and examine difficult cases separately. Any simulation should document its distributions, routing and resource assumptions and be checked against observable operating behaviour. A favourable result remains conditional on those assumptions. The controller should retain both the overall measure and the experience of relevant subgroups so that a rapid standard route does not conceal an unworkable exception route.

Specify acceptance through demanding operating situations

Acceptance design translates those conditions into demonstrations capable of changing a decision. Maya and Alex can construct scenarios that start with an identifiable invoice and follow it through receipt, routing, authorisation, settlement and reconciliation. Each scenario states the starting data, people available, event sequence, expected result and evidence retained. Some should examine ordinary work; others should include an unavailable approver, incomplete information, repeated messages and an interrupted external connection. The point is to select situations that challenge consequential assumptions, rather than accumulate a large count of easy scripts. Lena traces each scenario back to the operating need and forward to the person who can judge its result. A technical test report may supply part of the evidence, but the receiving owner must be able to explain whether the demonstrated service supports the intended work under the tested conditions.

Consider a payment instruction whose acknowledgement is lost. A retry should not create a second payment, and the operator should be able to determine the transaction's state without an uncontrolled workaround. The specification can require a defined recovery route, an identifiable audit history and reconciliation of the eventual settlement. The test should examine what happens to the same case across the relevant systems, including the failure and recovery sequence. A failed critical scenario needs a documented disposition before the dependent use is authorised. Successful execution of a limited set of scenarios does not prove that all future

failures are impossible, so continuing monitoring remains part of the design. A readiness percentage can be misleading here: many passed convenience tests should not offset one unresolved payment-authorisation failure. Samira needs a decision that names the material condition and its consequence, with authority and ownership attached to any permitted restriction.

A GAO review of preparations for the 2010 US Census examined handheld computers used in the 2007 address-canvassing dress rehearsal. Field evidence identified transmission and device problems, while the Bureau's assessment reported inadequate help-desk responses. GAO found few expected-performance benchmarks and said the June 2008 testing plan did not specify how or when operational readiness would be determined (United States Government Accountability Office, 2008, pp. 12–13, 16, 19, 21–22). Its recommendation distinguished operational needs from contractor product acceptance. The design implication is to specify field tasks, data transfer, support responses and measurable performance together. This is our application of the findings, not a GAO prescription for the handbook's method. The report assessed preparations at that time; it did not determine the accuracy of the completed 2010 Census or establish that every subsequent operating problem followed from the identified gaps.

Design for local variation and continuing benefit evidence

Greenhalgh and colleagues (2004) synthesise diverse research on innovation in service organisations, particularly health care. Their review distinguishes a generally receptive organisation from readiness for a particular innovation and examines the fit with skills, resources, technologies and working practices. It does not establish a universal readiness threshold or a manufacturing effect. For Maya, its useful implication is to investigate how the proposed process will fit the receiving setting. Users need to perform ordinary and exception work, support must be available when required, and the organisation needs a means of recognising unintended consequences. The operating tests proposed here apply that reasoning without claiming validation by the review. A successful rehearsal can support a bounded judgement under its observed conditions. Whether the practice becomes routine, remains workable and contributes to the intended benefit requires continuing evidence after release.

An international operating state should distinguish necessary consistency from justified variation. A common definition of confirmed settlement may be essential for comparable reporting, while supported contact channels, local working calendars and approval arrangements can differ. The specification should identify the invariant requirement and the permitted variation, with an owner for assessing whether a local design still meets the underlying need. Otherwise, standardisation can suppress information that matters, while unrestricted localisation can make the service impossible to govern. Maya might accept two support channels provided both preserve case identity, access and response obligations. Lena should record the reason for the variation and its effect on evidence, capacity and value. The assessment should include people who may struggle with the preferred channel, because an apparently efficient process can shift effort or exclusion onto suppliers whose experience is absent from the programme's usual reporting.

The benefit claim also needs a route through time. Chapter 12's ten per cent late-payment baseline remains attached to its own defined cohort; a future target belongs to a specified comparable population and observation period. The existing Measurement Agreement should preserve the definitions, data and ownership needed to judge that target, while distinguishing observed improvement from attribution to the programme. It should also retain adverse effects such as duplicate payments, supplier effort and additional operating cost. A faster average accompanied by more costly rework may change the value judgement. Where staff time is released, finance and Maya must explain whether it becomes additional capacity, avoids future expenditure or supports an actual cash reduction. These are different benefit mechanisms. Lena connects each to the relevant operating change and later review, making clear which evidence exists at design approval and which remains a future obligation.

English hospital e-prescribing research illustrates the importance of fit with real work. In a programme reported by Sheikh and colleagues (2022, scientific summary), six longitudinal qualitative hospital cases examined procurement, implementation and use. Researchers found that implementation took longer than teams expected, particularly around integration and interfaces. Limited configurable processes in packaged systems did not always fit hospital procedures, prompting customisation requests and workarounds. These observations suggest that a future-state design should expose process assumptions, boundaries and exception handling before defining acceptable operation. That is our interpretation, rather than a tested prescription from the study. The cases were historical research published in 2022, not a contemporaneous assessment of all English hospitals. They do not establish that every workaround was harmful or quantify a general causal effect on patient safety. The relevant lesson concerns the questions the design must make examinable.

Govern the design before committing the next wave

Design control begins by making the agreed position identifiable. The Future State Evidence Specification should have a version, scope, assumptions, owners and a decision record. Its acceptance for planning means that it is a sufficient basis for the next bounded commitment, with remaining conditions visible; it does not mean that every future operating result has been demonstrated. Samira authorises the commitment, Maya accepts the operating obligations she can support, and Alex confirms the delivery implications. Lena reconciles the evidence and exposes differences that require a decision. Unresolved questions should have a consequence as well as an owner and date. An uncertain interface can affect the sequence of a wave, while an unconfirmed operating resource may constrain the eligible population. Recording both as generic actions obscures their different implications. The specification becomes useful when those implications reach the person who can change scope, resources or timing.

Traceability then supports change assessment in both directions. Suppose a proposed amendment allows an exception to bypass an approval step to improve speed. Starting from that design change, Lena can identify affected authorisation requirements, operating responsibilities, tests and benefit assumptions. Starting from the payment-control objective, she can check whether another arrangement still provides the necessary protection. The assessment should consider consequences across the service rather than accepting a local reduction in handling time

as a sufficient justification. Where evidence is incomplete, the response may be a restricted trial or further investigation. The earlier design remains available so that later reviewers can understand what was authorised and why the position changed. This is particularly important when a supplier proposes a simpler technical implementation whose consequences appear in operations long after the delivery team has counted the associated work as complete.

The same reasoning makes scope negotiation more transparent. Some requirements protect the integrity of the service and cannot sensibly be traded for convenience. Others concern performance levels or features for which alternative arrangements may be acceptable. The team should explain the consequence of omission rather than rely on labels such as essential or desirable without supporting reasons. A requirement to preserve payment identity through a retry addresses a different exposure from an additional dashboard view. Samira can compare options when those differences are clear, while Maya identifies the continuing work that a deferred feature would create. Lena should retain the cost and capacity of that work in the investment appraisal. A smaller delivery package can still require more operating support, and a design that appears affordable within the project budget can transfer an unexamined commitment to the receiving organisation.

Before carrying the design into Chapter 14, the team reviews whether its principal claims can be followed from the investigated need to the proposed operating behaviour and planned evidence. The review examines the service boundary, critical interfaces, capacity assumptions, required practice and the ownership of continuing measurement. Its depth should reflect the exposure of the next commitment. A limited, reversible wave may proceed with some matters reserved for investigation, provided those matters do not undermine the immediate decision and have explicit containment. A larger commitment may require stronger demonstrations. Lena should distinguish disagreement about evidence from disagreement about priorities: the former may require investigation, while the latter requires an authorised choice. The review ends with a recorded design position and conditions for use, amendment or further development. It does not end merely because the participants have attended a workshop or completed a template.

Carry an accountable operating state into investment planning

Samira can now see the manufacturer's future service as a set of connected obligations. Alex knows which capabilities and interfaces must be delivered, Maya can identify the practices and capacity she must sustain, and Lena can show how evidence will support the relevant decisions. The absence calculation has exposed a recovery assumption that an average staffing figure concealed. The interface scenarios have clarified what a completed payment means and how the organisation will respond when acknowledgements fail. The proposed evidence specification connects those findings with the Value and Evidence Map without claiming that a convincing design has already produced a benefit. Where the design remains uncertain, the uncertainty has a location, an owner and a consequence for the next commitment. That is enough to make further investment planning more disciplined while leaving room for the operating state to improve through testing and use.

Chapter 13 turns the investigated proposition into an accountable description of how the organisation is expected to work. It carries forward the need and value claims, makes operating conditions explicit and attaches evidence to capabilities, practices and outcomes. This gives project controls a basis for challenging scope that is locally complete but operationally insufficient, as well as scope whose contribution to value is unclear. Chapter 14 places this design within the next wave of investment. That decision must connect the intended operating difference with feasible scope, schedule, cost, resources and exposure. The specification developed here supplies the service conditions and dependencies that such a commitment must respect. It will remain open to authorised change as evidence develops, with the original decision preserved and continuing responsibility for benefits visible beyond the work of the delivery team.

References

- Burton-Jones, A. and Grange, C. (2013). From use to effective use: A representation theory perspective. *Information Systems Research*, 24(3), 632–658. doi: [10.1287/isre.1120.0444](https://doi.org/10.1287/isre.1120.0444)
- Clegg, C. W. (2000). Sociotechnical principles for system design. *Applied Ergonomics*, 31(5), 463–477. doi: [10.1016/S0003-6870\(00\)00009-0](https://doi.org/10.1016/S0003-6870(00)00009-0)
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P. and Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly*, 82(4), 581–629. doi: [10.1111/j.0887-378X.2004.00325.x](https://doi.org/10.1111/j.0887-378X.2004.00325.x)
- Little, J. D. C. (1961). A proof for the queuing formula: $L = \lambda W$. *Operations Research*, 9(3), 383–387. doi: [10.1287/opre.9.3.383](https://doi.org/10.1287/opre.9.3.383)
- Nuseibeh, B. and Easterbrook, S. (2000). Requirements engineering: A roadmap. *Proceedings of the Conference on the Future of Software Engineering*, pp. 35–46. ACM. doi: [10.1145/336512.336523](https://doi.org/10.1145/336512.336523)
- Sheikh, A., Coleman, J., Chuter, A., Williams, R., Lilford, R., Slee, A., Morrison, Z., Cresswell, K., Robertson, A., Slight, S., Mozaffar, H., Lee, L., Shah, S., Pontefract, S., King, A., Wiegel, V., Watson, S., Salema, N.-E., Bates, D., Avery, A., Girling, A., McCloughan, L. and Watson, N. (2022). Electronic prescribing systems in hospitals to improve medication safety: a multimethods research programme. *Programme Grants for Applied Research*, 10(7). NIHR Journals Library. doi: [10.3310/KSRS2009](https://doi.org/10.3310/KSRS2009)
- United States Government Accountability Office. (2008, 31 July). 2010 Census: Census Bureau's Decision to Continue with Handheld Computers for Address Canvassing Makes Planning and Testing Critical. GAO-08-936. www.gao.gov/assets/gao-08-936.pdf

CHAPTER 14

Committing the Next Wave of Investment



IN THIS CHAPTER

Committing resources to a coherent and achievable next wave.

Make the next commitment a coherent decision

The manufacturer's future operating state now describes what must work, for whom and under which conditions. Chapter 14 turns that design into a bounded investment decision. We remain in the early planning story developed in Chapters 12 and 13, before the later programme commitments described elsewhere. Samira needs to know what authority she is being asked to exercise, Alex needs a deliverable scope and sequence, Maya needs operating obligations that her teams can sustain, and Lena must reconcile the evidence connecting them. Approving a wave creates commitments to money, people, suppliers and service change. Those commitments should describe the same intended result. A schedule based on unavailable specialists, or a budget that excludes the receiving team's work, cannot provide a coherent basis for authorisation. The chapter shows how to expose these differences before approval and how to state the conditions under which a commitment remains valid.

A wave should be large enough to produce an intelligible operating change or answer a consequential learning question, while limiting exposure that the organisation is not yet ready to accept. Its boundary therefore concerns more than expenditure during a reporting period. The decision should identify the operating population, capability, interfaces, acceptance conditions and subsequent evidence obligation. It should also distinguish commitments made now from choices reserved for later, including any contract that could make a nominally deferred choice costly to reverse. Samira may reasonably approve a limited release while leaving expansion conditional on experience. That decision becomes meaningful only when the organisation can identify what has actually been committed and what remains avoidable. Lena should examine the practical consequences of authorisation across the whole service, including transition support and obligations that continue after the delivery team reports completion.

The transformation lifecycle

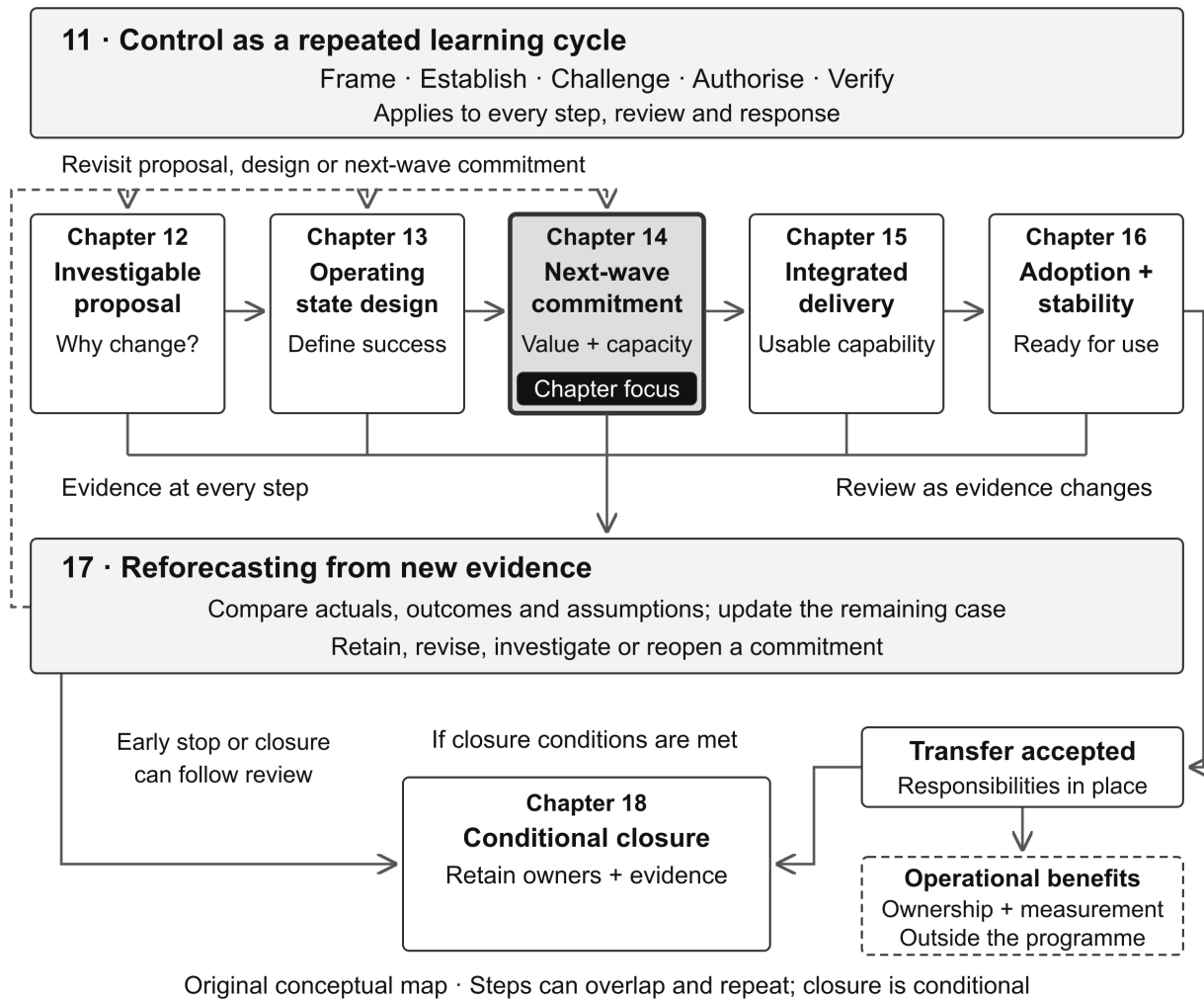


Figure 14.1 The shared lifecycle highlights the next wave commitment. The proposed operating state must support a bounded, affordable and resource-feasible authorisation. Delivery and operating evidence can reopen that commitment as knowledge changes.

The Wave Commitment Record introduced in Chapter 7 remains the principal working record. Here it brings together the Future State Evidence Specification, the Sequencing and Benefits Map and the Decision under Uncertainty Memorandum, rather than replacing them with another framework. These connected records form the integrated commitment brief presented for authorisation. The Credible Control Method organises the work: Frame the proposed wave, Establish a consistent plan and estimate, Challenge resources and exposure, Authorise stated commitments and Verify their subsequent performance. The record is a practical aid, not an empirically validated instrument. Its value depends on the quality of the evidence and the willingness to act when that evidence changes. The detailed models can remain with their owners, provided the decision record identifies their versions, assumptions and implications clearly enough for a later reviewer to reconstruct why this particular commitment was made.

Bound the operating result and the resources it requires

Cooper (2008) describes stages that develop information and gates that decide whether further resources should be committed. His perspective on product innovation addresses weak gate authority, excessive bureaucracy and the need for adaptable processes; it is not a controlled evaluation of transformation programmes. Applied here, Samira's approval should identify the next bounded commitment, the evidence supporting it and the people authorised to release resources. Passing a meeting date should not itself release the entire programme budget. A decision may require further work, defer commitment or discontinue an unsuitable proposal. The chapter applies this reasoning through the existing Wave Commitment Record, rather than adopting a product-development lifecycle. A gate can support learning and reassessment without making the work rigidly sequential. Its practical value depends on consequential decisions, rather than documentation alone.

For the manufacturer, the immediate scope might cover a defined supplier population and supported invoice routes, together with the interfaces, exception handling and controls required to operate them. A list of configuration tasks would omit part of that service. Alex and Maya should identify the work needed to demonstrate the Chapter 13 conditions, including usable information, available approvers, support and recovery arrangements. The wave must also state its exclusions and what those exclusions mean. Deferring a reporting enhancement may be acceptable if an agreed manual arrangement supplies the required evidence; deferring payment identity across retries would affect a critical control. Lena records the resulting obligations and their resource implications. Where an operating team supplies effort, the plan needs an agreed allocation and timing, rather than an assumption that the work will fit around existing duties without consequence.

Blichfeldt and Eskerod (2008) investigated portfolio management through 128 in-depth interviews in thirty companies. Their qualitative findings identify a practical boundary problem: projects outside the formally managed portfolio still consume resources promised to recognised projects. This does not establish a universal effect size or an optimal number of simultaneous initiatives. It does, however, challenge a capacity plan that counts only the work visible at the approval meeting. Before committing the next wave, Maya and Alex should identify other demands on the same specialists, including smaller changes and continuing operations. The handbook's application is to reconcile those demands with named availability and explicit priorities. A resource allocation is credible only when the organisation can honour it, or identifies what must move. Portfolio inclusion itself does not make scarce capacity expand automatically.

Reconcile the sequence with the estimate

Consider a new fictional candidate wave, separate from the numerical examples in earlier chapters. Table 14.1 contains six activities with a common working-day calendar and finish-to-start dependencies. Confirming the operating design takes three days. Interface configuration takes five days and data preparation four, both following design confirmation. Integrated testing takes three days after both are complete, followed by two days of operating rehearsal and one day for controlled release. Starting at day zero, the dependency network alone gives fourteen working days: three, then the longer of five and four, then three, two and one. This is a useful first calculation because its logic is visible. It remains conditional on resources and on what each completion event means. The durations assume the specified work and acceptance evidence; removing required tests to preserve the date would change the scope rather than improve the same plan.

Table 14.1 Establish the common scope for sequence and cost

Activity	Predecessor	Working days	Direct cost US\$000
A Confirm operating design	None	3	12
B Configure interface	A	5	30
C Map and prepare data	A	4	16
D Integrated test	B and C	3	24
E Operating rehearsal	D	2	8
F Controlled release	E	1	5

Fictional finish-to-start logic with no lags, task splitting or multitasking. B and C need the same single specialist full time; all other resources are assumed available. Direct costs total US\$95,000 and exclude separate coordination at US\$1,000 per working day.

The configuration and data activities both require the same specialist full time. Their apparent overlap is therefore infeasible under the stated resource limit. Sequencing configuration first places it between days three and eight, followed by data preparation between eight and twelve. Testing then runs to day fifteen, rehearsal to seventeen and controlled release to eighteen. Figure 14.2 compares the two plans. Reversing the two specialist activities gives the same duration in this simplified case, but a different dependency or availability window could change the choice. All other required resources are assumed available, and no task splitting or multitasking is allowed. Lena should retain those assumptions when presenting the eighteen-day plan. Obtaining a second specialist would constitute a different option, requiring evidence about availability, competence, coordination and cost; it should not be inserted silently to make the earlier date appear achievable.

Include the shared specialist in the wave commitment

Logic: $A \rightarrow B$ and $C \rightarrow D \rightarrow E \rightarrow F$ · B and C need the same specialist

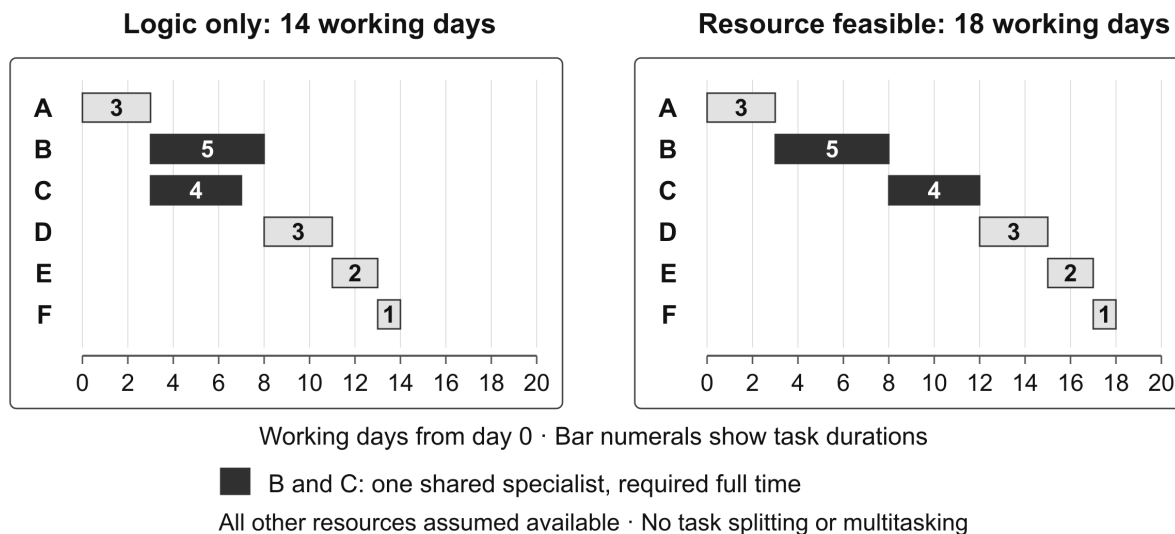


Figure 14.2 Fictional wave sequence. Logic alone permits fourteen working days, but B and C require the same specialist full time. Sequencing those activities gives eighteen days with the stated resources. Additional resource would be a different option requiring its own evidence.

The resource correction also changes the cost basis. The six activities have direct resource costs totalling US\$95,000. Wave coordination costs US\$1,000 for each working day from start through controlled release and is excluded from those activity amounts. A fourteen-day assumption would therefore give US\$109,000, while the feasible eighteen-day sequence gives US\$113,000. This example counts the specialist's planned activity work once and adds only the extra duration-dependent coordination cost. A live estimate would examine other affected costs, calendars and contractual provisions. The lesson is to reconcile the cost model with the sequence before discussing an allowance for uncertainty. The four-day difference is an identified resource constraint in the base plan, not a probabilistic event that can be hidden in contingency. Adding money without correcting the infeasible overlap would leave the investment decision supported by incompatible assumptions.

Jørgensen and Sjøberg (2004) studied customer expectations in a controlled estimation experiment involving thirty-eight students and twelve software professionals. Information about the customer's expected cost influenced estimates even when participants were told that it was not valid estimating input. The experiment concerned a specified software task, not the realised performance of transformation programmes. Its relevance is to preserve a distinction between an evidence-based estimate and the amount a sponsor hopes to spend. Lena can record an initial scope-based assessment before discussing a preferred ceiling, then examine feasible changes where affordability differs. Genuine constraints still belong in the decision; shielding judgement from an irrelevant anchor does not mean concealing required scope. An independently prepared estimate also needs transparent assumptions, historical comparisons and challenge before becoming the basis of an authorised commitment.

Challenge the funding and time limits together

Suppose Samira is considering a resource-cost ceiling of US\$125,000 and a completion limit of twenty working days. These are proposed authorisation limits for this example, not estimates generated by a confidence calculation. The US\$113,000 base leaves US\$12,000 of controlled allowance, while the eighteen-day sequence leaves two working days before the limit. A mapping-rework scenario adds two days and US\$6,000 of direct work. Its additional coordination costs US\$2,000, producing twenty days and US\$121,000 in total. A separate environment-delay scenario adds four days and US\$4,000 of direct cost, plus US\$4,000 of coordination. It also totals US\$121,000 but finishes at twenty-two days. Table 14.2 keeps the two dimensions together. Equal total cost does not mean equal acceptability: the second scenario stays within the funding ceiling while exceeding the proposed completion limit.

Table 14.2 Examine each scenario against both proposed limits

Scenario	Working days	Total cost US\$000	Numerical limits
Base plan	18	113	Both met
Mapping rework	20	121	Both met
Environment delay	22	121	Time exceeded
Both events	24	129	Both exceeded

Proposed limits are twenty working days and US\$125,000. The combined case assumes both events occur without overlap. Scenarios have no assigned probabilities; meeting these numerical limits does not replace resource and operating acceptance conditions.

If both events occur sequentially without overlap, the added duration is six days and the extra direct work costs US\$10,000. Coordination then totals US\$24,000, producing a twenty-four-day wave costing US\$129,000. Both limits are exceeded, as Figure 14.3 shows. These are deterministic teaching scenarios, not equally likely states, an exhaustive risk model or a basis for claiming a success probability. The same disruption can affect time, direct work and support cost together, so assessing separate cost and schedule margins would miss part of the exposure. Lena should show which assumptions cause the commitment to fail and what response remains feasible. A response might change scope, secure an alternative environment or seek different limits, but its consequences require analysis. The controller should not treat the unspent allowance as evidence that every delay or design change can be accommodated.

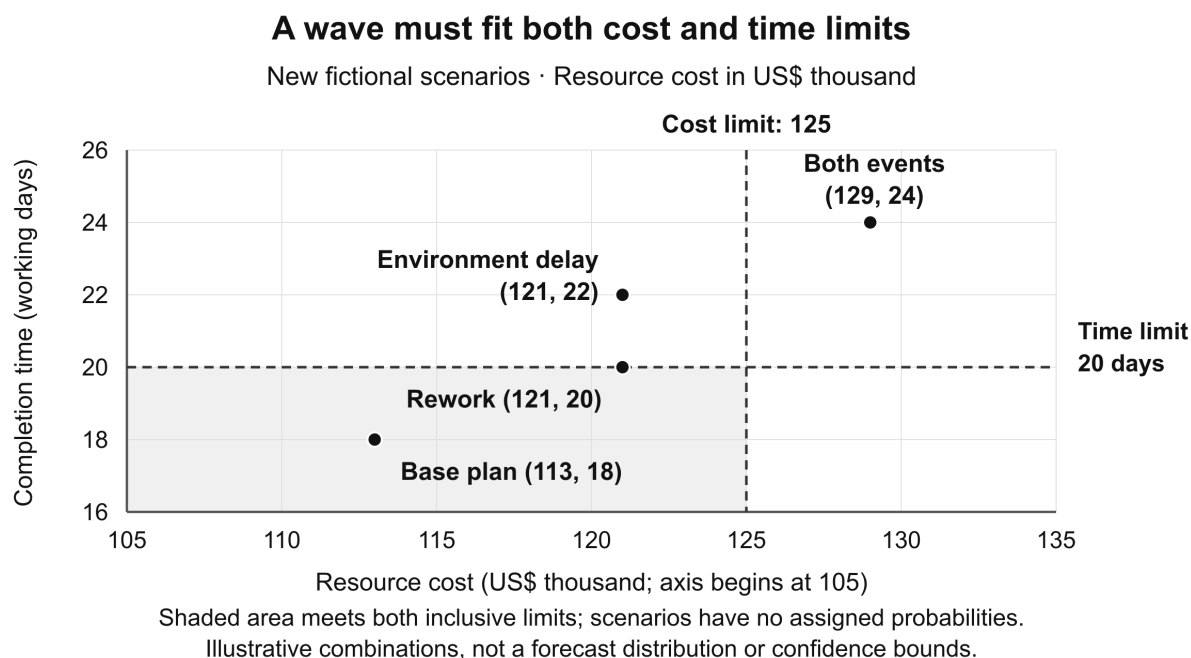


Figure 14.3 Original comparison of illustrative cost and time limits. Base and rework scenarios meet both numerical limits; environment delay exceeds time, and both events together exceed both limits. These tests do not establish overall readiness or a probability of successful delivery.

Touran (2003) develops a probabilistic cost-contingency model using assumptions about the frequency and cost of project changes, with numerical illustration informed by construction records. Its modelled probabilities depend on those assumptions; it does not establish a standard reserve percentage for transformation programmes. The useful question for Lena is what exposure the allowance is intended to cover and how its proposed size was derived. Planned work, explicit risk responses and uncertainty already included in component estimates need consistent treatment so that the same exposure is not funded repeatedly under different labels. The US\$12,000 in our example is an illustrative limit, not an application of Touran's model or a statistically justified reserve. Authority to release it is a separate governance choice. A reserve can support an authorised response while leaving a changed completion date, scope or operating condition subject to further approval.

Preserve the choices that later evidence should inform

Resource cost, cash availability and contractual commitment answer different questions. The US\$113,000 estimate uses a consistent valuation of the example's resources; it is not a payment schedule. Internal effort may consume scarce capacity without creating the same amount of new cash expenditure, while a supplier deposit may require cash before the associated work is performed. The authorisation brief should therefore reconcile the cost basis, expected payment timing and obligations created by orders or contracts. It should identify cancellation terms and any later work committed in advance. Lena does not need to reproduce every contract clause in the record, but material exposure must be visible to the decision maker. A wave described as small can still carry substantial future liability if its procurement arrangements commit expansion before the evidence intended to justify that expansion can be observed.

Huchzermeier and Loch (2001) model managerial flexibility in research and development projects, including continuation, abandonment and corrective action. Their analytical results show that more uncertainty does not always make flexibility more valuable; the effect depends on the source and timing of uncertainty and whether decisions can still respond. This is a model-based insight, not a measured return from staging a transformation. For Samira, the practical task is to preserve a specific choice that later evidence can inform. Lena should identify the expected information, the commitment still avoidable at that time and the cost of retaining or exercising the choice. Dividing expenditure into approvals will not create useful flexibility if contractual obligations have already removed the alternative. The assessment must concern available actions and their consequences, including the cost of waiting, rather than the existence of another review meeting.

NASA's Inspector General described milestone-based development agreements alongside separate commercial resupply contracts in a 2013 audit. The auditors found that resupply payments to Orbital had not been adjusted promptly as launch dates slipped and recommended aligning work plans, payments and authorised lead times with the revised schedule. NASA agreed with the recommendation while disputing the auditors' assessment of financial risk. Officials also reported that authority to proceed with Orbital's seventh resupply mission had been delayed (NASA Office of Inspector General, 2013, pp. 2–4, 15, 23–25). For our wave, the lesson is to examine each commitment and its dependencies: a control in one agreement does not automatically govern another. This dated audit does not demonstrate that concurrent development and production are always inappropriate, or that milestone payments alone ensure value. Agreement with corrective action should also remain distinct from verified completion.

The learning timetable determines whether a reserved choice is real. If controlled release is planned for day eighteen, reliable evidence about a mature invoice cohort may arrive later. The twenty-day completion limit in this example concerns the defined wave through release, not the final benefits judgement. Maya must retain ownership of the subsequent observation, while the plan identifies its resources and timing outside the worked cost boundary. Those costs still belong in the wider investment appraisal and must not disappear because they fall beyond this wave. Where expansion needs a long-lead order before the evidence matures, Samira faces an explicit trade-off between delay, a smaller reservation and commitment under uncertainty. Lena

should show the remaining exposure of each course. An approval labelled conditional should not conceal an obligation that cannot practically be withdrawn when the condition fails.

Keep the operating capacity and benefit obligations visible

Chapter 13's exception-service analysis remains relevant to investment even though its numbers belong to a different operating boundary. Its eighty per cent nominal load ratio does not establish spare capacity for programme work or resilience to absence. Maya must confirm the staff time needed for design decisions, data preparation, testing and rehearsal, as well as the capacity to receive the change. If the same people are also expected to keep the existing service running, the plan should account for that demand and any agreed cover. Alex needs those allocations at the points where the sequence depends on them. Lena should distinguish a named person from a confirmed contribution with a calendar, competence and accountable manager. The commitment review must resolve material conflicts or narrow the work. A budget can be available while the scarce capacity required to use it remains unavailable.

Canada's 2023 audit of Benefits Delivery Modernization examined prioritising migration from ageing systems while deferring further service transformation. The Auditor General considered migration first reasonable for continuity, while warning that delays and cost pressure could jeopardise transformation. During the audit, the department had not determined the combined effects of programme changes on scope, schedule and cost, and lacked a clear understanding of staffing gaps and needs (Office of the Auditor General of Canada, 2023, paras. 8.17–8.18, 8.21, 8.32). The application here is that narrowing immediate scope still requires reconciling the commitments and future value that remain. A revised wave should identify deferred obligations and who owns the next decision. This historical finding is not proof that migration first was the wrong choice or that later transformation was abandoned. It distinguishes a defensible priority from a complete authorisation basis.

Benefits should enter the wave decision through their mechanisms and conditions. A limited release may demonstrate that staff can resolve particular exceptions, while the wider payment outcome requires sustained use across a defined population. The Value and Evidence Map and Measurement Agreement retain those distinctions. The brief should identify the benefit contribution expected from this wave, the later changes on which it depends and the consequences of deferring them. It should also expose operating costs or adverse effects that could weaken the value claim. If a manual arrangement is used to make an early release possible, its effort, duration and owner belong in the appraisal. Lena should avoid assigning the whole programme's benefit to every wave, then adding the repeated claims. The relevant account is the additional contribution and continuing obligation associated with the commitment actually under consideration, including dependencies shared with other waves.

State the authority and the conditions for reconsideration

The proposed authorisation can now be written concretely. It identifies the defined supplier population and six-activity scope, the resource-feasible eighteen-day base plan, the US\$113,000 base resource estimate and the proposed limits of US\$125,000 and twenty working days. It names the evidence required for critical operating conditions and the owners of the resources on which the sequence relies. In this example, Samira retains release authority over the US\$12,000 allowance; Alex cannot treat it as an unrestricted spending balance. The record should state who may approve changes within the wave and which changes require the sponsor to reconsider the commitment. It should also identify continuing measurement and support obligations outside the worked boundary. These particulars give the approval operational meaning. The same document can record refusal, deferral or approval of a narrower scope when the evidence does not support the proposal as submitted.

Reconsideration conditions should allow time for action. If the environment owner reports a four-day delay, the scenario already shows that the proposed completion limit is threatened even though cost remains below the ceiling. Lena should bring the forecast consequence to the authorised decision maker when that information becomes credible, rather than wait until day twenty has passed. The decision may concern an alternative environment, a revised sequence, a change in scope or an explicit amendment of the limit. Each response needs an owner and a feasible implementation path. A reserve release addresses funding authority; it does not by itself approve the other changes. Chapter 11's control rhythm provides the recurring review, while event-driven escalation handles material evidence that cannot wait. The resulting decision must preserve both the prior commitment and the reason for changing or retaining it.

Baseline discipline makes that history usable. Once authorised, the agreed scope, dates, resource assumptions and cost basis form a dated reference for accountability. The current forecast then describes what the team expects on the latest evidence. If the wave moves from eighteen to twenty-two forecast days, the forecast should change promptly even while the approved limit remains twenty. A subsequent authorised amendment should be identifiable as a new decision, with the original retained. Quietly substituting the forecast for the baseline would erase the very variance the sponsor needs to understand. Conversely, preserving a baseline does not justify continuing to report an expectation that the evidence no longer supports. Lena should reconcile actual work, remaining effort, commitments and exposure at the same reporting cut-off, so that the next decision compares statements with consistent boundaries and dates.

Reauthorisation should consider the remaining course and its consequences, rather than treating expenditure already incurred as sufficient reason to continue. Sunk expenditure cannot be recovered by approving more work, although stopping may create cancellation costs, service disruption or unavoidable obligations that must be examined. The brief should compare feasible alternatives using their future consequences and the strategic need still to be addressed. A programme may remain justified while a particular wave needs reshaping, and a wave may be affordable while its contribution to value has weakened. Samira's judgement should therefore remain connected to the investigated need and future operating state, not only the delivery team's preference to finish the current package. Lena supports that judgement by retaining

evidence, alternatives and obligations clearly enough that continuation is an explicit decision rather than the default result of momentum.

Carry the authorised commitment into delivery

The worked example leaves Samira with a decision she can examine. The fourteen-day sequence depends on an unavailable parallel resource arrangement, so it does not support the original estimate. The corrected eighteen-day plan costs US\$113,000 before controlled allowance. Mapping rework remains within both proposed limits, an environment delay breaches time despite the same US\$121,000 cost, and the combined scenario breaches both. These comparisons do not decide the investment automatically. They show what the proposed commitment can accommodate under stated assumptions and where another decision would be required. Before approval, Alex and Maya must substantiate resource availability and operating obligations, while Lena confirms that the supporting records use the same scope and versions. If those conditions cannot be established, a smaller commitment or further investigation can be recorded without pretending that the full wave has acquired a credible basis merely because funding is available.

Chapter 14 connects a testable operating design with a bounded commitment to delivery and continuing value. The Wave Commitment Record now links the intended result to a feasible sequence, consistent cost basis, resource allocations, exposure and conditions for further authority. It preserves later choices only where contracts, evidence timing and response capacity allow them to remain usable. This provides the starting point for Chapter 15, Integrating Delivery without Losing the Purpose. Delivery control will examine what has actually been completed, what remains and whether the forecast still supports the operating commitment. The authorisation developed here supplies the reference and decision boundaries for that work. Its purpose is to make change accountable as knowledge improves, while keeping the investment connected to the service and benefits that justified committing the organisation's resources in the first place.

References

- Blichfeldt, B. S. and Eskerod, P. (2008). Project portfolio management – There's more to it than what management enacts. *International Journal of Project Management*, 26(4), 357–365. doi: [10.1016/j.ijproman.2007.06.004](https://doi.org/10.1016/j.ijproman.2007.06.004)
- Cooper, R. G. (2008). Perspective: The Stage-Gate® idea-to-launch process—Update, what's new, and NexGen systems. *Journal of Product Innovation Management*, 25(3), 213–232. doi: [10.1111/j.1540-5885.2008.00296.x](https://doi.org/10.1111/j.1540-5885.2008.00296.x)
- Huchzermeier, A. and Loch, C. H. (2001). Project management under risk: Using the real options approach to evaluate flexibility in R&D. *Management Science*, 47(1), 85–101. doi: [10.1287/mnsc.47.1.85.10661](https://doi.org/10.1287/mnsc.47.1.85.10661)
- Jørgensen, M. and Sjøberg, D. I. K. (2004). The impact of customer expectation on software development effort estimates. *International Journal of Project Management*, 22(4), 317–325. doi: [10.1016/S0263-7863\(03\)00085-1](https://doi.org/10.1016/S0263-7863(03)00085-1)
- NASA Office of Inspector General. (2013, 13 June). Commercial Cargo: NASA's Management of Commercial Orbital Transportation Services and ISS Commercial Resupply Contracts. Audit Report IG-13-016. National Aeronautics and Space Administration. oig.nasa.gov/docs/IG-13-016.pdf

Office of the Auditor General of Canada. (2023). Report 8—The Benefits Delivery Modernization Programme. Reports of the Auditor General of Canada to the Parliament of Canada.
publications.gc.ca/collections/collection_2023/bvg-oag/FA1-27-2023-1-8-eng.pdf

Touran, A. (2003). Probabilistic model for cost contingency. *Journal of Construction Engineering and Management*, 129(3), 280–284. doi: [10.1061/\(ASCE\)0733-9364\(2003\)129:3\(280\)](https://doi.org/10.1061/(ASCE)0733-9364(2003)129:3(280))

CHAPTER 15

Integrating Delivery without Losing the Purpose



IN THIS CHAPTER

Connecting delivery performance to the purpose of the transformation.

Keep the operating commitment visible in delivery

The wave has moved from authorisation into work. For this teaching continuation, assume that Samira approved Chapter 14's fictional candidate on its stated basis: an eighteen-working-day base sequence, US\$113,000 of base resource cost, a US\$125,000 ceiling and a twenty-day completion limit. These figures concern that candidate wave, not the manufacturer's earlier programme-wide account. Alex now reports activity progress, suppliers submit evidence and finance records costs. Maya needs to know whether the resulting capability will support the intended payment service. Lena's task is to connect these accounts so that the next decision reflects the same scope, reporting date and operating purpose. Delivery control requires an intelligible explanation of what has been accomplished, what remains and what the organisation can responsibly do next. A favourable percentage or a balanced cost report cannot provide that explanation without its underlying evidence and dependencies.

The most useful delivery review follows a claim across its boundaries. A supplier may report that a data-mapping file is complete, while the agreed package also requires reconciliation of records and control totals. That package may eventually meet its own acceptance conditions while the integrated payment route still needs testing. Even a usable route does not establish that suppliers experience a sustained improvement or that finance can verify a saving. These are related claims with different evidence and timing. Lena should preserve their distinctions while connecting them through the Value and Evidence Map. The work performed matters, including work awaiting formal acceptance, but it should be described under the appropriate status. Otherwise, the organisation can inadvertently convert evidence about an intermediate output into confidence about a service that has not yet been demonstrated under the conditions needed for use.

The transformation lifecycle

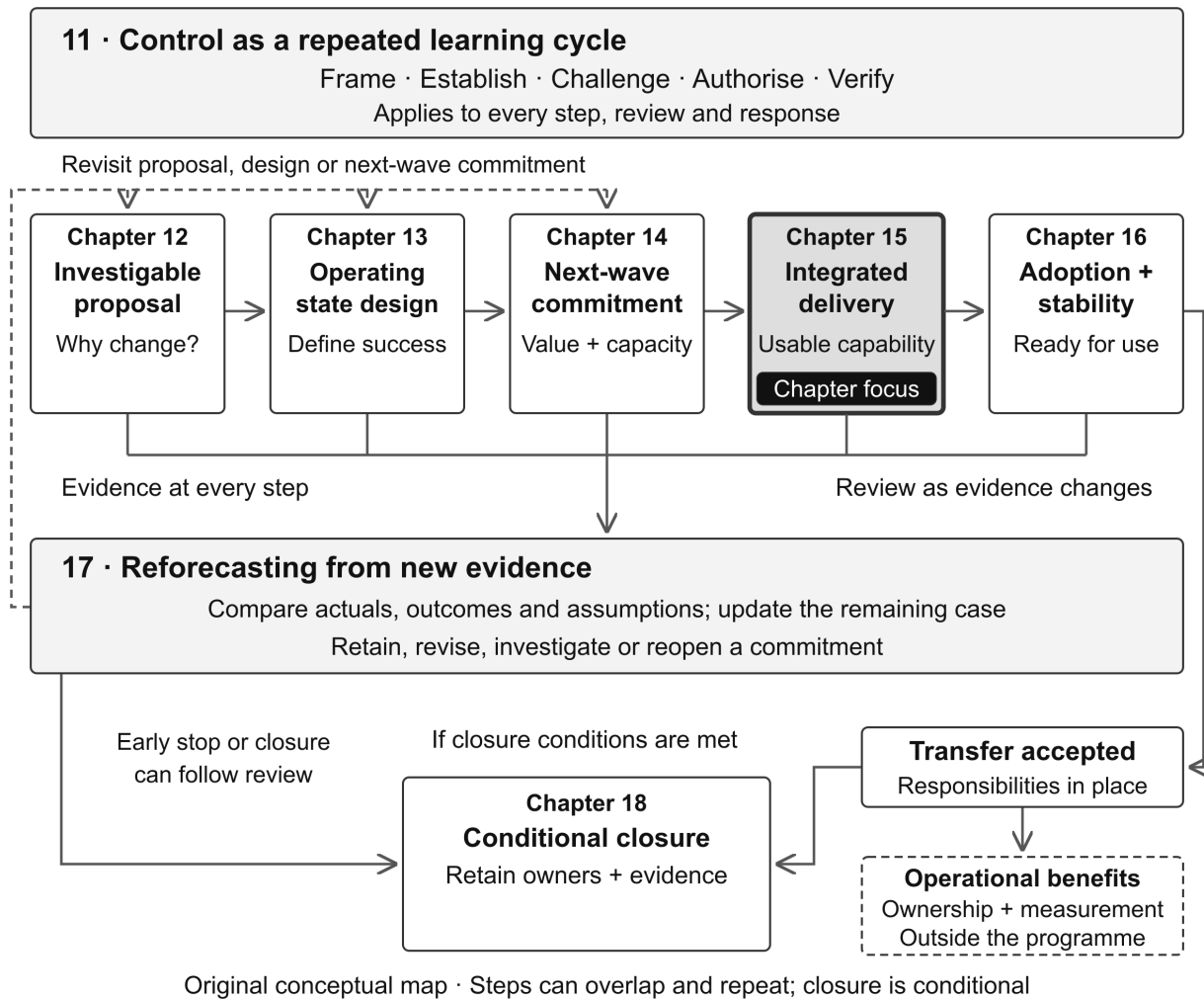


Figure 15.1 The shared lifecycle highlights integrated delivery. Progress, remaining work and current forecasts must stay connected to the operating purpose. Delivery evidence can reopen earlier commitments and supplies the basis for later readiness and adoption decisions.

The Wave Commitment Record supplies the reference for the authorised scope and conditions. The Future State Evidence Specification identifies what the operating design requires, while the Evidence and Learning Record connects discoveries with actions and subsequent verification. Chapter 15 brings their evidence together through the existing Credible Control Method. Frame the delivery decision, Establish a reconciled account, Challenge the inference and remaining forecast, Authorise an appropriate response and Verify what changed. This is an application of the established method, not a new assurance scale. The controller should be able to trace a consequential claim to its source, owner, date and limitations without reproducing every technical record in a board report. That connection allows Samira to understand the commitment, Alex to organise the remaining work and Maya to judge the operating implications. Integration is achieved through those usable connections, rather than the volume of information collected.

Measure accepted work without overstating capability

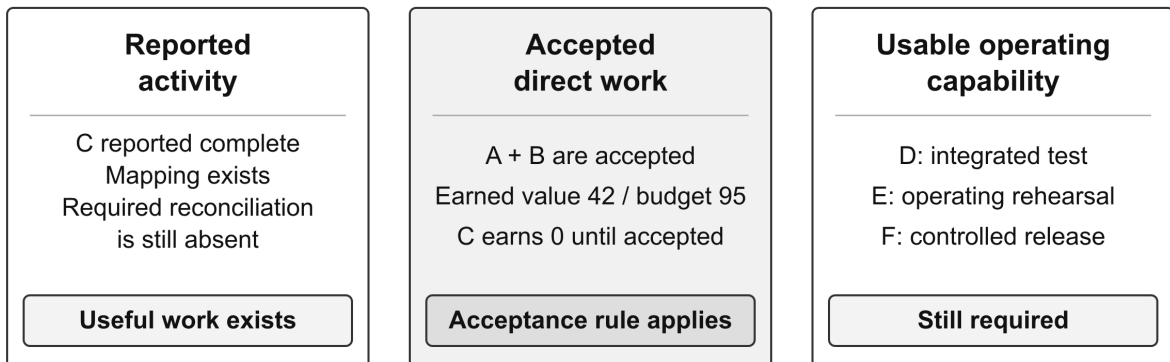
Kim, Wells and Duffey (2003) developed an earned value management implementation model from a practitioner survey and six organisational case studies. Their four groups of factors concern users, methodology, project environment and implementation process. This supports treating progress measurement as an organisational discipline rather than a calculation added to a reporting template. Applied here, Lena needs an agreed basis for recognising completed work, owners who can substantiate that recognition and cost records aligned to the same reporting boundary. These are the chapter's practical requirements, not a claim that the study validated this particular procedure. The survey's low response rate and reliance on practitioner views limit generalisation. An apparently favourable index still requires examination of the work behind it; implementation acceptance does not establish forecast accuracy or demonstrate realised operating benefits.

At the end of working day twelve, the original resource-feasible plan would have completed activities A, B and C. Their direct-work budgets are US\$12,000, US\$30,000 and US\$16,000 respectively. For this example, the agreed progress rule earns each package's budget only when its defined acceptance conditions are met, commonly called a zero-or-one-hundred rule. A and B have met those conditions. C has produced its mapping file but still lacks required reconciliation, so its reported completion is not supported for earned-value purposes. C therefore earns zero at this cut-off, although work has been performed and cost incurred. The rule was established before reporting; it is not a retrospective penalty for an unwelcome result. For longer packages, a pre-agreed series of objective milestones could provide more informative intermediate evidence without inventing a subjective percentage after the work has begun.

Figure 15.2 separates that reporting correction from the wider capability judgement. Acceptance of A and B concerns their own defined scopes, including the interface configuration, while end-to-end performance remains to be demonstrated through D, E and F. Table 15.1 keeps the direct-work account visible at one cut-off. The six packages have a total budget of US\$95,000; planned value is US\$58,000 and earned value is US\$42,000. The resulting 44.2 per cent is the share of the direct-work budget earned under the stated rules, not a measure of physical completion, operating readiness or realised benefits. Coordination is excluded from these measures and accounted for separately. Keeping that boundary explicit prevents elapsed management effort from concealing incomplete deliverables. Lena can report work in progress alongside earned value while preserving the difference between effort expended and the evidence required for acceptance.

Separate reported progress, acceptance and capability

Day 12 cut-off · Agreed 0/100 rule for direct-work packages · US\$ thousand



Under this rule, EV recognises accepted packages; it is not a benefit measure.

Useful activity can exist before acceptance.

Acceptance of a package alone does not demonstrate usable operating capability.

Figure 15.2 Original distinction between reported output, accepted direct work and usable capability. C retains useful work in progress but earns no budget until its agreed acceptance is met. Acceptance of A and B does not establish the integrated service or its benefits.

Table 15.1 Direct work at the end of working day twelve

Activity	Budget	Planned value	Earned value	Actual cost	Remaining estimate
A Design	12	12	12	14	0
B Interface	30	30	30	33	0
C Data	16	16	0	14	6
D Test	24	0	0	0	24
E Rehearsal	8	0	0	0	8
F Release	5	0	0	0	5
Total	95	58	42	61	43

All values in US\$ thousand. Six direct-work packages use an agreed zero-or-one-hundred acceptance rule. Coordination is excluded. Actual cost includes the stated accrual; the remaining estimate includes unperformed ordered work without adding it twice.

Reconcile progress and cost at the same reporting date

The direct-work account records actual incurred cost of US\$61,000. Earned value minus actual cost gives a cost variance of minus US\$19,000, while earned value minus planned value gives a schedule variance of minus US\$16,000. Both variances are expressed in budget currency; the latter is not sixteen days of delay. The cost performance index is 42 divided by 61, approximately 0.69, and the schedule performance index is 42 divided by 58, approximately 0.72. Each index describes a relationship within this defined reporting account. Neither identifies the delivery cause, determines the remaining critical path or proves a benefit shortfall. The zero-or-one-hundred rule also makes earned value move in steps. Lena should explain that behaviour so that a reader does not mistake the recognition pattern for a continuous physical measure or disregard the unresolved work merely because the index is sensitive to acceptance timing.

Actual cost also needs a consistent recognition basis. In the fictional account, US\$56,000 is posted and a further US\$5,000 has been incurred but not yet invoiced; the agreed control basis therefore includes that accrual in the US\$61,000. Cash paid is US\$48,000, a settlement measure rather than another cost to add. An open-order report shows US\$18,000 because it includes the accrued US\$5,000 and US\$13,000 of unperformed work. These are specified system behaviours for the exercise, not universal accounting rules. Lena reconciles them with finance and the supplier records, retaining the reporting cut-off and treatment of later corrections. When the accrued item is invoiced, the accounting treatment must prevent duplicate recognition. A cost report that omits incurred work would improve the apparent index without improving delivery; adding every outstanding order to actual cost would create a different distortion.

Alex's assessment of remaining direct work is US\$43,000: six for completing C, twenty-four for D, eight for E and five for F, all expressed in thousands of US dollars. The US\$13,000 of unperformed ordered work is already included in that estimate; it must not be added again. Actual direct cost plus remaining direct work therefore gives US\$104,000. By comparison, mechanically projecting the observed cost performance index across the full US\$95,000 direct-work budget gives approximately US\$138,000. That projection assumes the cumulative cost-to-earned-value relationship continues. The difference is material and needs explanation, especially because C's incurred cost currently earns nothing under the agreed milestone rule. The remaining-work assessment must demonstrate complete scope and credible resources, while the index projection retains its own assumptions. Selecting the lower number because it fits available funding would replace investigation with preference.

Batselier and Vanhoucke (2015) evaluate earned value time and cost forecasts using fifty-one real projects, most from construction, with supporting Monte Carlo analysis. Their study considers accuracy, timeliness and project network structure. It supports examining forecasting methods empirically without establishing that one index predicts every transformation wave reliably. For Lena, an index-based projection provides a useful comparison with the current assessment of remaining work. A difference can reflect the recognition rule, changed scope, discovered rework or a different mix of future activities. These explanations need evidence; they should not become convenient reasons to discard an uncomfortable result. Alex should substantiate the remaining

work and its assumptions, while the decision record preserves the comparison and the reasons for adopting a particular forecast. Neither this research nor a spreadsheet formula removes the need to understand the delivery conditions.

Forecast the remaining route to the operating commitment

The schedule forecast must be rebuilt from the same status. C requires two more working days after the day-twelve cut-off, followed by D for three days, E for two and F for one. With the stated resources available and no overlap, release is forecast at day twenty. The forecast is two days later than the eighteen-day base sequence and sits exactly at the authorised limit. Coordination costs US\$12,000 to date and a further US\$8,000 over the remaining duration. Table 15.2 therefore reconciles total wave actual cost of US\$73,000 with US\$51,000 remaining, giving a US\$124,000 forecast. Figure 15.3 connects this bridge with the remaining sequence. The calculation uses one scope and calendar throughout; it does not add the cost-index projection to a separately constructed schedule or count coordination twice within the direct packages.

Table 15.2 Reconcile actual and remaining resource cost

Cost element	Actual	Remaining	Forecast total
Direct work	61	43	104
Coordination	12	8	20
Whole wave	73	51	124

All values in US\$ thousand. Coordination costs US\$1,000 per working day, with twelve days elapsed and eight remaining. The US\$124,000 forecast is US\$11,000 above the base estimate and US\$1,000 below the ceiling; it does not authorise reserve release.

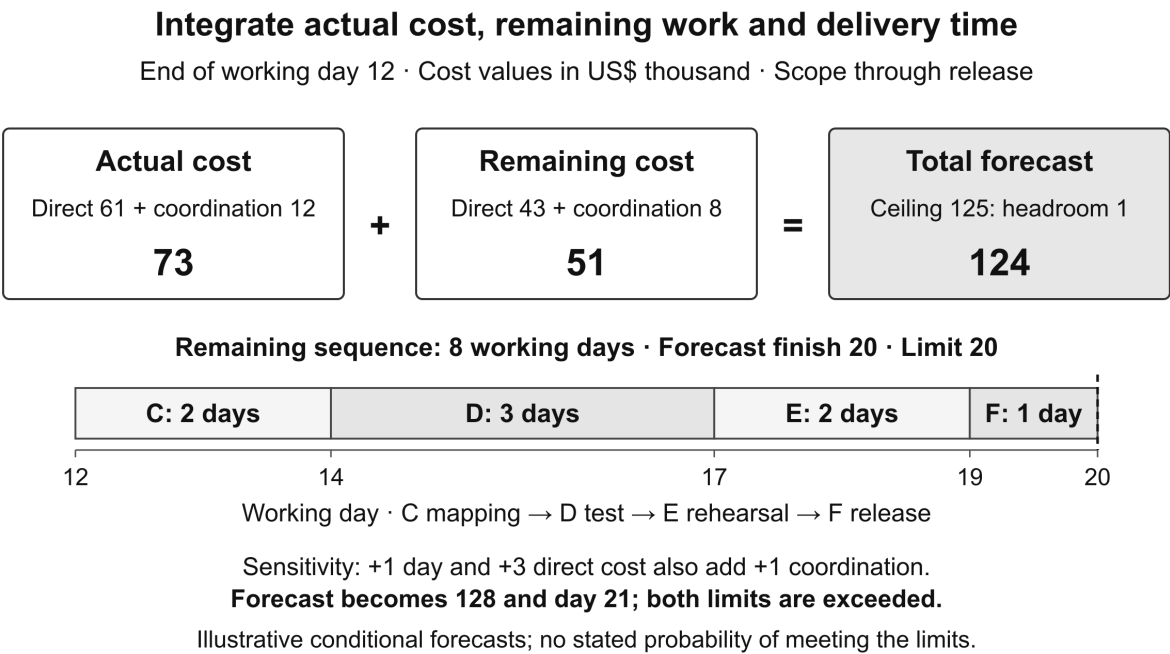


Figure 15.3 Fictional reconciliation at the day-twelve cut-off. Actual resource cost of US\$73,000 plus US\$51,000 remaining gives US\$124,000, with release at day twenty. Both forecasts depend on the remaining scope and resources; one additional day with the specified direct cost breaches both limits.

The US\$124,000 forecast is US\$11,000 above the base resource estimate and leaves only US\$1,000 below the funding ceiling. The schedule has no remaining margin to its twenty-day limit. These facts do not create an automatic authority to spend the allowance or accept unmet operating conditions. Samira retains the reserve-release authority established in Chapter 14. A simple sensitivity test makes the exposure clearer: one further working day with US\$3,000 of additional direct work also adds US\$1,000 of coordination, producing US\$128,000 and release at day twenty-one. Both limits would be breached. No probability is attached to that scenario. Its purpose is to identify a consequential assumption that deserves attention before the team describes the wave as comfortably within control. Lena should connect the remaining test and reconciliation uncertainties with the response options and decisions that would be needed if the forecast worsens.

Davies and Mackenzie (2014) studied systems integration in the London 2012 construction programme through interviews and documentary evidence. They distinguish coordination of the overall system from integration within its constituent systems, with explicit attention to interfaces. Their case provides a reason to ask who understands relationships across otherwise well-managed packages. For Alex, completing a payment component should therefore prompt an examination of the receiving interface, outstanding dependencies and evidence that the combined service works. This is an application of the case's organising insight, not a claim that construction findings establish the effectiveness of this fictional delivery arrangement. The study does not isolate integration as the cause of programme success. It supports examining responsibility across boundaries while leaving operational adoption and sustained benefits as separate claims requiring their own appropriate evidence.

Make changes and rework part of the remaining account

Ford and Sterman (2003) use a dynamic model informed by semiconductor development experience to examine concealed rework. Their simulations show how temporarily favourable progress reports can coexist with later iteration and weakened delivery performance. These are conditional modelled consequences, not a universal estimate of the effect of late disclosure. The practical implication for Alex and Lena is to distinguish work performed, accepted work and discovered repair obligations. A defect should lead to an assessment of affected dependencies, repeated tests and remaining resources. It should not automatically be treated as misconduct, because learning and changed requirements can also create legitimate rework. Equally, pressure to preserve a completion percentage can obstruct an honest forecast. The book's application is to make repair obligations visible and govern the response, while retaining the evidence needed to determine what actually caused the discovery.

A change request should carry its delivery and operating consequences together. Suppose a proposed recovery removes a reconciliation check so that C can be declared complete. The schedule might improve on paper, but the proposal changes an acceptance condition that supports payment integrity. Maya must examine the operating consequence, Alex must identify affected design and test work, and Samira or the appropriate delegated authority must decide whether any alternative control is acceptable. Lena traces the change through the Future State Evidence Specification and the relevant forecast. A genuine alternative may exist, but its implementation cost, effectiveness and continuing owner require evidence. The team should not remove the test from the schedule while leaving its assumed protection in the benefits narrative. That would preserve the appearance of progress by making the scope and value account inconsistent with each other.

Dvir and Lechler (2004) examined 448 projects using retrospective questionnaire evidence and structural equation modelling, distinguishing changes to goals from changes to plans. Their results associate goal changes with adverse success assessments and show why initial planning quality cannot settle subsequent delivery decisions. The observational design does not establish that all change damages performance, or that retaining an obsolete goal improves results. Applied here, Alex should distinguish a changed route towards the agreed operating result from a proposal to change that result itself. Lena can then connect each proposal to consequences for remaining work, interfaces, cost, timing and benefits. Samira authorises the consequential decision within the existing governance arrangements. Preserving the previous commitment and the reason for its revision supports accountable adaptation; these are the chapter's recommendations, not tested interventions.

The National Audit Office's 2016 review of Great Western modernisation illustrates the need for an integrated delivery account. Governance arrangements introduced in February 2015 brought together organisations responsible for infrastructure, trains and services. Nevertheless, the industry programme board still lacked a critical path showing how interacting projects affected delivery dates, alongside clear high-level cost and schedule performance indicators (National Audit Office, 2016, paras. 2.15–2.16). The report recognised management improvements while retaining concerns about the information supporting decisions. For our wave, the implication is

to connect completed work with the remaining route to usable service and to identify the owners of interface commitments. This is an application of historical audit findings, not a claim that nobody held a schedule or that one reporting technique would have prevented the difficulties. A common meeting can support integration, but cannot substitute for the evidence it needs.

Protect the purpose when local recovery looks attractive

Local optimisation can transfer difficulty to another part of the service. Accelerating data preparation by relying on unverified manual corrections may move work into operations or make later reconciliation harder. Assigning Maya's reviewers full time to testing may shorten a delivery activity while reducing the capacity of the existing payment service. A recovery option should therefore identify the work displaced, the additional cost, the new exposure and the effect on the intended benefit. Lena can compare the options using the same scope and operating conditions, including the feasibility of returning to the previous arrangement. The comparison should also explain what remains unknown and what evidence could resolve it. A faster package completion date is useful only in relation to the larger commitment; it should not be rewarded automatically when the response makes the receiving service less reliable or more expensive to sustain.

The Value and Evidence Map keeps that wider commitment available for challenge. Earned value can help assess budgeted delivery work, while the map asks whether the delivered result enables the required capability and practice. A payment interface accepted for its configuration may still fail with real data volumes, unfamiliar exceptions or a missing support arrangement. Maya's operating assessment therefore remains distinct from the supplier's completion certificate, even where one record supplies evidence to both decisions. Lena should identify which critical conditions have been demonstrated and which remain planned, without blending them into a reassuring average. The Claim and Challenge Map from Chapter 9 can structure examination of the strongest unsupported inference. It helps distinguish a claim with incomplete evidence from a claim already contradicted by findings, directing proportionate investigation or corrective work rather than another general request for confidence.

GAO's February 2022 review of the US Department of Veterans Affairs' electronic health record modernisation distinguished migration progress from usable information. Selected data had been migrated and validation testing performed for the initial October 2020 deployment. Yet subsequent evidence identified accessibility, accuracy and appropriateness problems, and the programme lacked specific measures and goals for assessing migrated-data quality (United States Government Accountability Office, 2022, pp. 19–23). The delivery-control application is to retain unresolved quality work in the completion assessment and remaining forecast, even when the transfer activity has been performed. This does not mean that no testing occurred or that all migration failed. It demonstrates why the evidence for carrying out an activity and the evidence for its intended use must remain distinguishable. The case supplies no quantified estimate of patient injury and makes no claim about every later deployment.

Turn the reconciled account into an authorised response

The delivery review should begin with the decision that the evidence must support. At the day-twelve cut-off, Samira needs to understand the proposed use of allowance, the lack of schedule margin and the work still required before release can be considered. Alex supplies the remaining sequence and resource commitments, Maya identifies operating conditions and finance confirms the cost reconciliation. Lena brings these contributions into a dated account that distinguishes observed facts from estimates and unresolved assumptions. The review should name the material differences between records and explain how they were resolved or bounded. This is more useful than inviting each function to present an independent percentage. Chapter 11's recurring control cycle provides the rhythm, while material discoveries can require an earlier decision. The meeting should end with a response whose authority, scope, owner and verification point are understood by the people who must act.

A proposed recovery should be compared with continuing the current plan and with any feasible narrower commitment. The assessment should retain the time required to obtain resources, agree supplier changes and validate affected work; a response cannot begin instantly merely because a report recommends it. If extra specialist time is proposed, Alex must demonstrate that it will address the limiting work rather than create another handoff or divert attention from integration. If scope is deferred, Maya should accept the residual operating obligations and the investment case should retain their costs and benefit consequences. Lena should distinguish the expected effect of the response from the evidence that the response has been implemented. An approved overtime arrangement, for example, is an input. Its effect on accepted work, remaining duration and quality must be examined rather than assumed from the additional hours authorised.

The forecast should change when evidence changes, while the baseline remains an identifiable account of what was authorised. Recording a day-twenty forecast does not erase the eighteen-day base sequence, and forecasting US\$124,000 does not by itself approve reserve use or revise the budget. If an authorised change establishes a new reference, retain the earlier version, decision and reason. The progress history also needs a clear correction policy. Where evidence invalidates an earlier acceptance, explain the correction and associated work rather than quietly removing the problem or earning the same scope twice. That situation differs from C in this example, whose agreed acceptance had never been demonstrated. Distinguishing delayed recognition, defective accepted work and new scope allows the team to interpret trends responsibly. It also improves later learning about estimates and controls without converting every change in knowledge into an accusation of poor performance.

Verification closes the response through appropriate evidence. Completing C's reconciliation should produce the agreed records and permit its package status to change under the established rule. It should also update the remaining forecast with actual effort and any new findings. Completion of D's integrated tests concerns a broader claim, while E examines relevant operating work and F remains subject to the required release authority. Lena should keep those decision boundaries connected so that an earlier acceptance is not mistaken for unrestricted permission to proceed. Open findings need an owner, consequence and planned disposition; closing a tracking item administratively is insufficient where its effect remains unverified. The same discipline

applies to recovery actions and lessons. The Evidence and Learning Record should identify what changed in practice and what observation supports the claimed improvement, allowing later reviews to distinguish an implemented action from a demonstrated result.

Carry usable evidence towards operational readiness

The learner can now reconstruct the manufacturer's day-twelve position. A and B have earned their direct-work budgets, while C remains unaccepted despite incurred effort. The direct account therefore shows planned value of US\$58,000, earned value of US\$42,000 and actual cost of US\$61,000. The remaining-work assessment supports a conditional wave forecast of US\$124,000 and day twenty, while the cost-index projection raises a challenge that needs explanation. An additional day with the specified direct cost would exceed both authorisation limits. The decision is consequently richer than whether the latest dashboard should appear favourable. Samira needs a defensible response to limited margin and uncertainty; Alex needs complete remaining work and available resources; Maya needs the operating conditions protected. Lena's contribution is the reconciled account and the traceable connection from evidence to authority, followed by verification of what the decision changes.

Chapter 15 makes delivery control a continuing examination of the connection between work and purpose. It reconciles progress rules, incurred cost, remaining scope and dependencies without treating any one indicator as proof of usable capability. It also makes change and rework visible enough for an authorised response, while protecting the operating conditions and benefits that justified the investment. Chapter 16 examines adoption, readiness and operational stability in greater depth. Its decisions need the evidence developed here: identifiable accepted outputs, demonstrated integration, a credible account of unresolved work and receiving owners who understand the conditions attached to use. Delivery control should carry that evidence forward with its limits intact. The transition into operations then begins from a defensible account of capability and remaining obligations, rather than a completion percentage whose meaning has been lost between teams.

References

- Batselier, J. and Vanhoucke, M. (2015). Empirical evaluation of earned value management forecasting accuracy for time and cost. *Journal of Construction Engineering and Management*, 141(11), 05015010. doi: [10.1061/\(ASCE\)CO.1943-7862.0001008](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001008)
- Davies, A. and Mackenzie, I. (2014). Project complexity and systems integration: Constructing the London 2012 Olympics and Paralympics Games. *International Journal of Project Management*, 32(5), 773–790. doi: [10.1016/j.ijproman.2013.10.004](https://doi.org/10.1016/j.ijproman.2013.10.004)
- Dvir, D. and Lechler, T. (2004). Plans are nothing, changing plans is everything: The impact of changes on project success. *Research Policy*, 33(1), 1–15. doi: [10.1016/j.respol.2003.04.001](https://doi.org/10.1016/j.respol.2003.04.001)
- Ford, D. N. and Serman, J. D. (2003). The Liar's Club: Concealing Rework in Concurrent Development. *Concurrent Engineering*, 11(3), 211–219. doi: [10.1177/106329303038028](https://doi.org/10.1177/106329303038028)

Kim, E., Wells, W. G., Jr. and Duffey, M. R. (2003). A model for effective implementation of Earned Value Management methodology. *International Journal of Project Management*, 21(5), 375–382. doi: [10.1016/S0263-7863\(02\)00049-2](https://doi.org/10.1016/S0263-7863(02)00049-2)

National Audit Office. (2016, 9 November). Modernising the Great Western railway. HC 781, Session 2016–17. Report by the Comptroller and Auditor General. www.nao.org.uk/wp-content/uploads/2016/11/Modernising-the-Great-Western-railway.pdf

United States Government Accountability Office. (2022, 1 February). Electronic Health Records: VA Needs to Address Data Management Challenges for New System. GAO-22-103718. www.gao.gov/assets/gao-22-103718.pdf

Adoption Readiness and Operational Stability



IN THIS CHAPTER

Establishing real adoption, service continuity and operational stability.

From delivered work to permission to operate

The delivery meeting has become quieter, but the receiving operation has become busier. Alex can explain which packages have been accepted and which integration questions remain; Maya must now consider what happens when ordinary staff encounter an unfamiliar exception while customers expect the service to continue. These are connected responsibilities, but they require different evidence. Chapter 15 ended with a conditional forecast of US\$124,000 and release at working day twenty, with unresolved work and little margin. It did not establish that release occurred or that anyone could operate the resulting service dependably. This chapter carries that distinction forward. A delivery forecast cannot become an adoption result merely because the reporting date moves on. The next decision concerns permission for specified people to use a specified configuration within conditions that the organisation can support, observe and, where necessary, change. Readiness is therefore a claim with boundaries and consequences, rather than a general description of confidence.

The shared lifecycle places adoption after integrated delivery while retaining routes back to design and commitment. Its position does not make readiness a late inspection: Maya's operating requirements should already have shaped the Future State Evidence Specification and the Wave Commitment Record. What changes now is the strength and proximity of the evidence. Planned training becomes observed performance; promised support becomes an available person with the necessary access; a recovery description becomes an exercised response. Later examples introduce separate fictional datasets for rehearsal, an operator cohort and daily operating performance. Their clocks and denominators are independent. Where live use is discussed, a new hypothetical assumption of bounded authorisation is stated; it does not rewrite Chapter 15's unresolved position as historical fact. This separation lets the controller connect the evidence without manufacturing a continuous success story from observations that concern different people, tasks or periods.

The transformation lifecycle

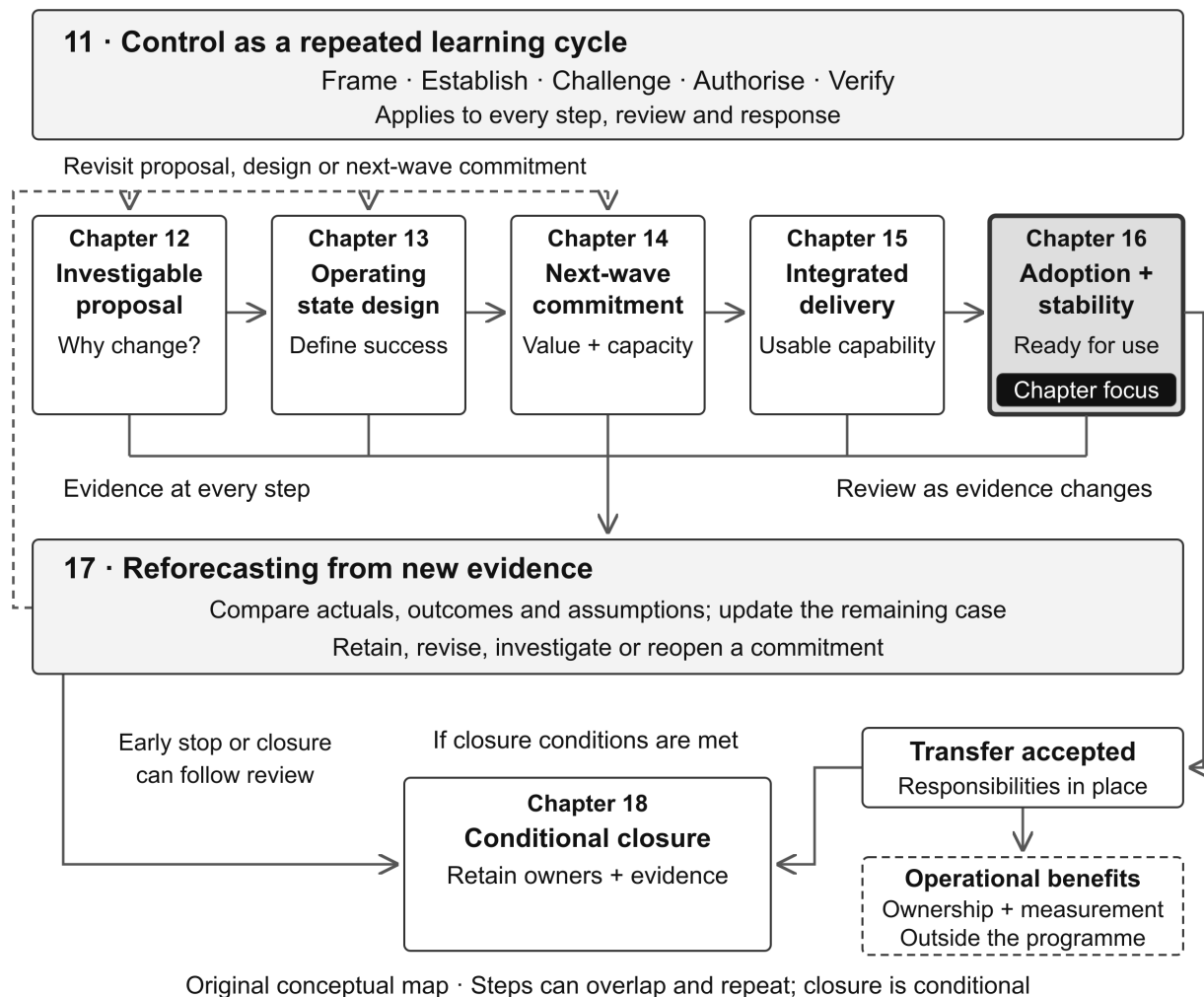


Figure 16.1 The shared lifecycle highlights adoption readiness and operational stability. Defined permission to operate depends on evidence of capability, support and recovery, followed by observation in use. Evaluation can reopen design, investment or delivery decisions.

Scaccia et al. (2015) distinguish motivation, general organisational capacity and the capacity required for a particular innovation. Their practical heuristic is illustrated through three behavioural health organisations; it is not a validated formula for approving operational release. Read alongside Chapter 4's discussion of commitment and confidence, it directs attention to the conditions supporting those perceptions. Maya may support the change while lacking available reviewers, suitable access or a workable escalation route. The practical response is to examine those conditions separately and identify what must be supplied or demonstrated. A favourable average should not conceal a missing critical resource. The chapter uses this distinction within its existing evidence records, without calculating a readiness score or implying that a capacity checklist predicts successful adoption. Any unresolved condition retains an owner and consequence.

Make readiness a bounded operating claim

The practical starting point is a sentence that could be contradicted by evidence. Maya might propose that named payment operators can process specified domestic invoices, using release version A, within an agreed workload, during supported business hours, while retaining a tested route for excluded transactions. Each qualification matters. Evidence for domestic invoices does not support international tax exceptions; a demonstration with vendor administrators does not establish what ordinary access permits; a weekday support arrangement does not cover a weekend payment run. Lena connects this claim to the Future State Evidence Specification, then records the observation required for each material condition. The specification remains the source of the requirement, while the Assurance and Decision Record captures whether the available evidence supports the present permission. A configuration change can reopen that judgement. Evidence is not permanently valid simply because its original test report carries an approval signature.

Figure 16.2 applies the Credible Control Method to this decision without introducing another competing governance system. Frame the proposed operating boundary, establish the relevant observations, challenge their representativeness and limits, authorise a bounded action, then verify performance in use. Some requirements can be improved progressively; others must hold before a particular activity is permitted. A missing critical reconciliation cannot be compensated by excellent training attendance. The record should explain which conditions prevent release, which allow a narrower scope, and which can remain open with an owner, resource and review date. These classifications require the responsible operational and technical judgement, rather than a controller assigning numerical weights. Samira's authorisation must be explicit about scope and conditions, with the relevant operating authority accepting its responsibilities. A high average across unrelated checks should never conceal the condition on which safe, reliable execution actually depends.

Bound operational use to the evidence

Practical application within the Credible Control Method

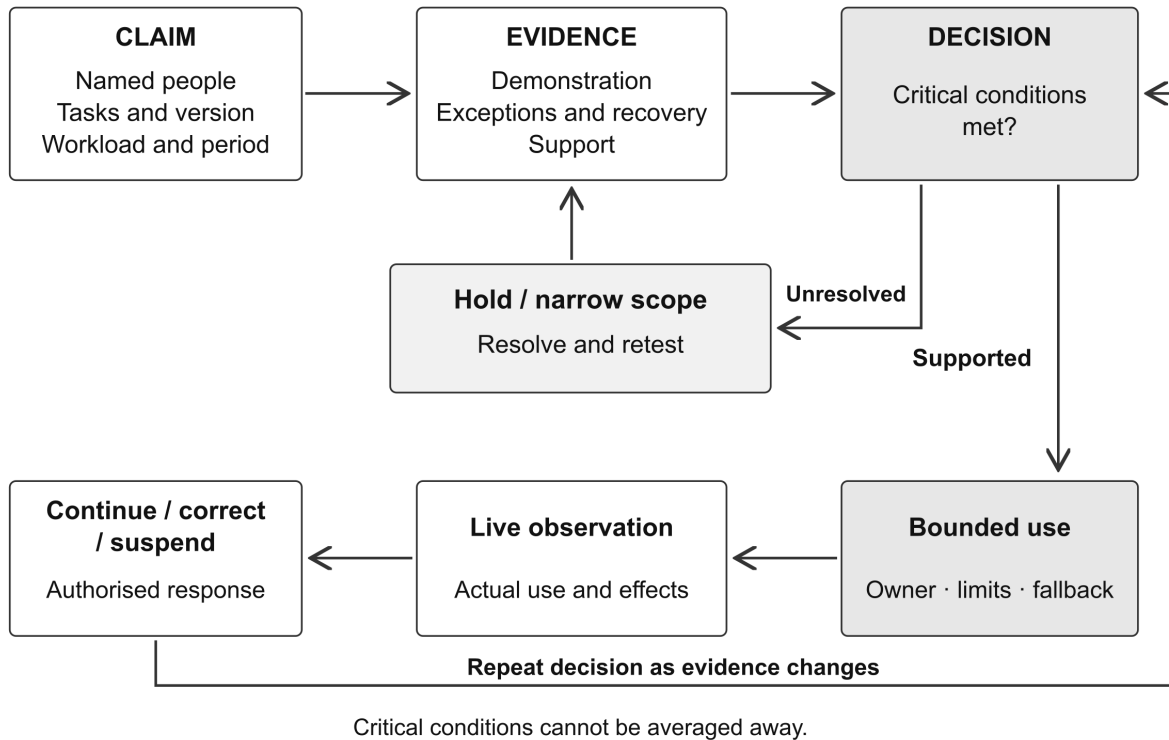


Figure 16.2 Original application of the Credible Control Method to operating authorisation. Critical conditions require a supported disposition before use; bounded permission retains ownership, limits and fallback. Live observations inform the next decision. This working aid is not a validated readiness score.

Bounded permission also requires a response that can be executed when its assumptions fail. Before use begins, Maya needs to know who can suspend new transactions, how work already accepted will be reconciled, which service will carry the remaining demand and how customers will be informed. Alex establishes the technical sequence and its constraints; returning software to an earlier version may not reverse payments already made or restore the previous meaning of changed data. Consequently, fallback can involve controlled manual processing, diversion or a temporary restriction rather than an automatic reversal. Lena records the decision triggers, evidence source, responsible authority and verification needed before restarting. The purpose is not to predict every incident. It is to prevent the organisation from discovering, during disruption, that its advertised contingency requires unavailable staff, incompatible data or an approval that nobody has authority to give.

Rehearse the receiving operation and its fallback

A useful rehearsal makes ordinary operating constraints visible. The receiving team should perform the tasks with its intended access, instructions, staffing and support arrangements, while observers distinguish normal assistance from specialist rescue. Deliberately selected exceptions test the difficult pathways; representative routine cases examine recurring work. These designs answer different questions and should retain separate results. Maya can introduce an absent reviewer, a delayed interface or a disputed invoice, then observe whether the team recognises the condition, routes the work and preserves the necessary records. Success includes reaching the correct controlled response when completion is inappropriate. Lena records what was observed, which configuration was used, who intervened and whether the exercise changed the system being tested. Repeating a task after coaching can demonstrate learning, but the second attempt does not erase the first failure or become an independent observation of initial readiness.

Hong Kong's new airport began operating on 6 July 1998. The Legislative Council Select Committee recorded operational and facilities trials, including some testing of fallback procedures. It nevertheless found that many identified deficiencies remained unresolved and that fallback arrangements did not cope adequately with opening-day problems. For the flight-information contingency, it questioned whether staff understood execution when the procedure had not been clearly documented and rehearsed jointly. The case concerns the adequacy and coordination of readiness evidence, rather than the complete absence of testing (Hong Kong Legislative Council Select Committee, 1999, paras. 5.141–5.142 and 5.154–5.157). The book's application is to require the receiving operation to demonstrate how people will use fallback arrangements together, under stated conditions, before treating the existence of backup equipment or an informal workaround as evidence of operational resilience.

The most useful rehearsal conclusion is therefore a disposition of evidence and remaining exposure. Suppose routine invoices pass, but the team cannot reconcile transactions diverted during an interface interruption. Maya cannot infer recovery readiness from the routine success percentage. She may recommend correcting and repeating the affected exercise, or restricting the initial service so that the unproven pathway cannot be entered. A restriction is credible only if it can be enforced and observed in practice. The controller links each material finding to the action, accountable owner and subsequent verification, using the Evidence and Learning Record already established in the book. A retest should identify what changed and which earlier conclusions remain applicable. This creates a traceable argument for the next decision, while preserving the possibility that a newly discovered dependency requires further design or investment before the proposed operating boundary is justified.

Measure adoption without changing the question

Burton-Jones and Straub (2006) argue that system use should be defined through the user, system and task, then measured for the question being examined. Their spreadsheet-task study shows that the choice of usage measure can alter the relationship with performance. It does not establish an adoption threshold for operational services. Applied here, a login count cannot by itself demonstrate that payment staff complete the required work correctly. Maya needs evidence about the relevant activities, operating conditions and exceptions, while Lena preserves the population and observation period attached to each measure. Training attendance and access provision remain useful preparation evidence, but answer different questions. The chapter applies the paper's measurement reasoning to the existing Future State Evidence Specification; it does not treat the student study as validation of the manufacturer's service.

Consider a separate fictional cohort of eighty operators whose roles and assigned work make them eligible for the change. Assume bounded use has now been authorised for those meeting the specified competence conditions; others retain an authorised existing route. Seventy-two attended training and sixty demonstrated the required tasks before use. During the first ten operating days, fifty-four used the new process and forty-five completed all their assigned in-scope tasks correctly, on at least five occasions, without unplanned specialist rescue. For this teaching dataset, these are nested groups, with complete records and relevant assigned work for everyone. Table 16.1 retains the original cohort denominator. Reporting forty-five as a share of the fifty-four users gives 83.33 per cent; reporting the same forty-five against the eighty eligible operators gives 56.25 per cent. Both calculations are valid, but answer different questions. Neither demonstrates improved payment outcomes or attributable financial benefit.

Table 16.1 Different questions within the same fictional operator cohort

Evidence	Operators	Share of cohort
Eligible for the change	80	100.00%
Attended training	72	90.00%
Demonstrated required tasks	60	75.00%
Used the new process	54	67.50%
Repeated correct use without rescue	45	56.25%

All shares use eighty eligible operators. The groups are nested for this illustration. Training and demonstration precede use; the final two rows concern the first ten operating days. The last row requires all assigned in-scope tasks to be correct, on at least five occasions, without unplanned specialist rescue.

A further, independent teaching dataset concerns rehearsal cases, not those eighty operators. Suppose a sample fixed in advance contains one hundred representative, independent task attempts, of which ninety meet the defined success condition without specialist rescue. Assume an effective assessment and a common success probability. The observed proportion is 0.90. To calculate a two-sided Wilson interval with nominal 95 per cent confidence, use $n = 100$ and $z =$

1.96. Set $d = 1 + z^2/n$; the centre is $(0.90 + z^2/(2n))/d$, and the half-width is $z\sqrt{[0.90(1 - 0.90)/n + z^2/(4n^2)]/d}$. The resulting interval is approximately 82.56 to 94.48 per cent. If an illustrative, pre-agreed rule requires the lower endpoint to reach 85 per cent, this evidence does not satisfy it, despite the ninety per cent observed success. The threshold is an invented decision requirement, not an industry standard.

Brown, Cai and DasGupta (2001) examine binomial intervals and show why the familiar Wald approximation can give unreliable coverage. Their analysis supports using an appropriate method, including the Wilson alternative, while retaining its assumptions and approximate coverage. It does not determine an acceptable operational success rate. Here, the confidence procedure describes repeated-sampling performance, not a 95 per cent probability that this fixed success probability lies inside the reported interval. Shared operators, dependent cases, selective exclusions or changing instructions may invalidate the simple binomial account. Deliberately difficult recovery scenarios also cannot be pooled uncritically with random routine cases. Lena should preserve the sample design and unresolved failures alongside the calculation. A wider or narrower interval changes the strength of one claim; it does not establish that every critical operating condition has been demonstrated.

Read operating performance as a process

Once bounded use begins, the question changes from what the team demonstrated to how the service behaves over time. A further fictional example follows the daily mean handling time for one defined class of payment exception. These are operating days one to eight, independent of Chapter 15's delivery clock. For teaching, assume independent, approximately normal daily means from a suitable reference process, with known mean twenty minutes and standard deviation four minutes. Four describes variation between daily means, not variation between individual cases; the daily sampling arrangement is held comparable. The reference parameters are assumed, not estimated from these eight observations. Daily workload, case mix, support and unresolved cases must also remain visible. A chart of completed work can look healthy while difficult cases accumulate outside its denominator, so Maya needs the waiting queue and its oldest cases alongside the handling-time series.

An exponentially weighted moving average retains information from earlier observations while giving more weight to recent ones, as explained in the NIST/SEMATECH handbook (n.d., section 6.3.2.4). Let X_t denote today's daily mean and Z_t the monitoring statistic. For this example, choose $Z_0 = 20$ and calculate $Z_t = 0.3X_t + 0.7Z_{t-1}$. These illustrative settings are not a universal chart design. With the assumed independent observations, calculate the statistic's variance recursively as $v_t = 0.3^2 \times 4^2 + 0.7^2 v_{t-1}$, starting with $v_0 = 0$. The time-specific limits are $20 \pm 3\sqrt{v_t}$. This retains the narrower limits near the start instead of immediately using their long-run values. Table 16.2 provides enough detail to reproduce the calculation. The chart monitors a process relative to a reference; its limits are not contractual service requirements or evidence of benefit realisation.

Table 16.2 Reproduce the fictional EWMA calculation

Day	Daily mean	EWMA	Lower limit	Upper limit
1	20	20.00	16.40	23.60
2	22	20.60	15.61	24.39
3	19	20.12	15.26	24.74
4	23	20.98	15.11	24.89
5	25	22.19	15.03	24.97
6	29	24.23	14.99	25.01
7	30	25.96	14.98	25.02
8	31	27.47	14.97	25.03

All values except day are minutes, rounded to two decimals after calculation. Reference mean 20; standard deviation of daily means 4; lambda 0.3; limit multiplier 3. The first upper-limit signal is day seven. These eight observations do not estimate or validate the assumed reference parameters.

On day two, the calculation is $0.3 \times 22 + 0.7 \times 20 = 20.60$ minutes. By day six, the statistic reaches 24.23, below its upper limit of 25.01. On day seven it reaches 25.96 against 25.02, producing the first upper-limit signal; day eight adds further evidence of deterioration. Figure 16.3 makes this accumulation visible. The signal calls for investigation, rather than proving that the new software caused the change. Maya should examine case composition, staffing, support withdrawal, data quality and unfinished work before choosing a response. Equally, she should not wait for a chart signal when a critical incident or an agreed service limit already requires intervention. The displayed series illustrates what would be observed if monitoring continued through day eight; it is not permission to postpone action after the first relevant trigger or evidence of actual stability in the manufacturer's service.

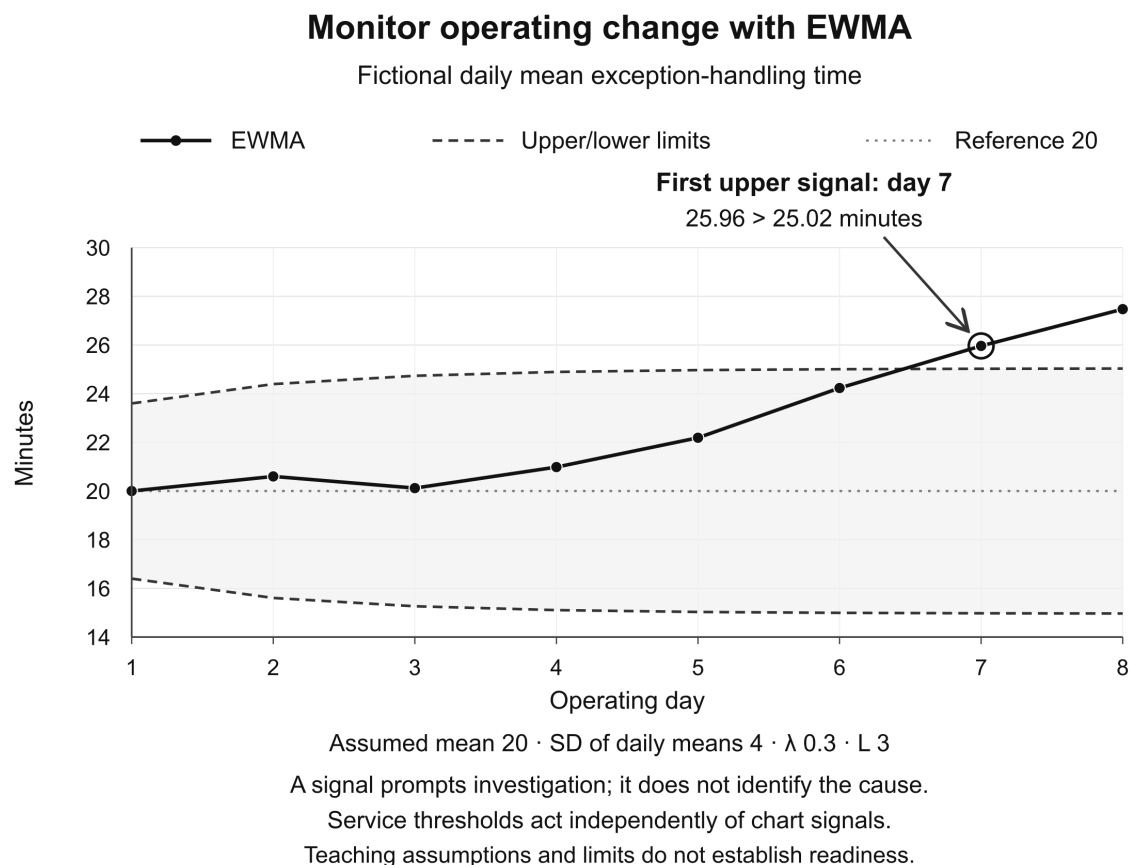


Figure 16.3 Original fictional EWMA example using the stated assumed reference parameters. The first upper-limit signal occurs on operating day seven. The statistic and time-varying limits concern daily means; a signal prompts investigation, while operational requirements can demand action independently.

Jones, Champ and Rigdon (2001) analyse exponentially weighted moving average charts when process parameters are estimated rather than known. Their analysis shows that estimation can increase false alarms and reduce sensitivity to change. It does not establish a universal minimum baseline length or validate this chapter's chosen settings. In practice, Lena must document the reference period, comparability and estimation method, obtaining statistical support where the decision warrants it. Learning effects, serial dependence and changing daily sample sizes can require a different analysis. Recentring the chart after every adverse observation would also hide the very change under examination. Descriptive monitoring remains useful when a defensible statistical reference is unavailable, provided its limitations are explicit. A quiet chart can coexist with consistently unacceptable performance; a signal can occur without breaching a service requirement. The operating decision needs both forms of evidence.

Stabilise practices as well as systems

Stabilisation requires attention to what people actually do when the prescribed process becomes inconvenient. A workaround may be a sensible response to a genuine design gap, an unauthorised control bypass or evidence that the intended activity is poorly specified. Counting all workarounds as resistance prevents learning; accepting all of them as local flexibility can destroy the control the transformation was meant to create. Maya should examine the task, reason, consequence and conditions of each material departure, then decide whether to change the design, improve support or enforce an essential requirement. Lena preserves the relationship between the observation and that decision rather than presenting an unexplained adoption percentage. Temporary experts should record their interventions so that apparently successful work does not hide continuing dependence. As support reduces, the organisation needs evidence about the resulting operation, because performance achieved with intensive assistance does not automatically transfer to ordinary staffing.

New Zealand's Auditor-General examined Inland Revenue's Business Transformation programme in 2020. The audit reported workshops after each stage or release and identified customer preparation as an area strengthened after Release 1. Before later releases, Inland Revenue expanded seminars and customer contact. Following Release 3, it also changed automatic income-tax assessments so that customers nominating a tax agent were removed from the automatic process; Inland Revenue told the auditors that customer feedback prompted this change (Office of the Auditor-General, 2020, paras. 3.33–3.37). This documents adjustment after deployment, without proving that implementation was trouble-free or final benefits had been secured. The book's application is to retain support and monitoring after release, use users' experience to identify necessary operating changes, and verify those changes through appropriate service evidence before withdrawing transition support.

Pluye et al. (2004) examined programme sustainability through organisational routines in five Quebec community health centres. Documents and interviews distinguished activities that continued informally from activities embedded in organisational arrangements. Their assessment described routines at a particular time; it did not establish that continued activities achieved programme objectives or would endure indefinitely. For Maya, this directs attention beyond the release to who performs the work, maintains its resources and updates the operating arrangements. A service sustained through temporary project support may need a different decision from one supported through ordinary operating responsibilities. Lena should therefore preserve the continuing obligations and the evidence needed after support changes. This is the book's application, not a tested transfer rule. Continued use still requires separate examination of quality, adverse effects and the intended benefit.

Transfer responsibility with its resources and obligations

The decision to reduce transition support should be designed before the support team is dismantled. Maya needs an agreed observation period that includes relevant operating conditions, such as a payment cycle, planned absence or a known demand peak. Its duration follows the claim and exposure, rather than an arbitrary number of green reporting periods. The exit criteria should cover performance, the remaining queue, incident response, demonstrated recovery and the receiving team's capacity. Where evidence remains incomplete, the available responses include retaining support, restricting scope or accepting a defined residual obligation through the appropriate authority. That acceptance records why the remaining exposure is tolerable, what resources support it and when it will be reconsidered. It does not convert an untested condition into a passed test. The Ownership and Capacity Framework keeps commitment, authority, capacity, evidence and continuity connected when responsibility moves into ordinary operation.

Financial control remains part of this transition. Chapter 15's US\$124,000 forecast concerned the specified delivery wave; it did not supply an unlimited budget for later operating support. If the hypothetical release requires additional experts, monitoring or corrective work beyond that boundary, Lena makes those requirements visible in the appropriate forecast and decision. Moving expenditure to an operating cost centre does not make the transformation cheaper or its benefits larger. Maya confirms which continuing activities have funded capacity, and Alex retains clear ownership of unresolved technical obligations until they are accepted or completed. The same discipline applies to data, supplier support and maintenance responsibilities. A signed handover that leaves essential work without resources merely relocates the problem. Any revised commitment follows the existing authority arrangements, with its effect on expected benefits and operating costs carried forward for Chapter 17's reforecasting discussion.

Benefit ownership must survive this transfer. The Value and Evidence Map connects usable capability to changed practice, operating outcomes and benefits, but those arrows remain propositions to be examined. The cohort's forty-five operators provide evidence about one definition of correct use, while the handling-time chart concerns one process characteristic. Neither establishes fewer late payments, better supplier relationships or realised cash savings. Maya retains the relevant Measurement Agreement, including its eligible population, event definition, timing, exclusions, source and adverse effects. Where benefit observation extends beyond transition support, its owner and review dates must remain explicit. The controller should also preserve the baseline and explanation of other changes that could influence the outcome. Operational stability creates conditions in which value may be sustained and examined; it does not remove the need to test the benefit claim or justify the next commitment of resources.

Turn observations into the next operating decision

At the transition review, Lena should be able to reconstruct the argument without compressing it into one colour. The authorised population and configuration define the scope; demonstrations and unresolved findings establish what is known before use; cohort evidence shows who is using the process and under what conditions; monitoring reveals how performance is developing. Maya explains the receiving operation's capacity and proposed response, Alex explains remaining technical dependencies, and Samira exercises the relevant authority over continued commitment. If the day-seven signal appears while expert support is being reduced, the immediate question is what changed and whether the operating boundary remains justified. The group may hold the reduction, investigate and verify a correction while maintaining permitted work. It should neither declare the whole programme unsuccessful from one signal nor ignore the signal because the delivery packages were previously accepted.

The worked examples can also be used as a practical review exercise. Ask a controller to reconcile the claim that adoption is ninety per cent with the cohort evidence: ninety per cent describes training attendance, while demonstrated competence covers seventy-five per cent and repeated correct use without unplanned specialist rescue covers 56.25 per cent. Ask whether ninety successes from one hundred rehearsal attempts satisfy the illustrative lower-bound rule: the Wilson lower endpoint is 82.56 per cent, below the required eighty-five. Finally, ask when the monitoring example first requires investigation under its stated rule: day seven, when the statistic exceeds its upper limit. Each answer corrects a different reporting error. The exercise does not create a combined readiness score. Its purpose is to show that the same apparently favourable story can conceal a changed denominator, limited evidence or emerging deterioration, each requiring a different response.

Chapter 16 makes movement into use a continuing responsibility for evidence, judgement and response. The operational claim must remain specific enough to test, the permission narrow enough to support, and the observation strong enough to reveal when the original conditions no longer hold. The controller's contribution is to connect those elements without taking ownership away from the people accountable for the service and its benefits. Readiness demonstrations, adoption measures and process monitoring become useful when they change a decision, preserve a necessary restriction or direct corrective work. They become misleading when treated as interchangeable proof of completion. The next chapter examines how such observations change forecasts and continued justification across the lifecycle. For now, the essential achievement is a receiving operation with understood responsibilities, available resources and a working means of recognising and responding to evidence, while the outstanding benefit questions remain visible.

References

- Brown, L. D., Cai, T. T. and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16(2), 101–133. doi: [10.1214/ss/1009213286](https://doi.org/10.1214/ss/1009213286)
- Burton-Jones, A. and Straub, D. W., Jr. (2006). Reconceptualizing system usage: An approach and empirical test. *Information Systems Research*, 17(3), 228–246. doi: [10.1287/isre.1060.0096](https://doi.org/10.1287/isre.1060.0096)

- Hong Kong Legislative Council Select Committee. (1999). Report of the Legislative Council Select Committee to inquire into the circumstances leading to the problems surrounding the commencement of the operation of the new Hong Kong International Airport at Chek Lap Kok since 6 July 1998 and related issues. Volume I, Chapter 5, paras. 5.141–5.142 and 5.154–5.157. www.legco.gov.hk/yr98-99/english/sc/sc01/papers/chapter5.pdf
- Jones, L. A., Champ, C. W. and Rigdon, S. E. (2001). The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*, 43(2), 156–167. doi: 10.1198/004017001750386279
- NIST/SEMATECH. (n.d.). EWMA control charts. e-Handbook of Statistical Methods, section 6.3.2.4. Accessed 19 September 2026. www.itl.nist.gov/div898/handbook/pmc/section3/pmc324.htm
- Office of the Auditor-General. (2020). Inland Revenue Department: Benefits management for the Business Transformation programme. B.29[20f], paras. 3.33–3.37. oag.parliament.nz/2020/ird-benefits/docs/ird-benefits.pdf
- Pluye, P., Potvin, L., Denis, J. L. and Pelletier, J. (2004). Program sustainability: Focus on organizational routines. *Health Promotion International*, 19(4), 489–500. doi: 10.1093/heapro/dah411
- Scaccia, J. P., Cook, B. S., Lamont, A., Wandersman, A., Castellow, J., Katz, J. and Beidas, R. S. (2015). A practical implementation science heuristic for organizational readiness: R = MC². *Journal of Community Psychology*, 43(4), 484–501. doi: 10.1002/jcop.21698

Reforecasting from New Evidence



IN THIS CHAPTER

Revising the remaining case as knowledge and performance change.

Change the forecast without rewriting the commitment

Samira has received a revised cost forecast, but the number does not tell her what to approve. Alex can explain the additional delivery work; Maya is less confident that the intended operating benefit will arrive at the rate previously assumed. Lena must bring these observations into one decision without pretending that they concern the same evidence or accounting boundary. A transformation can remain worth completing after its delivery forecast worsens, or lose its justification while its expenditure remains within budget. Reforecasting therefore concerns the remaining relationship between resources, operating change and value. Its purpose is to explain what the organisation now expects, why that expectation differs from the earlier account, and which action the evidence supports. The original commitment remains necessary for accountability, while the current forecast informs decisions that can still change the future. Neither number can responsibly perform both roles without its meaning being made explicit.

The shared lifecycle shows this work as a continuing activity across the transformation. Evidence can arise from investigation, design, procurement, delivery, operating rehearsal or actual use; the controller should not wait until the programme reaches a final evaluation stage. This chapter returns to Chapter 15's day-twelve cut-off for a clearly stated hypothetical revision, then adds a separate teaching extension for prospective resource benefits. Chapter 16's operator cohort, rehearsal sample and monitoring series remain illustrations rather than observed results for this financial account. Their methodological lesson carries forward: a signal can justify investigation without establishing its cause or the value of every consequence. The chapter applies the existing Credible Control Method and Chapter 10's decision reasoning to a dated reforecast. A revised expectation becomes an input to authorisation, while any change to scope, funding or performance commitment requires the appropriate decision in its own right.

The transformation lifecycle

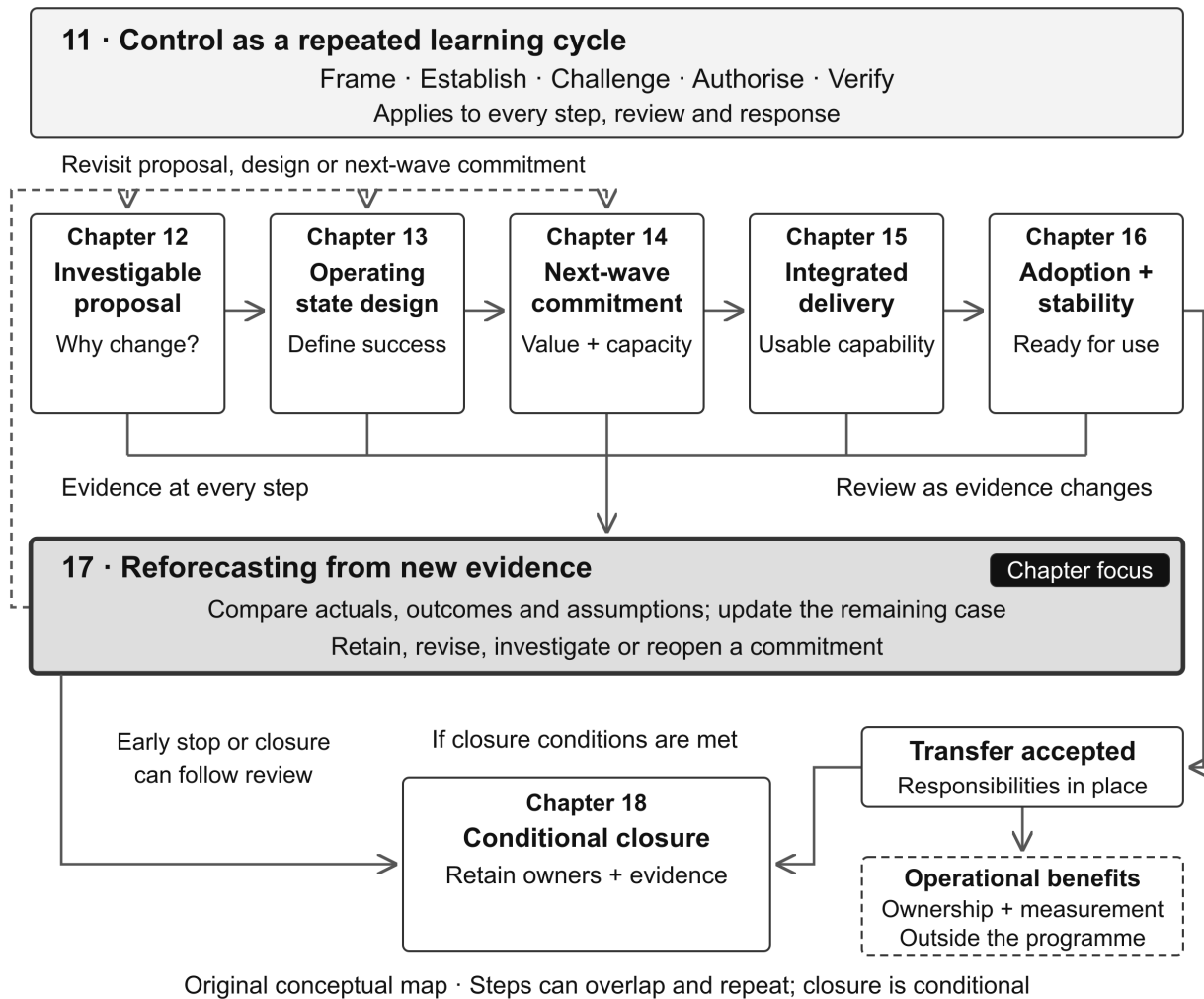


Figure 17.1 The shared lifecycle highlights continuing evaluation and reforecasting. Evidence from any lifecycle area can change the current assessment and reopen a proposal, design or commitment. A revised forecast informs an authorised response; it does not itself change the baseline.

Nordhaus (1987) examines whether forecasts incorporate information already available to their makers. For a fixed target under a squared-error criterion, an efficient forecast's next revision should not be predictably related to earlier revisions. This concerns an estimate of the conditional mean, rather than a cost ceiling or conservative allowance. It provides a diagnostic idea, not a rule that every worsening sequence demonstrates concealment. New findings, changed scope and a short observation history require separate consideration. Lena can preserve dated versions and ask whether the same known difficulty has been admitted repeatedly in small increments. Alex should distinguish genuinely new evidence from information available at the previous review, while Samira protects the reporting of unwelcome findings. The chapter uses this reasoning to improve challenge and traceability; a few project forecasts do not constitute a statistical test of efficiency or dishonesty.

Reconcile what changed at the same reporting cut off

Begin with the boundary of the new information. At the end of working day twelve, Chapter 15 recorded actual direct cost of US\$61,000 and coordination cost of US\$12,000. Remaining direct work was estimated at US\$43,000 and coordination at US\$8,000, giving a total forecast of US\$124,000 and completion on day twenty. For this chapter, assume Alex's technical review now identifies additional work in package C requiring one extra working day and US\$3,000 of direct resource cost. This converts the earlier sensitivity into a new hypothetical finding; it does not assert that the finding actually occurred. At the unchanged coordination rate of US\$1,000 per day, the additional duration adds another US\$1,000. Actuals have not changed because the cut-off has not advanced. The revision concerns the remaining work and its consequences, rather than a newly recorded expenditure or an authorised expansion of purpose.

Table 17.1 reconciles the two versions line by line. Actual resource cost remains US\$73,000. Remaining direct cost rises to US\$46,000 and remaining coordination to US\$9,000, giving US\$55,000 still required and a total forecast of US\$128,000. The increase from the previous forecast is US\$4,000. Keeping actuals fixed makes the source of the change visible and prevents normal spending since an earlier date from being mistaken for a forecast deterioration. When the cut-off does move, Lena should separately reconcile new actuals, work removed from the remaining estimate, changes in quantities or rates, newly identified scope and corrections. Purchase commitments require the same discipline: an order already represented in the remaining estimate is not added again merely because a supplier report now displays it. The bridge should explain the changed estimate, with its source, rather than become another total assembled from unreconciled reports.

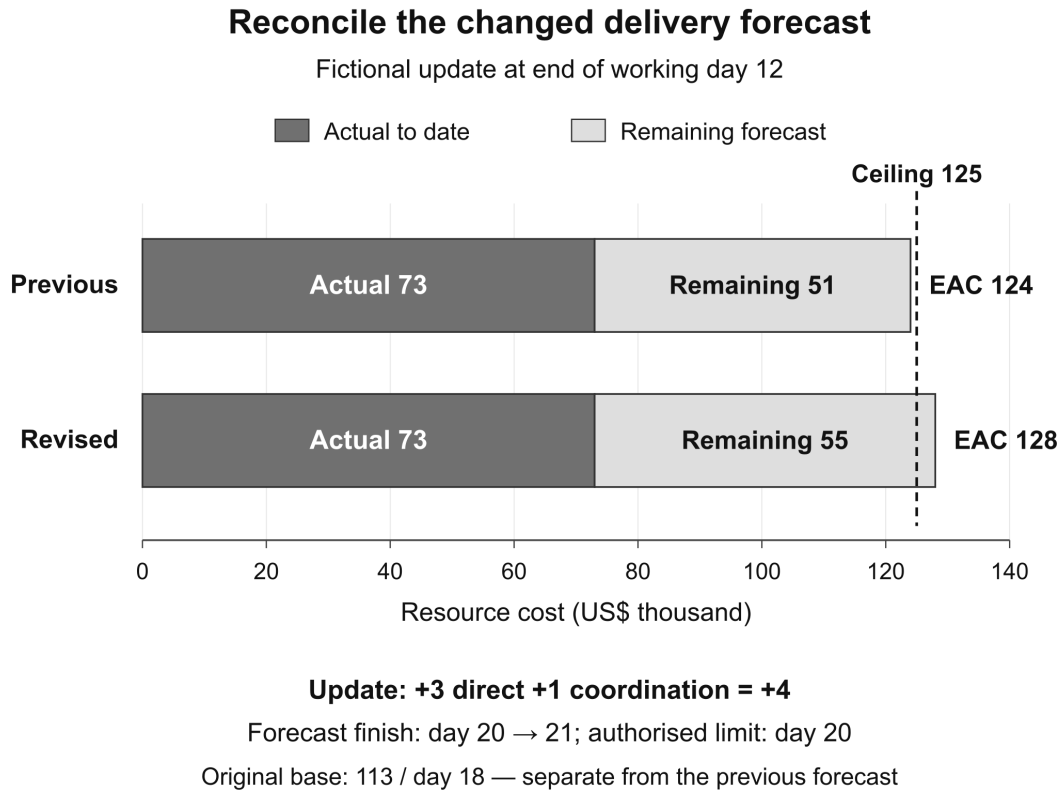
Table 17.1 Reconcile two forecasts at the same cut-off

Cost element	Previous	Revised	Change
Actual direct work	61	61	0
Actual coordination	12	12	0
Remaining direct work	43	46	+3
Remaining coordination	8	9	+1
Forecast total	124	128	+4

All values in US\$ thousand. The cut-off remains the end of working day twelve. Actual resource cost is unchanged at 73; remaining cost rises from 51 to 55. The revised finish is day twenty-one, exceeding the authorised day-twenty limit.

The original wave base remains US\$113,000 with completion on day eighteen; the authorised limits remain US\$125,000 and day twenty. Consequently, the new US\$128,000 forecast is US\$15,000 above the original base, US\$4,000 above the previous forecast and US\$3,000 above the authorised ceiling. These are three different comparisons. The revised completion on day twenty-one is likewise later than both the original plan and the authorised limit. Figure 17.2 keeps the references visible. Lena should report the latest credible assessment promptly even

when it breaches a limit; suppressing it until approval would confuse knowledge with permission. Samira then considers the response through the established authority arrangements. If a revised baseline is subsequently approved, the original commitment and change history remain retrievable. Approval can establish a new control reference, but cannot make the earlier variance or the resources already consumed disappear from the programme's account.



Revised forecast exceeds the cost ceiling by 3; it does not authorise the overrun.

Figure 17.2 Original fictional reconciliation at the same day-twelve cut-off. Actual resource cost remains US\$73,000. New evidence increases remaining direct cost by US\$3,000 and coordination by US\$1,000, changing the total forecast from US\$124,000 to US\$128,000. The original base and authorised limits remain visible.

Rebuild the benefit expectation through its operating mechanism

A new teaching extension makes the benefit revision equally inspectable. Suppose the earlier annual model covered 12,000 eligible cases, assumed effective use for ninety per cent, and assigned ten minutes released per effectively processed case. At an illustrative resource value of US\$30 per hour, with all released time productively redeployed, this gives 1,800 hours worth US\$54,000; subtracting US\$4,000 of incremental operating cost leaves US\$50,000. Now assume prospective effective use is revised to sixty per cent, with only seventy-five per cent of released time expected to support useful alternative work and incremental operating cost revised to US\$2,000. The result is 1,200 released hours, 900 productively redeployed hours, US\$27,000 of gross resource value and US\$25,000 net annually. These are explicit hypothetical assumptions about a future year, not ratios inferred from Chapter 16's independent datasets. Each revised input needs its own evidential basis before the model can support an actual investment recommendation.

This model values a proposed use of capacity; it does not establish a cash saving. Payroll expenditure could remain unchanged while staff use released time to clear a backlog or improve another service. Maya must identify the alternative work, demonstrate that the capacity is available when needed and prevent its value from being claimed again elsewhere. An hour fragmented into unusable intervals may not support the same proposition as a scheduled hour of available capacity. The forecast should also retain quality, rework and adverse effects, because faster processing can create additional work downstream. To recognise realised benefit, Lena needs observed outcomes and an appropriate comparison with what would otherwise have happened, using the Measurement Agreement and the evaluation methods already developed. A revised model can describe a more credible expectation without proving either the forecasted improvement or the transformation's causal contribution to it.

Breese (2012) examines benefits management through reflection on government-funded regeneration programmes in the United Kingdom. He describes targets that became untenable when external dependencies changed and an instance where replacement projects broadly restored outputs while longer-term regeneration impact was reduced. This is contextual analysis, not a controlled estimate of the effect of benefits management. For Maya and Lena, the lesson is to re-examine the causal proposition when revising the forecast: preserving a deliverable count does not necessarily preserve the intended operating improvement. The updated record should distinguish changed evidence, changed assumptions and an authorised change of objective. Retaining the previous version makes that distinction inspectable. A revised target can be justified, but should not erase the original commitment or turn a smaller expected benefit into evidence that the original promise has been fulfilled.

In its March 2019 investigation of GOV.UK Verify, the National Audit Office reported that the Government Digital Service (GDS) had reduced expected benefits for 2016–17 to 2019–20 from £873 million, estimated in 2016, to £217 million. It attributed most of that revision to lower take-up than assumed. GDS classified the benefits as non-cash releasing, and the NAO could not replicate or validate them from the evidence provided (National Audit Office, 2019, paras. 1.18–1.19). This is a revision of expected benefits, not a measured loss of cash savings. The book's

application is to rebuild the link from service use to the proposed benefit, retaining the forecast horizon, calculation basis and evidence status. A newer number becomes useful when the changed assumptions and their support are available for challenge.

Challenge adjustments and dependencies together

Fildes et al. (2009) examined judgemental adjustments to statistical demand forecasts in four companies. Adjustments improved average accuracy in three companies, but smaller changes often reduced accuracy and upward adjustments were less successful than downward ones. The result does not establish that project managers should avoid small corrections or rising cost forecasts: an upward demand adjustment has a different business meaning from acknowledging additional expenditure. It does support treating an override as a proposition that can later be examined. Lena should retain the estimate before adjustment, the revised estimate, the new information and the outcome when it becomes observable. Alex can explain why a technical finding changes remaining work, while Maya explains any corresponding benefit effect. The purpose is to learn which judgements add information, rather than reward the appearance of an unchanged forecast.

The challenge should follow the mechanism linking delivery and operations. If slower adoption extends specialist support while delaying useful redeployment, those consequences belong in the same scenario. It would be inconsistent to reduce benefit for weak adoption while automatically removing all transition support on the original date. Conversely, the existence of a plausible dependency does not justify adding an arbitrary percentage to every cost line. Alex and Maya should explain which activity, quantity, duration or operating condition changes, and Lena should ensure the effect appears once in the appropriate account. The additional coordination day in Table 17.1 is already included; a separate general delay allowance must not silently include it again. A central scenario and coherent alternatives can expose what matters without assigning unsupported probabilities. A percentile requires a defensible probability model, and a labelled expected value requires more than choosing the number that the team hopes will occur.

Different evidence also operates on different clocks. Delivery actuals may be complete to a month-end date while adoption observations cover only the previous week and benefit outcomes require a full operating cycle. Lena can still assemble a useful current assessment if each component states its period, population and status. Missing observations should remain visible rather than being replaced with the target. Currency, price basis, operating scope and treatment of contingency must be consistent before totals are compared. Where they differ, the reconciliation should show the conversion or explain why the values cannot yet be combined. The forecast record should distinguish evidence-led changes from management actions proposed to improve the outcome: an unapproved recovery measure belongs in a conditional alternative, not silently inside the main forecast. This allows Samira to see both the present trajectory and the additional commitment on which a better trajectory would depend.

Compare the value of decisions that remain available

The continuation decision concerns consequences that can still be changed. HM Treasury's Green Book (2026) distinguishes sunk costs from the opportunity costs of resources that retain alternative uses; it also distinguishes monetised value from cash effects. Apply that logic here without importing a jurisdiction's prescribed appraisal settings. The US\$73,000 already incurred remains in the historical resource account but is excluded from the prospective option ranking. Unpaid settlement of past work remains a funding obligation; excluding sunk resource cost does not cancel the invoice. For the following simplified organisational appraisal, compare incremental resource costs and monetised benefits against continuing the existing service. Use a common real price basis, a three-year horizon and an illustrative real discount rate of five per cent. This is a teaching assumption, not a recommended organisational rate. Any valuable alternative use of existing assets would also need recognition rather than being dismissed merely because the assets were purchased earlier.

Table 17.2 compares three hypothetical responses. Continuing requires US\$55,000 of future resource cost and provides US\$25,000 of annual net resource benefit. Redesign requires US\$65,000 and provides US\$35,000 annually. Stopping requires US\$8,000 of close-out resource cost and has no additional monetised benefit in this model. For simplicity, all future implementation costs occur at the decision date, benefits occur at each of three common year-ends, and residual values are zero. Annual net benefits already deduct incremental operating costs. Discounting one unit received at those year-ends gives approximately 0.95238, 0.90703 and 0.86384, whose sum is 2.723248. Multiply each annual net benefit by that factor and subtract its future implementation or close-out cost. Prospective net values are consequently US\$13,081 for continuation, US\$30,314 for redesign and minus US\$8,000 for stopping. These conditional values are neither realised returns nor probability-weighted expectations.

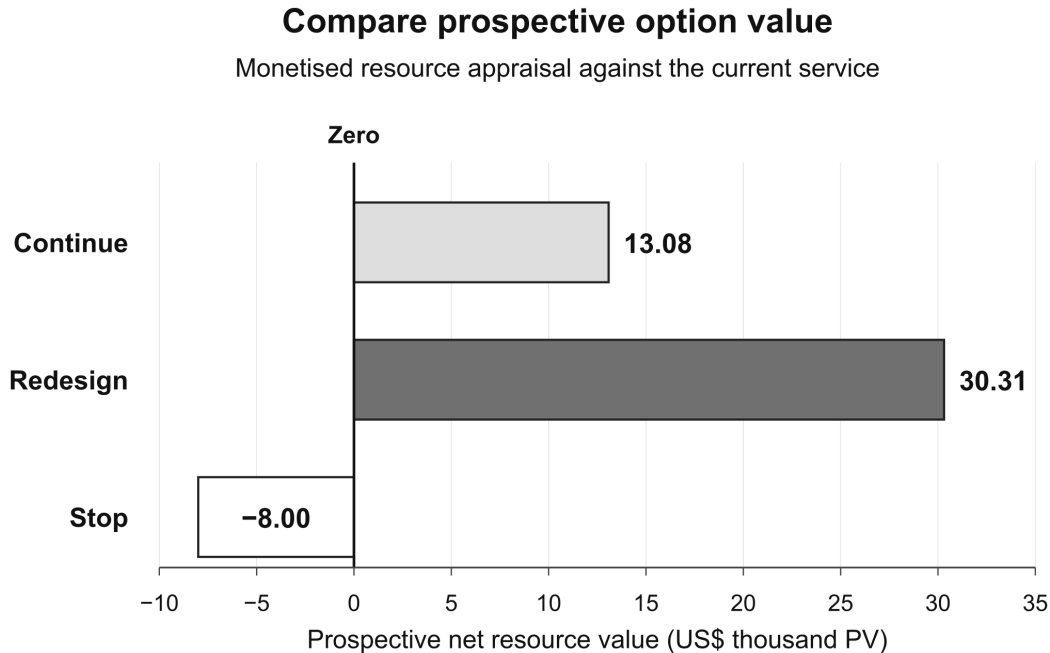
Table 17.2 Compare the value of future options

Option	Future cost	Annual net benefit	PV of benefits	Net value
Continue	55.00	25.00	68.08	13.08
Redesign	65.00	35.00	95.31	30.31
Stop	8.00	0.00	0.00	-8.00

All values in US\$ thousand of incremental monetised resource value against the current service. Future costs are placed at the decision date; net benefits occur at each year-end for three years, discounted at an illustrative real 5%. No residual value is assumed. Past cost of 73 is excluded from this prospective comparison.

Under these assumptions, redesign exceeds continuation by approximately US\$17,232 of prospective resource value. This does not establish positive value for the whole investment measured from its inception. Its remaining cost of US\$65,000 combines with the existing US\$73,000 to give a candidate whole-wave resource cost of US\$138,000, beyond the current ceiling. A favourable ranking does not supply funding or authority. Samira must also establish

whether the option meets essential operating requirements and whether its proposed benefit mechanism is credible. The comparison therefore supports a conditional recommendation, rather than permission to spend. Stopping has consequences too: the assumed US\$8,000 must represent the relevant close-out work, while the existing service continues under the common comparator. If cancellation liabilities, retained capability or service obligations differ materially from these assumptions, the option must be rebuilt. Figure 17.3 shows the ordering of prospective value within the stated resource model while leaving those feasibility, affordability and governance judgements explicit.



Redesign sensitivity: annual 28 → value 11.25 (same option)

Future 3-year appraisal · assumed 5% discount rate · not a cash budget

Past 73 excluded from ranking; funding obligations remain.

Redesign whole-wave resource cost: 138; requires authority.

Fictional conditional comparison; no probability or confidence claim.

Figure 17.3 Original conditional comparison of prospective resource value. Redesign ranks highest under the central assumptions, but its separate downside sensitivity falls below continuation. Past cost remains in the historical account. The chart establishes neither cash affordability nor authority to change the commitment.

The most useful sensitivity asks when the preferred course changes. Redesign costs US\$10,000 more than continuation. Over the same horizon, it must therefore provide approximately US\$3,672 more annual net benefit to offset that difference: divide ten thousand by 2.723248. With continuation at US\$25,000 annually, the switching value for redesign is approximately US\$28,672. If redesign yields only US\$28,000, its prospective net value falls to approximately US\$11,251, below continuation's US\$13,081. This is a separate sensitivity, not another observed outcome or an assigned probability. Maya can now focus further investigation on whether the redesigned operation credibly supports benefit above the switching value, while Alex examines any associated cost or timing change. If redesign delays the benefit years, the discount factors

must change too. Chapter 10's information-value reasoning can then determine whether additional evidence is worth obtaining before the relevant commitment becomes irreversible.

Reopen the commitment when the route no longer fits the purpose

Pan et al. (2006) analyse commitment during an e-government project in a United Kingdom local council, using Lewin's change theory to interpret movement from escalation towards an alternative course of action. Their account highlights the possibility that stakeholders interpret facts through previously accepted beliefs. This single-case analysis does not prove that a particular intervention reliably produces de-escalation. Applied here, a revised forecast needs a decision process able to consider its implications, not simply a more accurate number. Samira should examine feasible continuation, redesign and stopping choices against current evidence, while identifying whose assumptions and responsibilities would change. Lena can preserve conflicting interpretations and the evidence that would resolve them in the existing decision record. The comparison must include future obligations and consequences; an adverse variance alone does not establish that stopping is preferable.

The US Air Force's Expeditionary Combat Support System illustrates a reassessment reaching beyond a revised completion date. GAO reported that the Air Force restructured the programme in early 2009 after finding that it had underestimated its size and complexity. During a second critical change in 2012, the Air Force determined that the system was no longer a viable route to the statutory 2017 financial-auditability requirement: scope was contracting while costs and delays increased. The Department of Defense cancelled the programme in December 2012 (United States Government Accountability Office, 2013, pp. 62–63). The book's application is to re-examine whether the remaining scope and timing can still deliver the required outcome. The report does not establish an optimal cancellation date or the benefits of an alternative; those require their own evidence.

The existing Decision under Uncertainty Memorandum can hold the resulting recommendation. Lena states the current decision, dated evidence, forecast differences, feasible alternatives, unresolved assumptions and consequences of delay. Alex confirms the remaining delivery basis; Maya confirms the operating and benefit assumptions; Samira identifies the authority required for the proposed action. Where the current forecast breaches an authorised limit, the record should specify what work remains permitted while the issue is resolved. Possible responses include bounded investigation, recovery, redesign, deferral or a controlled stop, depending on the actual circumstances. The recommendation needs a verification condition and a review date, so that approval does not become the end of the control cycle. A decision to continue can be justified despite adverse evidence, but its reasoning must address that evidence. A decision to stop likewise needs a plan for obligations and value that remain after delivery activity ceases.

Learn from forecast performance without moving the goalposts

Hyndman and Koehler (2006) show why an apparently convenient forecast-accuracy measure can mislead. Absolute errors preserve the outcome's units, while percentage measures can become undefined or unstable when actual values are zero or close to zero. Measures also differ in how they weight large errors. Applied here, Lena should distinguish the direction of error from its magnitude and avoid combining pounds, working days and adoption percentages into one unexplained score. A mean signed error can reveal the direction of the errors in the observed set, but positive and negative errors can cancel. Mean absolute error describes their size without that cancellation. Neither establishes a forecasting process's reliability from three observations. Comparable scope, forecast horizon and outcome definitions remain necessary chapter safeguards when interpreting the evidence and deciding what to investigate next.

A separate fictional example makes the measurement question concrete. Three comparable outcomes were forecast one month ahead at 110, 125 and 85 units; the eventual observations were 100, 120 and 80. Retaining Chapter 8's convention of actual minus forecast gives errors of minus ten, minus five and minus five. Their mean is minus 6.67 units, indicating overprediction in these three observations, while mean absolute error is 6.67 units. The signed measure describes direction; the absolute measure describes magnitude without cancellation between positive and negative errors. Three observations do not establish persistent bias or a reliable correction factor. Lena should examine the definitions, information available at issue and reasons for the differences before adjusting future forecasts. These outcomes are separate teaching data, not measurements of the manufacturer's annual benefits. The exercise evaluates prediction against observation, not whether the transformation caused the observed result.

Forecast versions must therefore be retained with their issue dates, target periods, scope and conditions. Comparing a forecast made two days before completion with one made a year earlier cannot by itself demonstrate improved forecasting capability. Where a management action changes the eventual outcome, the original conditional forecast and the action should remain visible; otherwise an effective intervention can be misclassified as forecasting failure. Lena should compare like horizons, examine revisions as well as final errors, and retain a simple benchmark appropriate to the task. She should also preserve revisions to the measured actuals, because improved accounting or corrected source data can change apparent error without changing what the forecaster knew. Learning then concerns a reproducible process for using information, rather than rewarding whichever team produced the most favourable final number or revised its forecast most frequently.

Make the next commitment traceable to the evidence

A usable reforecast account can be read forwards and backwards. Starting from the latest recommendation, Samira can trace the option values to operating assumptions, remaining activities and dated observations. Starting from a new finding, Maya or Alex can trace which estimate changes, which dependency carries its effect and which decision may need reopening. Lena maintains those connections through the existing Wave Commitment Record, Measurement Agreement, Evidence and Learning Record and decision memorandum, without creating duplicate registers for the same fact. The account should show what changed since the previous version, why, who owns the judgement and what future observation would challenge it. It should also state what did not become known: a new schedule does not resolve adoption uncertainty, and a narrower estimate does not establish that omitted work has disappeared. Those distinctions make the forecast usable by the people who must act on it.

The chapter's examples provide three practical questions for a review. Does the US\$128,000 forecast authorise expenditure beyond the US\$125,000 ceiling? It does not; it reveals a decision that requires attention. Does the revised US\$25,000 annual resource benefit establish a cash saving? It does not; the model depends on useful redeployment and a defined valuation, with observed benefit still to be demonstrated. Does the US\$73,000 already incurred make continuation preferable? It does not determine the prospective ranking, although it remains essential to accountability and may leave funding obligations. These answers do not minimise cost, benefit or delivery performance. They place each item in the question it can answer. The controller can then challenge the assumptions that genuinely alter the next choice instead of allowing a single headline total to conceal several incompatible interpretations.

Chapter 17 makes reforecasting a disciplined connection between changing knowledge and continuing responsibility. The controller preserves original commitments, reconciles the latest evidence, updates the remaining delivery and benefit proposition, and explains which alternatives remain feasible. The decision maker then authorises an action with understood conditions, resources and a route to verification. This process can support continuation, a revised design or a conclusion that the programme should end; none follows automatically from one variance or one favourable calculation. The next chapter examines that ending and the responsibilities that must survive it. Before reaching closure, however, the organisation should be able to explain why its chosen course remains justified by the evidence now available. A credible forecast does not promise certainty. It gives the next decision a transparent basis and preserves the means to reconsider it when that basis changes.

References

- Breese, R. (2012). Benefits realisation management: Panacea or false dawn? *International Journal of Project Management*, 30(3), 341–351. doi: [10.1016/j.ijproman.2011.08.007](https://doi.org/10.1016/j.ijproman.2011.08.007)
- Fildes, R., Goodwin, P., Lawrence, M. and Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: An empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25(1), 3–23. doi: [10.1016/j.ijforecast.2008.11.010](https://doi.org/10.1016/j.ijforecast.2008.11.010)

- HM Treasury. (2026). The Green Book. Sections 6–8. www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government/the-green-book-2026
- Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. doi: 10.1016/j.ijforecast.2006.03.001
- National Audit Office. (2019, 5 March). Investigation into Verify. HC 1926, Session 2017–2019, paras. 1.18–1.19. www.nao.org.uk/wp-content/uploads/2019/03/Investigation-into-verify.pdf
- Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *The Review of Economics and Statistics*, 69(4), 667–674. doi: 10.2307/1935962
- Pan, G., Pan, S. L., Newman, M. and Flynn, D. (2006). Escalation and de-escalation of commitment: A commitment transformation analysis of an e-government project. *Information Systems Journal*, 16(1), 3–21. doi: 10.1111/j.1365-2575.2006.00209.x
- United States Government Accountability Office. (2013, 28 March). Major Automated Information Systems: Selected Defense Programs Need to Implement Key Acquisition Practices. GAO-13-311, pp. 62–63. www.gao.gov/assets/gao-13-311.pdf

Ending the Programme while Protecting Its Value



IN THIS CHAPTER

Ending temporary governance while preserving ownership and value.

End the temporary arrangement with continuity in place

A closure paper can look reassuring because the remaining work has become somebody else's responsibility. Samira needs to know whether that responsibility has actually been accepted, whether the receiving people can perform it and what will reveal a problem after the programme board stops meeting. Alex may have completed the intended delivery contribution while Maya still needs specialist support, reliable records and resources for benefits review. Lena's task is to make those boundaries visible before the temporary organisation disperses. Closure is a consequential commitment about how the organisation will govern what follows. It should explain why the programme can end, what has been achieved, what remains unresolved and how the continuing responsibilities will be exercised. The existence of a final report, an exhausted budget or a calendar deadline does not answer those questions. An ending becomes credible when its evidence and receiving arrangements support the decision being proposed.

Chapter 17 left the manufacturer's continuation choices conditional. This chapter therefore examines closure arrangements without assuming that a particular option was implemented or that its forecast became an actual result. A separate fictional closure account supplies the numerical teaching example. The shared lifecycle highlights the point at which temporary programme governance can end, while retaining the route for early stopping and the continuing operational benefits beyond closure. Those relationships matter because delivery acceptance, administrative closure, payment settlement, benefits observation and eventual retirement of an asset can occur at different times. The chapter brings the earlier controls disciplines together through a Closure and Continuing Obligations Record. Its purpose is to preserve the decision, evidence and accepted responsibilities, using the existing records rather than creating another disconnected archive. The receiving organisation should be able to understand the account and act on new information after the programme's original authors are no longer available.

The transformation lifecycle

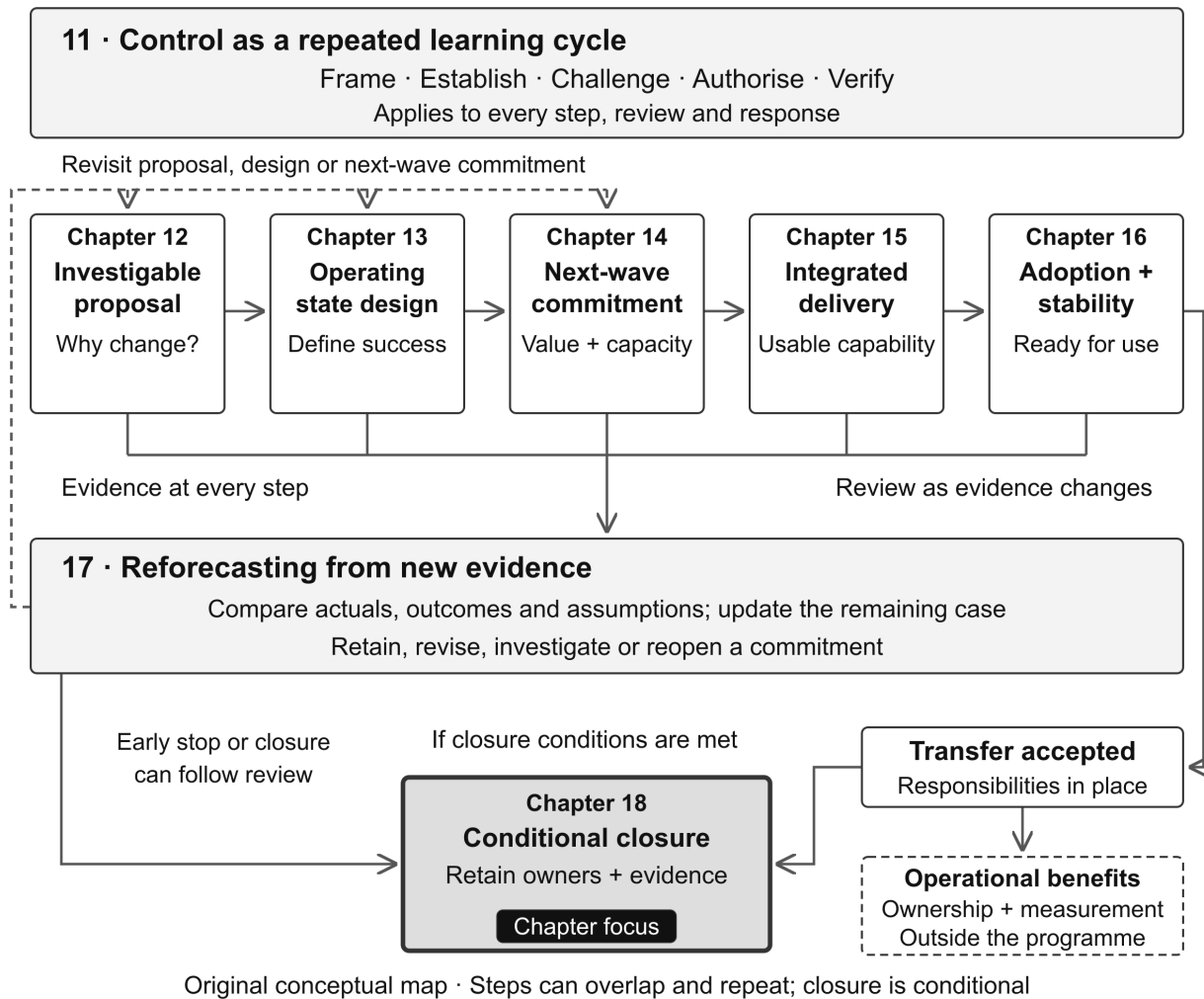


Figure 18.1 The shared lifecycle highlights conditional programme closure. An authorised ending can follow completed contribution or early stopping. Accepted transfer preserves continuing responsibilities, while operational benefits remain subject to ownership, measurement and response after temporary governance ends.

Whyte and Nussbaum (2020) study the transition between project delivery and operations in the London megaproject setting. Their account shows different understandings of time meeting as delivery teams seek an ending while operators prepare for continuing work. They identify documents, procedures, gradual transitions and tests as ways of maintaining continuity across that change. This interpretive evidence does not establish that one handover sequence works everywhere. Applied here, Samira should distinguish the date when programme governance ends from the dates when individual obligations transfer or benefits become observable. Maya needs the operating responsibilities to remain usable after Alex's team disperses. Lena can preserve the evidence, unresolved conditions and receiving owners in the existing records. A closure decision can then end a temporary arrangement without implying that every related activity or value claim has reached its endpoint.

State what the closure decision authorises

The United Kingdom's GovS 002 standard requires controlled closure for both completed scope and premature termination. Its closure provisions address achievement to date, unfulfilled requirements, accepted transfer of ongoing responsibilities and arrangements for later reviews; its subsequent-outcomes provision retains the question of benefits and operating performance (Government Project Delivery, 2025, clauses 6.4.8–6.4.9). These are expectations within that standard's government context, rather than evidence that a checklist guarantees successful closure internationally. The useful discipline here is to state the basis of the ending and preserve the work that must continue. Lena should distinguish what the programme has completed from what has been transferred, withdrawn or left conditional. The receiving authority needs to accept the relevant obligations before the temporary arrangement disappears. A closure report should make those distinctions inspectable, rather than describe every item as complete because it has acquired a new owner.

Planned completion and early stopping consequently require different narratives. For planned completion, the record compares the authorised contribution with accepted outputs, operating capability and outcomes demonstrated to date, retaining explanations for changes. For an early stop, it identifies what will not be delivered, why continuing is no longer justified and which consequences must be managed. Undelivered scope does not become an achievement through careful wording. In either case, a critical operating weakness cannot be made acceptable simply by transferring its entry to a business risk register. The responsible authority must understand the exposure and the conditions under which the service may continue. Some residual work can be transferred with adequate resources and oversight; other gaps prevent the proposed closure. Lena should explain the consequence of each material gap, so that Samira can decide what must be resolved and what can responsibly remain with a continuing organisation.

Figure 18.2 applies the Credible Control Method to this final commitment. Frame the proposed ending and its scope; establish achievement, remaining obligations and receiving arrangements; challenge their adequacy; authorise the appropriate action; and verify the continuing work through its new governance. The Closure and Continuing Obligations Record connects that reasoning to the Wave Commitment Record, Measurement Agreement and Evidence and Learning Record. It is an original practice aid, not a validated readiness score. A high percentage of completed actions should not compensate for one necessary responsibility that nobody can perform. Equally, an unresolved minor administrative task need not keep the entire programme structure alive if a competent successor has accepted it with the required resources and authority. The decision concerns whether the remaining organisation can govern the work, rather than whether the temporary team can produce a report with no open lines.

End temporary governance; preserve responsibility

Practical application within the Credible Control Method

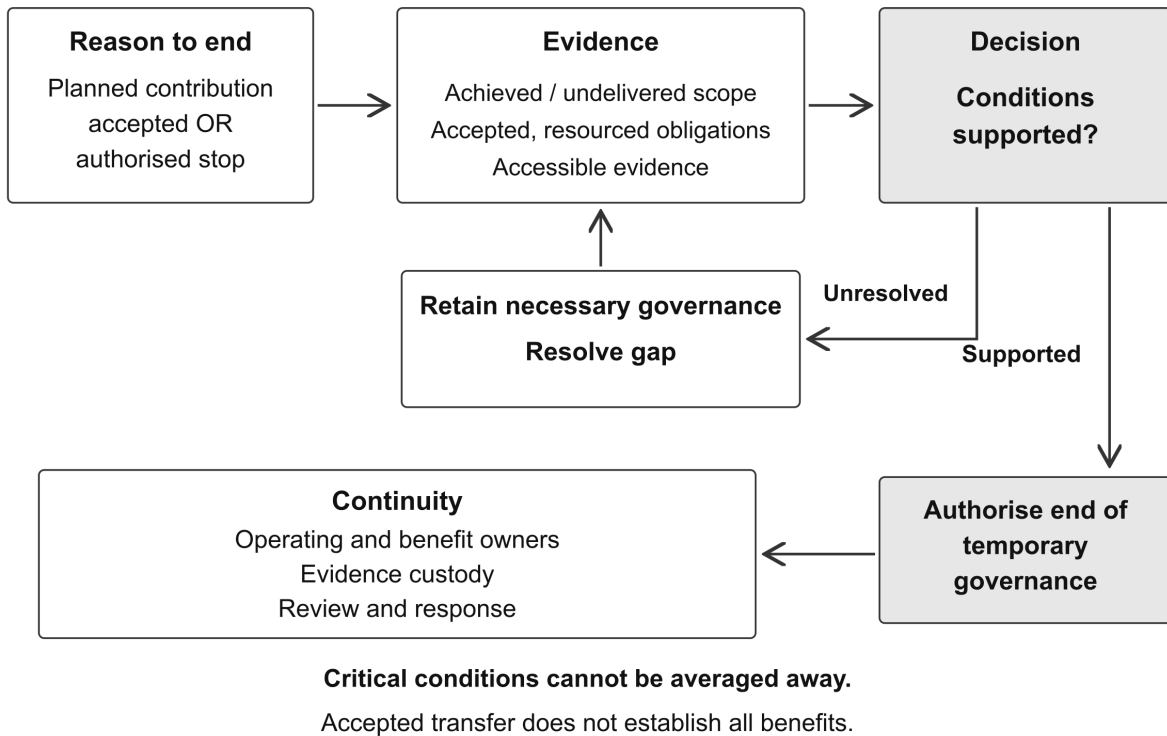


Figure 18.2 Original application of the Credible Control Method to closure. The evidence must support the stated ending and the receiving arrangements. An unresolved necessary condition retains the governance needed to resolve it; supported conditions inform authorisation and continuing ownership, not an automatic declaration of benefit success.

Obtain a transfer that the receiving organisation can perform

Zhang et al. (2024) examine how Beijing Daxing International Airport organised its transition into operations. Their longitudinal case connects continuity with organisational arrangements spanning delivery and operations, including coordination mechanisms and people moving between the two settings. The findings arose within a strong owner organisation and do not establish a universally superior structure. For the manufacturer, the practical question is whether the receiving organisation can carry the obligations being transferred, rather than whether a handover meeting has occurred. Maya should have access to the people, information and authority needed to respond when an operating condition changes. Alex should identify any specialist support that must continue and its agreed end condition. Lena should make those arrangements traceable to accepted obligations. This is an application of the case's continuity reasoning, not a prescription to reproduce its airport organisation.

Consider the separate fictional proposal in Table 18.1. Assume the defined service contribution has been accepted and the receiving operation has confirmed its ordinary staffing, authority and budget. The finance lead accepts US\$7,000 of remaining contract close-out work, and the records custodian accepts US\$5,000 for an archive migration while source records remain accessible.

Maya also accepts responsibility for continuing benefits evaluation, but the resources needed to perform those reviews have not been confirmed. The proposal is therefore not ready for closure on its stated terms. Writing Maya's name in the final row does not create capacity, and the US\$12,000 reserved for the two residual work packages does not fund an unspecified evaluation service. Lena should retain the gap, identify the required decision and obtain evidence of the receiving arrangement before recommending the proposed ending. The table is a proposal for transfer, not a record that closure has already happened.

Table 18.1 A fictional proposal for continuing obligations

Obligation	Receiving owner	Position before closure	Planned follow-up
Contract close-out work	Finance lead	Accepts scope and US\$7,000 funding	Review by day 30
Archive migration	Records custodian	Accepts US\$5,000; source records remain accessible	Verify retrieval by day 45
Operating service	Maya	Accepts authority, staff and operating budget	Routine review continues
Benefits evaluation	Maya	Resources unconfirmed; closure condition unmet	Proposed reviews on days 90 and 180

This is a pre-closure proposal, not evidence that closure has occurred. Days are elapsed calendar days after any authorised closure. The US\$12,000 covers only the two residual work packages. Continuing operations and benefit-review resources have separate funding arrangements; the benefit-review gap must be resolved before this proposed closure is supported.

The Canterbury Earthquake Recovery Authority ceased operating in April 2016 while recovery responsibilities moved to other bodies. New Zealand's Auditor-General reported that CERA had established a change-management office to organise the handover and support staff. Information, knowledge and skills had been effectively transferred, and the receiving agencies' 2016 annual audits had identified no significant issues with the transfer of roles and responsibilities (Office of the Auditor-General, 2017, paras. 1.23, 2.40 and 2.45). This favourable finding concerns the transfer, not proof that all recovery objectives or benefits had been achieved. The book's application is to close a temporary organisation through an explicit transfer of work, capability and evidence. Continuing responsibility belongs with identified successors, while closure needs a record of what was handed over and what remains to be verified.

Acceptance should identify a continuing role and organisation, not depend solely on the availability of one named individual. Maya may move to another position, the supplier may change its account team, or the receiving service may be reorganised. The transfer therefore needs a route for succession, access to the relevant records and an escalation authority that survives those changes. The Ownership and Capacity Framework already provides the necessary questions about commitment, authority, capacity, evidence and continuity. Lena can test the arrangement by asking the recipient to explain a realistic exception: who acts if a remaining supplier deliverable is rejected, an operating measure deteriorates or an expected benefit fails to

appear? An answer that requires the dissolved programme board to decide reveals an incomplete transfer. The response may use existing business governance or a bounded residual arrangement, with its purpose and end condition made explicit.

Reconcile the financial account without hiding remaining work

Table 18.2 introduces a separate fictional closure account in thousands of US dollars; it is not an outturn for Chapter 17's options. Assume received work has an actual resource cost of 240. Cash paid for that work is 220, with 20 still unpaid and already included in the actual cost. The remaining contract close-out services cost seven and archive migration five, giving twelve of future resource work. Those future services are separate from settlement of the existing unpaid twenty. The forecast final resource cost is therefore 240 plus twelve, or 252. Calling 240 the final outturn would omit known remaining work, while adding the unpaid twenty again would count received work twice. The forecast should retain the scope, owner and estimate basis for each residual item, even if the temporary programme reporting arrangements are about to end.

Table 18.2 Reconcile a separate fictional closure account

Account item	Amount	Interpretation
Actual resource cost	240	Received work, including 20 unpaid
Cash already paid	220	Settlement of received work
Unpaid received work	20	Already included in actual cost
Future resource work	12	7 close-out work plus 5 archive migration
Forecast final resource cost	252	240 actual plus 12 future work
Future cash requirement	32	20 unpaid plus 12 future work

All values in US\$ thousand. This is a separate teaching account, not an outturn for Chapter 17. Future work is assumed to be paid at its stated cost, with no other cash adjustments. The 12 excludes settlement of the existing unpaid 20. A resource budget of 260 gives arithmetic headroom of 8, subject to the relevant release decision.

The cash question produces a different number. Under the example's assumption that the future work is paid at its stated cost and there are no other cash adjustments, future payment requirements are twenty for received work plus twelve for future work, giving thirty-two. The organisation has paid 220 and expects to pay another thirty-two, reconciling to the forecast resource total of 252 in this simplified account. Real accounts require their actual payment, tax, currency and funding basis, rather than assuming that every resource cost follows this pattern. Lena should keep the cost and cash views connected without treating them as interchangeable. Payment of the existing twenty reduces the outstanding settlement balance; it does not create another twenty of resource cost. Completion of the future twelve, in contrast, moves work from the remaining forecast into actual cost as the relevant service is received and recognised.

Suppose the authorised resource budget is 260. The arithmetic headroom against the forecast of 252 is eight, but that difference is not an automatic cash release or proof that all financial exposure has ended. The appropriate owner must consider remaining commitments, unresolved claims and the authority for returning or reallocating provision. Moving the twelve of future work to an operating cost centre does not reduce the whole programme's forecast cost; the receiving account and the programme history need a traceable connection. Continuing operations and benefits-review resources sit outside this residual-work boundary and require their own funding decisions. Lena should also distinguish a management forecast from the final contractual and accounting position. Administrative closure can transfer responsibility for finishing that work, provided the receiving arrangements are accepted, while the financial record continues to explain what remains estimated, unsettled or subject to a further decision.

Keep benefits observable after the programme ends

Williams et al. (2025) examine benefits through three United Kingdom public-sector cases, showing why delivery and subsequent value cannot be treated as the same responsibility. Their analysis identifies uncertainty over who should track benefits after implementation and recommends continuing review by the receiving organisation. The study offers contextual findings and recommendations, not proof that a particular review schedule produces better outcomes. Applied here, programme closure should leave Maya with a defined benefit, a usable Measurement Agreement, resources for observation and a route for acting on adverse evidence. Lena should preserve the original expectation alongside subsequent revisions and realised observations. The review period should follow the benefit's operating mechanism and relevant assumptions, rather than end automatically with the programme team. Delivering a capability can enable future value while leaving its extent and attribution to be examined.

Figure 18.3 shows a conditional timetable for the teaching proposal. Day zero means authorised closure after the necessary transfer gaps, including benefits-review resources, have been resolved. It does not mean that the present proposal has passed. Contract close-out work is reviewed by day thirty, archive migration and retrieval by day forty-five, and benefit reviews are proposed for days ninety and one hundred eighty. These are elapsed calendar days, separate from the working-day clocks in earlier examples. The dates are illustrative: a benefit dependent on seasonal demand, a full payment cycle or a longer behaviour change may require another observation period. Operating ownership and evidence custody continue beyond the displayed horizon. Each review should have a question, source, accountable recipient and response route, so that a disappointing result can influence the service. Scheduling an observation is not evidence that the expected benefit will have materialised by that date.

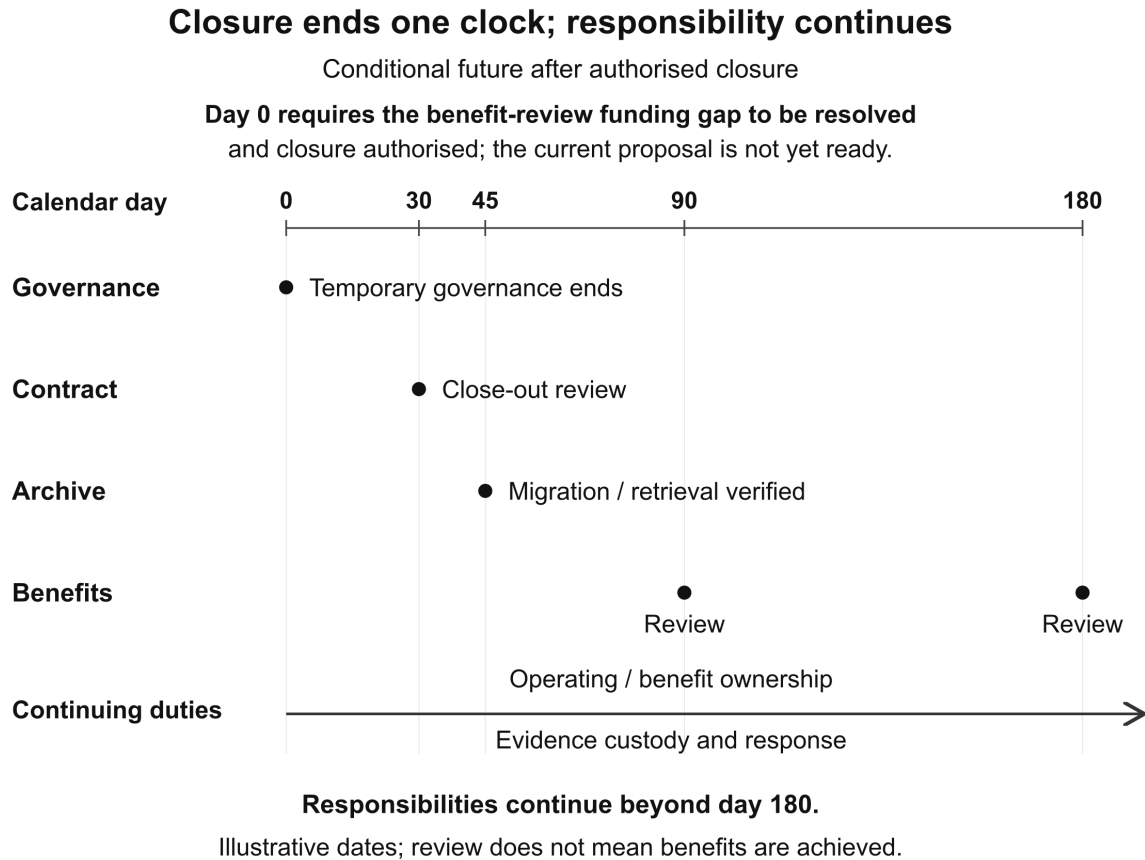


Figure 18.3 Original fictional timetable measured in calendar days from authorised closure. Day zero is conditional on resolving necessary transfer gaps. Residual work and later benefit reviews have different dates; operating ownership and evidence custody continue. Review dates are illustrative and do not certify benefit attainment.

Nicolaou (2004) studied postimplementation review in two organisations using enterprise resource planning systems. The exploratory analysis developed a concept of review quality, distinguishing it from conditions during implementation and from subsequent outcomes. It proposed relationships for further investigation rather than demonstrating that conducting a review causes improved financial performance. Applied to programme closure, this distinction prevents a signed evaluation report from being treated as proof that benefits have been secured. Maya needs a continuing route to examine operating results, unresolved changes and the assumptions supporting expected value. Lena should identify which questions can be answered at closure and which require later observations, with a named owner, usable evidence and a review date. Those arrangements preserve the possibility of learning after the temporary team disperses, without postponing every closure decision until all long-term outcomes are observable.

The Australian National Audit Office's November 2017 review of Reinventing the ATO found closure reports for 41 of 57 projects closed, cancelled or transferred into business as usual. Under the revised approach, closure reports recorded benefits to date, while later reviews were meant to record outcomes. None of the 21 projects closed under that approach had completed a post-implementation review at the time of the audit (Australian National Audit Office, 2017, paras. 4.20–4.26). This does not establish that all 21 reviews were overdue or that no benefits

occurred. The book's application is to assign, fund and schedule continuing measurement at closure. A report ending a project cannot certify benefits requiring later observation; the intended outcome and evidence trail must remain usable by whoever inherits that responsibility.

Leave knowledge that the next team can use

Newell et al. (2006) examined thirteen projects across six United Kingdom organisations, using 137 interviews over two years alongside available project records. They found limits to transferring project learning through stored lessons: potential recipients could regard the material as unhelpful or remain unaware of relevant knowledge. Their exploratory cases also show the importance of social connections; they do not establish that repositories are inherently ineffective. For closure, Lena should therefore preserve context and connect each significant lesson to a person or process able to use it. Alex can explain the conditions under which a technical remedy worked, while Maya identifies what the receiving operation needs to practise. A retrievable document and an identified contact support transfer, but the practical test is whether the recipient can interpret and apply the knowledge appropriately.

The evidence package should support reconstruction of a consequential claim. A future reviewer of the closure account needs the cut-off, received-work records, payment status, remaining estimates and links to the relevant acceptance decisions. A future benefit reviewer needs the original population, baseline, definitions, exclusions, source lineage and explanation of revisions. Preserving a final presentation without those relationships leaves an attractive account that cannot be checked. Lena can ask the receiving controller to reproduce one selected reconciliation or outcome measure from the retained sources, using the access and tools that will remain available. This is a bounded test of usability, not proof that every record is complete. The result should identify missing access, unexplained adjustments or dependence on a departing specialist while those gaps can still be addressed. Later archive migration then preserves an already usable evidence trail rather than postponing its creation until after closure.

Learning should also reach the activity it is intended to improve. Chapter 14's scheduling example exposed a resource conflict that changed a fourteen-day dependency calculation into an eighteen-day feasible sequence. A useful retained lesson describes the shared specialist, the conflicting work and the method used to test availability. Its recipient might be the owner of the estimating or planning procedure, who agrees how the next commitment review will examine comparable constraints. Merely recording that resources should be considered adds little. Equally, forcing the exact original sequence onto unrelated work would ignore context. Lena should record the agreed adaptation, its owner and the evidence expected from subsequent use. The lesson can remain open for verification after the originating programme closes, provided responsibility has transferred. Organisational learning becomes visible when the receiving practice changes and its consequences are examined, rather than when the last lesson is uploaded.

Authorise the ending and retire controls with a clear purpose

The final review should make the proposed decision concrete. Samira receives an account of the programme's authorised contribution, achievement to date, remaining obligations, financial position and accepted receiving arrangements. Alex confirms the delivery and technical evidence; Maya confirms operating and benefit responsibilities; finance and records owners confirm their respective commitments. Lena reconciles the statements and identifies where evidence remains insufficient. If a necessary condition is unresolved, the record should state the work and governance retained to resolve it, rather than dispersing the team and treating a conditional signature as completed transfer. An approval can specify an effective closure date after named conditions are verified. Where all necessary arrangements are supported, Samira can authorise the ending and communicate its meaning to affected parties. The record then identifies which authority will receive subsequent exceptions, so that later accountability does not depend on recreating the programme's last meeting.

Retiring a control is itself a change to how the organisation sees and responds to its work. A programme dashboard may no longer be needed once the receiving service has incorporated the relevant measures into its routine review, but removing it before that happens can create an observation gap. Lena should identify each continuing reporting purpose and the receiving process before ending the temporary report. Access changes require similar care: departing users and temporary permissions can be retired while authorised custodians retain the ability to retrieve, interpret and protect necessary records. Retention and contractual arrangements should follow the organisation's applicable requirements rather than an invented universal duration. Supplier contacts, support rights and remaining obligations need a usable home. The aim is to remove unnecessary temporary administration while keeping the information, resources and authority required for ongoing decisions available to the people who now own them.

The teaching proposal can be tested through three linked questions. Is closure supported because every obligation has a named owner? The benefits-review row shows why it is not: the required resources remain unconfirmed. Is twelve the remaining cash requirement? It is only the future resource work; the assumed unpaid received work makes future cash thirty-two. Does a benefit review on day ninety establish that value has been realised? It establishes a planned opportunity to examine evidence, with the conclusion still dependent on what is observed. These answers expose different gaps in an apparently tidy ending. Maya must secure a workable review arrangement, the finance lead must retain both cost and settlement responsibilities, and the receiving governance must be able to act on later findings. Resolving those gaps makes a closure recommendation stronger without pretending that every uncertainty or future obligation has disappeared.

Leave the continuing organisation able to govern value

A credible Closure and Continuing Obligations Record preserves both the ending and the unfinished questions. It states whether the programme completed its authorised contribution or was stopped, distinguishes observed achievement from expectation and identifies the accepted arrangements for what follows. It links to original commitments and subsequent decisions so that an altered objective cannot silently replace the history. It also records the scope and limits of the closure judgement: which responsibilities have transferred, which residual estimates remain and when further evidence will be reviewed. The receiving organisation should be able to find this account without navigating the programme's former personal folders or relying on informal recollection. Its usefulness rests on continuity of meaning and action. A shorter, maintained record with accessible sources can serve that purpose better than a comprehensive final report whose assumptions, owners and supporting evidence are no longer recoverable.

The end of temporary governance does not remove the organisation's ability to change its mind. If later evidence reveals a deteriorating service, an unsupported benefit claim or an emerging opportunity, the responsible business authority can commission corrective work, revise the operating arrangement or authorise another bounded initiative. That response should use the current evidence while preserving what was known at closure. It need not imply that the previous ending was improper, because circumstances and knowledge can change. Conversely, evidence that a material known obligation had no capable recipient should prompt examination of the closure decision itself. The control discipline is to retain a route from observation to accountable action across organisational boundaries. The programme's existence is temporary; the organisation still needs to understand the value, cost and exposure of the service it now operates and the commitments it makes next.

Chapter 18 brings the handbook back to the purpose of transformational control. Schedules, cost accounts, benefit models and assurance reports matter because people use them to make and honour consequential commitments. Throughout the lifecycle, the controller connects those commitments to evidence, preserves the meaning of uncertainty and makes the next decision understandable to those responsible for it. At closure, that contribution is expressed through an organisation able to continue the work without losing its history or its obligations. The achievement is an ending whose conditions can be explained and whose continuing responsibilities can be performed. Value will still require observation, judgement and, sometimes, correction. By leaving those capabilities with the receiving organisation, project controls supports the transformation beyond the life of the temporary programme and gives future decision makers a usable foundation for the work that follows.

References

- Australian National Audit Office. (2017, 22 November). Costs and Benefits of the Reinventing the ATO program. Report No. 15 2017–18, paras. 4.20–4.26. www.anao.gov.au/sites/default/files/ANAO_Report_2017-2018_15.pdf
- Government Project Delivery. (2025). Government Functional Standard GovS 002: Project delivery. Version 2.1, clauses 6.4.8–6.4.9. projectdelivery.gov.uk/library-products/government-functional-standard-govs-002-project-delivery/

- Newell, S., Bresnen, M., Edelman, L., Scarbrough, H. and Swan, J. (2006). Sharing knowledge across projects: Limits to ICT-led project review practices. *Management Learning*, 37(2), 167–185. doi: [10.1177/1350507606063441](https://doi.org/10.1177/1350507606063441)
- Nicolaou, A. I. (2004). Quality of postimplementation review for enterprise resource planning systems. *International Journal of Accounting Information Systems*, 5(1), 25–49. doi: [10.1016/j.accinf.2004.02.002](https://doi.org/10.1016/j.accinf.2004.02.002)
- Office of the Auditor-General. (2017). Canterbury Earthquake Recovery Authority: Assessing its effectiveness and efficiency. B.29[17a], paras. 1.23, 2.40 and 2.45. oag.parliament.nz/2017/cera/docs/cera.pdf
- Whyte, J. and Nussbaum, T. (2020). Transition and temporalities: Spanning temporal boundaries as projects end and operations begin. *Project Management Journal*, 51(5), 505–521. doi: [10.1177/8756972820919002](https://doi.org/10.1177/8756972820919002)
- Williams, T., Vo, H., Bourne, M., Bourne, P., Kirkham, R., Masterton, G., Quattrone, P. and Toczycka, C. (2025). Benefits realisation: Case studies in public major project delivery with recommendations for practice. *Production Planning & Control*, 36(1), 127–147. doi: [10.1080/09537287.2023.2256287](https://doi.org/10.1080/09537287.2023.2256287)
- Zhang, X., Denicol, J., Chan, P. W. and Le, Y. (2024). Designing the transition to operations in large inter-organizational projects: Strategy, structure, process, and people. *Journal of Operations Management*, 70(1), 107–136. doi: [10.1002/joom.1275](https://doi.org/10.1002/joom.1275)

About the author

Professor Amgad Badewi is Chief Executive Officer of Kent Business College and a researcher and educator in project controls, project management and benefits realisation. He previously held a Readership at Kent Business School, University of Kent. His work examines how organisations connect project delivery with the operational changes and benefits that justify investment.

He completed his PhD at Cranfield University in 2016, investigating benefits realisation in enterprise resource planning systems. He also holds a Postgraduate Certificate in Higher Education from the University of Kent and is a Fellow of Advance HE. His research addresses project benefits governance, organisational capability, enterprise systems and transformation, with publications in the International Journal of Project Management and Business Process Management Journal.

At Kent Business College, he combines educational leadership with teaching on the Project Controls Professional Level 6 and Associate Project Manager Level 4 programmes. His professional service has included finance and professional development leadership roles at PMI UK. Across research and teaching, he focuses on the relationship between successful delivery, organisational readiness and lasting value.

Professor Amgad Badewi

Kent Business College
School of Project Control

How to cite this book

Recommended reference

Badewi, A. (2026). *Project Controls in Transformational Projects*. Kent Business College.

Citing a particular passage

Use the author and year together with the printed page number for a quotation or a specific argument. Identify the relevant chapter when this helps the reader locate the discussion. The chapter references acknowledge the original academic papers, official reports and professional guidance supporting the text.

First edition · 2026

