

Model Management under Finite Conditions

A Diagnostic and Decision Architecture for Validity, Friction, and Revision

Abstract

Finite cognitive systems organize experience, stabilize reference, and guide expectations and action without fully surveying their presuppositions and consequences. Stabilized orders support orientation but can be overextended or continued inappropriately. This paper develops a programmatic diagnostic and decision architecture for managing such orders, particularly their elaborated forms as models. Examining the claim architecture connects units of reference, claims, domains, scope, test conditions, and couplings. Validity is understood as domain-bound viability under strain; friction denotes an epistemically readable mismatch of an activated stabilized order under its operating conditions. The architecture separates assessments of validity from decisions about continued use. Rules for opening, redistribution, and abandonment structure exploration, attribution of findings, and the justification of correction, revision, suspension, replacement, or abandonment. The social conditions for testing in distributed epistemic and model management are also considered.

A boundary case of personal identity examined through epistemic critique and an internal, LLM-assisted comparison involving the Allegheny Family Screening Tool challenge different parts of the program. The comparison with a strong disciplinary baseline establishes no general superiority. LLM judges rated certain additional contributions concerning claim specification, model identity, transfer of findings, and justification of revision as substantive. Their soundness by disciplinary standards, generalizability, and practical utility have not been independently confirmed. The architecture is not a decision algorithm and replaces neither disciplinary procedures nor normative justification. Its supplementary aim is to make the testing, use, and revision of stabilized orders and models more controllable and correctable across everyday, scientific, technical, and institutional contexts of cognition.

Keywords

Model management; finitude; cognitive systems; claim architecture; validity; domains; friction; revision; falsification; distributed model management

Stefan Rapp

Independent Researcher

Last revised: 10 October 2026

<https://orcid.org/0009-0004-0847-9164>

<https://doi.org/10.5281/zenodo.23278445>

Epistemic Reality Project: <https://www.epistemics.de>

© 2026 Stefan Rapp — Licensed under CC BY-NC-ND 4.0

Table of Contents

1. Problem and Aim.....	3
2. Inside, Outside, and the Boundary Rule.....	6
3. Domains and Validity.....	9
3.1 Validity and Validity Mapping.....	9
3.2 Domains and Modes of Challenge.....	9
3.3 The Three Domains.....	10
3.4 Transition Functions and the Transfer of Validity.....	11
3.5 Status of the Triad.....	12
4. Stabilization, Units of Reference, and Models.....	13
4.1 Units of Reference and Identity Rules.....	13
4.2 Ontologization and Absolutization.....	14
4.3 Models as Organized Forms of Order.....	14
4.4 Selection, Omission, and Tentative Stabilization.....	15
4.5 Stabilization Costs and Testability under Strain.....	16
5. Friction and Diagnosis.....	17
5.1 Friction as Relational Mismatch.....	17
5.2 Where Friction Becomes Readable and How It Develops.....	18
5.3 Scope and Indicators.....	19
5.4 Friction as a Signal.....	20
6. The Decision Architecture.....	21
6.1 Examining the Claim Architecture.....	21
6.2 Proportional Forms of Response.....	22
6.3 Retention and Immunization.....	23
6.4 The Opening Rule.....	23
6.5 The Redistribution Rule and Falsification.....	24
6.6 The Abandonment Rule.....	25
6.7 Relations among the Three Rules.....	25
7. Shared Epistemic Reality.....	28
7.1 Shared Reference and Intersubjective Validity.....	28
7.2 Trust, Institutions, and Power.....	29
7.3 Social Stability and Loss of a Shared Basis for Correction.....	29
7.4 Interplay of the Domains and Science.....	30
7.5 Distributed Model Management.....	30
8. A Boundary Case and a Comparative Architecture Test.....	32
8.1 Personal Identity: The Limit of Exclusive Continuation.....	32
8.2 AFST: Comparative Stress Test of the Model Management Architecture.....	33
9. Self-Limitation.....	36
9.1 Scope and the Limit Drawn by Epistemic Critique.....	36
9.2 Limited Coordinating Function.....	36
9.3 Self-Application and Non-Applicability.....	37
9.4 State of Testing and Corroboration.....	38
9.5 Assessment.....	40
Conceptual Canon of This Paper.....	41
References.....	50
Appendix A: Illustrations of the Diagnostic and Decision Architecture.....	52

1. Problem and Aim

Finite cognitive systems must orient themselves without being able to fully survey the conditions and consequences of their orientation. They distinguish, compare, anticipate, act, and correct, even though neither all relevant information is available to them at once nor all the effects of their own orders are foreseeable. Cognition therefore begins not with a complete overview, but with the task of enabling viable orientation under limited conditions. Every orientation remains selective: it foregrounds differences, brackets others, and stabilizes continuations before their full scope can be known.

This task concerns more than scientific theories. Everyday expectations, medical classifications, technical systems, legal categories, political interpretations, and self-descriptions also depend on selecting differences, stabilizing points of reference, and ordering possible continuations. Without such orders, experiences could not be compared, expectations tested, or actions coordinated. Collective and technical cognitive systems likewise depend on such selective orders, even where their forms of access, processing, and correction are organized differently.

Stabilized orders enable cognition but also create limits. They help determine what appears relevant, which deviations become visible, and which forms of correction are possible. An order can remain viable in one context and fail in another. It can successfully provide orientation and, precisely through that success, obscure its presuppositions, costs, or omissions. Strain therefore does not necessarily mean that the entire order must be abandoned; equally, local success must not be confused with unlimited validity.

Models are elaborated forms of such orders. They stabilize reference, connect features and relations, generate expectations, and enable testing. Yet they do not automatically specify their own scope, limits, or conditions of appropriate application. A model can perform well locally while also being overextended. It can fail in a particular application even though its basic structure remains viable, or persist even though recurring deviations are accommodated only through additional assumptions, altered test conditions, or the exclusion of competing interpretations.

The central problem therefore concerns not only the formation of models, but their management. What claim is being made, under what conditions is it viable, and what accounts for an emerging strain? The strain may concern the model itself, but equally its application, scope, units of reference, couplings, or test conditions. This determines whether further observation, local correction, scope restriction, reordering, or a more fundamental revision is appropriate. Model management thus encompasses not only change, but also the justified retention, restriction, suspension, replacement, or abandonment of a model's use in a guiding role.

Scientific and philosophical approaches address central component functions of the task examined here. AGM theory formalizes contraction and revision of logically represented belief states, while Bayesian confirmation theories investigate changes in degrees of belief in response to new evidence (Alchourrón, Gärdenfors, and Makinson 1985; Earman 1992). The architecture developed here is not a competing formal theory of revision. In its proposed use, it first clarifies, before a revision procedure is chosen, whether a strain concerns the model, its application, its units of reference, the domain assignment, a coupling, or the test conditions. A proportional response therefore need not mean minimal change within an already fixed space of beliefs or hypotheses. It means a response

whose depth of intervention is justified by the scope and structural significance of the diagnosed strain and the relevant consequences. Such a response may also consist in retention, scope restriction, recoupling, reconfiguration, suspension, replacement, or abandonment.

Quine's holistic account of empirical testing likewise combines openness in the attribution of findings with a conservative orientation when changing the web of beliefs (Quine 1951, Section VI). The proportionality principle therefore makes no claim to novelty simply by calling for limited change. Here it is developed as a requirement to justify different forms of response: the scope of an established loss of validity, the depth of structural intervention, and the scope of practical responses must be specified separately. Extensive restructuring or precautionary suspension may thus be justified without turning the finding into a more comprehensive refutation.

A further point of contact is Wimsatt's philosophy of limited and fallible cognitive systems and the literature on models in philosophy of science. Wimsatt examines cognition under conditions of limited overview, error-prone heuristics, and piecemeal approximation; Morgan and Morrison, as well as Giere, characterize models as mediating or perspectively situated cognitive instruments (Wimsatt 2007; Morgan and Morrison 1999; Giere 2006). The present approach does not adopt these insights as a complete theory of scientific representation. It focuses on the management of stabilized orders that can be managed as models: the assignment of their claims, the diagnosis of frictions and costs, and the distinction between possible forms of response.

The analysis of shared epistemic reality also connects to Longino's theory of critical social interaction and Goldman's investigation of social epistemic practices and institutions (Longino 1990; Goldman 1999). The present paper links this perspective to the distinctions between distributed intersubjective order and its local reconstruction, between social stability and intersubjective validity, and between validity status and guiding status. The distinctive contribution claimed by this approach lies not in the isolated novelty of all its components, but in their combination within a shared architecture of model management. This combination is intended to make claim assignment, units of reference and identity conditions, scope, test conditions, friction diagnosis, and proportional revision jointly controllable. Whether it thereby provides a substantive additional benefit over strong disciplinary procedures is not presupposed, but is itself subject to testing. This comparison must also consider existing tools that already combine several of these functions (§9.4).

This paper consolidates central arguments of the Epistemic Reality Project (ERP) to date, clarifies their conceptual relations, and revises some earlier specifications. It brings concepts, distinctions, and diagnostic instruments previously developed in several separate works into a shared programmatic framework. The earlier works remain relevant for their respective detailed treatments; the present paper, however, specifies the architectural context within which their mutual connections and conceptual scope are to be assessed.

Functional reconstructions and the methodological organization of possibilities for testing and response make different claims. A reconstruction must be assessed by whether it adequately captures the cognitive achievements under investigation; the model management architecture by whether it improves their testing, use, and revision. A methodological benefit in application therefore does not automatically confirm claims about the emergence or actual organization of cognitive processes.

The paper examines the combination of unit of reference, model, claim type, domain, mode of challenge, scope, test conditions, couplings, and continuation rules. This combination is called the claim architecture. Examining it means asking whether its elements remain jointly viable or whether a locally successful order is misassigned, overextended without notice, or shielded from strain. The examination therefore concerns not only the internal structure of a model, but also the conditions of its activation, application, and continued use.

Four concepts form the operational core of this broader architecture. Validity is the domain-bound viability of a model or claim under strain. Domains distinguish spaces of ordering with different forms of stabilization, testing, and correction. Friction designates an epistemically readable mismatch of an activated stabilized order under its operating conditions. It may become apparent in expectations, connections, or the model-based organization of that order without already determining its cause or the appropriate response. Revision designates the restructuring, under epistemic pressure, of a model, an order that can be managed as a model, or a claim to validity.

The proposed architecture is neither a general decision algorithm nor a competing disciplinary methodology. It replaces neither disciplinary methods and empirical investigations nor normative justification or practical judgment, and it does not itself generate the disciplinary testing procedures required in each case. Its function is to distinguish conditions and transitions that must be considered in the use, transfer, and revision of models and findings. A subjective experience, a medical finding, an institutional rule, and a physical prediction can be closely connected without being viable under the same conditions. Nor does the performance of a description establish a warrant for intervention.

The paper develops the architecture step by step. Chapter 2 reconstructs inner form and outer form, system-bound perspectivity, and the boundary rule. Chapter 3 specifies domains, validity, and transitions between orders of validity. Chapter 4 examines stabilization, units of reference, and models; Chapter 5 friction and diagnosis. Chapter 6 develops the proportional decision architecture, including the distinction between validity status and guiding status. Chapter 7 elaborates shared epistemic reality and distributed model management. Chapter 8 first challenges the architecture through a boundary case for epistemic critique concerning personal identity and then through a comparative stress test using the Allegheny Family Screening Tool. Chapter 9 specifies its scope, the extent of its testing to date, and its self-limitation.

The aim remains deliberately limited. The program provides no ultimate justification, resolves no disciplinary or normative conflicts on behalf of the relevant fields, and promises no frictionless cognition. Its possible benefit lies in more precisely assigning the claim, unit of reference, test conditions, and scope of findings. Generalizing a loss of validity must be supported by evidence; the depth of revision must additionally be justified by the structural significance of the strain. Practical restrictions or precautionary suspensions require a separate justification in light of consequences and uncertainty.

2. Inside, Outside, and the Boundary Rule

Everyday and scientific orientation usually presupposes an already differentiated situation: a knowing and acting system here, objects, other systems, and an external world there. This order is practically indispensable. A reconstruction that critically examines the conditions of cognition, however, must not simply take it as its starting point. It must clarify what sustains the distinction between inside and outside without already presupposing them as originally separate regions.

The following account is not an empirical developmental history. It does not claim to identify the temporal stages through which a particular cognitive system develops these distinctions. It proposes a functional reconstruction of the conditions under which determinations can be treated as an inner or outer order within the forms of orientation examined here. Whether individual conditions are necessary and how far the reconstruction extends beyond these forms require separate justification.

This chapter establishes the context of epistemic critique within the program; it does not provide a complete derivation of its decision rules. The distinction between validity status and guiding status, the continuity requirement for transferring findings, and the separate justification of practical precaution can also be assessed through explicitly specified model claims, test conditions, and consequences of use, without adopting the entire reconstruction of inner and outer form. The boundary rule concerns the scope of additional claims made in epistemic critique or ontology. An objection to the reconstruction challenges these methodological distinctions only insofar as it affects a presupposition actually required to justify them; conversely, their applicability does not confirm the reconstruction as a whole.

Whatever becomes determinable for a cognitive system is accessible to it only within an act of determination. For humans, this dependence on acts of determination can be illustrated through consciousness: a person distinguishes themselves from an external world, yet both the determination of themselves and the determination of the world occur within their acts of perception, remembering, thinking, and interpretation. Consciousness is an illustrative human instance of the context of determination intended here, not its general definition.

Even what later counts as an external object is perceived, registered, remembered, compared, designated, measured, or modeled. A cognitive system can examine and alter its conditions of determination, but cannot step outside every condition of determining in order to grasp them additionally from a presuppositionless standpoint. Any attempt to describe such a standpoint positively would itself again be an act of determination.

It does not follow that objects initially lie “inside” or are merely produced by the cognitive system. Such an inference would already employ the inside–outside distinction whose emergence is to be reconstructed. Initially, one can state only that determinations are not all equally available, alterable, or resistant.

Some determinations allow a comparatively wide range of variation. Imagined or simulated scenarios can be changed, possibilities tentatively combined, and stored connections reordered. This flexibility remains limited by forms of access, processing capacities, and already stabilized structures. Nevertheless, it differs functionally from determinations that cannot be altered correspondingly through internal variation.

Other determinations exhibit a greater degree of recurrence, unavailability, and resistance. They constrain expectations or compel correction. A determination later conceptually marked as “sun,” and thus made easier to take up again, may appear differently and be described differently under changing conditions. Its regular appearance and observable effects, however, cannot be changed simply by a cognitive system varying its representation of it.

These features do not necessarily occur separately or in clearly distinct forms. Determinations stand in contexts in which different profiles of variability, availability, dependence, recurrence, and resistance develop. Resistance provides an early distinguishing feature, but does not determine classification on its own. When such profiles become consolidated across multiple determinations and acts, contexts of ordering can be treated as inner form or outer form.

Inner form and outer form therefore designate not original classes of individual determinations, but functional consolidations of such contexts. They order determinations according to how these jointly become available for further operations, constrain expectations, and compel corrections.

In humans, these already established orders can subsequently be differentiated further. Impressions are then interpreted, for example, as visual, auditory, tactile, or bodily perceptions and attributed to an environment, one's own body, or one's own states. These sensory and bodily interpretations do not necessarily first generate the inside–outside structure; rather, they specify, connect, and alter an already stabilized order.

Perspective, too, is not a presuppositionless starting point. It designates the ordered dependence of determinations on a cognitive system's particular possibilities for access, processing, distinction, and action. Perspectivity is therefore a consequence of finite and system-bound determination, not a property of a particular form of biological organization. That something is determined differently under changed conditions neither makes it arbitrary nor removes the limits of the relevant access.

Outer form is particularly relevant to the architecture developed here. It brings together orders stabilized as comparatively repeatable, not arbitrarily available, and resistant to mere variation. In the form of finite orientation reconstructed here, it enables, in particular, functional distinctions between internal variation and observation, expectation and disappointment, and action and limit.

This does not yet show that an outer form in this sense is necessary for every possible finite cognitive system. The reconstruction initially specifies a function that sustains the forms of orientation examined. Whether this function is universally required, can be fulfilled by differently organized structures, or arises only under certain conditions of realization remains a further question.

Nor may such an outer form be equated with a positively determinable outside in itself. This is where the boundary rule applies:

What can be marked only as a limit of positive determinability must not simultaneously be treated as a positively determined region beyond that limit.

A cognitive system can make viable statements about stones, stars, technical systems, or physical processes. It can observe, measure, compare, intervene, and correct models. Such determinations can be precise, robust, and technically effective. It does not follow, however, that it could additionally determine the complete constitution of the objects concerned independently of every

possible determination. Such a statement, too, would have to be formed and justified within an act of determination.

The boundary rule therefore opposes neither empirical research nor realist discourse. It does not deny that a stone is hard, a measurement deviates, or a technical system fails. Rather, it prevents the transition from viable, correctable determinations to an ultimate claim about a fully positively accessible region of the world beyond all conditions of cognition. Dependence on acts of determination alone does not prohibit statements about objects independent of cognition; what must be examined in each case is whether the reasons offered support the additional claim being made.

The rule thereby limits two opposing overextensions. The first absolutizes outer form into a world fully describable independently of every determination. The second infers from cognition's dependence on acts of determination that reality is merely subjectively produced or arbitrary. A cognitive system's inability to leave all its conditions of determination does not eliminate resistance. Conversely, the repeatability and resistance of orders do not establish a complete ontology of an outside in itself.

Realist discourse therefore remains legitimate, but limited. Outer form designates a viable order within which determinations can be repeated, tested, and corrected. It is neither mere invention nor the immediate representation of a fully accessible world in itself.

This reconstruction, too, must not be absolutized. A cognitive system's inability to completely negate the act of positive determination without performing it again limits its self-negation. It does not establish an ultimately valid ontology of the cognitive system or of the act of determination.

For the form of finite orientation reconstructed here, inner form and outer form thus function as stabilized forms of ordering through which different, including gradual and mixed, profiles of determination can be distinguished functionally. It follows neither that precisely this differentiation is necessary for every possible cognitive system nor that it designates presuppositionless ontological regions. Its theoretical status depends on which achievements of orientation, distinction, and correction it actually explains.

Models can therefore be treated neither as immediate representations of a fully accessible outside nor as arbitrary products of a closed inside. The next step is to clarify the different conditions under which claims within such orders remain viable, come under strain, and can be corrected.

3. Domains and Validity

The proposed reconstruction treats inner form and outer form as functionally stabilized orders. It distinguishes profiles of availability, recurrence, controllability, and resistance without already deriving a general mechanism of emergence. The conditions under which statements, expectations, and models within such orders remain viable have not yet been clarified.

Not every claim is tested and corrected in the same way. The statement that someone feels pain requires a different assessment from the statement that a drug causes a physiological effect. A legal obligation remains viable under different conditions from an astronomical prediction. Before a claim is assessed, its type must therefore be clarified, together with the order within which it claims validity and the forms of testing and strain through which its limits can become visible.

3.1 Validity and Validity Mapping

The viability of a model or claim under the conditions relevant to it is called validity:

Validity is the domain-bound viability of a model or claim under strain.

Validity is neither mere agreement nor mere usefulness, nor is it an undifferentiated substitute for truth. A claim is not valid simply because it is expressed, believed, socially accepted, or institutionally established, and it is not valid in every context merely because it has proved viable in one. Its viability is bound to presuppositions, claim type, scope, conditions of testing and strain, and possibilities of correction. Validity status records the assessment justified by the available tests; it remains fallible. Insufficient testing or an inconclusive finding constitutes neither corroboration nor an established loss of validity.

Pain illustrates this distinction. The claim “I feel pain” can possess high subjective viability with regard to the occurrence of an experience. It does not follow that the claim “This injury causes the pain” is also viable; it requires additional observation and investigation. The further statement that the person is therefore unable to work introduces an institutionally and normatively regulated context. The claims can concern the same unit of reference without being viable under the same conditions. The immediacy of experience does not establish a certain medical cause; an unremarkable measurement does not cancel the experience; an institutional rule does not determine whether pain is felt.

The preliminary mapping of a claim to the context and conditions in which it can meaningfully be assessed is called validity mapping. This mapping does not yet judge the claim's validity. Many conflicts arise not because a claim is clearly unviable within its own context, but because its local validity is transferred unnoticed to another context.

3.2 Domains and Modes of Challenge

The project calls functional spaces of ordering in which claims are stabilized, tested, challenged, and corrected in characteristic ways domains. Domains are not separate parts of the world or independent substances. They organize conditions of validity without assigning each object unambiguously to one region. Different claims about the same unit of reference can be made in different domains.

Claim type, domain, and mode of challenge answer different questions. Claim type specifies what is asserted or claimed; the domain designates the functional space of ordering in which the claim is to be viable; the mode of challenge specifies how it is tested or otherwise subjected to strain. These specifications are related but distinct. The linguistic form of a claim does not automatically determine its domain, and a testing procedure does not by itself establish which conditions of validity it actually addresses. Within one domain, a claim can be tested or challenged in several ways.

Four modes of challenge can be distinguished as a compact cross-cutting axis:

- subjective-phenomenal testing: connection to experience, capacity for differentiation, and phenomenal fit;
- dialogical justification and taking up again: contrasting perspectives, justification, availability to be taken up again, and correctability;
- observation, measurement, and intervention: observation, measurement, reproduction, prediction, intervention, and technical success;
- conceptual-formal testing: consistency, boundary cases, countermodels, and competing reconstructions.

These modes do not constitute additional regions of the world; they make claims sensitive to error in different ways. A scientific model can be challenged through observation, measurement, and intervention while also being tested conceptually and formally. An institutional rule belongs to an intersubjective context of ordering, but can be justified and disputed dialogically, tested for consistency conceptually and formally, and examined through observation and measurement of its effects. None of these tests replaces the others. Normative warrant does not constitute a fourth domain. Normative claims can be examined through several of these modes, but their warrant cannot be reduced to any one of them alone. An independent theory of normative validity is not developed here.

The misunderstood promise (Appendix A.5) illustrates the distinction. Evidence that a sentence was uttered, its function as a mutually binding commitment, and the disappointment experienced can each be the subject of a statement. A recording may establish the wording, but cannot by itself settle its binding force within the shared order or the experience involved. Domain assignment asks which order supports the respective claim and what transition from wording to obligation is being claimed. A further measurement of the same wording may therefore miss the actual problem requiring correction; clarification of shared expectations and agreements may instead be needed. The example illustrates the intended diagnostic function, but does not yet demonstrate an additional benefit over another sufficiently differentiated analysis.

3.3 The Three Domains

The subjective domain comprises the individual organization of experience, memory, expectation, meaning, and self-reference. It designates not a closed inner space, but an order in which immediate concern and individual connection to experience are decisive. Strains become visible here when experiences can no longer be coherently situated, expectations repeatedly fail, or self-descriptions become incompatible. Immediate experience does not automatically establish the validity of its causes or of decisions derived from it.

The intersubjective domain comprises orders of shared reference, taking up again, and correction among several cognitive systems. The reconstruction so far initially concerned the formation of order within one cognitive system. Other cognitive systems can be stabilized as units of reference that respond independently, contradict, and are not freely available. Where determinations can be taken up, continued, disputed, and corrected among them, an intersubjective order of validity emerges. Language, rules, promises, institutional roles, and shared concepts belong to this order. Contrasting perspectives, taking things up again together, and correction are decisive, rather than mere agreement. Social stability alone therefore does not establish intersubjective validity. An interpretation may persist through habit, power, or institutional inertia despite restricted correctability.

The functional-empirical domain comprises orders stabilized through repeatable relations of resistance and effect. Claims are viable here when they enable reproducible observations, measurements, predictions, or interventions under determinable conditions. This domain, too, provides no immediate access to a world in itself; observation and measurement remain bound to instruments and conditions. Its distinctive achievement lies in challenging connections through resistance, reproduction, and effects that cannot be arbitrarily altered either through individual imagination or social agreement.

The three domains can be reconstructed from different patterns of strain and correction: the subjective domain from disturbances in contexts of experience and meaning, the intersubjective domain from failures to take something up again and blocked shared correction, and the functional-empirical domain from reproducible deviation, resistance, and failure of effects. These patterns will later be specified more precisely as epistemic friction; here they serve only to distinguish orders of validity provisionally.

3.4 Transition Functions and the Transfer of Validity

The domains are analytically distinguishable but closely coupled in practice. Subjective experience is communicated linguistically, situated intersubjectively, and possibly investigated functional-empirically. Empirical results are interpreted conceptually, processed institutionally, and experienced subjectively.

Such transitions do not transfer determinations unchanged. Experienced pain becomes a report, the report becomes medical information, and this may in turn become a problem of measurement, classification, or decision. With every transition, the relevant features and the conditions under which the respective claim remains viable change.

Couplings that make determinations connectable across different orders of validity are called transition functions. They enable transfer without eliminating the differences between the domains involved. Determinations and units of reference can be continued across orders of validity without remaining unchanged. The same unit of reference can be taken up again under different claim types and test conditions; which features are relevant and what constitutes its viability change with the transition.

Transitions can fail, however. Subjective concern may be treated as empirical proof, measurability confused with complete reality, or social consensus with substantive confirmation. Conversely, a

subjective interpretation may reject every external correction. Such problems do not necessarily concern a simple error within a domain. Often, a claim is assigned a viability it would possess only under different conditions.

3.5 Status of the Triad

The distinction between subjective, intersubjective, and functional-empirical domains is not a final division of reality. It is a functional, reconstructive, and revisable working hypothesis. Its purpose is to make different conditions of validity, testing, strain, and correction visible. Not every cognitive system must develop all three domains internally. The triad distinguishes orders of validity for claims, not necessary internal modules of every cognitive system.

Further subdivisions may be required. Overlaps and boundary cases do not automatically refute the triad either. The decisive task is not to classify every phenomenon unambiguously, but to clarify the claim type, the order in which it is to be viable, the relevant modes of challenge, and the transitions claimed. Purely formal claims must not be assigned to the intersubjective domain solely because they can be tested jointly. Whether they require a further domain specification remains to be examined through boundary cases. The triad may offer an additional benefit by explicitly examining transitions between orders of validity. If a more precise specification of claim type and mode of challenge already captures the same distinctions and unwarranted transfers just as reliably and with less effort, an additional domain assignment is unnecessary for that task. The triad therefore claims neither completeness nor an independently demonstrated additional benefit.

The domain distinction initially marks epistemic asymmetries: claims managed within different orders of validity are not mutually substitutable merely because they concern the same unit of reference. Subjective experience, intersubjective meaning, and functional-empirical determinations possess different conditions of access, testing, strain, and correction. Their respective claims therefore cannot be confirmed, canceled, or fully taken over on the basis of another domain's conditions without specified transitions.

This establishes neither an ontological separation of the domains nor their irreducibility in principle. A complete reduction or ontological identification may be possible, but is not accomplished by mere functional coupling or by explanatory successes within one domain. It requires an additional step of justification and testing. The domain distinction thus initially separates different conditions of validity without deciding in advance the ontological relation between the contexts they capture.

The boundary rule limits the metaphysical overextension of outer form; the domain distinction limits the uncontrolled transfer of validity. The modes of challenge prevent a claim from being confirmed or rejected using an unsuitable testing procedure.

This has not yet explained how the units of reference toward which such claims are directed emerge. For a model to be tested, it must be possible to determine what it refers to, which features count as relevant to identity, and under which changes something is taken up again as the same. The next section therefore examines stabilization, units of reference, and model formation.

4. Stabilization, Units of Reference, and Models

For a claim to be tested, it must be possible to determine what it refers to. Observations can be compared, measurements repeated, and statements corrected only if something can be taken up again as the same despite changing conditions. Cognition therefore presupposes stabilization.

Stabilization does not mean keeping something unchanged. It orders differences so that particular features and connections are continued while other changes are permitted or set aside. A tree can count as the same tree despite growth and changing appearance; an institution can persist despite changes in membership. Stabilization thus consists in selective continuation.

This operation is to be understood analytically, not necessarily temporally. Not every stabilization already forms a unit of reference, and not every unit of reference forms a model. Nor do units of reference emerge entirely before the models within which they are managed. They enable model formation and are themselves shaped by models that organize relevant differences, identity features, and tolerated changes.

4.1 Units of Reference and Identity Rules

A unit of reference is something stabilized under specific conditions as identifiable, available to be taken up again, and referable. It need not be a material thing. Events, processes, persons, institutions, symptoms, rules, or theoretical parameters can also be managed as units of reference.

What counts as the same unit depends on which features are treated as relevant to identity and which changes are tolerated. An identity rule therefore specifies under which changes a unit of reference is continued as the same, and when reassignment or dissolution becomes necessary. Such rules need not be explicitly formulated. They can reside in perceptual habits, linguistic practices, technical procedures, institutional arrangements, or scientific classifications.

Identity rules often become visible only when boundary cases arise. New observations can require earlier criteria to be specified, restricted, or changed. The epistemic attribution of continuing identity thus proceeds according to rules across permitted changes. Which ontological determination of identity follows from this is not thereby decided.

This also applies when the same unit of reference is described differently in different domains. What matters is not complete identity of features, but sufficiently connectable continuation under specified identity conditions. Where this breaks down, it must be examined whether the same unit is still intended or merely the same name continues to be used.

This applies particularly to models and technically or institutionally realized systems that change under a continuing name. Changes in model structure, data basis, interface, conditions of deployment, or institutional embedding can alter the unit of reference relevant to a test so substantially that findings concerning an earlier state cannot simply be transferred to a later one. If such a transfer is proposed, it must be specified which properties or mechanisms relevant to the particular finding persist and why this continuity is sufficient for the claim being made. This requirement applies equally to supporting and challenging findings.

4.2 Ontologization and Absolutization

The project calls the operation through which something is stabilized as an identifiable, referable unit available to be taken up again ontologization.

Ontologization is a particular form of stabilization, not an ontology and not in itself a problematic pattern. Without it, cognition could return to nothing, attribute no properties, track no changes, and form no models. Even establishing that a measurement differs from an earlier one presupposes a sufficiently stable unit of reference.

Ontologization therefore does not mean attributing a final metaphysical mode of existence to a unit. It designates the operative formation of a unit to which further determinations can be related. A symptom, institution, electron, or self-model is ontologized in different ways; what they share is their continuation as determinable units of reference.

The operation becomes problematic when its presuppositions, selections, and limits become invisible. A unit formed for a particular purpose may then appear as a wholly independent and conclusively determined thing. A classification is confused with the thing itself, a measured quantity with a complete property, or an institutional category with a person's identity.

This problematic pattern is called absolutization. Ontologization enables reference; absolutization treats the resulting order as immediate, exhaustive, or without alternatives. Cognition cannot dispense with ontologization, but must keep the conditions of its units of reference revisable.

4.3 Models as Organized Forms of Order

A unit of reference does not yet form a model. A model emerges when units of reference, features, relations, expectations, and continuation rules are connected into a sufficiently determined order. The order must be capable of being marked as such, taken up again, delimited, compared, tested, and corrected. Continuation rules specify how relations, expectations, operations, or further model steps proceed; unlike identity rules, they do not primarily concern whether a unit of reference remains the same.

Models can be expressed linguistically, formalized mathematically, depicted visually, implemented technically, or realized in routines. What matters is not their representational form, but the structured connection between units of reference and expectations. A model specifies which differences are relevant, which relations are expected, which changes count as possible, and which observations can challenge its continuation.

Units of reference and models mutually determine one another. Units of reference enable stable relations; models influence which units become visible, which features are attributed to them, and how their identity is continued. A medical model may specify which symptoms count as belonging together, while the resulting disease category orders further observations. This interaction is not arbitrary: repetition, resistance, comparison, and correction constrain models.

Models are therefore neither mere representations nor arbitrary inventions. They do not reproduce an already fully determined reality unchanged, but neither do they depend solely on their creators' decisions. They select, simplify, and order within positive determination. The questions are

therefore which units of reference and relations a model stabilizes, for which tasks this order remains viable, and under what conditions its limits become visible.

4.4 Selection, Omission, and Tentative Stabilization

Every model formation rests on selection. A model can function only by not considering everything at once. It foregrounds some differences and brackets others. This omission is not merely a deficiency, but a condition for an order to be managed as a model: reducing complexity makes relations visible, expectations formulable, and tests possible.

What is omitted does not thereby disappear. It is simply not activated within the respective model function. Selection becomes problematic when a model can no longer treat its omissions as omissions or when it is transferred to tasks for which precisely what has been bracketed is decisive. The performance of an order therefore also depends on whether its selections remain visible and revisable.

Not every stabilization already makes an elaborated claim to validity. Cognitive systems can provisionally impose an order on a diffuse field to make possible relations, consequences, and strains visible. This intermediate form is called tentative stabilization.

A hypothesis, thought experiment, provisional classification scheme, or speculative formal relation can serve as a working structure without already being treated as conclusively viable. Its status must remain explicit. A tentative order must not pass unnoticed into a canonical or ontological commitment.

Tentative stabilization must be sufficiently determined to generate consequences and strains, while remaining provisional enough to be changed or abandoned without excessive commitment. It forms the transition between open exploration and model management in which claims to validity can be meaningfully assessed.

4.5 Stabilization Costs and Testability under Strain

Every stabilization has benefits and costs. It facilitates taking up again, comparison, and orientation, but restricts possible interpretations, commits resources, and can obscure relevant differences. Such costs may appear as increasing complexity, additional assumptions, maintenance effort, rigid identity rules, or the persistent exclusion of particular strains.

Costs do not automatically refute a model. Every viable order requires commitments, simplifications, and effort. What matters is whether these costs are proportionate to the orientation provided, remain visible, and can be corrected. A model may provisionally continue to be used despite rising costs if no viable alternative is available. Conversely, an apparently simple model can be epistemically expensive if it shifts strains elsewhere or makes them invisible.

A model must therefore be both stable and flexible. Without stability, it cannot be taken up again and tested; without flexibility, it cannot respond to changed conditions. Testability under strain here does not yet designate established validity, but the capacity of a sufficiently determined and open order to be meaningfully tested and corrected. Selective stabilization thus makes it possible to manage an order as a model and allows excluded differences, unsuitable identity rules, or overextended expectations to emerge later as mismatches.

This establishes the basis for a more precise specification of friction. The following chapter examines when such a mismatch becomes epistemically readable, where it can be located within a claim architecture, and what diagnostic scope it possesses.

5. Friction and Diagnosis

Models and other stabilized orders do not become problematic simply because they are selective and incomplete. Every order that can be managed as a model foregrounds particular differences, generates expectations, and brackets other things. Its limits become epistemically relevant when a mismatch becomes visible under concrete conditions and can be attributed to an activated epistemic order. A strain that can be interpreted in this way is called friction below.

Friction is an epistemically readable mismatch of an activated stabilized order under its operating conditions.

Expectation, connection, and model-based organization do not necessarily designate separate types of structure, but overlapping functional aspects of stabilized orders. The same order can sustain expectations, enable or constrain certain continuations, and be organized in a model-based way. Simple formations of expectation or connection need not presuppose an elaborated model.

5.1 Friction as Relational Mismatch

This definition contains four necessary elements.

First, a stabilized order must be activated. An order that guides no perception, expectation, action, communication, or further reference cannot exhibit friction in this sense. Only in operation does it become visible whether the intended continuation remains viable.

Second, a mismatch must occur. An expected course of events fails to materialize, a connecting operation fails, an observation deviates, or an action produces consequences other than those intended.

Third, the mismatch occurs under concrete operating conditions. The same order may be under strain in one context while remaining viable in another. Application, instrument, context, unit of reference, coupling, and test condition therefore belong to the diagnosis.

Fourth, the mismatch must become epistemically readable. Not every irritation reveals which activated order, or which of its functional aspects, is affected. Friction begins where a deviation can be related to an activated order and taken up as a diagnostic problem.

Resistance is more general than friction. Something may not be arbitrarily available or alterable without challenging a particular expectation. Friction arises only in the relation between an activated stabilized order and the conditions of operation under which its continuation fails.

Error and friction are not identical either. An error usually presupposes a standard of correct execution or assignment. Friction, by contrast, can indicate that it remains unclear whether an instrument, application, unit of reference, coupling, test condition, or model is affected. An irritation or disturbance can occur without already being diagnostically related to an activated stabilized order. An anomaly initially designates a conspicuous result within a context of observation or modeling; only its diagnostic assignment to an activated stabilized order and its conditions of operation makes it a sufficiently determined friction.

Falsification and revision lie on different levels. Falsification designates a particular loss of validity under suitable testing, revision a possible response to epistemic pressure. Neither does every friction entail falsification, nor does every falsification immediately entail a particular revision.

This openness guards against two shortcuts: attributing every deviation to the model itself encourages premature model abandonment; shifting every deviation to application, boundary conditions, or measurement error immunizes models against serious challenge.

5.2 Where Friction Becomes Readable and How It Develops

Friction diagnosis begins with the context in which a mismatch first becomes epistemically readable, and with the way it develops. Both dimensions guide the investigation; neither already identifies the cause.

In the subjective domain, friction may appear as a disturbance in contexts of experience, expectation, or meaning. In the intersubjective domain, it manifests as failures to take something up again, incompatible references, or blocked correction. In the functional-empirical domain, it emerges as reproducible deviation, declining predictive performance, or failure of effects. Friction can also become visible in transitions between domains when relevant differences are lost in a transfer or different conditions of validity are mixed.

The context in which friction becomes readable is not necessarily the level at which it originates. An institutional conflict may, for example, result from an unclear rule, an unsuitable empirical indicator, or a faulty transfer of subjective experience.

The course of a strain describes how it unfolds over time:

- In a gradual course, possible overextension, increasing stabilization costs, drift, or creeping shifts in coupling should be examined.
- In an abrupt course, thresholds, context changes, and measurement, coupling, or assignment errors are particularly relevant.
- In a recurring course, it should be examined whether local repairs conceal a structural mismatch, blocked feedback, a conflict of goals, or an incorrect domain assignment.

The course of friction is a search heuristic, not a causal diagnosis. The same course can arise from different problems; different courses can reflect the same structural strain. The form of the course does not automatically determine a response.

For initial orientation, the context in which friction becomes readable can be combined with its course:

Where friction becomes readable	Gradual	Abrupt	Recurring
subjective	increasing incoherence or interpretive costs	collapse of an expectation	similar conflicts despite reinterpretation
intersubjective	increasing effort required for mutual understanding	overt breakdown of a rule or connection	repeated misunderstandings or blocked corrections
functional-empirical	declining performance or increasing additional effort	marked deviation or failure	reproducible failure
transition/coupling	increasing translation effort	breakdown of a transfer	systematic misclassification

The matrix identifies typical manifestations. It neither replaces disciplinary testing nor permits automatic assignment of causes or domains.

5.3 Scope and Indicators

It must additionally be determined how far a friction extends within the claim architecture. Local friction concerns a single expectation, application, measurement, or assignment. Consolidated friction occurs when several local mismatches form a connected pattern. Systemic friction concerns central presuppositions, identity rules, transitions, or conditions for continuing an order.

Course and scope constitute different diagnostic axes. Recurring friction is not systemic simply because it recurs; the diagnosis must first show whether the mismatches form a structural pattern. Specifying scope does not yet determine the appropriate response either.

Friction is often inferred from indicators. Three groups are particularly relevant:

- declining performance: increasing error or failure rates, declining predictive accuracy, or reduced reproducibility;
- increasing stabilization costs: more additional assumptions, parameters, exceptions, maintenance effort, or coordination effort;
- loss of correctability: objections without consequences, exclusion of serious alternatives, or an inability to represent deviations.

No single indicator establishes a fundamental friction. More parameters can indicate necessary differentiation or defensive stabilization; institutional stability can mean a viable order or blocked correction. Diagnostic weight arises from the conjunction of several indicators, their consolidation over time, and their persistence under improved test conditions.

5.4 Friction as a Signal

Friction indicates a need for diagnosis but does not determine the appropriate response. The same mismatch can result from faulty application, unsuitable test conditions, overextended scope, problematic coupling, an inadequate unit of reference, or a central model assumption.

The absence of visible friction does not establish validity either. An order may be so closed that it fails to register deviations, its test conditions exclude counterexamples, or its costs are shifted elsewhere. Correctability therefore requires not freedom from friction, but the ability to make relevant mismatches visible and keep their significance open.

The quality of a model order consequently lies not in completely avoiding strains. It lies in keeping friction perceptible, not prematurely treating the initially affected context of ordering as the cause, and appropriately determining the scope of the strain.

Diagnosis thus leads to the decision architecture. Only in the next step can it be clarified which response corresponds to the strain that has been located and classified.

6. The Decision Architecture

Friction indicates that an activated order does not remain viable as expected under particular conditions. It determines neither what constitutes the mismatch nor how to respond to it. A distinct task of decision therefore lies between diagnosis and change. It must clarify which component of the claim architecture is under strain, which depth of intervention corresponds to the finding, and which new costs or frictions the response may generate.

The decision architecture follows a principle of proportionality:

Change should be as narrow as possible, but as extensive as necessary.

The depth of intervention is not measured solely by the number or extent of observed deviations. Even a limited finding can concern a supporting presupposition and justify extensive restructuring. Narrower responses are preferable insofar as they adequately address the diagnosed problem; their possibilities need not first be exhausted in practice if the testing already counts against them. Repeated local repairs must also be assessed in terms of their overall effect and cumulative costs.

The proportionality principle thus specifies requirements for justifying a response, rather than thresholds independent of disciplinary judgment. The level affected, the scope and structural significance of the finding, and the reasons why a narrower response is or is not sufficient must be stated. Consequences and uncertainty may justify more extensive precaution; this must not be presented as a correspondingly extensive loss of validity. Standards determined within the relevant discipline remain necessary, but the connection between finding, standard, and response must be explicitly open to scrutiny.

6.1 Examining the Claim Architecture

The first step is not to ask whether a model must be abandoned, but which claim is actually under strain.

Examining the claim architecture means investigating which component of the combination of unit of reference, model, claim type, domain, mode of challenge, scope, test conditions, couplings, and continuation rules is affected. The purpose of the claim and its evaluative standards must be explicitly included: the purpose specifies the performance claimed; the standards specify the test conditions and conditions of failure. The architecture does not itself justify their selection or normative warrant.

The examination typically proceeds from the narrower to the broader level:

1. Execution, data, instrument, and context;
2. coupling, transition, and validity mapping;
3. model structure, units of reference, and identity rules;
4. problem framing, goals, and evaluative standards.

This sequence is not a rigid prescription. It is intended to avoid both premature model criticism and the persistent displacement of problems onto boundary conditions. If several assignments of error

remain compatible with the findings, this indeterminacy must be recorded as a diagnostic result; it must not be concealed by forcing an unambiguous assignment.

Conditions of failure and expected results should, where possible, be specified before testing. Subsequent adjustments can be legitimate, but must be marked as changes to the test and challenged again under the new conditions. Otherwise, it is difficult to distinguish whether a model has been specified more precisely or shielded against contrary indications.

Multiple tests or targeted variations are not independent merely because they are conducted separately or produce concordant results. Variation initially shows only how a result behaves under the conditions actually changed. Stability across several variants therefore establishes robustness with respect to those variants, not independence from presuppositions retained across them.

Additional testing power arises insofar as relevant presuppositions, access to data, error mechanisms, or evaluative assumptions differ. Far-reaching claims may therefore require a counter-test that does not merely repeat the same claim architecture under slightly altered conditions, but can actually change the problem framing, model status, domain assignment, criterion of failure, or claim to validity.

6.2 Proportional Forms of Response

Forms of response differ in the object and depth of a change and in its consequences for further use. They do not constitute a sharply separated ladder of stages: recoupling or reconfiguration may also be a revision. A separate question is whether the model is retained, restricted, suspended, replaced, or abandoned in its previous claim to a guiding role (§6.6).

Further observation is appropriate when the data situation, reproducibility, or scope remains unclear. A testing task and a time or occasion for reassessment must be specified for such observation. It must become neither indefinite postponement nor a reason to delay necessary precautions until diagnosis is complete.

Local correction changes data, parameters, execution rules, or limited assumptions without restructuring the supporting model structure. Not every update is already a revision. When new data are fed into an unchanged procedure, the model order remains essentially intact.

Scope restriction preserves a model's validity within a narrower range and abandons a more extensive claim. It is legitimate if the boundary is substantively justified. It becomes immunization if the scope of validity is shifted after every contrary indication so that the model can no longer fail in principle.

Recoupling changes the connection between models, units of reference, or orders of validity that remain viable. What is under strain is then not necessarily their content, but their translation, assignment, or joint activation.

Reconfiguration changes what is available for epistemic processing and how it is structured. Units of reference can be split or combined, model boundaries shifted, and relations reordered. It can prepare a revision or coincide with one, but is not identical to it.

Revision is the restructuring, under epistemic pressure, of a model, an order that can be managed as a model, or a claim to validity.

Revision can concern model components, central relations, identity rules, scope, or domain assignment. It differs from updating in that the order that can be managed as a model itself changes; reorganization, by contrast, can also encompass broader orders that are not yet sufficiently model-based. Scope restriction, recoupling, or reconfiguration become revision when they deliberately restructure the order that can be managed as a model or its claim to validity. Whether a revision is justified must be assessed separately. A proportional revision considers the scope and structural significance of the strain, available alternatives, and the costs and consequences of restructuring. Even excessive or poorly justified restructuring remains a revision.

6.3 Retention and Immunization

Provisionally retaining a model under strain can be rational. Despite known limits, it may remain the best-performing available order as long as its remaining performance is high, friction is limited, further testing is promising, and no viable alternative is available. Its strain must remain visible, its claim restricted where necessary, and subsequent revision possible. The absence of a replacement does not suffice to justify continued use; its justification and consequences must be examined separately.

The boundary with immunization does not depend on the number of auxiliary assumptions. What matters is whether additions can be tested independently, generate new expectations, specify conditions transparently, and preserve genuine conditions of failure. If contrary indications are regularly dismissed as exceptions, sources relevant to testing are excluded, or boundaries shifted after the fact, stabilization becomes defensive. Immunization occurs when hardly any possible result can still seriously challenge the claim to validity. This distinction is close to Popper and Lakatos without coinciding with their approaches (cf. Popper 1935; Lakatos 1970).

6.4 The Opening Rule

Finite cognitive systems cannot pursue all alternatives at once. They must allocate resources between consolidating an existing model and opening new spaces of search.

The opening rule designates the point at which further consolidation of stability should give way to renewed exploration.

Recurring relevant frictions can change not only individual contents, but also the firmness of expectations, the breadth of the search space, and readiness for revision. Repeated corroboration under varying conditions, conversely, permits justified narrowing.

Exploration becomes plausible when several of the following conditions combine, or when one occurs with sufficient strength and diagnostic relevance: stagnating gains in robustness, disproportionately increasing consolidation costs, persistent relevant friction, or a serious viable alternative or one developable with reasonable effort.

None of these criteria decides the matter simply by being present. High friction may justify further work on the existing model if substantial gains in robustness are expected. High costs may reflect necessary precision. An attractive alternative is insufficient if it does not take on the relevant strains.

Opening changes only the allocation of search resources. It proves neither that the existing model is false nor that an alternative is better. Comparability requires no uniform metric, but does require a transparent comparison of performance, costs, and unresolved frictions.

If deliberately different variations no longer generate a new relational structure, viable alternative, relevant friction, or changed problem formulation, provisional consolidation becomes plausible. It does not prove that the search space has been exhausted.

6.5 The Redistribution Rule and Falsification

A failed test often leaves open which component of the claim architecture loses validity. Predictions depend not only on a model assumption, but also on data, instruments, auxiliary assumptions, boundary conditions, and transitions. This distribution problem is close to the Duhem–Quine thesis, according to which assumptions are usually not tested in isolation (cf. Duhem 1906/1954; Quine 1951).

The loss of validity established by a test should be assigned to the level, or connection between levels, that was actually challenged.

This redistribution rule requires a double proportionality: as narrow as possible, so that local strain does not lead to premature model destruction; as extensive as necessary, so that consolidated strain is not persistently distributed among peripheral errors. If the test reveals only the failure of a compound claim, assignment to individual components remains provisionally open.

Falsification here designates the domain-relative loss of validity of a sufficiently determined model or claim under fair, suitable functional-empirical testing. It presupposes explicit claim conditions, test conditions, and conditions of failure, does not designate every friction, and does not automatically lead to model abandonment. For other modes of challenge, the more general term loss of validity is initially used; whether and under what conditions the concept of falsification can be extended to them remains open.

Contextual falsification concerns a model's claim for a particular application, subdomain, or condition. A mere execution error, by contrast, does not already falsify the model applied. The scope of the loss of validity must be distinguished from the object challenged.

A model is globally falsified when its general claim to validity persistently fails to remain viable within the claimed range under fair, domain-appropriate test conditions, and the deviation exceeds an independently justified tolerance range ϵ . No alternative is required for this. For quantifiable deviations, ϵ can be determined from measurement uncertainties, approximation conditions, or margins of error justified in advance and, where appropriate, relative to a purpose. For statistical claims, the measured quantity, data basis, uncertainty, and decision rule must additionally be specified. The criteria must not be adjusted after the fact to protect the model; switching costs or institutional inertia do not justify a better assessment of validity.

Contextual and global thus distinguish the scope of a loss of validity. Its object must be specified independently: for example, a model claim, a claimed application performance, a model–domain assignment, or a claim to universality. An architectural loss of validity concerns the assignment or connection of claims and test conditions. It counts as falsification only if justified through suitable

functional-empirical testing in the sense specified above; a conceptually demonstrated misassignment is insufficient. “Architectural” therefore does not designate a third level of scope.

Even globally falsified models can remain useful historically, didactically, heuristically, as approximations, or within limited regimes. This does not restore their impaired general claim to validity. Falsification determines validity status; it does not by itself decide future guiding status.

6.6 The Abandonment Rule

The abandonment rule concerns the justification of a model's existing claim to a guiding role. What is meant is its guiding role within a particular orientation or use, not its complete removal from all contexts of cognition. The claim to a guiding role must be restricted, suspended, or abandoned insofar as its continued exercise is not sufficiently justified under the specified epistemic and practical conditions.

Retention continues this role; restriction limits it. Suspension temporarily sets it aside and requires conditions for renewed testing. Abandonment ends the previous claim to a guiding role.

Replacement occurs when another order assumes the guiding role. A sufficiently viable alternative is required for justified replacement, but not for suspension or abandonment. Justified withholding of judgment or refraining from a use that is no longer justified are also possible outcomes.

Simply pointing to friction, costs, or an impaired validity status does not determine guiding status. Remaining performance, the justification of the intended use, uncertainty, possible harms, reversibility, and the consequences of continuing and refraining must be examined. If a model is to be replaced, the alternative must be able to assume its relevant functions more effectively, taking account of maturity, comparable test conditions, transition conditions, and reasonable costs.

Historical investments and infrastructure can influence the form and speed of a transition, but must not count as epistemic performance of the established model. The relative comparison of performance touches on questions also emphasized by Lakatos and Laudan concerning changes in theories and research (cf. Lakatos 1970; Laudan 1977).

A model can continue to be used in a limited way despite global loss of validity if this use remains independently justified. The absence of a replacement alone does not justify it. Its use must be marked as limited, provisional, or pragmatic and must not appear as epistemic rehabilitation. Conversely, a claim to a guiding role can be suspended even before error localization is complete if the consequences and uncertainty involved justifies this. Such a precautionary decision does not establish a more extensive loss of validity.

Abandonment need not be total. A model can lose its claim to a guiding role while continuing in limited areas as an approximation, tool, or historical reconstruction.

6.7 Relations among the Three Rules

The three rules fulfill different functions:

- The opening rule governs the weighting of consolidation and exploration.
- The redistribution rule locates the loss of validity established by a test.

- The abandonment rule examines the justification of continued guidance, including suspension, replacement, and abandonment.

Falsification is not a fourth decision rule, but a possible validity status resulting from suitable testing and redistribution. Neither global falsification nor suspension or abandonment necessarily requires a replacement model. Only justified replacement requires a sufficiently viable alternative and a transparent comparison.

Opening decides neither falsification nor abandonment. Redistribution decides neither search resources nor further use. A decision about guidance does not retrospectively determine which test failed or how far its loss of validity extends.

Together, the rules form a skeletal control structure:

Diagnose friction → locate the challenged level or record an unresolved assignment → assign loss of validity appropriately → justify revision and, where necessary, open the search space → examine guiding status in light of performance, consequences, and uncertainty.

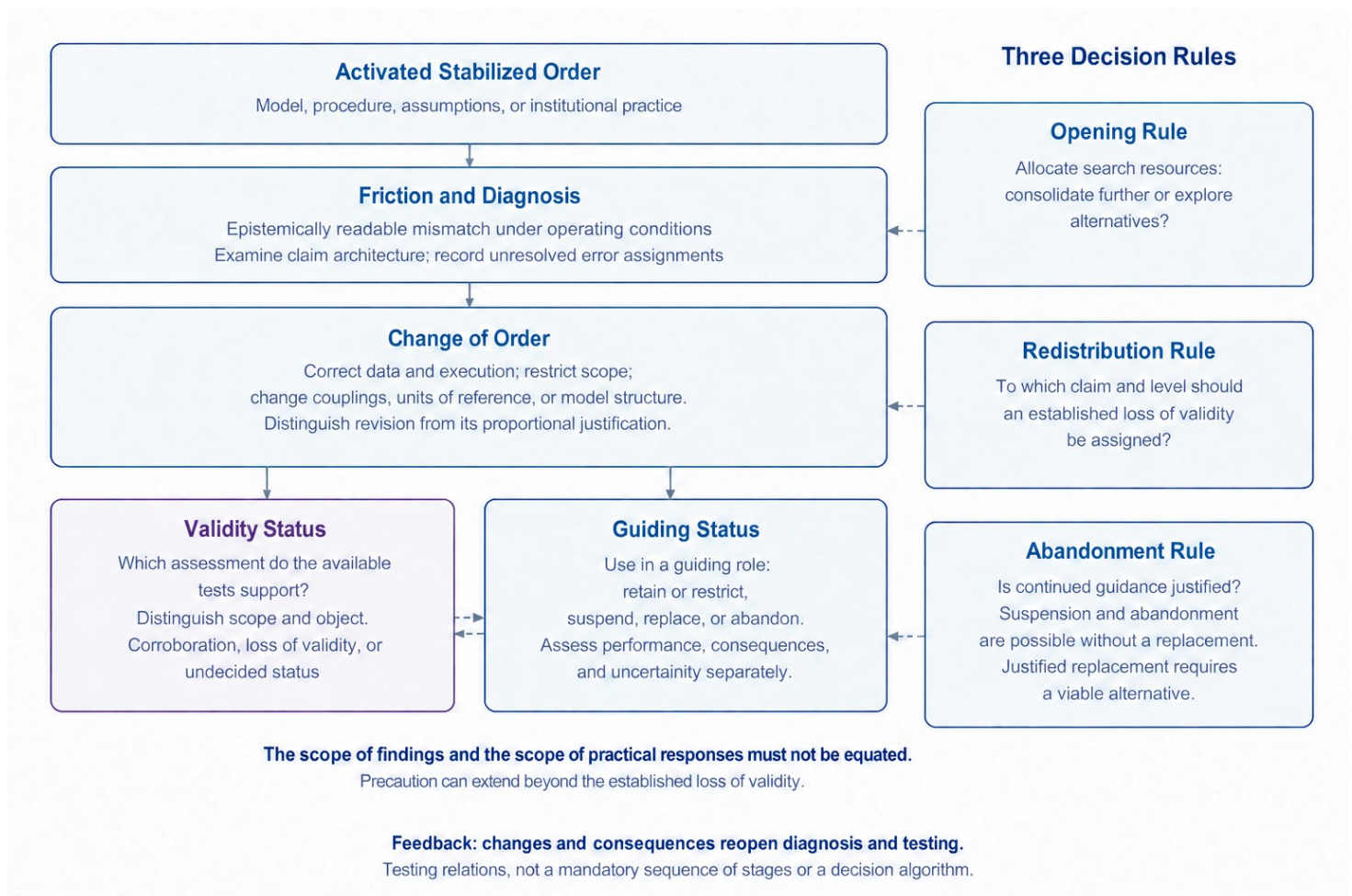


Figure 1: Diagnostic and decision architecture of model management. Validity status, justified change, and further use are examined jointly but not equated. The three rules fulfill different functions; suspension and abandonment do not presuppose a replacement model. The arrows designate testing relations and feedback, not a mandatory sequence of decisions.

This sequence is recursive and not an automated decision procedure. Responses change test conditions, generate costs, and can make new frictions visible. Assessing untested alternatives under uncertainty remains an open problem of finite search. The benefit of the architecture lies in making decisions decomposable, justifiable, and correctable.

Model validity does not automatically establish a warrant for action or intervention. Consequences, reversibility, normative reasons, and the capacity for feedback must additionally be examined. The scope of the established loss of validity limits what may be claimed about the model. The scope of practical precaution can differ: a limited finding can justify extensive suspension where possible harms are severe. Its reasons and conditions of reassessment must be specified separately. In complex orders, such examinations and decisions are distributed across persons, institutions, instruments, and procedures. The following chapter examines the conditions under which this distributed model management preserves intersubjective validity and correctability.

7. Shared Epistemic Reality

The architecture so far might suggest that model management is primarily a function performed by a single cognitive system. In fact, complex orders are formed, tested, and continued across persons, institutions, instruments, and procedures. Units of reference, models, and corrections do not remain individual. They are communicated, contested, stored, and taken up again under changing conditions.

Shared epistemic reality emerges through stable and correctable forms of shared reference.

It requires neither a uniform worldview nor general consensus. Dissent can show that a shared point of reference and a shared form of engagement persist. Conversely, consensus can be epistemically weak if stabilized through habit, power, exclusion, or the absence of alternatives. What matters is not sameness of results, but the possibility of continuing shared reference amid difference.

7.1 Shared Reference and Intersubjective Validity

A unit of reference is continued intersubjectively when several cognitive systems can relate their contributions to one another so that something is taken up again despite changing participants, situations, and descriptions. Complete agreement on what the unit is or how it should be assessed is unnecessary. A dispute does, however, presuppose that it remains sufficiently recognizable what contributions refer to, where the differences lie, and which contributions are relevant to their joint examination.

For a single cognitive system, such a shared order remains determinable only through its own forms of access and reconstruction. Within this reconstruction, it can distinguish its own position from positions attributed to others and from a presumed shared, societal, or institutional position. These attributions must not be equated with the intersubjectively distributed order itself. They remain correctable through statements, responses, documents, decisions, and other traces available to be taken up again.

Shared reference is supported by language and symbols, but not guaranteed by identical words. The same expression can designate different things; different expressions can address the same context. What matters is whether recognition, comparison, and correction remain possible.

Intersubjective orders also affect persons who do not agree with them. Legal categories, diagnoses, or institutional roles can structure expectations and decisions without individual endorsement. Being affected does not presuppose agreement. Their effects do not, however, already establish intersubjective validity.

Intersubjective validity designates correctable viability and requires at least three elements:

1. Availability for taking up again: contributions and deviations must leave traces that others can return to.
2. Articulability of difference: objections and competing interpretations must remain representable within the shared order.
3. Traceability of correction: changes must be related to understandable reasons, rules, or procedures.

These conditions guarantee neither truth nor ideal mutual understanding. They nevertheless distinguish intersubjective validity from mere enforcement or a formal simulation of correction in which objections are acknowledged but remain without consequences in principle.

7.2 Trust, Institutions, and Power

Shared model management would hardly be possible if every connection had to be secured entirely afresh. It therefore rests on trust:

Trust is the expectation that the processes of taking up, taking up again, and correction will not break off arbitrarily.

Trust means neither harmony nor the expectation of a particular result. It permits concepts, procedures, and roles to be used provisionally without retesting every presupposition. It thereby reduces effort while creating dependence on stored orders whose possible correction must be preserved.

Institutions stabilize such conditions of connection across situations and participants. They store not only contents, but regulate who can contribute, which forms of evidence and justification count, how disagreement is processed, what makes decisions binding, and under which conditions they can be examined again. They thus enable duration, scope, and division of labor.

Their epistemic authority rests not solely on status, title, or formal responsibility, but on the performance and correctness of their procedures. An order can be socially stable yet epistemically weak if relevant frictions become invisible, inconsequential, or inarticulable.

Power influences who defines problem boundaries and units of reference, who can contribute and make decisions, which contributions are documented, which evidence receives weight, and who bears the benefits and costs of a model order. Unequal roles are not automatically illegitimate or epistemically deficient. Expertise, division of labor, and responsibility generate necessary asymmetries. They become problematic when they can no longer be justified or corrected, or when relevant frictions experienced by those affected cannot enter the shared examination.

7.3 Social Stability and Loss of a Shared Basis for Correction

Shared epistemic reality does not disintegrate simply because dissent is strong. It can withstand conflicts as long as references can be continued, differences articulated, and corrections transparently justified.

A loss of a shared basis for correction occurs when there is no longer a sufficiently shared order within which a contrary indication can be taken up as a possible challenge and a correction pursued in a way others can follow.

Such a loss can proceed gradually and asymmetrically. Concepts can drift apart, sources be acknowledged only within separate groups, and objections treated as signs of insufficient belonging. Shared words and institutions can persist even though their conditions of connection have already diverged. Social stability can therefore outlast the loss of intersubjective validity.

Restoration does not necessarily require a new overall consensus. Limited shared points of reference, data, problems, translations, or paths of correction can suffice to reopen joint examination.

7.4 Interplay of the Domains and Science

The three domains interact within shared epistemic orders without forming a developmental ladder or hierarchy. Subjective orders contribute concern, experience, and individual differentiation; intersubjective orders enable shared continuation and store conditions of correction; functional-empirical orders challenge claims through repeatable relations of resistance and effect.

A community cannot determine by agreement how a substance behaves under particular conditions. It nevertheless helps determine which observations are repeated, which deviations documented, and which alternatives tested. Conversely, empirical robustness does not by itself decide meaning, institutional appropriateness, or normative warrant.

Science is an institutionalized form of shared epistemic reality. Empirical sciences additionally face functional-empirical testing and strain; in the formal sciences, conceptual-formal testing takes precedence.

Scientific validity therefore arises neither through consensus nor through institutional status alone. Science requires specialized institutions, methods, and responsibilities organized through a division of labor. Titles, status, or institutional affiliation do not themselves establish epistemic validity; the authority of these orders remains bound to the performance and correctability of their procedures.

7.5 Distributed Model Management

In complex orders, the decision architecture of Chapter 6 is enacted in a distributed manner. Different participants have different information, experience, responsibilities, and perspectives on consequences. The benefits and costs of a model order can also be unequally distributed.

Distributed model management therefore requires the examination not only of models but also of their social conditions of testing: Who can introduce relevant frictions? Which deviations are documented? Who specifies scope, test conditions, and criteria of failure? Which consequences return as feedback to the examining order, and which experiences, data, or ways of being affected remain outside its model boundary?

A distributed model order remains revisable if its units of reference, test conditions, and decisions can in principle be made visible and alterable again. Not every participant must survey all

presuppositions. Procedures must nevertheless exist through which latent orders can become visible, justifiable, and correctable again under strain.

Shared epistemic reality thus designates not a conflict-free community, but an order in which shared references can be taken up again, challenged, and revised despite finite perspectives, unequal roles, and continuing differences. The following chapter challenges the architecture developed so far in two methodologically different ways: first through a boundary case for epistemic critique concerning personal identity, and then through a comparative stress test against a strong disciplinary analysis of a real application case.

8. A Boundary Case and a Comparative Architecture Test

The following tests have different methodological statuses. Personal identity serves as a boundary case for epistemic critique. The vignette condenses arguments from the separate investigation of personal identity (Rapp 2026a) without replacing its detailed examination. It permits an examination of whether unit of reference, identity rule, claim to validity, and boundary rule can be distinguished so that the functional performance of a stabilized order does not coincide with a more extensive claim about its status.

The subsequent test using the Allegheny Family Screening Tool asks a stricter question. Here, plausibly organizing a complex case through the ERP architecture is insufficient. What is tested is whether its application provides a substantive additional benefit over a strong disciplinary baseline generated in separate LLM runs. The two tests therefore neither serve the same purpose nor possess the same evidential force.

8.1 Personal Identity: The Limit of Exclusive Continuation

Humans treat themselves as persons persisting over time. This self-stabilization enables memory, responsibility, planning, self-care, and social assignment. Forming the self as a unit of reference available to be taken up again is therefore not merely an error, but a necessary function of orientation.

A stronger claim arises when this practical viability is taken to imply that, at every later time, exactly one instance must exclusively be the numerical continuation of the currently experiencing person, and that this exclusivity must follow entirely from material or functional determinations.

The duplication case examines this transition and challenges the identity rule used: What happens if several later instances fulfill the relevant criteria equally? Branching and duplication cases in different forms are used by, among others, Williams and Parfit to test criteria of personal identity (cf. Williams 1970; Parfit 1984).

Suppose the structures relevant to memory, personality, behavior, and self-description were continued equivalently in two later instances. The case makes no claim about technical possibility. It asks only which identity criterion could sustain the required exclusivity under branching.

Qualitative, psychological, and functional sameness do not decide the question if several instances fulfill them equally. Organizational sameness is also insufficient insofar as the same organization can be realized more than once. Causal-historical or organismic continuity, by contrast, can privilege one line and is therefore a strong materialist candidate.

What claim this continuity fulfills must nevertheless still be clarified. It can establish which later person stands in the continued organismic or genealogical line of the earlier person. It does not automatically follow that it also determines the exclusive continuation of present experience. A viable identity rule and an additional metaphysical determination of exclusive identity must not be silently equated.

The friction thus concerns not the practical order of selfhood itself, but its absolutization. Functional sameness supports functional continuation; psychological continuity supports biographical

connectability; organismic continuity can single out a privileged process line. Whether any of these relations additionally determines the required exclusive subjective sameness remains a further claim.

It follows neither that materialism is false nor that personal identity is meaningless. The architecture identifies a possible underdetermination and distinguishes three directions of response. Personal continuation can be managed relationally or gradually. Causal-organismic continuity can be defended as a privileged identity rule provided that its precise claim to validity is specified. Alternatively, an additional metaphysical determination of exclusive identity can be assumed; it must then be marked as not reconstructed by the material-functional determinations used and as not counting as an established result within the present reconstruction.

The diagnostic benefit lies in distinguishing a functionally indispensable unit of reference from an additional ultimate ontological claim. The boundary rule prohibits neither self-stabilization nor organismic identity. It merely prevents the scope of viable criteria from concealing a determination that has not been accomplished.

8.2 AFST: Comparative Stress Test of the Model Management Architecture

A stronger test of the architecture must distinguish mere conceptual reordering from substantive additional performance. The use of the Allegheny Family Screening Tool in child welfare was therefore selected as the test case. What mattered was not whether the case could be described using the architecture's concepts, but whether its application generated additional tests, different model restrictions, changed institutional measures, or different consequences for action.

The test was conducted within the project in separate LLM work sessions using ChatGPT and Claude; the analyses and judges' assessments are not external expert reviews. The basis was the frozen master-paper test version revised on 28 July 2026. A case dossier was first compiled from six sources specified in advance; in a separate run, a theory-neutral framework with twelve criteria was created before the analyses were compared. Three disciplinary analyses (B1 and B3: ChatGPT; B2: Claude) were combined into the strongest integrated baseline. Two applications of the unchanged ERP architecture (E1: ChatGPT; E2: Claude) were consolidated into an ERP steelman. According to the task instructions, baseline runs received only the case dossier, while ERP runs additionally received the frozen architecture, but no comparative analyses; information from earlier conversations was excluded. The two judging runs J1 (ChatGPT) and J2 (Claude) assessed the combined analyses using the same framework. J1 received the baseline first, J2 the ERP analysis first. A difference counted only if it entailed an additional concrete test, model restriction, institutional change, or decision about action; specialized vocabulary and length were not to be assessed in themselves (Rapp 2026b, protocols 04b–04k). J1 refined its initial judgment following a follow-up request for renewed examination; the reported results concern the revised overall assessment. This reassessment within the same run is not an independent replication.

The fixed source basis comprised the methodology report for Version 2 (Vaithianathan et al. 2019), the official deployment description in the archived version of 6 August 2026 (Allegheny County DHS 2026), the Phase 2 evaluation (Goldhaber-Fiebert and Prince 2023), the studies by Cheng et al. (2022) and Cheng and Chouldechova (2022), and the critical model analysis by Gerchick et al. (2023). They concern different model states, periods, and decision levels; the technical production

state in 2026 cannot be fully reconstructed from them. The selection rule, source files, case dossier, task instructions, analyses, and assessments are stored in the project's internal test documentation "Test-Vignette-3" (Rapp 2026b). The task instructions, comparative analyses, assessment framework, and judges' assessments relevant to the reported comparison are available from the author on request.

Separating the work sessions limits mutual knowledge of results, but does not establish independence from shared LLM, dossier, or evaluative assumptions in the sense of §6.1. Blinding remained partial because ERP vocabulary was recognizable. The change in presentation order coincided with a change of judge and therefore does not isolate an order effect. Furthermore, differently structured task instructions and three versus two initial analyses were compared; the test therefore does not isolate a causal effect of individual architecture components. ChatGPT runs B1, B3, E1, and J1 used GPT-5.6 Sol; Claude runs B2, E2, and J2 used Claude Opus 5. The finding concerns the documented comparison of analyses, not an effect of the proposed measures tested in child welfare practice.

The result first limits a strong claim to distinctive contribution. Both LLM judges attributed a strong, nonredundant additional benefit to the disciplinary baseline in concrete disciplinary tests and the options for use derived from them (Rapp 2026b, J1, revised overall result, §8; J2, §4). It generated additional tests concerning, among other things, actual model performance and error distribution, calibration, selection mechanisms, causal effects, drift, and institutional safeguards and auditing measures. The ERP architecture does not generate such disciplinary procedures by itself. The test therefore does not support the claim that ERP analyzes complex cases better overall than a strong combination of established disciplinary methods.

Both LLM judges also identified specific additional contributions of the ERP analysis which they classified as substantive under the predefined framework. These primarily concerned control of the transitions between unit of reference, model version, claim, test condition, strain, scope of findings, and justification of revision. Differences between the submitted analyses and the judges' assessments of them are thereby documented; neither independent disciplinary confirmation of their superiority nor practical usefulness has been established.

This was particularly apparent in the historical changes to the AFST. A name that remains unchanged over time does not guarantee that findings about earlier model, interface, staffing, or policy states may be transferred unchanged to a later production state. The ERP analysis therefore requires a specified bridge of identity or continuity before historical findings are transferred to a changed unit of reference. J2 judged this rule a substantive difference: parts of the disciplinary baseline applied older findings more extensively than their tested scope supported; measures based on them would first have to be made conditional on an examination of the current state (Rapp 2026b, J2, §§2–4). The continuity requirement applies symmetrically to challenging and supporting findings. A version change neither automatically invalidates earlier criticism nor transfers earlier performance evidence to the new state. What matters is the continuation of the properties or mechanisms relevant to the respective finding.

A second difference highlighted by both judging runs concerned test conditions (Rapp 2026b, J1, §§5 and 8; J2, §§2–4). The architecture requires that, where possible before testing, it be specified which claim is to count as challenged under which conditions. Without such conditions of failure,

assessments such as sufficient calibration, acceptable error distribution, or good fairness performance can remain unspecified and be refined only after the result is known.

The assessments thus emphasize a closer assignment between finding and claim. The narrow-response rule also assessed positively there nevertheless requires clarification: an error must not be generalized beyond what the evidence supports; the depth of restructuring additionally depends on which supporting presuppositions are affected. Practical precaution is further subject to considerations of consequences and uncertainty. Extensive suspension can therefore be justified even while the scope of the loss of validity remains open (§§6.6–6.7). This clarification belongs to the present revision, not to the result of the historical test.

The test suggests a complementary distribution of functions. However, a disciplinary analysis supplemented with ERP rules was not examined as a separate test arm; its expected integration benefit is an inference from the assessments, not a measured combination effect. The baseline specified investigations, metrics, causal designs, and institutional measures more concretely. The ERP contributions highlighted by the judges primarily concerned claim specification, model identity, transfer of findings, and justification of revision. Whether these contributions could also be achieved through a short general testing checklist with less effort remains open.

The AFST test therefore supports neither a general superiority claim nor a blanket classification of the submitted ERP analysis as redundant. It documents additional contributions assessed as substantive by LLM judges within a limited comparison of analyses. Their soundness by disciplinary standards, generalizability, and practical usefulness require further testing. The test also identifies a limit of the program: the architecture does not itself generate the required disciplinary testing procedures.

9. Self-Limitation

The program developed here assigns claims to validity to their conditions and scope, reconstructs the stabilization of units of reference and models, makes frictions diagnostically readable, and prepares proportional correction and revision. This combination does not constitute a complete epistemology. The diagnostic and decision architecture is itself a model under finite conditions.

It, too, selects problems, stabilizes concepts, and brackets differences. Its application generates effort, can overlay disciplinary descriptions, and can itself be overextended or immunized. It must therefore be assessed in terms of visible presuppositions, limited validity, additional diagnostic benefit, and revisability.

9.1 Scope and the Limit Drawn by Epistemic Critique

The guiding principle of epistemic critique is:

Clarifying limits without ultimate justification.

The boundary rule can reveal when positive determinations are extended beyond the range in which their transitions and conditions of validity are specified. This does not establish a positive ultimate ontology. Dependence on acts of determination, too, shows only that an act of determination cannot completely negate itself without being performed again. It does not decide what mode of being cognition, experience, or reality ultimately possesses.

Resistant and correctable orders can therefore be treated realistically without presenting their viability as presuppositionless access to a fully determinable world in itself. “Realism without an Outside” designates a methodological restriction of scope, not a metaphysical position.

A second guiding principle is:

Unity through structure rather than through being.

Connections can be reconstructed across different orders of description and validity through references, transitions, and couplings, without everywhere presupposing a uniform ontological foundation. Such connections are viable only insofar as they themselves remain testable. Similarity, recurring patterns, or identical concepts are insufficient.

Structure, too, must not become a universal metaphysical answer. The domain triad is a reconstructive working hypothesis. If another differentiation captures forms of strain and correction more precisely, the triad must be open to restriction, supplementation, or replacement. The concepts and distinctions of this architecture are also analytical instruments, not necessary components of every possible reality. Their claim depends on whether they make differences visible, refine tests, prevent misdiagnoses, and improve the proportionality of responses.

9.2 Limited Coordinating Function

The third guiding principle is:

Coordination rather than decision.

The architecture can examine claim architectures, clarify validity mappings, investigate transitions, locate frictions, and organize possible responses. As a decision architecture, it structures justifications and options without replacing the concrete judgment. It provides neither disciplinary data and explanations nor ultimate normative justifications, political or moral decisions, or answers to ultimate why-questions.

Model validity does not automatically establish intervention or normative warrant. Nor does the decision architecture replace domain-specific and case-specific judgment. What counts as substantial friction, disproportionate costs, a sufficiently viable alternative, or an independently justified tolerance range must be determined within the relevant disciplinary and practical orders. The assessment of untested alternatives under uncertainty remains open.

The program can structure conflicts, but cannot decide on behalf of the disciplinary, practical, or normative orders concerned. Its operative aim is to make unmarked transitions, overextended validity, and disproportionate responses more recognizable.

9.3 Self-Application and Non-Applicability

The central question of self-application is:

What additional test, model restriction, decision, or revision does the epistemic diagnosis enable compared with an appropriate alternative, or make available more reliably, more easily, or with less effort?

Conceptual reordering alone does not establish an additional operative benefit. Such a benefit can nevertheless also lie in selecting and assembling existing distinctions, operationalizing them as diagnostic questions and testing rules, and making them jointly available and appropriately activated for the situation. Insights already scattered across the literature are therefore not automatically an equivalent tool. Whether the architecture makes relevant tests easier to find, more reliably applicable, or less susceptible to omission must be investigated against appropriate alternatives.

Additional benefit is absent or limited if a simpler and comparably accessible disciplinary description enables the same analytical performance equally well, opposing results can retrospectively be classified equally well under the project's concepts, or no observation, alternative reconstruction, or application can seriously count against a diagnosis. It would also be problematic to adjust domains, criteria, or assignments of error without marking the change only after the result is known, or to absorb every contrary indication as confirmation of the architecture.

Subsequent changes remain permissible, but must be identified as revision and tested again under changed conditions. The program can lose additional benefit, become conceptually imprecise or redundant, be surpassed by a simpler differentiation, or remain practically inconsequential. These possibilities constitute not an additional typology of falsification, but conditions of its restriction, revision, or non-applicability.

Self-application therefore also includes examining the architecture's own costs. Conceptual, familiarization, search, application, maintenance, and translation effort are justified only insofar as they are matched by corresponding additional epistemic benefit over simpler descriptions or already

established procedures. The latter's actual accessibility and operationalization must also be considered in the comparison. A more differentiated architecture is not preferable simply because it makes more distinctions available.

Non-applicability is not a deficiency to be avoided at any cost. A disciplinary description or simpler distinction may suffice; some problems are not problems of model management or are not yet sufficiently determined. The architecture must allow cases in which it provides no relevant additional benefit. Claiming additional benefit everywhere would itself be an overextension.

9.4 State of Testing and Corroboration

The extent of the architecture's testing and corroboration to date must be differentiated. The personal identity boundary case shows that central distinctions of the architecture can be applied to a problem of epistemic critique and can distinguish unit of reference, identity rule, claim to validity, and possible overextension within it. Such an application demonstrates diagnostic usability, but does not itself establish additional benefit over a strong competing analysis.

The internal, LLM-assisted comparison using the Allegheny Family Screening Tool goes beyond illustrative application. The frozen ERP architecture was compared with a disciplinary baseline generated in separate work sessions; the assessment framework also permitted the absence of additional benefit. Both judges attributed a strong, nonredundant additional benefit to the baseline in concrete disciplinary tests and options for use, while also identifying limited additional contributions of the ERP analysis classified as substantive under the framework. Their agreement concerns this documented comparison and is not an independent disciplinary replication (Rapp 2026b; test design and limits in §8.2).

A general superiority claim is therefore unsupported. In particular, the architecture generated fewer concrete disciplinary tests and institutional options for action than the strong baseline. This finding should not be treated as a mere implementation gap to be closed by incorporating ever more specialized methods. Such an expansion would transform the model management architecture into a collection of individual disciplinary procedures and shift its object.

The positive finding is narrower. The judges particularly emphasized specification of claims and units of reference, model identity and historical scope, advance conditions of failure, and justification of revision. These distinctions can limit the untested transfer of earlier findings and the generalization of local strains. Whether they produce better disciplinary decisions and justify their additional effort has not been independently confirmed. The appropriate scope of practical precaution must also be assessed.

Existing tools must be considered as functional comparators in this examination. Model Cards document, among other things, model version, intended and out-of-scope uses, evaluation conditions, and performance differences across relevant groups (Mitchell et al. 2019, §4). The model risk management guidance in SR 11-7 already links model testing to assessment of use, documentation of limitations, ongoing validation, critical review, and possible restrictions on use (Federal Reserve and OCC 2011, Sections IV–VI). This historical version of the guidance is cited here as a methodological comparator. Specifying the limits of claims, versioning, and distinguishing testing from use are therefore not functions exclusive to ERP.

The potential distinctive contribution lies in making such functions available, together with explicit examination of units of reference, identity conditions, transitions between orders of validity, and justifications of responses, across different contexts of modeling and practice. This is a claim about usefulness as a tool, not evidence that existing procedures lack these functions. A concrete comparison must show which question, attribution of error, or response ERP helps to justify additionally or more reliably, and whether that benefit warrants its costs. Where an established tool already provides the same performance, this must not be counted as an additional benefit of ERP. Greater generality or more concepts do not by themselves establish superiority.

The comparison to date thus establishes a limited testing task: as a supplementary architecture, ERP is intended to make the use, transfer, and revision of specialized models and findings more controllable. Further applications should compare a disciplinary analysis without supplementation, with an established testing or documentation tool selected independently of ERP and suited to the task, and with ERP. A short general testing checklist may provide a further comparison condition, but does not substitute for a strong established alternative. The provenance, composition, and operationalization of the comparison tools must be specified: a short checklist derived from ERP may condense its performance and is not an independent alternative merely because it is short. Additional contributions with a sound disciplinary basis, misjudgments, more reliable activation of relevant tests, and the full familiarization, search, and application effort must be examined. An improvement in practical outcomes additionally requires investigations of actual use.

The architecture has thus undergone comparative testing, but has not yet been independently corroborated within a discipline or in practical use. The LLM comparison concerns the frozen test version of 28 July 2026. Revisions made after this test distinguish, in particular, abandonment from replacement, revision from its proportional justification, and the scope of findings from the scope of practical responses; they reorganize the falsification distinctions and refine the methodological claim to usefulness as a tool. These changes were not tested in the reported comparison and do not automatically inherit its testing status.

9.5 Assessment

Finite cognition cannot dispense with stabilization. The same orders that enable orientation can nevertheless conceal their presuppositions, omissions, and costs. Viable model management therefore requires neither maximum stability nor permanent revision. It requires limited validity, readable friction, and proportional change.

The program promises no final resolution of epistemic conflicts. Its contribution lies in an additional ordering of model management through which claim, unit of reference, conditions of validity, scope, and revision can be controlled jointly. The LLM-assisted comparison provides indications of specific additional contributions in this function. Whether they provide independently confirmed disciplinary and practically effective benefits over simpler procedures remains to be tested.

Its goal remains the preservation of correctability:

Clarifying limits without ultimate justification.

Unity through structure rather than through being.

Coordination rather than decision.

The project makes conflicts readable and their examination more controllable; it does not decide them on behalf of the orders within which they arise.

Conceptual Canon of This Paper

The following conceptual canon stabilizes the concepts, distinctions, and rules that support this paper. It does not claim to capture all concepts of the Epistemic Reality Project.

The canon is an explicitly stabilized but revisable reference basis. Subsequent changes, refinements, or extensions remain possible, but must be explicitly identified, justified, and made clear in their scope. Implicit shifts of meaning, silent extensions, and retrospective reinterpretations of earlier project texts are excluded.

Canonical Status and Relation to Earlier Canons

This paper consolidates and revises the previous basic canon of Epistemics and several local conceptual developments in later project texts. Upon its approval, the canon presented here becomes the programmatic master canon for new and subsequently revised project texts.

The previous basic canon of Epistemics remains documented as a historical stage of development. The local canons of existing module and application papers are not retrospectively changed either. They retain their published historical or local status until explicitly revised to connect them to the new master canon.

The omission of a previously used concept from this canon means neither its substantive rejection nor the loss of a possible local, historical, or exploratory status. It initially means only that this paper does not assign it project-wide master status.

Core Master Concepts

Epistemics

Brief definition: Programmatic designation for the investigation of model management under finite conditions.

Function: Designates the working context of epistemic critique developed within ERP in which stabilization, validity, friction, and revision of orders that can be managed as models are examined.

Distinction: Not a separate theory alongside the model management architecture and not identical to the Epistemic Reality Project as a whole.

Finitude

Brief definition: Structural limitation of cognitive systems with respect to access, processing, coordination, and overview of the consequences of their own orders.

Function: Explains why cognition must select, stabilize, form models, limit claims, and enable revision.

Distinction: Not merely a deficit of individual persons or institutions, and no assertion of identical limits for all cognitive systems.

Cognitive System

Brief definition: A functional context capable of taking up differences, stabilizing them, relating them to claims, testing them under strain, and correcting them.

Function: Designates the general bearer of the cognitive and model management achievements reconstructed in this paper.

Distinction: Not necessarily an individual, biological, or experiencing subject. Collective, technical, or institutionally distributed systems can also be examined as cognitive systems.

Orientation

Brief definition: Selective ordering of differences, expectations, and possible continuations under finite conditions.

Function: Enables comparison, expectation, action, and coordination before a complete overview of presuppositions and consequences is available.

Distinction: Neither mere subjective opinion nor already fully tested cognition.

Stabilization

Brief definition: Selective continuation through which particular features, differences, and connections are kept available to be taken up again while other changes are permitted or set aside.

Function: Enables identification, comparison, expectation, reference, and model formation.

Distinction: Neither immutable fixation nor a substitute for truth, and not necessarily already the formation of a unit of reference or a model.

Unit of Reference

Brief definition: Something stabilized under particular conditions as identifiable, available to be taken up again, and referable.

Function: Enables properties, changes, observations, and claims to be related to something and compared across changing conditions.

Distinction: Neither a presuppositionless given thing nor a mere name, and not yet a model.

Model

Brief definition: A sufficiently determined order connecting units of reference, features, relations, expectations, and continuation rules so that the order can be marked as such, taken up again, delimited, compared, tested, and corrected.

Function: Orders relevant differences, generates expectations, and enables robust orientation and testing.

Distinction: Neither immediate representation of a fully accessible reality nor arbitrary invention. Not every stabilization or unit of reference already constitutes a model.

Model Management

Brief definition: Justified retention, restriction, correction, recoupling, reconfiguration, and revision of models and their claims to validity, together with suspension, replacement, or abandonment of their use in a guiding role.

Function: Clarifies which claim remains viable under which conditions, which level is challenged by friction, and which response corresponds to the finding.

Distinction: Not restricted to model change and not a general decision algorithm.

Claim Architecture

Brief definition: The combination of unit of reference, model, claim type, domain, mode of challenge, scope, test conditions, couplings, and continuation rules.

Function: Makes it possible to test which component of a claim to validity is actually challenged and whether an order is misassigned, overextended, or shielded against contrary indications.

Distinction: Not equivalent to a model's internal structure alone. It also includes conditions of activation, application, transition, and testing. Purpose and evaluative standards must be explicitly included as specifications of the claim made and its test conditions; their warrant requires its own reasons. Conditions of failure should, where possible, be specified in advance. Separate execution or agreement of results does not guarantee independence of tests; additional testing power depends on differences in relevant presuppositions, access to data, error mechanisms, or evaluative assumptions.

Validity

Brief definition: The domain-bound viability of a model or claim under strain.

Function: Specifies whether and within what scope a claim can be continued under the conditions relevant to it.

Distinction: Neither mere agreement, social enforcement, nor usefulness, and no undifferentiated substitute for truth. Validity status is a fallible assessment based on available tests; an undecided status constitutes neither corroboration nor an established loss of validity.

Domain

Brief definition: A functional space of ordering in which claims are stabilized, tested, challenged, and corrected in characteristic ways.

Function: Organizes conditions of validity and prevents the uncontrolled transfer of local viability to differently structured contexts.

Distinction: Not an ontological region of the world, an independent substance, or an unambiguous assignment of objects to separate parts of reality.

Costs

Brief definition: Effort, commitment, or restriction associated with stabilization, model management, testing, coupling, correction, or transition.

Function: Makes visible which resources an order requires, which differences it obscures, and which strains it may shift elsewhere.

Distinction: Not restricted to money, energy, or quantitative measurability. Increasing costs do not automatically refute a model.

Friction

Brief definition: An epistemically readable mismatch of an activated stabilized order under its operating conditions; it may become apparent in expectations, connections, or the model-based organization of that order.

Function: Indicates a need for diagnosis by indicating that an intended continuation is not viable under concrete conditions.

Distinction: Not identical to resistance, irritation, disturbance, error, anomaly, falsification, or revision. Friction does not itself already determine its cause or the appropriate response.

Revision

Brief definition: Restructuring, under epistemic pressure, of a model, an order that can be managed as a model, or a claim to validity.

Function: Under epistemic pressure, changes the order that can be managed as a model. Its justification must be examined separately in terms of the scope and structural significance of the strain and the costs and consequences of restructuring.

Distinction: Not every change, data update, reorganization, scope restriction, or reconfiguration is already revision. Proportionality is a quality criterion; even an excessive revision remains a revision.

Foundational and Boundary Concepts

Dependence on Acts of Determination

Brief definition: Positive determinations are accessible to a cognitive system only within its respective acts of determination.

Function: Limits the claim that the conditions of cognition can be positively determined from a completely presuppositionless standpoint.

Distinction: Establishes neither the subjective production of reality nor a positive ontology of the cognitive system or the act of determination.

Inner Form and Outer Form

Brief definition: Functional consolidations of different profiles of variability, availability, dependence, recurrence, and resistance.

Function: Order determinations according to how they become available for further operations, constrain expectations, and compel corrections.

Distinction: Neither originally separate ontological regions nor a closed subjective inner space nor a fully positively determinable outside in itself. Their function within the forms of finite orientation reconstructed here does not establish the universal necessity of precisely this differentiation for every possible cognitive system.

Perspectivity

Brief definition: The ordered dependence of determinations on a cognitive system's possibilities for access, processing, distinction, and action.

Function: Makes visible that determinations are system-bound without thereby becoming arbitrary.

Distinction: Not an exclusively human or biological property and no proof that resistance is only subjectively produced.

Boundary Rule

Brief definition: What can be marked only as a limit of positive determinability must not simultaneously be treated as a positively determined region beyond that limit.

Function: Prevents the transition from viable, correctable determinations to an ultimate claim about a fully accessible region outside all conditions of cognition.

Distinction: Opposes neither empirical research nor limited realist discourse.

Ontologization

Brief definition: A particular form of stabilization through which something is formed as an identifiable, referable unit of reference available to be taken up again.

Function: Enables returning to something, attributing properties to it, tracking changes, and forming models.

Distinction: Not a complete ontology and not in itself a problematic pattern.

Absolutization

Brief definition: Treating a stabilized unit of reference, order, or model determination as immediate, exhaustive, or without alternatives.

Function: Identifies the problematic pattern in which a stabilization's presuppositions, selections, scope, and conditions of revision become invisible.

Distinction: Not every stabilization, ontologization, or realist reference is absolutization.

Transition Function

Brief definition: A coupling through which determinations become connectable across different orders of validity.

Function: Enables units of reference and determinations to be continued across domains without eliminating the differences between their conditions of validity.

Distinction: Neither identical, lossless, nor automatically valid transfer.

Canonical Distinctions and Rules

Validity Mapping

Brief definition: Prior clarification of which claim can meaningfully be tested within which order and under which conditions.

Function: Prevents local viability from being transferred unnoticed to another claim type, domain, or scope.

Distinction: Not yet a decision about whether the assigned claim is actually valid.

Domain and Mode of Challenge

Brief definition: The domain designates the functional space of ordering to which a claim belongs; the mode of challenge specifies how it is tested or otherwise subjected to strain.

Function: Distinguishes what is claimed, the order in which the claim is to be viable, and how it is tested or challenged.

Distinction: Claim type, domain, and mode of challenge are related but distinct specifications. The domain is determined neither by the linguistic form of the claim alone nor by the testing procedure.

Modes of challenge: connection to experience; dialogical justification and taking up again; observation, measurement, and intervention; conceptual-formal testing.

Domain Triad

Brief definition: Reconstructive distinction between subjective, intersubjective, and functional-empirical orders of validity.

Function: Makes different forms of stabilization, testing, strain, and correction visible.

Distinction: Neither a final division of reality nor a developmental ladder nor a necessary internal modular structure of every cognitive system. The triad is revisable; different conditions of access and testing initially establish epistemic asymmetries, not ontological separation or irreducibility in principle. Any additional benefit it offers beyond a sufficiently precise specification of claim type and mode of challenge must be examined separately.

Identity Rule and Continuation Rule

Brief definition: An identity rule specifies under which changes a unit of reference is epistemically continued as the same; a continuation rule specifies how relations, expectations, operations, or further model steps proceed.

Function: Distinguishes taking up a unit of reference again from the further management of a modeled order.

Distinction: The two rules can interact but are not identical. A continuing model name does not guarantee model identity relevant to testing; transferring supporting or challenging findings between model versions requires continuity of the properties or mechanisms relevant to the respective finding.

Tentative Stabilization

Brief definition: Provisional ordering of a field not yet sufficiently clarified, through which possible relations, consequences, and strains become visible.

Function: Connects open exploration to an already sufficiently determined working structure.

Distinction: Not yet a conclusively viable claim to validity and neither a canonical nor an ontological commitment.

Overextension

Brief definition: Continuing a model or claim beyond the range in which its presuppositions and performance are sufficiently viable.

Function: Diagnoses problems of scope without rejecting the model across its entire range of application.

Distinction: Neither a mere execution error nor automatic global falsification.

Immunization

Brief definition: Defensive stabilization of a claim to validity in which hardly any possible result can still seriously challenge the model.

Function: Marks the boundary between legitimate refinement or provisional retention and systematic shielding against contrary indications.

Distinction: Determined not by the mere number of auxiliary assumptions, but by the loss of genuine conditions of failure.

Principle of Proportionality

Brief definition: Change should be as narrow as possible, but as extensive as necessary.

Function: Relates the depth of intervention of a response to the diagnosed scope and structure of the strain.

Distinction: Neither a rigid sequence of stages nor a preference for local repairs at any cost.

Practical precaution can extend beyond the established loss of validity in light of consequences and uncertainty.

Recoupling and Reconfiguration

Brief definition: Recoupling changes the connection between models, units of reference, or orders of validity that remain viable; reconfiguration changes what is available for epistemic processing and how it is structured.

Function: Enables responses to problems of coupling and ordering without prematurely abandoning all models involved.

Distinction: Neither is automatically revision, but either may develop into a revision.

Falsification

Brief definition: The domain-relative loss of validity of a sufficiently determined model or claim under fair, suitable functional-empirical testing.

Function: Specifies the scope within which a claim is no longer viable under fair, domain-appropriate test conditions.

Distinction: Explicit claim conditions, test conditions, and conditions of failure are required. Not every friction or loss of validity is falsification, and falsification does not automatically mean model abandonment. Extension to other modes of challenge remains open.

Scope and object: Contextual falsification designates loss of validity in an application, subdomain, or under particular conditions. Global falsification designates the persistent loss of the general claim to validity within the claimed range above an independently justified tolerance range. Independently of this, architectural loss of validity designates the challenge to a model–domain assignment or a connection between claims. It is not a third level of scope and counts as falsification only under suitable functional-empirical testing.

Validity Status and Guiding Status

Brief definition: Validity status designates the assessment of the viability of a model or claim justified by available tests; guiding status designates its role within further orientation and model use.

Function: Distinguishes establishing a loss of validity from asking whether a model's use in a guiding role should be retained, restricted, suspended, replaced, or abandoned.

Distinction: A globally falsified model can continue to be used in a limited way under independent justification without being epistemically rehabilitated. The absence of a replacement is insufficient; suspension and abandonment are possible without a replacement.

Opening Rule

Brief definition: A rule for weighting further consolidation of an existing model against the opening of alternative search spaces.

Function: Shifts resources toward exploration when several reasons combine or one is sufficiently strong and diagnostically relevant: stagnating gains in robustness, disproportionately increasing consolidation costs, persistent relevant friction, or a serious available or developable alternative.

Distinction: Establishes neither the falsity of the existing model nor the superiority of an alternative.

Redistribution Rule

Brief definition: The loss of validity established by a test should be assigned to the level, or connection between levels, that was actually challenged.

Function: Prevents both premature model destruction and the persistent displacement of consolidated strain onto peripheral errors.

Distinction: Decides neither search resources nor the practical abandonment of a model.

Abandonment Rule

Brief definition: An existing claim to a guiding role must be restricted, suspended, or abandoned insofar as its continued exercise is not sufficiently justified under the specified epistemic and practical conditions.

Function: Examines the justification of use in a guiding role in light of performance, consequences, and uncertainty. For replacement, the alternative's viability, maturity, and transition costs must be compared.

Distinction: Suspension temporarily sets aside a claim to a guiding role; abandonment ends it; replacement transfers the role to another order. Only justified replacement requires a sufficiently viable alternative. No decision about guidance retrospectively determines validity status.

Intersubjective and Distributed Model Management

Intersubjective Validity

Brief definition: The correctable viability of a claim within shared reference, taking up again, and critical exchange.

Function: Enables claims to be continued, disputed, and corrected among several cognitive systems.

Distinction: Not mere consensus, social stability, institutional status, or functional-empirical robustness.

Minimum conditions: availability of contributions and deviations to be taken up again; articulability of difference; traceability of correction.

Shared Epistemic Reality

Brief definition: A stable and correctable form of shared reference.

Function: Enables units of reference, expectations, and corrections to be continued across changing participants, situations, and descriptions.

Distinction: Neither a conflict-free community nor a uniform worldview nor a necessary overall consensus.

Trust

Brief definition: The expectation that the processes of taking up, taking up again, and correction will not break off arbitrarily.

Function: Reduces the effort of continuing shared testing and enables model management through a division of labor.

Distinction: Neither harmony nor unconditional agreement nor a guarantee of a particular result.

Loss of a Shared Basis for Correction

Brief definition: A state in which there is no longer a sufficiently shared order within which a contrary indication can be taken up as a possible challenge and a correction pursued in a way others can follow.

Function: Describes the loss of shared possibilities for epistemic processing despite persisting concepts, institutions, or social stability.

Distinction: Strong dissent or a loss of trust alone does not establish such a loss.

Distributed Model Management

Brief definition: Formation, testing, continuation, and revision of models across several persons, systems, institutions, instruments, and procedures.

Function: Applies the diagnostic and decision architecture to cognitive orders organized through a division of labor and makes their social conditions of testing visible.

Distinction: Not merely the addition of individual cognitive achievements.

Concepts Not Recanonized across the Project

Earlier project texts use further concepts, including orientation structure, stopping structure, relational graspability, connection structure (Anschlusstruktur), modelability, cost blindness, apparent stability, malfunction, stability space, boundary signal, cost profile, friction competence, and various concepts of contingency mechanics.

This canon neither adopts nor rejects these concepts wholesale. Until a separate canonical decision, they retain their historical, local, metaphorical, or exploratory status. Their later inclusion in the master canon presupposes demonstrable additional functional benefit and an explicit canonical decision.

Canonical Status and Scope of Validity

The present canon forms the programmatic reference basis of this paper and, upon its approval, the project-wide master canon for new and subsequently revised project texts.

Local module and application concepts may supplement it or operationalize it more narrowly for specified purposes. They may not, however, silently redefine its master concepts, distinctions, rules, or typologies.

Every subsequent change to the master canon must be explicitly identified, justified, and examined for its consequences for existing project texts. Earlier publications remain interpretable within their historical conceptual state; their future revision is guided by the master canon stabilized here.

References

- Alchourrón, Carlos E., Peter Gärdenfors, and David Makinson. 1985. “On the Logic of Theory Change: Partial Meet Contraction and Revision Functions.” *The Journal of Symbolic Logic* 50 (2): 510–530. <https://doi.org/10.2307/2274239>.
- Allegheny County Department of Human Services (DHS). 2026. “Allegheny Family Screening Tool.” Official deployment description. Archived webpage version of August 6, 2026. Source document 03b of the test documentation (Rapp 2026b).
- Cheng, Hao-Fei, et al. 2022. “How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions.” In *CHI Conference on Human Factors in Computing Systems (CHI ’22)*. 22 pp. <https://doi.org/10.1145/3491102.3501831>.
- Cheng, Lingwei, and Alexandra Chouldechova. 2022. “Heterogeneity in Algorithm-Assisted Decision-Making: A Case Study in Child Abuse Hotline Screening.” Preprint, arXiv:2204.05478v1, April 12, 2022. Identified as a CSCW 2022 preprint.
- Duhem, Pierre. 1906/1954. *The Aim and Structure of Physical Theory*. Translated by Philip P. Wiener. Princeton, NJ: Princeton University Press.
- Earman, John. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- Federal Reserve and Office of the Comptroller of the Currency (OCC). 2011. *Supervisory Guidance on Model Risk Management*. Attachment to SR 11-7, April 4, 2011. Used as a historical methodological comparator. <https://www.federalreserve.gov/boarddocs/srletters/2011/sr1107a1.pdf>.
- Gerchick, Marissa, et al. 2023. “The Devil is in the Details: Interrogating Values Embedded in the Allegheny Family Screening Tool.” In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’23)*. <https://doi.org/10.1145/3593013.3594081>.
- Giere, Ronald N. 2006. *Scientific Perspectivism*. Chicago: University of Chicago Press.
- Goldhaber-Fiebert, Jeremy D., and Lea Prince. 2023. *Impact Evaluation of a Predictive Risk Modeling Tool for Allegheny County (Phase 2)*. Stanford University. Report dated March 26, 2023, provided by Allegheny County DHS.
- Goldman, Alvin I. 1999. *Knowledge in a Social World*. Oxford: Clarendon Press.
- Lakatos, Imre. 1970. “Falsification and the Methodology of Scientific Research Programmes.” In *Criticism and the Growth of Knowledge*, edited by Imre Lakatos and Alan Musgrave, 91–196. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139171434.009>.
- Laudan, Larry. 1977. *Progress and Its Problems: Towards a Theory of Scientific Growth*. Berkeley/Los Angeles: University of California Press.
- Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press.

- Mitchell, Margaret, et al. 2019. "Model Cards for Model Reporting." In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 220–229. <https://doi.org/10.1145/3287560.3287596>.
- Morgan, Mary S., and Margaret Morrison, eds. 1999. *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Popper, Karl R. 1935. *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*. Vienna: Julius Springer.
- Quine, Willard Van Orman. 1951. "Two Dogmas of Empiricism." *The Philosophical Review* 60 (1): 20–43. <https://doi.org/10.2307/2181906>.
- Rapp, Stefan. 2026a. *The Limits of the Self in Ontological Materialism: On the Indeterminacy of Exclusive Personal Identity*. Zenodo. <https://doi.org/10.5281/zenodo.18219333>.
- Rapp, Stefan. 2026b. "Architekturtest AFST." Internal project test documentation "Test-Vignette-3," including the frozen master-paper test version (revision date: July 28, 2026), source-selection rule, source list and six source documents, protocols 04a–04k, and "05_Schlussauswertung." Unpublished working documentation of the Epistemic Reality Project. Materials consulted for the paper version of September 20, 2026.
- Vaithianathan, Rhema, Emily Kulick, Emily Putnam-Hornstein, and Diana Benavides Prado. 2019. *Allegheny Family Screening Tool: Methodology, Version 2*. Allegheny County Department of Human Services, April 2019.
- Williams, Bernard. 1970. "The Self and the Future." *The Philosophical Review* 79 (2): 161–180. <https://doi.org/10.2307/2183946>.
- Wimsatt, William C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.

Appendix A: Illustrations of the Diagnostic and Decision Architecture

Status and Function of This Appendix

The following appendix serves exclusively to illustrate the concepts, distinctions, and connections developed in the main text. It introduces no new canonical concepts, changes no definitions, and establishes no additional theses. The examples are deliberately simplified and should not be understood as exhaustive case analyses. The main text and the paper's conceptual canon remain authoritative.

A.1 The Jammed Door

A person wants to open a door. They expect to be able to press the handle down and then pull the door open. The door does not move, however. The person can imagine the door standing open and mentally alter their planned sequence of movements. This does not open the door. Only further pushing, pulling, or examination of the lock can establish why the movement fails.

The situation first illustrates the functional difference between inner form and outer form. Imaginings and plans for action permit a comparatively wide range of variation. The door, by contrast, appears as a recurring order that is not arbitrarily available and resists mere imagination. This resistance constrains the expectation and compels correction of action.

It does not follow, however, that the person has now determined the door's complete constitution independently of every possible perception, investigation, or description. They can establish that the lock is blocked, the wood warped, or the door locked. Yet each determination again arises within an act of perception, testing, or action. Outer form enables viable realist determinations without thereby providing presuppositionless access to a fully positively determinable outside in itself.

The emerging resistance is not yet friction in the narrower sense either, as long as it is not related to an activated expectation structure. Only because the person expects the door to open through ordinary operation of the handle does the absence of movement become readable as a mismatch. Friction thus lies not simply in the door, but in the relation between an activated expectation and the conditions of its operation.

The example shows how realist discourse, resistance, and the boundary rule interact. The door can be reliably examined and determined as blocked. This determination need neither be devalued as mere subjective production nor absolutized into a complete statement about the door independently of all conditions of cognition.

A.2 The Delayed Train

According to the timetable, a train is to depart from a particular platform at 10:15. At 10:20, it has not yet arrived. For the waiting person, a recognizable mismatch emerges between the activated expectation and the observed course of events.

This friction does not yet specify its cause. Possibilities include an ordinary delay, a last-minute platform change, cancellation, a misread display, or confusion between two similarly designated services. The timetable itself may also have changed. The first deviation therefore falsifies neither the entire timetable system nor the general model of rail transport.

For diagnosis, the claim architecture must be examined more closely. The unit of reference may be the particular train, the train number, the service, or the announced departure. The claim made is not merely that trains generally run, but that this service is available at a particular time and place under the stated conditions. The data basis, the timeliness of the information, the platform assignment, the concrete operating process, or the reliability of the information system used may be under strain.

The principle of proportionality initially requires a narrow response. The person checks the departure board or a current announcement. If a delay is confirmed, updating the expectation suffices. If the platform has changed, the connection between timetable information and spatial assignment must be corrected. Only if deviations occur repeatedly, systematically, and beyond a tolerance range justified for this purpose can the information system's more general claim to validity be challenged.

The situation may also concern different domains. Whether the train has arrived can be tested functional-empirically. Whether an announcement was understandable and available to passengers in time additionally concerns intersubjective and institutional conditions. The waiting person's irritation or uncertainty belongs to the subjective domain. These claims concern the same situation but are not viable under the same conditions.

The example shows why friction initially opens a need for diagnosis. An unfulfilled expectation does not itself already specify either the challenged level or the appropriate response.

A.3 The Kitchen Thermometer

A kitchen thermometer shows only 82 degrees Celsius in vigorously boiling water. Its user expects a substantially higher value and suspects that the underlying physical model has been refuted.

The deviation is functional-empirically relevant, but its assignment is initially open. The thermometer may be miscalibrated, improperly positioned, or damaged. The measurement may take place under unusual environmental conditions. The unit of reference may also be unclear: Is the device actually measuring the water's temperature, the steam above the surface, or a region affected by the vessel? Only once these conditions are specified can it be assessed which claim was tested.

The claim architecture here comprises at least the thermometer as measuring instrument, the water as unit of reference, the measurement procedure, environmental conditions, the theoretical model, and the expectation concerning the displayed value. A deviation between measured value and expectation does not automatically challenge all these components simultaneously.

The redistribution rule requires that the established loss of validity be assigned to the level actually tested. If comparison with a reliable reference instrument shows that the thermometer systematically produces values that are too low, its measuring performance is initially challenged. The underlying physical expectation is not thereby falsified. If, by contrast, several mutually independent measurement procedures support the same unexpected finding under controlled conditions, the need for diagnosis shifts to other parts of the claim architecture.

A proportional response would still be required. Conditions, measurement procedures, or model assumptions might first be specified more precisely. Such an adjustment remains legitimate as long as it can be tested independently and preserves genuine conditions of failure. If, instead, every possible deviation were retrospectively accommodated through untestable additional assumptions, immunization would result.

The example shows the function of the redistribution rule. An observation can be robust without thereby refuting the most comprehensive theory involved. What matters is which unit of reference, instrument, assignment, and claim to validity were actually tested under which conditions.

A.4 The Sick Houseplant

Several yellow leaves are discovered on a houseplant. Its owner immediately suspects a particular plant disease and begins treating it accordingly. A few days later, further changes occur that do not fit the initially assumed disease pattern.

Initially, there is a field of experience that has not yet been clarified. Yellow leaves, dry patches, altered growth, and soil conditions may be connected, but need not be. To process the observations, particular features are selected and combined into a possible unit of reference: a common process of damage or disease.

Such a provisional order can be useful as tentative stabilization. It makes expectations possible: Which further symptoms would be expected? Which changes should occur after treatment? Which observation would count against the assumption? Its provisional status must nevertheless remain visible.

Ontologization occurs as soon as the assumed process is managed as an identifiable unit available to be taken up again. The owner now speaks, for example, of “the disease” and relates new observations to it. This operation is not already problematic. Without it, observations could hardly be ordered and compared. It becomes absolutization only when the provisionally formed unit is treated as conclusively determined and deviant features are reinterpreted solely to preserve the original diagnosis.

The new observation generates friction. It must now be examined whether only an individual assumption was wrong, several causes interact, the initially connected symptoms must be separated, or the unit of reference as a whole was formed unsuitably. Merely supplementing the existing model may suffice. Reconfiguration is also possible: some changes are no longer treated as expressions of the same cause, but assigned to different processes.

Only when the supporting relations of the diagnostic model must be restructured does a more extensive revision occur. The principle of proportionality guards against two errors: the first deviation must not immediately destroy the entire model, but neither may it be fitted arbitrarily into an already absolutized diagnosis.

The example shows how tentative stabilization, ontologization, and revision are connected. Cognition needs provisional units of reference, but must keep their identity and assignment rules alterable under strain.

A.5 The Misunderstood Promise

At the end of a meeting, two people discuss an upcoming move. One says, “I’ll come over on Saturday.” Later, they do not appear. The other person declares that a binding promise has been broken. The first replies that they had merely expressed a nonbinding intention.

Several claims must be distinguished. Whether the sentence was actually said can be challenged functional-empirically through memory, messages, or witnesses present. Whether it counted as a promise in the particular situation, by contrast, concerns an intersubjective order of language use, expectations, earlier agreements, and mutual recognition. How certain or disappointed the participants felt additionally belongs to the subjective domain.

Validity mapping prevents these questions from collapsing into one another. An established wording does not by itself decide which social obligation arose. Conversely, disagreement about the obligation does not necessarily refute that the sentence was uttered. Nor does the intensity of disappointment establish either the wording or the binding force of the statement, although it remains real as subjective concern.

The connection between wording, situational expectation, and social obligation forms a transition function. It transfers a linguistically observable statement into an obligation managed intersubjectively. This transfer is neither identical nor automatic. Tone, context, follow-up questions, and prior agreements may be decisive.

A proportional response must therefore first clarify which reference can be taken up again jointly. The participants may be able to agree that the sentence was uttered but its function understood differently. This can lead to correction of future communication rules, such as explicit confirmation of binding agreements. Such recoupling between statement and obligation differs from revising all mutual expectations or completely abandoning trust.

The example shows that the same situation contains claims in several domains. Conflicts often arise not because each side experiences a wholly different reality, but because transitions between experience, wording, and jointly managed obligation are stabilized differently.

A.6 The Inadequate Forecasting Model

A company uses a simple model to forecast demand for a product. The model was sufficiently useful over an extended period. After a market change, however, its predictions repeatedly fall well outside a tolerance range independently justified beforehand. Several adjustments of individual parameters fail to improve forecasts durably.

The recurring deviations initially challenge the model's validity status. It must be examined whether only particular applications are affected or whether its general claim to predict demand reliably within the previously claimed range is no longer viable. If the model fails only for a new product group, contextual falsification may apply. If, by contrast, its general claim persistently fails across the relevant applications, its previous claim to validity is globally falsified.

It does not automatically follow that the model must immediately be removed from every practical use. It may serve as a comparison basis or rough initial estimate if this limited use remains justified. The absence of a viable alternative is not, by itself, a reason for further use in a guiding role. Validity status and guiding status must remain separate.

The opening rule now concerns the allocation of search resources. If further consolidation produces hardly any gains in robustness, the costs of additional adjustments increase, and relevant friction persists, this favors greater exploration of alternative models. The rule does not yet establish which alternative is superior. Initially, it justifies only why further resources should not flow exclusively into stabilizing the existing model.

The abandonment rule now requires a separate examination of further use. If demand forecasts threaten to produce substantially mistaken decisions, their use in a guiding role can be provisionally suspended or abandoned even without an available replacement. Suspension requires a time or occasion for reassessment. If another model is to replace it, that model must be able to assume its relevant functions more effectively under reasonable transition conditions and costs.

The example shows why falsification, opening the search space, suspension, replacement, and abandonment are different decisions. An impaired validity status determines none of them on its own. Justified replacement requires a viable successor order; justified suspension or abandonment does not.

A.7 The Deadlocked Working Group

A working group is to investigate the cause of recurring errors in a technical process. Initially, all participants use the same reports, measurements, and categories. After several unsuccessful meetings, however, two camps emerge. Each side trusts only the analyses of its own members. Deviant data are treated either as faulty or as the result of the other side's lack of competence.

Strong dissent alone does not yet mean a loss of shared epistemic reality. Joint examination can remain conflictual as long as references can be taken up again, differences articulated, and corrections transparently traced to earlier assumptions or decisions. Contradiction can in fact improve shared model management.

The situation becomes problematic when these conditions disintegrate. If the same concepts are used differently and the divergences can no longer be translated, contributions lose shared connectability. If contrary indications are interpreted solely as expressions of group affiliation, they can no longer challenge the respective model order. Formally, the working group, its meetings, and its documents persist. The social order remains stable while its intersubjective validity declines.

Trust here means neither harmony nor personal sympathy. It designates the expectation that the processes of taking up, taking up again, and correction will not break off arbitrarily. Without this expectation, the effort required for every shared test increases. Data must not only be examined, but first defended against the suspicion that they have been strategically selected or are fundamentally unreliable.

Institutional responsibilities can stabilize examination, but do not guarantee epistemic authority. A leader can allocate speaking time, specify measurement procedures, or make decisions. If they systematically block particular deviations or distribute the costs of possible errors one-sidedly, they can weaken correctability despite formal order.

A loss of a shared basis for correction occurs when there is no longer a sufficiently shared order within which a contrary indication can be taken up as a possible challenge. Restoration does not necessarily require a new overall consensus. Agreeing on a limited shared dataset, a jointly testable experiment, or a documented translation rule between the categories used may suffice.

The example shows that distributed model management has social conditions of testing. Shared concepts and institutions are insufficient if objections, data, and consequences can no longer be traced back into a transparent correction.