FRAMEWORK | ABRIDGED WORKING PAPER

Galapagos Transformation Risk Model

Why transformation risk is a relationship problem.

Brian Harkin | Version 1.1 | Working paper | 2026

WHAT THIS PAPER IS

Six external cases. 238 evidence observations. Seventeen failure mechanisms frozen before scoring. Two success cases chosen because they should have broken the model. The abridged version of the research paper, written for people running live transformations.

Capacity absorbs load. Nothing absorbs coupling. Nothing survives a defensive culture.

Galapagos / DigitalXform | Galapagos is a registered trade mark in the UK

EXECUTIVE SUMMARY

The argument in one page

Transformational risk is a relationship problem, not a size problem.

Most of what the industry says about why transformations fail is aimed at the wrong variable. Sponsorship, culture, scope, governance and pace each explain something. None of them explains the outcome.

Transformation risk is governed less by the absolute size of change than by the dynamic relationship between Change Load, Coupling, Execution Capacity and Defensive Culture. Successful organisations can execute very broad and very deep change when they sequence the load, constrain coupling, build capacity and shorten evidence-to-action loops. Failure becomes more likely when the opposite happens, particularly because deteriorating execution can itself consume capacity and make the system progressively less recoverable.

The evidence behind that sentence is six external transformations: the Care Quality Commission, Birmingham City Council's Oracle programme, the Emergency Services Network, Target Canada, the Netflix cloud migration and Capital One's cloud transformation. Four failed or ran into serious trouble. Netflix and Capital One were chosen because they should have broken the model, and the conditions that would count as breaking it were written down before scoring began.

6	238	17	2
External cases	Evidence observations	Failure mechanisms	Success cases chosen to break the model

Five propositions to remember

- → Large transformations do not fail simply because they are large. Risk rises when Change Load and Coupling outrun Execution Capacity.
- → Parallelism is never free. Every additional concurrent change increases coordination demand and the surface area across which unforeseen dependencies emerge.
- → Accelerate learning, not organisational overload. Fast evidence loops can coexist with a deliberately paced transformation.
- → Transformation can consume the very Execution Capacity required to complete it. Deterioration feeds itself.
- → A defensive culture multiplies everything. It hides coupling, slows the response and drains capacity, so it sits outside the brackets.

RESEARCH STATUS

Evidence-informed and prospectively testable. Not a statistically validated predictor, and not an estimate of the probability that any transformation will fail. The six cases are enough to reject the simple explanations and sharpen the model. They are not enough to calibrate it, and this paper does not pretend otherwise.

THE PROPOSITION

The Amplification Hypothesis

$$\text{Transformational Risk} \propto \left[\frac{\text{Change Load}}{\text{Execution Capacity}} \times \text{Coupling} \right] \times \text{Defensive Culture}$$

∝ *means "is proportional to". Risk is not a fixed score. It moves as these terms move against one another.*

Everything the industry blames is in it. Size sits inside Change Load. Sponsorship and skills sit inside Execution Capacity. Culture is outside the brackets, multiplying everything inside. Assessing them in isolation misses the impact of their interaction, and the interaction is what does the damage. Load and coupling multiply: the same load, tightly coupled, turns every slip into several. And when capacity falls, the same load and coupling produce more risk than they did before, with nothing new added. That is why Target Canada and CQC failed with visible sponsorship in place.

The relationship is conceptual, not a calibrated equation. Load over capacity is the absorbed load. Coupling sits outside that fraction and multiplies whatever is left, which is why no amount of capacity removes it. Defensive Culture then multiplies the whole of that. Across the six cases it scored 9, 9 and 10 in CQC, Birmingham and ESN, 6 in Target Canada, and 1 and 2 in Netflix and Capital One. In all three UK failures the mechanism was the same: uncomfortable information could not be surfaced or challenged. That hides coupling, lengthens the response delay and drains capacity, which is why it sits outside the brackets rather than inside Execution Capacity, where the register records it as H1. The expression does not produce a probability.

What follows from it

- → Breadth and depth are exposure conditions, not verdicts. Netflix and Capital One carried both at the extreme and succeeded.
- → Execution Capacity is relative to the demand placed on it. A capable organisation can be overloaded. A constrained one can recover by sequencing demand and building capacity.
- → Trajectory matters as much as the snapshot. Rising load and coupling are far more dangerous when capacity is falling and response is slowing.
- → A defensive culture is not necessary for failure. Target Canada scored 6 and failed on pace and concurrency.

> **CAPACITY ABSORBS LOAD. NOTHING ABSORBS COUPLING. NOTHING SURVIVES A DEFENSIVE CULTURE.**
>
> No amount of capacity removes the dependency between two things changing at the same time. Coupling has to be designed down, sequenced out or contained. A defensive culture multiplies the whole relationship: added capacity, careful sequencing and contained coupling are all discounted by the culture they operate in. At its worst that culture becomes a toxic one.

Where the model came from

The model was written before it met the cases. The constructs are in the 2024 book: sequential transformation, the principle that parallelism is never free, and the human, organisational and delivery factors that recur in every failed programme. The seventeen hypotheses were frozen before any case was scored, so contradictory evidence counts as a result rather than being explained away.

THE TERMS

What the relationship is made of

VARIABLE	COMPONENTS	WHAT IT MEANS
Change Load	Breadth. Depth. Concurrency. Pace.	The amount and speed of material change demanded of the organisation. High load is not unhealthy on its own. The question is whether it is proportionate to coupling and capacity.
Coupling	Technical, data, process, resource, decision, organisational, supplier and regulatory dependencies. Known, knowable but unmanaged, or emergent.	The degree to which the success, failure or timing of one changing element affects the others. Coupling is interdependency between the things changing. It is not a description of how change spreads.
Execution Capacity	Human. Organisational. Delivery.	The available ability to absorb, coordinate, learn from and execute the change. Capacity is dynamic. It can be built, protected, depleted or redirected while the programme runs.
Defensive Culture	H1 in the register. The degree to which the culture inhibits challenge, experimentation and the open surfacing of problems.	Acts on all three terms at once: it hides coupling, lengthens the response delay and drains capacity. It multiplies the whole relationship, which is why it is drawn outside the brackets.

Why the relationship is dynamic

Execution Capacity is not a fixed asset. In every troubled case it was consumed while the programme ran. Execution problems create rework. Rework creates remediation and workarounds. Workarounds create attrition and governance burden. Each step diverts capacity from delivery into recovery, so the denominator shrinks while the load stays where it was. The programme starts eating the capacity it needs to finish.

Coupling moves too. Some dependencies are known at the outset. Some are knowable and left unmanaged. Some only appear when the change meets the real organisation. The model does not assume you will map them all. It asks how fast you expose and contain the ones you missed.

Culture moves last, and under pressure it moves in the wrong direction. As a programme deteriorates, defensiveness rises, and the multiplier grows at the exact time when the transformation can least afford it.

The evidence loop

Two measures make the loop visible.

- → How long until you know. The time from a material assumption to credible evidence about whether it was right.
- → Response delay. The number of days between credible evidence arriving and the programme materially changing what it does. An acknowledgement is not a response. A meeting is not a response. Only a change in scope, sequence, resourcing, design or action counts.

In the research register these are D4 Validation Lag and D5 Strategic Response Lag.

WHY RESPONSE DELAY MATTERS

Fast learning with slow response is still dangerous. Unresolved evidence accumulates, and a programme can know exactly what is wrong for months without changing anything. The dashboard turns red after the delay has already run.

THE SIX CASES

Same size, opposite outcomes

The comparison is useful because it highlights differences. Forcing every case into the same pattern would prove nothing.

Care Quality Commission

TROUBLED / FAILED

Very high load and coupling met weak and deteriorating human, organisational and delivery capacity. Rework, technology friction, loss of expertise and delayed response consumed what remained. The Dash review (2024) set out the result.

Birmingham City Council Oracle

FAILED; RECOVERY OPEN

A broad ERP and operating-model change went live with testing, readiness, skills and assurance gaps that were known beforehand. Recovery improved when validation strengthened and decisions started following the evidence.

Emergency Services Network

TROUBLED; RECOVERY IN FLIGHT

National mission-critical technology, supplier, user and operating-model dependencies created extreme coupling. Warned by the NAO in 2016 and 2023 and by the Public Accounts Committee in 2019. Slow strategic response amplified the risk.

Target Canada

COMPLETED FAILURE

Clear strategy. Visible sponsorship. 124 stores opened in a year while supply chain, systems, data, workforce and operating routines were still being built. Target's own verdict: too much, too fast.

Netflix cloud migration

SUCCESSFUL

Extreme breadth and depth. Seven years to complete. Sensitive workloads sequenced, service coupling engineered for containment, and real-scale testing shortening the evidence loop throughout.

Capital One cloud transformation

SUCCESSFUL OVERALL; MATERIAL CONTROL FAILURE

Eight years of staged migration, skills investment and reusable patterns. Regulators identified serious risk-assessment and response weaknesses in one critical control domain (OCC, 2020).

What the six cases show

- → Breadth and depth multiply risk. They do not cause failure on their own.
- → Concurrency hurts because of coupling, not because parallel work is wrong in itself.
- → Coupling can be engineered downward. Netflix and Capital One both did it deliberately.
- → Execution Capacity is dynamic, and it was consumed in every troubled case.
- → Learning speed and programme duration are different variables. Netflix took seven years and learned in days.
- → A serious local failure need not sink the whole transformation. The unit of analysis matters.
- → Governance volume and governance effectiveness are different things.
- → Not every cause belongs inside the model. Target's pricing problem stays visible as Not Explained by Current Model.
- → A defensive culture was present in all three UK failures and absent in both successes. In each failure the result was the same: bad news could not be delivered.

THE COMPARISON

Where the cases differ

CASE	OUTCOME	CHANGE LOAD	COUPLING	EXECUTION CAPACITY	CULTURE	EVIDENCE AND RESPONSE
CQC	Troubled / failed	Very High	Very High, poorly coordinated	Weak, and deteriorating	9	Warnings and feedback did not change decisions fast enough.
Birmingham	Failed; recovery open	Very High	Very High	Weak at original go-live	9	Late validation and weak response. Recovery shortened the loop.
ESN	Troubled; recovery open	Extreme	Extreme	Weak historically	10	Repeated warnings, resets and supplier dependencies.
Target Canada	Failed; exited	Extreme	Very High	Constrained for the launch pace	6	Learning arrived after a large live footprint had amplified the problems.
Netflix	Successful	Very High	High, actively contained	Strong and growing	1	Rapid real-scale validation and response over a long horizon.
Capital One	Successful overall	Very High	High, actively simplified	Strong, uneven by domain	2	Fast application learning. Slow risk and control remediation in one domain.

The pattern is not that failures score high everywhere. It is that the two successes carry the same extreme breadth and depth as the failures without the rest of the failure pattern. Netflix and Capital One both score ten out of ten on Scope Breadth and Scope Depth.

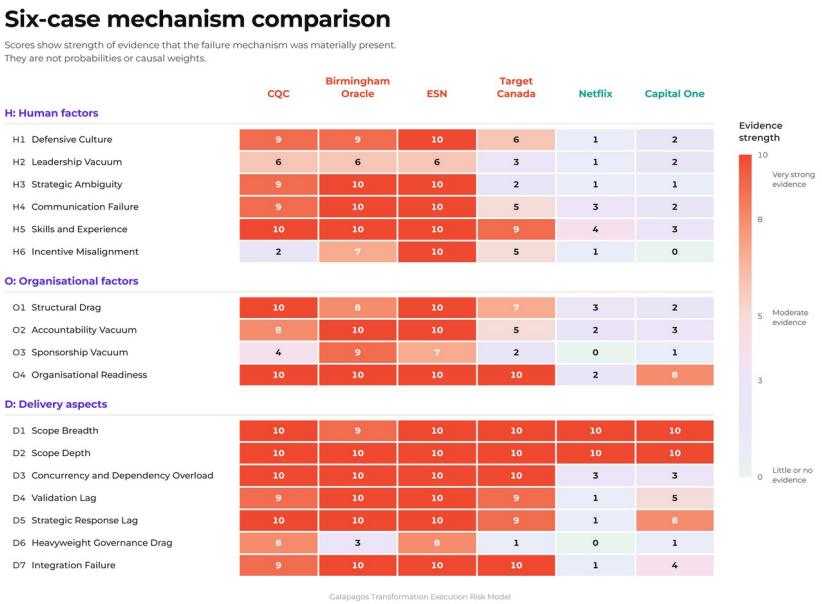


Figure 2. Evidence-strength scores for the seventeen mechanisms across the six cases. Scores show the strength of evidence that a mechanism was materially present, not the probability of failure.

THE DISCRIMINATING PATTERN

Breadth and depth score very high in both successes. What separates the cases is Defensive Culture, Concurrency and Dependency Overload, Organisational Readiness, the two evidence-loop measures, Integration Failure and the sufficiency of Execution Capacity.

CLAIM BOUNDARIES

What the model does not claim

Credibility depends as much on what the model does not claim as on what it does.

USE

An evidence-informed, prospectively testable model of transformation execution risk. It identifies the conditions and mechanisms that raise or lower execution risk before and during a transformation. It can be proved wrong, and this paper says how.

DO NOT USE

A prediction of the probability that a transformation will fail. A statistically validated result across transformations generally. A calibrated numerical risk score. A maturity rating. Evidence that Galapagos has already been validated.

LIMITATION	WHAT IT MEANS FOR THE READER
Small, purposive sample	Six cases are enough to sharpen and challenge the theory. They cannot establish population-level validity or universal effect sizes.
Selection bias remains	Four cases were chosen because they failed. The two pre-registered successes improve discrimination but do not remove the effect.
Source visibility is uneven	Public-sector failures generate forensic official evidence. Private-sector successes are mostly described by the companies themselves.
Scores are evidence strength	The 0 to 10 scale measures how strongly the evidence shows a mechanism was present. It is not a likelihood, a weight or a severity.
The equation is conceptual	It expresses interaction. It is not calibrated and must not be used to calculate a numerical probability of failure.
Capacity must not become circular	Capacity is evidenced independently through skills, readiness, coordination, learning and delivery conditions. Failure is not proof that capacity was insufficient.
Culture must not become circular	A defensive culture is evidenced through contemporaneous findings, such as inspection reports, reviews and staff surveys. It is not inferred from the failure.
The model is not universal	Market fit, macroeconomic conditions, regulatory shocks and technology obsolescence can decide outcomes. They stay visible as Not Explained by Current Model rather than being forced into a mechanism.

The distinction is substantive. The model exists to sharpen executive judgement and intervention, not to replace it with false precision. Its near-term test is prospective: whether deterioration in the relationship between Change Load, Coupling, Execution Capacity and Defensive Culture can be seen before conventional measures make failure obvious.

THE INVITATION

Prove it wrong

A model that cannot be proved wrong is not worth your time. This one can. Two kinds of case would do it.

- → A large transformation that failed with low coupling, strong Execution Capacity, an open culture and a short response delay.
- → A tightly coupled, overloaded programme with weak capacity that succeeded inside a defensive culture, without sequencing, containment or a fast evidence loop.

Send me the case. I will score it against the frozen model, using the same evidence-first protocol applied to the six above, and publish the result whichever way it falls.

What happens next

The six cases explain failure after the fact. The next test is application before and during delivery: establish the relationship at the start, write down what deterioration would look like, and track whether Load and Coupling are becoming more or less proportionate to Execution Capacity, and whether the culture around them is becoming more or less defensive, while the programme runs. That is the evidence the current cases cannot provide, and it is the evidence this work now needs.

Provenance

Harkin, B. (2024). Evolving from Digital Transformation to Digital Acceleration Using The Galapagos Framework. CRC Press / Taylor & Francis.

Galapagos Transformation Failure Model v1.0. Frozen research standard, August 2026.

Galapagos Baseline & Trajectory Assessment Protocol v1.0. Working research protocol, August 2026.

Galapagos Research Workbook v1.0. 238 evidence observations across Cases 001 to 006, with observation-level provenance, stance and confidence.

Selected sources

Dash, P. (2024). Review into the operational effectiveness of the Care Quality Commission. | Grant Thornton (2025). Public Interest Report, Birmingham City Council Oracle implementation. | National Audit Office (2016, 2023). Emergency Services Network reports. | Public Accounts Committee (2019). Emergency Services Network: further progress review. | Target Corporation (2013, 2014). Form 10-K filings; Alvarez & Marsal (2015). Monitor pre-filing report, Target Canada Co. | Netflix (2016). Completing the Netflix Cloud Migration; Netflix TechBlog (2010 to 2015). | Capital One (2017, 2018). Engineering and transformation publications. | Office of the Comptroller of the Currency and Federal Reserve Board (2020). Enforcement actions against Capital One.

THE FULL PAPER

The full research paper, the seventeen-mechanism register and the assessment protocol are available on request. Contact Brian Harkin via LinkedIn or digitalxform.co.uk.

This abridged working paper condenses the Galapagos Transformation Risk Model research paper v1.1. Galapagos is a registered trade mark in the UK. © Brian Harkin 2026.