



PASCALISM

A Protocol for Testing Public Authority Claims

Version 3.0 - 2026



A Protocol for Testing Public Authority Claims

Version 3.0 | 16-09-2026

Author: Pascal Becker Hoff

Original Pascalism v1.0 deposited as i-DEPOT No. 155199 (BOIP/Benelux)

I do not know the truth. I do not tell the truth. I do not believe in it. I test it.

1. Definition

Pascalism is a testing protocol. It is not a theory of truth, and it does not produce truth. It asks whether a claim made by a public authority can be tested, applies that test, and publishes what happened. Including when the result strikes the tester, or the author of the framework.

Version 2.0 described itself as a truth-validation framework. That description was too wide, and the formula that followed from it did not hold. Version 3.0 claims only what the protocol actually does. The pretence is narrower. The stake is not.

The protocol exists because public authority has a habit of offering its own text as evidence that the text can be tested. That exchange is the object. Pascalism does not exist to declare institutions good or evil. It exists to show where a body can bind others while remaining structurally insulated from equivalent scrutiny: whether the individuals and populations subject to an institutional claim still hold an independent path on which that claim could be challenged, corrected, or broken. The step before determining evil and good.

A performance audit asks whether a body meets a criterion. This asks something else: whether the party bound by that criterion can start the path by which the body's claim could fail. That is a contestability audit. Stated as one architecture: a claim of public authority is split into what it states and the failure path that would test it, after which the path is examined from the position of the subject. Startable, and measurable.

The protocol is not neutral in what it examines. It is outcome-blind in what it finds. The test package is locked before observation, and the decision rule does not bend to the result. That is a property of procedure, not a claim to objectivity. Operationalising a claim, choosing a source, cutting stages, and naming a reference class all take judgement, which is why every one of them is published. Which of the two conditions a given authority turns out to be in is settled by the test, not by the protocol, and not by whether the authority recognises the method.

Private life and public authority are held apart throughout. Individual belief is not tested here. Only what claims public authority, produces public consequences, or holds power over others falls inside the scope.

2. Origin and Inheritance

A protocol that demands traceable origin from others has to state its own.

2.1 What is inherited

The following is Karl Popper's, drawn from *Logik der Forschung* (1934) and *The Open Society and Its Enemies* (1945). It is inherited in substance and is not claimed as original. The wording is this document's; the thinking is his.

- Falsifiability as the line that separates a claim from an assertion.
- The prohibition on ad-hoc immunisation: a claim rescued by adjustment after failure has not survived.
- Intersubjective testability: a result that only its author can obtain is not a result.
- The productive value of refutation.
- Critical publicity: knowledge becomes objective through exposure to criticism.
- The criterion that a political system is judged by whether power can be removed without violence.

Popper published no numbered protocol. His four lines of deductive testing provide part of the underlying logic used in steps 1 through 4 below. The sequence is this framework's contribution. The logic underneath it is his.

2.2 Adjacent work

Four traditions sit close to what follows. None of them is claimed here.

Meyer and Rowan (1977) described the gap between formal structure and actual operation as decoupling; contemporary governance research meets the same gap as the distinction between de jure and de facto accountability. The text/loop split below is a procedure built on that distinction. It is not a restatement of it.

Pettit holds that power is legitimate only where those subject to it can contest it, and that popular influence must be individualised, unconditioned, and efficacious. That requirement sits close to what startability operationalises. The idea is his. The test condition is not.

Bovens defines accountability as a relation in which an actor must explain, a forum may question and judge, and consequences may follow. Pathway measurement sits inside that frame.

Platt (1964) required that competing hypotheses and the observation distinguishing them be named before the experiment. The locked test package applies that requirement to claims of public authority. The requirement itself is Platt's; preregistration in empirical research does comparable work against adjustment after the fact.

Those texts describe the terrain. They are not this protocol. Citing them is provenance, not an application for membership in their field.

2.3 What is not inherited

- **Publication as a failure condition.** For Popper, publicity is the social condition under which knowledge becomes objective. Here an unpublished test is not a valid test. That is stronger, and it is different.
- **The origin and attribution rule.** No claim may reach past what its source supports. Observation is recorded at the observer and is never attributed as intent to the observed. Popper examines the logical form of a claim; he does not require traceable provenance and does not restrict inference about motive.
- **The private/public boundary as a limit on the method itself.** Popper limits what counts as a scientific statement. This framework limits who may be subjected to testing at all.
- **The refusal rule.** Refusal to submit a claim to testing withdraws its admissibility unless three conditions are met (step 3).
- **Pathway measurement.** Where the failure condition consists of a staged procedure, signal loss across the stages is measured against a reference class (step 4).
- **The text/loop split.** Every claim of public authority separates into what it states and whether its failure path is startable by the party subject to it (step 1). Popper examines the form of a claim; he does not ask who administers the procedure that would test it.
- **The locked test package.** Claim, rival, discriminator and decision rule are published as one dated block before any observation is made (step 2). Popper requires a theory to forbid something in advance; this fixes what it forbids in a form a third party can check afterwards.

3. The Boundary

The limit is set before the method, because a testing protocol without a limit becomes an instrument of discipline.

Pascalism does not test private life. Individuals hold beliefs without owing anyone a falsification. That is not an exemption this framework grants. It is the condition the framework is built to protect.

The boundary is crossed when private belief becomes public authority.

Private domain, exempt	Public domain, subject to test
Personal religious belief	Religious authority dictating law or policy
Individual emotional life	Government or institutional claims about public welfare
Private aesthetic or cultural preference	AI systems making decisions at scale
Personal financial decisions	Corporate claims about data, safety, or social impact

A private enterprise operating inside a market is not subject to testing unless it enters the public domain of influence. "Public authority" is otherwise a matter of judgement, so it is made countable here.

A claim falls within scope when at least two of the following four hold:

1. **Involuntary reach.** Non-consenting third parties bear the consequence.
2. **Scale.** The decision operates at population level rather than at the level of a contract.
3. **Closed observation.** The party controls the only data that could break the claim.
4. **Normative use.** The claim is used to justify an obligation, an access condition, or a sanction.

Two of four is deliberately low. Requiring one would draw too much private life inside the boundary; requiring all four would leave too much public authority outside it. The boundary protects private life. It is not a shelter for power.

The floor. Two of four is necessary and not sufficient. Scope also requires that the body hold power over others. Influence is not power, and neither is visibility.

The distinction that carries the weight is what a body's closure protects. Closure that shields the body from being tested counts toward scope under the third criterion, and it is the condition this protocol was built for. Closure that shields individuals from the power of others does not: the protection of sources, of members, of whistleblowers, of clients under privilege. Such a body remains testable on what it claims and on the power it does hold, and its closure is a ground under step 3, where it is stated and verified like any other.

This is set down because the obvious misuse of a testing protocol is to aim it downward. Demanding that those who shield people from power open their files, in the name of transparency, is older than any of this. It is not what this instrument is for, and a test that does it is not Pascalist testing under section 8.

4. The Function

Version 2.0 stated $\text{TRUTH}(C) = D \wedge I$: a claim is true if and only if it is destructible and internally consistent. That formula does not hold. "The moon is made of cheese" is falsifiable and internally consistent, and would therefore be true under it. What the formula expressed was admissibility, not outcome. It contained no term for the test itself, and the paragraph beneath it said something different from the formula above it.

Version 3.0 splits the function in two.

C is the claim under test, in the operationalised form produced at step 1. Not the source statement, and not the party making it. Where C is a claim of public authority it separates into C-text and C-loop, and each is carried through the protocol as its own C.

4.1 Admissibility

$$\text{TESTABLE}(C) = D \wedge I$$

D, destructible. At least one identifiable observation exists that would break C.

D is a condition on the claim, not on access. Whether that observation can actually be reached is settled at step 3, not here. An unjustified block withdraws admissibility. A justified one leaves the claim admissible and undecided. Keeping the two apart is what stops a valid ground for confidentiality from reading as a failure of the claim.

I, integrity. C contains no self-contradiction, and extends no further than its source supports.

A claim that fails this gate is not false. It is not an admissible claim under this protocol. It carries no weight here, and the test stops. Whether it remains a meaningful proposition elsewhere is not a question this protocol answers or needs to answer.

4.2 Validation

$$\text{VALIDATED}(C) = \text{TESTABLE}(C) \wedge S$$

S = applied $\wedge$ survived

applied. The failure condition defined in advance has been carried out. Where that condition consists of a staged procedure, application is complete only when signal loss across the stages has been measured against a reference class fixed in advance.

survived. No rival explanation covers the observation better than C does.

Validation is provisional and revocable. It is a status held until the next test, not a terminal state.

4.3 Asymmetry

Failure is final for C as defined in that locked package. A collapsed claim does not return under that package unless the application itself is shown to have been faulty.

A single discrepant observation is not automatically a reliable failure. Where the observation carries measurement error, sampling error, or a dependence on auxiliary assumptions, the decision rule set at step 2 states what would count as a reliable failure rather than a stray one. Logically, one counter-instance destroys a universal claim. Evidentially, an instrument can be wrong. The rule is fixed in advance so that this distinction cannot be invoked after the fact by whichever side the result inconveniences.

A revised C' is a new claim and re-enters at step 1 with its own package. This is not a route back for a failed claim: C' must differ in what it asserts, not merely in wording, and the failed version stays published. Restating a broken claim until a package is found that it survives is ad-hoc immunisation, and is prohibited under 5.2.

Validation is not final. A surviving claim stands until the next test. That asymmetry is why failed results must be published and surviving results may only be published as provisional.

5. Core Principles

5.1 Testability over belief. Without a defined failure condition there is no claim, only assertion. (inherited, §2.1)

5.2 Self-destruction as a purity criterion. A claim that no conceivable condition could falsify is not truth. It is dogma. (inherited, §2.1)

5.3 Reproducibility as the ethical baseline. A result obtainable only by its author is not a result. The test package at step 2 exists so that replication is possible for anyone. A second tester is not required in order to publish, because such a requirement would hand the method to whoever has staff, which is the class of party this protocol exists to test. (inherited, §2.1)

5.4 No exemptions for power. Tradition, authority and scale are not grounds for exemption. This applies to the author of this framework and to the framework itself. Recognition by the party under test is not a criterion of success.

5.5 Destruction is validation of the method. When a claim collapses under test, the protocol has worked. This includes claims made by the author.

5.6 Publication as verification. An unpublished test is not a valid test. (original, §2.3)

5.7 Attribution restraint. Observation is recorded at the observer. Intent is never attributed to the observed on the basis of an observed effect. A result that establishes what happened does not license a statement about why. (original, §2.3)

6. The Protocol

Five stages. The sequence is binding. No stage may be skipped because the outcome appears known in advance.

Step 1. Formulate

Operationalise the claim so that its failure condition is legible in the same sentence. A source statement is rarely admissible as it stands. A body that represents, protects, serves, or acts on behalf of some group has defined nothing, and no observation can break it. Under 5.1 that is assertion, and the test stops.

Operationalisation is the act of restating the claim in terms that name what would break it. If no such restatement is possible without changing what was claimed, record that and stop.

The split. Every claim of public authority separates into two claims, and both are carried through the protocol. No split, no test.

C-text is what the rule, model, procedure or policy states. Its test is a source check. C-text may survive, and often will.

C-loop is fixed in the same words regardless of domain:

The failure path of C can be started by the party subject to C and brought to one of the three outcomes, without the party behind C controlling the only observation that could close it.

C-loop operationalises whether the observation identified under D can be reached through a failure path available to the party subject to C, in the case where the party making the claim also administers the procedure that would test it. It is not a second D. D asks whether such an observation exists; C-loop asks whether there is a route to it that the subject can take. The last clause is the refusal rule at step 3; it is not a separate condition.

This fixes what a failed C-loop means. It does not make C-text false, and it does not break it. It means the claim of public authority cannot obtain validated status under this protocol, because the path required to test it is unavailable or has failed. C-text may still survive its source check, and that survival establishes only what the text states. Reporting it as more than that is the layer error below.

A surviving C-text is cheap. It establishes that a rule says what it says. Where the party behind the claim also holds the path by which the claim could fail, that establishes nothing about whether the claim can be tested at all.

Startable means that the party subject to C can set the failure path in motion without leave from the party behind C. Difficulty is not the criterion. A two-thirds threshold, a long procedure, or a high evidentiary bar is heavy, and heavy is startable. What is not startable is a path that cannot be entered at all without the permission of the party whose claim it would test. Representation counts as starting only where the subject can mandate that representative on this claim and remove them over it. A chamber that may amend a rule the subject cannot put on its agenda is not, by that fact alone, a startable path for the subject.

This is a design choice and is stated as one. It does not follow from logic. Voting in a general election is real influence, and a protocol could count it as starting the path. This one does not, because it asks whether the subject holds the first move on this claim, and a vote across a whole programme does not carry a particular claim into a procedure. A reader who rejects that choice can reject the finding that rests on it. The choice is published here so that it can be.

A gate written into a public text is an observation, not a gap. Where a published rule places a given subject-matter outside the reach of a given remedy, bars a given body from reviewing a given class of decision, or withholds from an external party the right to define the test condition, the measurement is complete. C-loop is failed, or inadmissible under the refusal rule. It is not undecided. Undecided is reserved for the case where no such gate is stated and the observation is genuinely missing.

Layer discipline. A finding on C-loop is a finding about the path, not about the text. Where C-loop fails, C-text is not thereby broken: the rule still says what it says, and no observation of ¬C-text has been made. The reverse holds equally. Publishing a result on one layer as though it settled the other is the error this split exists to prevent.

The same mould applies across domains:

Domain	C-text	C-loop
AI laboratory	"the model does X within Y"	an external party may define Y and observe the run
Complaint or FOI procedure	"there is a six-stage procedure"	signal reaches the end, measurable against a reference class
System-actor company	"this data, drug, or infrastructure is safe"	an independent failure condition is possible on the data bearing the consequence
Regulator	"we enforce norm N"	files pass the published stages at the rate of the regulator's own reference class
Constitutive rule	"this rule may be revised under the procedure it sets out"	a party subject to the rule can start a revision or review that does not die at the gate

Step 2. Define the failure condition

State what would break C.

The test package consists of five statements, published as one dated block before any observation is made:

1. **C in falsifiable form.** One sentence in which the failure condition is legible.
2. **Rival R.** The account that predicts the same observations as C except at the discriminator.
3. **Discriminator D*.** The one observation on which C and R must diverge.
4. **Decision rule.** What counts as failed, survived, and undecided.
5. **The tester's relation to the party behind C.** Independent, involved, or the party itself.

Nothing observed after the package is published may change it. A failure condition chosen after the data are seen is not a failure condition.

A party may test itself, and says so. Nothing here reserves testing to outsiders, and a rule that did would put the method beyond the reach of anyone without standing. But a body running the protocol on its own claims holds every position at once: it sets the failure condition, holds the observation, and reports the outcome. That is the conflict this protocol exists to make visible, and it does not disappear because the procedure was followed.

So it is declared. A self-test is published as a self-test and carries that label wherever its result is cited. A self-test may be a valid test, if it satisfies every other requirement of this protocol. The label buys nothing: a self-test that bends the source rules, the refusal rule or the publication conditions fails for the same reasons any other test would. What it can never be is an independent test, and the two are never reported as the same thing. An undeclared self-test is not Pascalist testing under section 8.

Selecting R. R is not the rival that damages C most. R is the rival that covers the same data with the fewest additional assumptions. Where the party making C has itself named a rival account, that account is R unless it is internally inconsistent. This removes the tester's freedom to defeat C by selecting a caricature.

Selecting D*. The discriminator is the observation under which C and R would differ. It is not the condition C is least likely to survive. A test selected for maximum lethality is an execution, not a test, and it establishes nothing about what C asserted. A weak condition that C passes easily is equally worthless.

If no failure condition can be formulated, C fails the gate at 4.1.

A package published before observation may be replicated. A second tester, without access to the first tester's observation, operationalises C and selects R and D* independently. Where both arrive at the same D* and read the observation the same way, the result is replicated. Where both arrive at the same D* and read the observation differently, both readings are published and C remains undecided until a third package resolves it. Where they arrive at different discriminators, the result is underdetermined: neither failed nor survived. The disagreement is itself the finding, and is published as such.

Replication raises the status of a result. It is not a condition of publication. A protocol that required two testers before anything could be published would be available only to parties with staff, which is the class of party this protocol exists to test.

Open. For claims of public authority, D and R for C-loop are prescribed at step 4 and inserted into the package before observation, which closes package-shopping on the loop. It does not close every dispute. Testers may still disagree on who the party behind C is, and on whether a published sentence is a gate or a threshold. Those disputes are published. They are the remaining underdetermination, and they are stated here rather than concealed. Elsewhere, selection of R constrains the tester without fully determining the choice.*

Step 3. Expose

Apply the test to publicly identifiable sources. Record source, date and method so that another person can repeat the test without access to the tester.

Refusal. Where the party making C controls the only observation that could break it and withholds it, the observation identified under D cannot be reached by anyone else. C therefore loses admissibility. It is not shown false, and no observation of $\neg C$ has been made. It is no longer a testable claim in practice, and may not be published as validated.

This applies unless the ground for refusal meets all three of the following conditions:

1. **Specified.** The refusing party names which part is withheld and which is not. An unbounded appeal to confidentiality is not a ground, because it presents nothing that can be tested.
2. **Externally verifiable.** The ground exists independently of the refusing party: a statutory provision, a patent, a contractual clause, a data-protection basis. Internal classification by the refusing party does not qualify, as that would make the party judge in its own case.
3. **Covering.** The ground extends across the full scope of the claim. Where part of the claim remains testable, that part is tested, and refusal on the remainder does not protect it.

A refusal meeting all three conditions leaves the claim untested without withdrawing its admissibility. The outcome is undecided, not survived.

Re-testing the ground. The ground for refusal is itself a claim, and is re-tested against the same three conditions at a stated interval. If it lapses, admissibility lapses with it. If it holds, the outcome remains undecided for as long as it holds. Each re-test is published, whether the ground holds or lapses.

Undecided has no expiry. A valid ground does not weaken through the passage of time, and treating it as though it did would penalise a party for refusing rather than test what was claimed. Undecided is a genuine third outcome, not a waiting room for falsification.

Undecided does not close the file. Losing admissibility is not nothing. The claim may no longer be presented as validated under this protocol. What it means under any other epistemic or legal system is not this protocol's jurisdiction and is not decided here.

Continuing to assert the claim in the public domain after that date is observable, and can be tested as its own C. Be exact about what that C is. It is not the original claim. It is a claim about the record: that party A continued to assert C after date T. Testing it establishes that the assertion continued. It establishes nothing about whether C is true, and it does not return the original claim to the protocol. Anyone reporting one as the other has broken layer discipline.

The undecided result stands. Later publications concerning the same claim carry the unresolved outcome and the date it arose. The open status attaches to the claim, not to the party. No inference about the party follows from it. This is bookkeeping, not a sanction.

Repeated undecided is its own C. A party may hold a valid ground more than once. A patent, a live investigation, the privacy of third parties: these are real, and each one on its own says nothing. But a run of undecided outcomes against one party is an observation, and it can be tested rather than insinuated.

The claim is not that the party is concealing anything. It is narrower and it is measurable: *the testability of party A deviates from that of comparable parties on comparable claims*. Of all tests published against A in a stated window, what proportion closed as undecided on refusal, measured against the reference class in the same window.

Fix the eligible claim class before counting. Comparable parties is not enough. A tester who puts privacy, live investigations and security material to A while putting routinely published material to B has measured the selection, not the parties. The claim class is stated in advance: the same kind of claim, the same domain, the same window, on both sides. This constraint protects the party under test, and it is the tester's own work that it disciplines.

This runs as a pathway measurement under step 4 and carries every constraint there. Five comparable parties minimum, two neighbouring classes, metric and threshold fixed before counting. No deviation means no finding, and the run of undecided outcomes stands as what it is: a party with grounds, in a sector where grounds are ordinary.

Deviation gives a finding on testability, not on the claims. The underlying claims remain undecided. Nothing here establishes that any of them is false, and under 5.7 nothing here licenses a statement about why the grounds were held. What it establishes is that this party can be tested less than its peers can, which is the asymmetry section 1 names as the object.

The rule governs the claim that was refused, not the party refusing. Withholding evidence is not evidence of falsity, and under 5.7 no inference about motive follows from a refusal.

Step 4. Observe

The outcome is one of three.

Failed. The observation breaks C, or a rival explanation covers the observation better than C does.

Survived. The failure condition was fully applied and did not occur.

Undecided. The failure condition could not be fully applied. Undecided is not published as survived.

Before recording a failure, state the best available alternative account of the same observation and show why it does not fit. A test that never names a rival has not been completed.

The loop. Where C is a claim of public authority, C-loop remains undecided until the path has been measured or fixed as not startable. A survived C-text may not be published as a survived C-loop. Where C-loop concerns a constitutive rule, one that governs the conditions of its own revision or review, pathway measurement is required and not optional. The reference class is revision trajectories of comparable constitutive rules in the same period.

Default discriminator. For any claim of public authority, D* for C-loop is fixed, and remains fixed until someone publishes a more economical R:

Can the party subject to C start the path, and does signal emerge from it when measured against the reference class?

R is prescribed with it, and it is not abolition of the procedure: the path is heavy, but startable and measurable.

The survive branch is reachable, including for constitutive rules. A binding citizens' initiative that does not exclude the rule in question, a referendum route open on that subject matter, or constitutional review that a subject can bring directly, all place the first move with the party subject to the rule. Where such a route exists and throughput does not deviate from the reference class, C-loop survives. A test whose outcome is settled in advance for an entire class forbids nothing in that class, and under 5.2 that is not a test. This branch is what keeps the discriminator discriminating.

- **Failed:** not startable, or startable with throughput indistinguishable from zero against the reference class.
- **Survived:** startable, with throughput distinguishable from zero against the reference class. A negative deviation beyond the fixed threshold does not change the outcome. It qualifies it, and is carried into the mark at step 5. Throughput at or above the reference class is not a deviation for this purpose.
- **Undecided:** the observation is missing, or the class contains fewer than five, or the result inverts under a neighbouring class.

Heavy is not a failure. A fixed discriminator removes the tester's freedom to choose the terms on which public authority is judged.

Pathway measurement. Where the failure condition consists of a staged procedure, the existence of the procedure is not the result. Signal loss across its stages is measured. Five elements are fixed before counting begins.

1. **Stages.** The procedure broken into publicly documented, countable steps.
2. **Reference class.** All comparable procedures in the same period, measured along the same stages. A throughput figure for a single subject carries no meaning in isolation and can be read to suit any conclusion.
3. **Window.** Stated before counting begins.

4. **Metric and threshold.** The measure of throughput, and the size of deviation that counts as a finding, both stated before counting begins. Otherwise "no deviation" and "indistinguishable from zero" are decided after the numbers are seen, which is the researcher freedom this step exists to remove.
5. **Decision rule.** No deviation from the reference class means no finding, and C proceeds unchanged. Deviation narrows C: the mechanism exists in form and performs measurably differently from comparable procedures.

Three constraints bind the count.

Minimum class. A reference class containing fewer than five comparable procedures in the same window yields no finding, only missing data. Five is a chosen threshold, not a derived one. It is set low enough to keep the method usable on small procedural populations and high enough that a single outlier cannot carry a result. Anyone who proposes a better-grounded threshold and publishes the reasoning has improved the protocol. A single case is an anecdote regardless of how it is presented. Report raw counts, not percentages, where the class is small.

Fragmentation. The minimum class can be used against the measurement. A party that defines its procedures narrowly enough that no reference class ever reaches five has removed the comparison rather than failed it. Where the stages are drawn from the party's own definitions and no comparable class can be built from them, that is the party controlling the only observation that could break the claim, and it goes to the refusal rule at step 3. Test the definitions against the three conditions there. Definitions that fail them do not shield the claim. Definitions that meet them leave the outcome undecided, not survived.

Sensitivity. The outcome must hold under at least two neighbouring reference classes: a different period, a different sector, or a different scale. If the conclusion changes with the class, C is not failed. The measurement is unstable, and the instability is the finding.

Codebook first. Stages are taken from published procedural texts, not from the tester's reading of them. Each stage definition cites the document the step comes from.

Pathway measurement records where signal is lost, not why. Under 5.7, no result of this measurement licenses an inference about intent. Absence of throughput is missing data unless the reference class shows that comparable procedures produced it. Silence is not falsification. Non-response by a party is not a result. The outcome is not adjusted after it is known.

Step 5. Publish

An unpublished test is not a valid test. Publication includes failed and undecided outcomes, and corrections to the author's own earlier tests. A protocol that publishes only its confirmations is a protocol that cannot fail.

Prior packages are named. Where more than one test package has been published against the same C, every earlier package is listed with the result of the test that followed it. Preregistration inside a single test does not prevent selection between tests. Running packages in succession until one fails is the same defect at one level up, and listing them is what makes it visible.

The finding and its significance are separated. The outcome is reported bare: the failure condition, the rival, the discriminator, the sources, the result. No evaluative term enters it. There are three outcomes and no fourth.

What the finding means for the party subject to C is stated after it, under its own heading, in plain language. A failed C-loop means one of two things, and the publication states which was observed: either the party bound by the claim cannot set the failure path in motion at all, or it can start the path and throughput along it is indistinguishable from zero against the reference class. The two are different findings and are never reported as one. Saying so is not an addition to the result. It is the result, restated for the person it concerns.

The C-loop result may be published as a mark. Four states, each tied to an outcome already produced by step 4. The mark adds no new empirical finding. It is a fixed communication layer derived from the measured result, in a form a reader can take in at a glance.

Mark	Outcome	What it means for the party subject to C
Green	Survived, no adverse deviation from the reference class	You can open this path, and it runs at least as comparable paths run
Amber	Survived, with negative throughput deviation beyond the fixed threshold	You can open this path, but measurably less comes through it than through comparable paths
Red	Not startable, or throughput indistinguishable from zero	You cannot open this path yourself, or you can open it and no measurable signal comes through it. The rule exists. What it would take to test it is out of your hands
Open	Undecided	Not measured, or the party has withdrawn on a stated ground. There is no result

The wording in the third column is fixed. It is published as it stands, in the reader's language, and no evaluative term is added to it. What a reader concludes about the body from a red mark is the reader's own. The mark states the measurement and stops.

Four conditions on publishing marks. Without them the scale becomes a reputation score, which is not what it measures.

1. **The mark attaches to a tested claim, not to a body.** A body may hold green on one claim and red on another. Every mark carries the claim it belongs to.
2. **The reference class is published with it.** A single mark on a single body is an opinion about that body. The same measurement across the comparable class in the same window is a position in a field. Amber additionally carries the deviation threshold set at step 4, since that boundary is the tester's choice and not a property of the path.
3. **Every mark carries its date.** Validation is provisional under 4.2, and so is a mark. A body that is green today may change its procedure tomorrow.

4. **Open is not a weak red.** It is set apart from the scale, because placed beside the others it reads as a verdict. It is not one.

The document is not submitted to the party behind C for recognition. If that party answers, the answer is source material for the next test, not a condition of this one.

A published test is itself startable. A protocol that measures whether the subject of a claim can open a failure path cannot leave its own results closed. Run the discriminator on this protocol and the answer has to be yes.

So: every published test carries a stated route by which a party it concerns can put a fault in the application to the tester. Not a right of reply on the finding, which would return the verdict to the party. A route against faulty application: a source misread, a stage taken from the wrong document, a reference class that excludes comparable cases, a package altered after observation.

Where such a fault is shown, the test is re-run and the correction is published with the original. Where it is not, that is published too, with what was claimed and why it did not hold. Either way the fact that the route was used goes on the record. A route that leaves no trace when it is used is not a route.

The tester is bound by 5.4 like anyone else. This is where that stops being a sentence.

7. Applications

The protocol is domain-independent. Any staged procedure operated by a public authority can be tested by it.

AI systems. The productive application is not bias testing, which is established practice and requires no new framework. It is the class of claims a laboratory makes about its own model where no external party can define the failure condition. Under step 3, a refusal to permit external definition of that condition withdraws the claim's admissibility unless the three refusal conditions are met; where they are met, the outcome is undecided.

Regulatory and institutional procedure. Complaint handling, freedom of information procedures, and audit routes are staged procedures with public documentation. Step 4 applies directly.

Constitutive rules. Rules governing the conditions of their own revision or review are staged and publicly documented, which makes step 4 applicable in full.

Applied cases in every domain are published separately under this protocol, including corrections to the author's own earlier tests. No worked case is reproduced here. A protocol carrying one is read as an argument about that case, and this document is not about any case.

8. Attribution and Integrity

The method is open. The name is conditioned.

Open. The protocol, the two functions, and the terminology are published for use, citation, implementation and modification by anyone, without permission and without licence. Reproducibility under 5.3 requires that a result be obtainable without access to the tester, and that requirement cannot coexist with a permission gate.

Conditioned. Pascalism is a name, and the name carries a condition. An implementation may be called Pascalism, or described as Pascalist testing, only if it carries all five stages of section 6 and the boundary of section 3. A partial implementation is a legitimate thing to build and an illegitimate thing to call by this name.

One case is named because it is the likely one. Where a claim of public authority has been split under step 1, a publication reporting a surviving C-text while omitting the result on C-loop is not Pascalist testing, whatever else it contains. Reporting the text and dropping the loop is the form selective adoption takes here: it lets a party present its own rule as evidence about the path that would test it. The layers are published together or the name does not apply.

This condition exists to address a specific risk: selective adoption, in which an institution takes the parts of the protocol that serve it and discards the conditions that constrain it. An audit gate at the author does not prevent that, because it is unenforceable in practice and, under 5.4, exempts the author from the standard applied to everyone else. A naming condition does address it, because a partial implementation carrying the name can be identified as such by anyone applying section 6 to it. The check is distributed rather than held centrally.

Version 2.0 required audit rights for the author and a licence for operational deployment. Those conditions are withdrawn. The concern behind them stands; the mechanism did not. Adoption by the party under test is not a substitute mechanism.

Attribution to the author and to i-DEPOT 155199 is requested wherever the protocol or its terminology is used.

9. Standing Falsification

This document is subject to its own protocol.

It fails if the protocol is shown to be underdetermined to the point where two independent testers applying it to the same text reach incompatible results with no way to identify which application was faulty. For claims of public authority the discriminator on C-loop is prescribed, so divergence there is not the expected site. The likelier sites are the attribution of "the party behind C", the attribution of "the party subject to C", which may be read as the citizen, the user, the complainant or an affected third party, and the classification of a published sentence as a gate or a threshold. A single underdetermined result does not falsify the protocol.

Underdetermination is a published outcome under step 2, and a protocol that can report its own indecision is functioning. What would falsify it is systematic underdetermination: divergence as the ordinary case rather than the exception.

It does not fail merely because a claim validated under it is later destroyed. Validation is provisional by construction, and a claim that survives a correct test today and falls to a better one tomorrow is the protocol working as specified. What would falsify it is a pattern of validations that fall to observations already available at the time of testing, with the test package correctly formed. That would show the gate admits claims it cannot discriminate.

The remedy in either case is retraction or revision, published as such.

Pascal Becker Hoff

Turnhout, Belgium

Version 3.0, 16-09-2026:

SHA-256: 5b1a20020cce9255856af7bd009c7dd00b84a6982a36c252b7c01cae498329f6

This hash is of the source file, not of this PDF. The source is available on request.

Version 2.0, 2025-2026: SHA-256 0e4361fed19937fc2753970ba733c242de134dc364644a0c13c7bd13dc964b79

Version 1.0, 2025: SHA-256 67e96fe766927f334fa3eed5b783fa6bc833b7a448d28d1004a280924a47ed9

References

Popper, K. R. (1934). *Logik der Forschung*. Vienna: Julius Springer. English edition: *The Logic of Scientific Discovery* (1959), London: Hutchinson. §3 on deductive testing; §6 on falsifiability as the criterion of demarcation; §§21-24 on falsifiability and ad-hoc immunisation; §8 on intersubjective testability.

Popper, K. R. (1945). *The Open Society and Its Enemies*, Vol. 1, ch. 7. London: Routledge. On the criterion that institutions permit removal of power without violence.

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, 83(2), 340-363.

Pettit, P. (2012). *On the People's Terms: A Republican Theory and Model of Democracy*. Cambridge University Press.

Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, 13(4), 447-468.

Platt, J. R. (1964). Strong Inference. *Science*, 146(3642), 347-353.

של