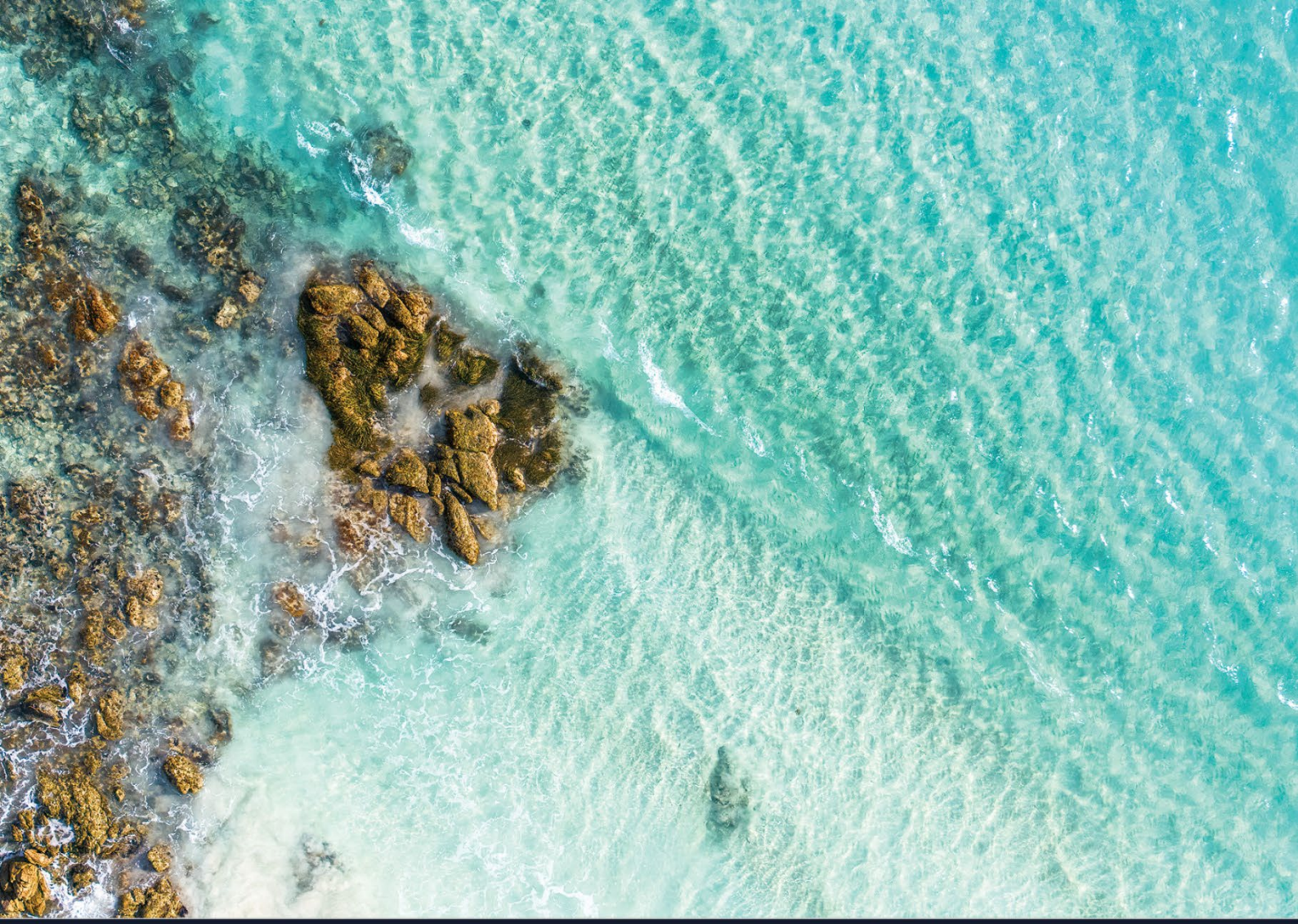


A diagnostic framework for robust and micro-calibrated policy design

Connecting the robustness of policy to the design choices that produce it

Professor Giliberto Capano, University of Bologna

ANZSOG Academic Fellow

July 2026

ANZ SOG Australia & New Zealand School Of Government

For further information on the ANZSOG Insights series and of this report: insights@anzsog.edu.au

This project is supported by an Academic Fellowship with the Australian and New Zealand School of Government (ANZSOG) to consider designing robust and micro-calibrated policy instruments.

Suggested citation

Capano, Giliberto. A diagnostic framework for robust and micro-calibrated policy design: Connecting the robustness of policy to the design choices that produce it. Melbourne: ANZSOG, July 2026.

Contents

Summary

This paper presents a diagnostic framework for assessing the robustness of a policy from the design of its instruments. Robustness is the capacity of a policy to keep delivering its objectives across unpredictable conditions. It is not a property a policy holds in the abstract, but one produced or undermined by concrete choices made when policy makers design policy instruments. The diagnostic framework has been applied to two Australian reforms, the Job-Ready Graduates Package and the Social Media Minimum Age Act, in a companion ANZSOG Insights paper, *Designing Robust and Micro-Calibrated Policy* (Capano, 2026). That paper presents the case evidence and derives five practical design principles from the analysis. This paper focuses on the framework itself and illustrates its use through two hypothetical examples (see section 5). Put simply, the case evidence is there; the diagnostic method is here.

The framework connects two bodies of work usually kept apart: the analysis of policy robustness and the micro-design perspective on policy instruments, and turns their connection into a usable diagnostic framework. In practical terms, the framework moves through four stages: mapping policy instruments, scoring their design details, reading a robustness profile, and finally identifying the appropriate response. The result is not a single robustness score but a structured profile that distinguishes the fragilities a design can fix from those that depend on institutions beyond the design. For a visual summary of the framework, see section 2.

When to use this framework

The framework is intended for use in policy planning, when design choices can still be changed. It can also be used to review an existing policy. Users who want to go straight to the procedure can begin at Section 3. The worksheets are in the Appendix.

Use this framework to:

- test a proposed policy before enactment
- support an internal design review or challenge process
- examine whether an existing policy can adapt to changing conditions
- distinguish problems that can be fixed through design from those requiring institutional action.

1. Designing robust policies for uncertainty

Governments increasingly design policies for conditions they cannot predict. Technologies change faster than the rules governing them, social expectations shift, and the institutions responsible for delivery span multiple levels and agencies. A policy that performs well under the conditions for which it was designed may perform poorly, or become impossible to correct, once those conditions change (Ansell et al., 2024; Capano & Woo, 2017).

The usual response is to call for a more robust policy. Robustness means the capacity of a policy to keep delivering its intended functions and objectives across a range of possible conditions, not only the single set its designers had in mind (Capano & Woo, 2017; Howlett, Capano, & Ramesh, 2018). It differs from resilience, which is about returning to a prior equilibrium after a shock, rather than continuing to reach objectives amid shocks and uncertainty (Capano & Woo, 2017). The difficulty is that robustness is usually discussed too abstractly to act on: treated as a property of the whole policy and judged only after a reform has succeeded or failed. That does not tell a policy team which of the choices in front of them, in the drafting and the configuration of the instruments, will make the difference (Howlett & Ramesh, 2023).

This framework closes that gap with a single claim: robustness is produced, or undermined, at the level of design detail, the concrete configuration of each instrument, its goals and its operational settings. Robustness does not float above the instruments; it is the cumulative result of how they are designed. Three examples:

1. A policy can be adjusted more easily if its instruments are not all locked together.
2. A policy can adjust and correct itself only if it includes performance-monitoring instruments, goals and timelines.
3. A policy can support local decision-making only if its instruments grant implementers a defined measure of discretion (polycentricity).

In each case the robustness property is built into the design.. The practical consequence is direct: to judge whether a policy will be robust, you examine its design detail at the stage when changes are still cheap, rather than waiting to infer robustness from performance.

Two clarifications keep this precise. First, the relationship is not one of perfect determination: design detail strongly shapes some robustness dimensions, partly shapes others, and barely affects a few produced mostly by institutions, resources and political authority. Section 2 sets out which is which. Second, the relationship is most dependable in one direction. Good design detail does not guarantee robustness, since factors beyond the design can still undermine it, but poor design detail reliably undermines it, because no institutional capacity can rescue instruments that are locked together and blind to their own performance. The framework is therefore at its most powerful as an early warning of fragility.

2. The analytical framework

This section provides the framework's vocabulary. Before the framework can be used, the user needs to know three things: the design details to be scored, the robustness dimensions to be read, and how far each robustness dimension can be diagnosed from design details alone.

2.1 The design details of policy instruments

Every policy instrument can be described in terms of two sets of micro-design dimensions: goals (specifications) and means (calibrations).

Together, these determine whether a policy can be held to account for results; a policy needs to state who it targets, what it expects, and by when, so it can be evaluated against these dimensions (Capano and Pritoni 2026, Capano and Howlett 2024). Table 1 below presents ten micro-design dimensions and the diagnostic questions to evaluate goals and means.

Table 1. The ten micro-design dimensions of a policy instrument (Capano and Pritoni, 2026)

Set	Dimension	Diagnostic questions
Target specifications (goals)	Designation of the target population	Does the instrument state explicitly who it is aimed at?
	Expected outcome of the intervention	Does it state what change it is meant to produce?
	Timeframe	Does it have a timeline for delivery?
Calibrations (means)	Primary recipients	Who directly receives the instrument or is bound by it?
	Additional resources	Are new resources dedicated to making it work?
	Organisation in charge of implementation	Is a body named to carry it out?
	Intrusiveness	How strongly does it constrain the behaviour of those it targets? (scored descriptively)
	Discretion of implementers	How much latitude do implementing actors have in applying it? (scored descriptively)
	Accountability rules	Are there mechanisms that attach consequences to performance?
	Monitoring and auditing	Are there mechanisms that generate evidence about whether it is working?

Specifications and calibrations do different work for robustness. Specifications fix what the policy is trying to do: without a stated outcome and timeframe, a policy cannot measure itself and so cannot adapt. Calibrations define how it behaves: whether it can be adjusted, monitored, and enforced. A robust policy needs both, which is why the framework scores all ten dimensions.

2.2 The dimensions of robustness

Robustness is described here through a set of dimensions drawn from the robust governance literature (Ansell et al., 2024; Howlett & Ramesh, 2023). These dimensions are grouped into three families by function, plus one foundation. The purpose of the grouping is practical: it helps the user see what kind of robustness each dimension is testing.

Variety of means concerns whether the design has more than one way to act. This family includes diversity, redundancy and bricolage, which move from having variety, to having backup options, to being able to recombine existing instruments and settings.

Distributed authority concerns where decision-making power sits. This family includes polycentricity, bounded autonomy and coordinated autonomy, which ask whether authority is placed where knowledge and responsiveness sit, without losing overall coherence.

Adaptive capacity concerns whether the design can change after enactment. This family includes modularity, prototyping, dynamism, scalability and strategic polyvalence. These share a logic of time: a robust design is not finished at enactment but keeps adjusting (Howlett & Ramesh, 2023).

Political-institutional capacity stands apart as a foundation rather than a family. It is the ability of the responsible institutions to steer, coordinate and enforce. It is considered foundational because it enables the operation of all other dimensions. A design can be modular, monitored and polycentric on paper, but if the institutions lack the authority, resources and expertise to act, none of those properties will function in practice (Wu, Ramesh, & Howlett, 2015).

2.3 How far the design details reach

Design detail produces robustness, but as Section 1 carefully notes this is uneven. This section shows how the framework is used: for each robustness dimension, we ask how much of it can be diagnosed from the design details alone?

The answer is one of three:

- Some dimensions are **design-determined**, meaning the design detail settles them almost completely.
- Some are **design-shaped**, meaning the design detail provides a reliable part of the answer, which must then be checked against facts about institutions.
- Some are **institution-dependent**, meaning the design detail tells you little, and the real assessment must be made elsewhere.

Table 2 applies this three-way test to each of the ten dimensions introduced Table 1.. It helps users decide which robustness issues can be diagnosed from the policy design itself, and which ones require a separate judgment about institutional capacity.

Table 2. Robustness dimensions, their families, and how far the design detail reaches

Family	Dimension	What it means in practice	Reach of design detail	Settings to read
Variety of means	Diversity	The presence of varied instruments, actors and resources.	Design-shaped	Instrument mix; organisations
	Redundancy	The presence of backup options that keep the policy working when one pathway fails.	Design-shaped	Instrument mix; organisation in charge
	Bricolage	The capacity to meet a new problem by recombining instruments the system has, rather than building new ones.	Design-determined	Instrument mix; intrusiveness; discretion
Distributed authority	Polycentricity	Coordination across multiple centres of authority and levels of government.	Institution-dependent	Organisation in charge
	Bounded autonomy	Local discretion exercised within a shared overall strategy.	Design-determined	Discretion; organisation in charge
	Coordinated autonomy	The working balance between central steering and decentralised flexibility.	Institution-dependent	Discretion; organisation in charge
Adaptive capacity	Modularity	The capacity to change one part of the policy without being forced to change all the others.	Design-determined	Interdependence of calibrations
	Prototyping	Iterative experimentation and feedback, through pilots or staged rollout, before the design is fixed.	Institution-dependent	Timeframe (staged / sunset)
	Dynamism	Continuous monitoring and self-correction once the policy is operating.	Design-shaped	Monitoring; accountability; timeframe
	Scalability	The capacity to expand or contract quickly as demand changes.	Design-shaped	Additional resources; timeframe
	Strategic polyvalence	The capacity to adjust the policy, including changing its instruments over time, as the context shifts	Design-shaped	Adjustable calibrations; review / sunset
Foundation	Political-institutional capacity	The ability of the responsible institutions to steer, coordinate and enforce.	Institution-dependent	Additional resources; organisation in charge

Design-determined dimensions can be diagnosed almost entirely from the design detail. Modularity is created or destroyed by whether the calibrations are locked together. Bounded autonomy lives in the discretion-of-implementers and organisation-in-charge settings. Bricolage is produced by the instrument mix and by whether the settings are loose enough to be recombined. For these three, if the relevant settings are well specified, the dimension is robust, and if they are poorly specified, it is fragile, and the user can be confident in that reading without looking beyond the design.

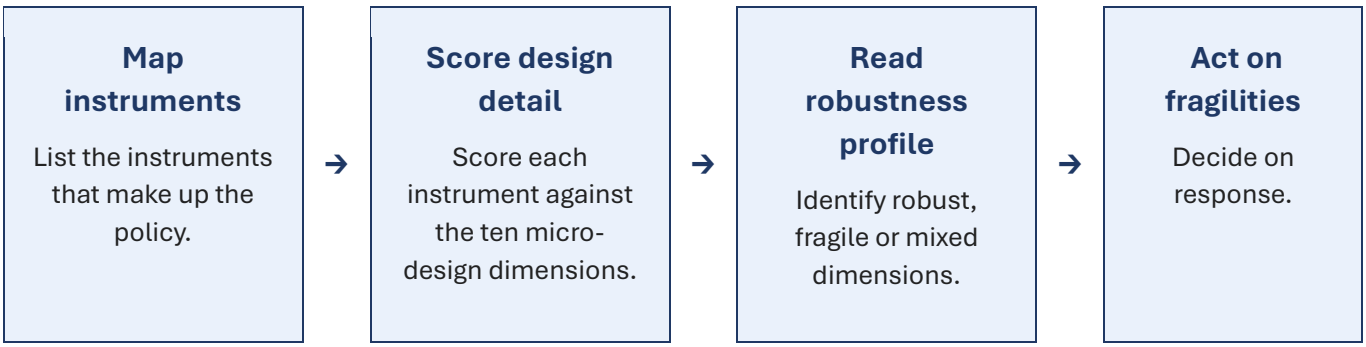
Design-shaped dimensions are reliably indicated by the design detail but not settled by it, so the framework produces a provisional reading that the user must confirm against institutional facts. Dynamism is shaped by the monitoring, accountability, and timeframe settings, and is then confirmed by whether the institution actually acts on what it monitors. Scalability is shaped by the resources and timeframe settings, then confirmed against real funding and administrative capacity. Strategic polyvalence is shaped by the presence of adjustable calibrations, review clauses, and rulemaking powers, and is then confirmed by whether the system has any orientation toward anticipating change. Redundancy and diversity are shaped by the instrument mix and the organisations involved, then confirmed against whether the apparent variety and backups are real.

Institution-dependent dimensions are produced mostly above the instrument level, so the design details are a weak signal at best, and the user is directed to assess them primarily through information outside the design. Political-institutional capacity is institution-dependent despite its foundational importance: traces appear in resource and organisational settings, but real capacity is money, expertise, mandate, and political backing. Polycentricity and coordinated autonomy are institution-dependent: the organisation setting shows how many bodies are named, but whether they coordinate well is a fact about institutional relationships. Prototyping through a staged or sunset timeframe leaves a faint trace, but the decision to pilot is a process choice made before the instrument is finalised.

This distinction is central to the framework. It prevents the user from treating the design as if it can answer every robustness question. The value of the framework is that it separates problems, so the user can see which weaknesses can be fixed by redrafting and which require institutional action.

Figure 1 Summarises the main dimensions of the proposed framework

Figure 1: Summary of framework



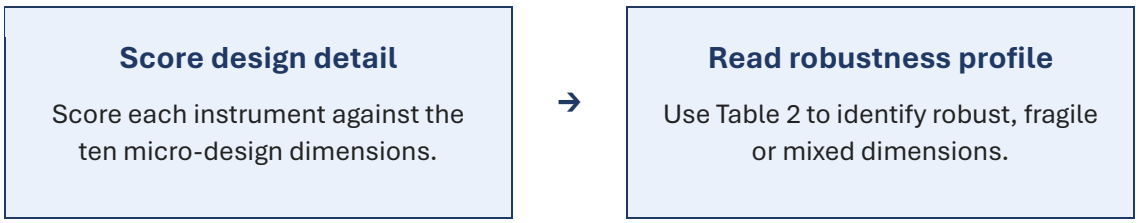
Output	Value for practitioners
A structured robustness profile, not a single score.	Shows which weaknesses can be fixed by design and which require institutional action.

3. The framework in use

Section 2 set out what the framework looks at. This section explains the diagnostic core of the framework: scoring the design detail and reading the robustness profile. The framework works in two steps, and the order of the steps matters.

It would be easy to assess robustness directly by forming an overall impression of how adaptable a policy appears to be. The framework deliberately refuses that shortcut. It requires the user to score the design detail first, and only then to read robustness from it, so that the diagnosis passes through the design choices that produce robustness rather than resting on a general impression. This discipline is what separates the framework from the vague, after-the-fact judgment it is meant to replace.

Figure 2: From classification to response



If the weakness is...	Treat the reading as...	Likely response
Design-determined	Strong diagnosis from the design	Revise design settings
Design-shaped	Useful but provisional	Verify facts, then revise or strengthen implementation
Institution-dependent	Weak design signal only	Strengthen capacity, authority, resources or coordination

3.1 Score the design detail

The user goes through the policy one instrument at a time and assigns each of its ten design dimensions a score from 1 to 5. The ten dimensions fall into two groups, each scored differently:

1. Eight dimensions are scored on **clarity.**

2. Two dimensions are scored on **level.**

The two groups use two separate scales, set out below.

The first group covers eight of the ten dimensions: the three target specifications (target population, expected outcome, timeframe) and five of the seven calibrations (primary recipients, additional resources, organisation in charge, accountability rules, monitoring and auditing). For each of these, the question is simple: is the setting there, and how clearly is it pinned down? A high score means the setting is present and fully defined; a low score means it is vague or missing. For this group, a higher score always means a stronger contribution to robustness. Table 3 gives the wording for each point on the scale.

Table 3. Scoring scale for the eight clarity dimensions

Score	What it means
5	**Fully set out.** The setting is stated explicitly and completely, with no real ambiguity about who, what or when.
4	**Mostly set out.** The setting is present and largely clear, with only minor gaps that do not threaten how it works.
3	**Partly set out.** The setting is present but has significant gaps, or leaves key content to be decided later.
2	**Barely set out.** The setting is mentioned but largely undefined, so its real content will be decided somewhere else.
1	**Absent**. The setting is not addressed at all.

The second group covers the two remaining calibrations: intrusiveness and implementers' discretion. These are not scored for clarity but for level, because they describe how much of a quality an instrument has rather than whether a setting is present. Intrusiveness is the degree to which an instrument constrains the people it targets. Discretion is the degree of freedom the implementing actors have in applying it. For these two, a high score does not mean better. A highly intrusive instrument can give a policy real force or can lock it rigidly in place. Wide discretion can let implementers adapt to local conditions, or leave the policy's content undefined and dependent on whoever happens to fill it. Which way it goes depends on the other settings, above all on the organisation in charge and its capacity. For this reason, the user simply records the level here, and the meaning of that level is worked out at step two, read together with the organisation and capacity scores. Table 4 gives the wording for each point on this scale.

Table 4. Scoring scale for the two level dimensions (intrusiveness and discretion)

Score	What it means (a higher score is not better, only higher)
5	**Very high.** The instrument is maximally intrusive or leaves implementers maximally wide discretion.
4	**High.** The instrument is strongly intrusive, or grants implementers broad discretion.
3	**Moderate.** The instrument is moderately intrusive, or grants implementers some discretion.
2	**Low.** The instrument is only mildly intrusive, or grants implementers little discretion.
1	**Very low.** The instrument is barely intrusive at all or leaves implementers almost no discretion.

3.2 Read the robustness profile

Once the design detail is scored, the user reads each robustness dimension from the relevant settings. Each dimension falls into one of three classes: design-determined (settled almost entirely by the design), design-shaped (a reliable partial answer, to be confirmed against institutional facts), or institution-dependent (the design tells you little; assess mainly through information outside it) — see Section 2.3 for the full reasoning behind each classification.

The scores are not averaged into a single robustness score. Instead, they provide structured evidence for a qualitative judgement about each robustness dimension. The reading procedure differs by class, and each dimension is explained in context.

Making the judgement

- Robust: The relevant settings support the robustness dimension, with no major gaps.
- Fragile: A critical setting is absent, unclear or works against the robustness dimension.
- Mixed: Some settings support the dimension, but important gaps or tensions remain.

For design-shaped and institution-dependent dimensions, record what further evidence is needed before confirming the judgement.

1. Design-determined dimensions are read as a near-complete diagnosis

Modularity is read from the interdependence of the calibrations, and it is the clearest case the framework offers. Modularity is the capacity to change one part of a policy without affecting the others. When the settings of different instruments are locked together so that the value of one depends on the others, the policy is a single block: any change requires changing everything, which raises the cost of correction so high that the policy effectively becomes fixed. Where the settings are separable, parts can be adjusted on their own, and the policy can be corrected step by step. Nothing about institutions enters this judgement, which is why modularity can be settled from the design alone.

Bounded autonomy is read from the discretion-of-implementers and organisation-in-charge settings together. Meaningful discretion, held by a named and capable body within a shared strategy, produces it. Discretion with no designated body to exercise it, or unlimited discretion with no strategy to bound it, does not. Both settings are explicit in the design, so the reading stands on its own.

Bricolage is read from the instrument mix and from whether intrusiveness and discretion leave room to recombine. A varied mix with adjustable settings is high in recombination potential; a single instrument fixed at maximum stringency, with no discretion, offers nothing to recombine. A practical caution applies here: a long list of instruments does not by itself create variety, because instruments that all work the same way provide no genuine alternative. What matters is functional variety, whether the instruments draw on different mechanisms, so that the failure of one mechanism does not disable the whole policy.

2. Design-shaped dimensions are read provisionally, then confirmed against institutional facts

Dynamism is inferred provisionally from the monitoring, accountability, and timeframe settings, which together create the capacity to detect problems and respond to them, and is then confirmed by whether the institution acts on what it monitors. The three target specifications matter here as much as the calibrations: a policy that states its expected outcome and a timeframe has given itself evaluation criteria, to assess whether it is succeeding or needs to adapt in a disciplined way.

Scalability is provisionally read from the additional-resources and timeframe settings, then confirmed against actual funding and administrative machinery.

Strategic polyvalence is inferred from the presence of adjustable calibrations, review clauses, sunset provisions, and rulemaking powers, which constitute the part of future adaptability written into how the instruments are set up, and are then confirmed against whether the system has any orientation toward anticipating change.

Redundancy and **Diversity** are read provisionally from the instrument mix and the organisations involved, then confirmed against whether the apparent backups and variety are real rather than nominal.

3. Institution-dependent dimensions are treated as a weak signal only

Political-institutional capacity is noted from the resources and organisation settings but assessed mainly through what is known about the named institutions such as budgets, staffing, mandate and record (Wu, Ramesh, & Howlett, 2015).

Polycentricity and **Coordinated autonomy** are evident in the organisational settings, but are mainly assessed through the actual coordination relationships among the bodies involved.

Prototyping is noted from any staged or sunset timeframe, but assessed mainly through whether a pilot or trial was in fact conducted before the design was fixed.

The output of step two is not a single number but a structured profile: for each robustness dimension, a reading, a level of confidence that follows from its class, and, where relevant, a note of what must be checked beyond the design. A policy can then be described precisely. It might be strong on the dimensions the design fully controls, modular, bounded in its autonomy, rich in recombination potential, yet fragile on the dimensions that depend on institutions the design cannot supply, genuine capacity, real coordination, real surge capability. That precise description is what the framework exists to produce.

4. Using the framework in practice

Section 3 explained how the diagnosis is produced. This section explains how to use that diagnosis in practice:

1. **Map the instruments.** List the instruments that make up the policy and identify, for each, the ten design dimensions. This is a reconstruction of the design: what each instrument targets and how it is calibrated.
2. **Score the design detail.** Score each dimension of each instrument from one to five, recording the evidence for each score. Use the clarity scale (Table 3) for the eight clarity dimensions, and the level scale (Table 4) for intrusiveness and discretion, where a higher score means only higher, not better.
3. **Read the robustness profile.** For each robustness dimension, apply the rule of its class: take design-determined readings as near-complete, treat design-shaped readings as provisional and note what to confirm, and treat institution-dependent readings as signals to be checked outside the design.
4. **Act on the fragilities.** Identify the dimensions on which the policy is fragile and separate those that the design can fix from those that need institutional action. Where a fragility lies in a design-determined dimension, the design itself can be changed. Where it lies in an institution-dependent dimension, the design cannot fix it, and the question becomes whether the supporting institutions can.

Why not a single score. The framework deliberately does not produce a single robustness score. A single number would collapse the very distinction the framework exists to draw, between robustness that the design controls and robustness that depends on the institutions around it. The structured profile keeps that difference visible, which is what makes the framework useful for deciding what to do rather than merely for grading.

What the framework can and cannot tell you. The framework diagnoses what the design can show. It is strongest where robustness is produced at the level of the instruments and weakest where robustness depends on institutions, resources and political authority that no legislative text can supply. A user who treats a clean design profile as a guarantee of robustness will be misled, because a well-designed instrument operated by an incapable institution will still fail. The three-way classification is the safeguard against that error: it states exactly where the design stops speaking and where judgment about institutions must take over. Read with that discipline, the framework does not overpromise. It identifies the fragilities that can be fixed by redrafting, separates them from those that cannot, and points the user to the appropriate remedy for each.

Where to act first

1. **Fix design-determined fragilities first. These are the clearest and can often be addressed through redrafting.**
2. **Investigate design-shaped fragilities. Confirm whether institutions can manage the risk indicated by the design.**
3. **Escalate institution-dependent fragilities. These may require changes to authority, resources, capability or coordination.**

This order helps teams focus first on weaknesses that are both clearly diagnosed and still relatively easy to change. Design-shaped risks then require verification, while institution-dependent risks may need escalation beyond the policy design team. Used this way, the framework changes the question a policy team asks. Instead of asking, after the fact, whether a reform turned out to be robust, it asks, before enactment, which design choices are making it fragile and whether those choices can still be changed. That shift, from looking back to looking ahead, is the practical purpose of connecting robustness to design detail; to show that robustness is not a quality policymakers can only hope for, but one they can design for, examine, and correct, choice by choice, while there is still time to do so.

5. The framework applied: Two illustrations

The two illustrations below show the framework running from start to finish for each policy instrument. Both examples are hypothetical, constructed to show the method rather than to describe any real reform. The first is a regulatory instrument, the second a service-delivery instrument. They are chosen because regulation and service delivery are the two most common instrument families, and they tend to fail in different ways. Each illustration provides a short description of the instrument, the ten design scores, each with a one-line justification, and the robustness profile derived from those scores. Scores for the eight clarity dimensions use the scale in Table 3; scores for intrusiveness and discretion use the level scale of Table 4, where a higher number indicates **a higher level, not a better one**.

5.1 A regulatory instrument: a “reasonable steps” duty

A statute requires online platforms to take reasonable steps to prevent users below a certain age from using their services, on pain of substantial penalties. The duty is legally binding and carries real force, but the statute does not define what constitutes reasonable steps, state by when the situation should change, or specify what level of compliance counts as success. A single regulator is named to enforce it, but the statute is silent on whether that regulator has the resources or expertise to do so. This is a common modern regulatory pattern: a strong obligation whose actual content is left to be worked out later.

Table 5. Design scores for the regulatory instrument

Design dimension	Score	Justification
Target population	**4**	The population (users below the age threshold) is named, though the boundary is left somewhat open.
Expected outcome	**2**	The statute does not say what level of result counts as success, only that platforms must act.
Timeframe	**1**	No date by which the situation should change is stated anywhere.
Primary recipients	**4**	The platforms bound by the duty are clearly identified.
Additional resources	**2**	No dedicated resources are attached to enforcement in the statute.
Organisation in charge	**3**	One regulator is named, but its remit for this duty is only partly specified.
Intrusiveness	**5**	Level scale: the duty is highly intrusive, imposing a binding obligation with heavy penalties.
Discretion of implementers	**5**	Level scale: discretion is very wide, since reasonable steps is left entirely undefined.
Accountability rules	**3**	Penalties exist, but the standard they attach to is undefined, weakening accountability.
Monitoring and auditing	**2**	No systematic monitoring of platform behaviour or outcomes is provided for.

How to read this profile

On the **design-determined dimensions**, the instrument is mixed. Bounded autonomy is fragile: discretion is very wide (5), but the organisation that should exercise it is only partly specified (3) and operates under no clear strategy, so wide latitude sits with a body that may not be equipped to use it. Bricolage is moderate: a single instrument fixed at maximum intrusiveness offers little to recombine. Modularity is not really tested by a single instrument, but nothing locks this duty to others, so it is not a source of fragility here.

The clearest signals are on the **design-shaped and institution-dependent dimensions**. Dynamism is fragile, because expected outcome (2), timeframe (1) and monitoring (2) are all low, so the policy has given itself almost nothing to measure and no trigger to adjust.

Political-institutional capacity is the decisive institution-dependent dimension, and the profile flags it for assessment outside the design: the whole instrument depends on a regulator whose capacity the statute does not establish. The remaining dimensions are not material to this single-instrument case and are left unscored in the reading.

Diagnostic takeaway

A legally forceful instrument can still be fragile if it does not define success, establish monitoring or secure the capacity required for enforcement.

Priority response: Clarify the expected outcome and compliance standard, establish monitoring and review arrangements, and verify the regulator's capacity.

5.2 A service-delivery instrument: A targeted benefit

A program pays a cash benefit to households below an income threshold to offset the cost of a basic service. The instrument is well specified: it names exactly who is eligible, states the payment amount, names the agency that will deliver it, and sets out clear rules for claiming and for accountability. It looks like a model of clear design. The fragility lies elsewhere. The eligibility threshold, the payment amount and the delivery channel are written into the design so that they are tightly bound together: the payment is calculated as a fixed function of the threshold, and the delivery channel is built around that exact calculation. Changing any one of the three requires re-engineering the other two. This is a well-intentioned, clearly specified program that cannot be adjusted in part.

Table 6. Design scores for the service-delivery instrument

Design dimension	Score	Justification
Target population	**5**	Eligible households are defined precisely by an income threshold.
Expected outcome	**4**	The intended relief (offsetting a service cost) is stated, though not quantified as a target.
Timeframe	**3**	The program runs on an annual cycle, but no end point or review date is set.
Primary recipients	**5**	The households receiving the benefit are clearly identified.
Additional resources	**3**	A budget is attached, but it is fixed and does not adjust to changing demand.
Organisation in charge	**5**	A single delivery agency is clearly named and mandated.
Intrusiveness	**2**	Level scale: the instrument is only mildly intrusive, offering a benefit rather than imposing a duty.
Discretion of implementers	**2**	Level scale: discretion is low, since eligibility and payment are fixed by formula.
Accountability rules	**4**	Clear rules govern claims, payment and audit.
Monitoring and auditing	**4**	Claims and payments are monitored, though outcomes for households are not.

How to read the profile

On clarity, this instrument scores high almost everywhere, which is exactly why it can mislead. The fragility is in a **design-determined dimension** that the high clarity scores conceal: modularity. Because the threshold, the payment amount and the delivery channel are written as a single locked calculation, none can be changed without re-engineering the others, so the scheme cannot be corrected in part. A government that wanted to raise the threshold while keeping the payment, or change the delivery channel without changing eligibility, would have to rebuild the whole instrument.

Scalability is the second fragility, a **design-shaped dimension**: the resources calibration (3) is fixed and the timeframe (3) sets no adjustment point, so the program cannot expand if demand rises. The **institution-dependent dimensions** are not the problem here; the named agency is clear and capable. The remaining dimensions are not material to this case and are left unscored in the reading.

Diagnostic takeaway

A clearly specified and well-administered instrument can still be fragile if its settings are locked together and cannot be adjusted separately.

Priority response: Separate the eligibility, payment and delivery settings and introduce clear review and adjustment mechanisms.

Appendix: The diagnostic worksheet

The worksheet is completed instrument by instrument. Worksheet A records the design-detail scores. Worksheet B translates them into the robustness profile, applying the rule for each dimension's class.

Worksheet A: Design-detail scores (one column per instrument)

Complete one column per instrument. Use the clarity scale (Table 3) for the eight clarity dimensions. Use the level scale (Table 4) for intrusiveness and discretion — a higher score means only higher, not better. Record the evidence for each score in a separate note.

Design dimension	**I-1**	**I-2**	**I-3**	**I-4**	**I-5**	**I-6**
Designation of target population						
Expected outcome of intervention						
Timeframe for achieving the aim						
Primary recipients						
Additional resources						
Organisation in charge						
Intrusiveness (descriptive)						
Discretion of implementers (descriptive)						
Accountability rules						
Monitoring and auditing						

Worksheet B: Robustness profile (read from Worksheet A)

Read each dimension from your Worksheet A scores using the classification in Section 2.3. In the Reading column, record whether the dimension is robust, fragile, or mixed. For design-determined dimensions, the reading stands on its own. For design-shaped dimensions, the reading is provisional, and the final column records the institutional facts that must confirm it. For institution-dependent dimensions, the reading is a signal only, and the final column records where the real assessment must be made.

Dimension	Reach of design detail	Reading	What to confirm beyond the design
Modularity	Design-determined		
Bricolage	Design-determined		
Bounded autonomy	Design-determined		
Diversity	Design-shaped		
Redundancy	Design-shaped		
Dynamism	Design-shaped		
Scalability	Design-shaped		
Strategic polyvalence	Design-shaped		
Polycentricity	Institution-dependent		
Coordinated autonomy	Institution-dependent		
Prototyping	Institution-dependent		
Political-institutional capacity	Institution-dependent		

References

Ansell, C., Sørensen, E., Torfing, J., & Trondal, J. (2024). *Robust governance in turbulent times*. Cambridge University Press.

Ansell, C., Trondal, J., & Øgård, M. (Eds.). (2017). *Governance in turbulent times*. Oxford University Press.

Bali, A. S., Capano, G., & Ramesh, M. (2019). Anticipating and designing for policy effectiveness'. *Policy and Society*, 38(1), 1-13. https://doi.org/10.1080/14494035.2019.1579502

Capano, G. (2026) *Designing Robust and Micro-Calibrated Policy.* ANZSOG. https://a.storyblok.com/f/287203581577192/x/043203c873/anzsog-insights-paper-giliberto-capano.pdf

Capano, G., & Howlett, M. (2020). The knowns and unknowns of policy instrument analysis: Policy tools and the current research agenda on policy mixes. *SAGE Open*, 10(1), 1-13. https://www.doi.org/10.1177/2158244019900568

Capano, G., & Howlett, M. (2024). Calibration and specification in policy practice: Micro-dimensions of policy design. *Policy Design and Practice*, 7(2), 115-128. https://doi.org/10.1080/25741292.2024.2353929

Capano, G., & Pritoni, A. (2026, January 20-23). Micro-dimensions of policy design in action: Unpacking and evaluating the NRRP reforms in Italy [Conference paper]. Conference on Policy Process Research (COPPR26), Bern, Switzerland.

Capano, G., & Woo, J. J. (2017). Resilience and robustness in policy design: A critical appraisal. *Policy Sciences*, 50(3), 399-426. https://doi.org/10.1007/s11077-016-9273-x

Howlett, M., Capano, G., & Ramesh, M. (2018). Designing for robustness: Surprise, agility and improvisation in policy design. *Policy and Society*, 37(4), 405-421. https://doi.org/10.1080/14494035.2018.1504488

Howlett, M., & Ramesh, M. (2023). Designing for adaptation: Static and dynamic robustness in policy-making. *Public Administration*, 101(1), 23-35. http://doi.org/10.1111/padm.12849

Howlett, M., & Rayner, J. (2007). Design principles for policy mixes: Cohesion and coherence in new governance arrangements. *Policy and Society*, 26(4), 1-18. https://doi.org/10.1016/S1449-4035(07)70118-2

Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *American Economic Review*, 100(3), 641-672. https://doi.org/10.1257/aer.100.3.641

Wu, X., Ramesh, M., & Howlett, M. (2015). Policy capacity: A conceptual framework for understanding policy competences and capabilities. *Policy and Society*, 34(3-4), 165-171. https://doi.org/10.1016/j.polsoc.2015.09.001

ANZSOG

anzsog.edu.au