

# Split-Domain Cognition

---

Canonical long-form articulation. Version 1.0, April 2026.

---

## Preface

---

**About this document.** This is the long-form articulation of Split-Domain Cognition. It begins from a single observation — that two unlike kinds of thinking are reliably collapsed into a single act, and that the collapse produces a measurable cost — and works that observation out at the length the empirical material requires. It is the source from which the shorter forms are condensed; it is not itself an introduction. Readers who want a shorter entry should begin with *SDC in brief* on the website. Readers who want to know why the principle is not only an AI-pipeline pattern should read the position essay, *A Principle, Not a Pattern*. Readers who want to apply the principle to a new domain should read *Process* and the *Derivation Protocol*. This document is for readers who want the empirical material itself.

**Where this sits in the corpus.** The canon home is [splitdomaincognition.org](https://splitdomaincognition.org). The condensed version is at [/sdc-in-brief/](#); the position essay at [/principle/](#); the operational methodology in *Process*; the protocol for deriving a domain-specific variant at [/derivation-protocol-v1/](#); the catalogue of declared variants at [/variants/](#). All of these refer back to the present document.

**Authority and version.** Canonical, v1.0. April 2026. The website is the source of record. If this PDF and the website disagree, follow the website.

**Use.** Openly citable. See [Governance](#) for the one-person canon and how variants are admitted.

**Author.** Prayas Abhinav.

---

## Introduction

---

Split-Domain Cognition begins as an observation about how mistakes get made. It is not a theory and not a model; it is a name for a pattern. The pattern is that two unlike kinds of thinking — describing what is there, and judging whether it is acceptable — are reliably collapsed into a single utterance, by a single faculty, in a single moment. The collapse is the failure. The principle is the discipline of not collapsing them.

The articulation that follows is empirical first. It names where the collapse appears, what it costs, and what holding the two domains apart looks like in practice. It is structural second. It says how the holding-apart can be built into a piece of software, an institutional process, a curriculum, a curatorial method, a labour market, a clinical practice. The argument is meant to be portable across domains — not because the domains are the same, but because the failure mode is.

What follows is the long form. The reader is being asked to slow down.

---

## The observation the idea starts from

---

A pattern shows up whenever a single agent — a person, a model, an institution, a form — is asked to do two unlike kinds of thinking in the same breath. The pattern is predictable enough that it deserves a name. The first kind of thinking is descriptive, generative, interpretive. It attends to what is there: what the student said, what the text states, what the image shows, what the menu offers, what the situation contains. The second kind of thinking is evaluative, verdictive, categorical. It closes on an outcome: whether the work is good, whether the claim is valid, whether the treatment is ethical, whether the candidate is suitable. The first kind we will call **language work**. The second kind we will call **judgement work**. The two kinds can look similar from the outside and feel like one continuous act from the inside. They are not one continuous act. When the two are fused in the same call — by the same faculty, the same model, the same moment of utterance — certain failures appear reliably.

The failures are empirical. They can be named, demonstrated, reproduced, and measured.

- Verdicts issued but not auditable. A decision was made; no one, including the decider, can reconstruct against what criteria.

- Descriptions that quietly smuggle evaluation. Prose that claims to report is already steering belief.
- Rubrics that cannot be taught. What is being assessed cannot be shown to the person being assessed.
- Rejection that cannot be learned from. The person on the receiving end cannot calibrate, because the criteria were never separated from the verdict.
- Drift under social pressure. What "counts" as acceptable shifts with whoever is loudest; there is no explicit standard to drift against.
- Sycophancy in machines. Large language models, asked to judge, default to confirming the interlocutor; the verdict inherits the fluency of the language layer instead of being held to its own criterion.
- Hierarchies of worth that collapse on inspection. Institutions asserting global rankings that cannot survive the first question about common units (see `/home/prayas/Dropbox/personal_projects/gabor.mate/worth-is-not-hierarchical.md` for the formal argument that global hierarchies of worth fail at the measurement level).

Each failure is a symptom of the same underlying collapse. The collapse is between two cognitive domains whose criteria are different in kind. **Split-Domain Cognition (SDC)** is the proposed architecture that refuses the collapse.

## The claim, stated plainly

---

When language work and judgement work are at risk of collapsing, build the work so the two stay distinct. Let the language domain be answerable to the criteria proper to language (coverage, interpretability, fluency, pattern recognition). Let the judgement domain be answerable to the criteria proper to judgement (consistency, explicitness, reproducibility, revisability). Do not let either domain borrow the other's authority.

This is a claim about architecture. It is not a claim that thinking must always be divided. Some acts of mind are rightly unified — a drawing, a first intuition, an improvisation, a poem being written. The principle applies only when language work and judgement work are both in play and are being forced to share one channel. In that situation, and only in that situation, the architecture matters.

The claim is also narrow. It does not say code is better than language models, or that humans should judge less, or that explicitness is always a virtue. It says: where language and judgement co-occur, keep them structurally unlike, because the failures of conflation are predictable and the benefits of separation are measurable.

## The mechanism — why the domains are different in kind

---

Language work and judgement work differ in the criteria that make them *go right*.

Language work goes right when it captures the shape of what is there. Its excellence is thickness: how much of the situation is represented, at what resolution, with what interpretive fidelity. A good description of a student's studio output names its moves, its references, its questions, its blind spots. A good qualification of a menu item names its ingredients, its substitutions, its claims, its silences. A good classifier reports what it found with the right probability attached. The test is not "is the description correct?" in a binary sense; it is "is the description adequate to what was there?" The criterion is coverage-under-interpretation.

Judgement work goes right when it issues a verdict that can be reconstructed. Its excellence is explicitness: the verdict connects to a rule, the rule is prior to the case, and another judge applying the same rule to the same facts would reach the same verdict. The test is not "does the verdict feel right?"; it is "can the verdict be reconstructed and contested?" The criterion is reproducibility-under-rule.

Coverage-under-interpretation and reproducibility-under-rule are different virtues. They can be optimised separately. They cannot be optimised together in the same act without one contaminating the other. When judgement work borrows language work's virtue, verdicts become impressive but incontestable (the rhetoric of "it just feels like a weak piece"). When language work borrows judgement work's virtue, descriptions become confident but thin (the rhetoric of "obviously this is unethical", where the evaluation has replaced the observing).

This asymmetry is the mechanism. It is why probabilistic systems that attempt to do both stages at once (a single prompt that asks a model to "decide whether this is good") produce outputs that are simultaneously over-confident and under-justified. It is also why human institutions that pretend to unified judgement (a critic, an interviewer, a curator, an

examiner) drift under pressure. The two virtues pull in different directions. A single channel cannot carry both.

## The empirical signature — what you see when SDC is present and absent

---

A good theoretical claim is one whose presence and absence leave different traces. SDC does. The traces are observable without being told what to look for.

When SDC is **absent** from a practice, you see:

- Verdicts that shift across cases without a rule to track. Two candidates with the same evidence receive different decisions from the same panel on different days.
- Critiques that are long on adjectives and short on criteria. The word "strong" or "weak" carries the work that should have been carried by a rubric.
- Writing that cannot be challenged without challenging the writer. The evaluation is fused to the description so tightly that disagreement with the evaluation reads as disagreement with the reporting.
- Institutions that explain themselves in brand language. "Our values", "our standards", "our excellence" — language that is impossible to check against any explicit commitment.
- Machines that flatter. A model asked to assess a draft that the user has clearly invested in will return warmer assessments than the same model returns on the same draft presented without investment.
- Drift. What was excluded last year is included this year, with no conversation about the change. The drift is not noticed because there was no explicit line to cross.

When SDC is **present** in a practice, you see:

- A document, however short, stating the criteria in advance. Rubric, editorial policy, curatorial thesis, decision matrix, rights framework — the form varies; the existence is constant.

- A separation, in the practice's own self-description, between what it attends to and what it judges. "We first record what the work is; then we assess against the criteria."
- Disputes that can be located. When two stakeholders disagree, they can name whether the disagreement is about the description (what was there) or the criterion (what should count).
- Refusal capability. The practice can say "this is outside our criteria" without producing a verdict. Silence is available where silence is honest.
- Traceable changes. When the criteria change, the change is a visible act rather than an invisible drift.
- Teachability. A newcomer to the practice can read the criteria and learn what will be assessed, rather than learning by absorbing the room's reactions.

The two lists are not value judgements. They are descriptive of the architecture. SDC predicts that practices that embed the separation will exhibit the second list; practices that refuse the separation will exhibit the first. Every Koher tool deployed so far is a test of this prediction in a specific domain. The prediction has held.

## The cross-domain catalogue

---

The pattern recurs far from its technical origin. Each application below is a real practice (some already deployed, some in research phase, some internal to Koher, some outside). Each instance shows the collapse, the separation, and the difference the separation makes.

### AI tools

The technical pipeline is the most compact expression of SDC. A language model produces structured signals from unstructured input (Stage 1). Deterministic code applies thresholds and rules and issues a verdict (Stage 2). A language model narrates the verdict in readable prose at temperature zero (Stage 3). The three stages share no memory and enforce no overlap. Coherence Diagnostic, Play Shape Diagnostic, Fragment Mapper, studioMeetingCompanion, and animalRightsLens are built on this pattern. Their

accuracy on held-out test cases, their auditability, and their ability to refuse verdicts outside their rubric are direct consequences of the separation.

## Design studio pedagogy

A design studio asks students to generate and faculty to evaluate, usually in the same room, often in the same conversation. When faculty generate alternatives inside the critique, or when students judge their work inside the act of making, the studio becomes anxious. Students cannot calibrate because they do not know what is being judged; faculty cannot judge fairly because they are mid-flow generating. Good studios — empirically, the ones students report as formative — separate the stages ritually: a phase for making and describing, a phase for rubric-bound assessment, a phase for narration of the assessment back to the student. *studioMeetingCompanion* is the software expression; the pedagogical move predates it.

## Curation

A curator can show or endorse. Endorsement and showing are different acts with different grammars. *Museum of Vestigial Desire* (2012–2021) separated generative fragment-making from curatorial refusal of assembly. *Of Games* at Khoj separated residency practice from the thesis that framed it. *Growing Down* (2009–2010) separated what musicians did on the night from what was curated by placing the event on a rooftop. Curation-as-SDC treats the description of each work and the thesis that positions it as distinct artefacts, both written down, neither substituted for the other. Curation that collapses them produces the tastemaker — a figure who generates and judges in one breath and leaves nothing the audience can re-enter.

## Public art

*The Murmur Engine* (FICA 2026–27 submission) is SDC staged for public encounter. A machine placed in a public site will refuse to issue confident verdicts, instead emitting its hesitations, retrievals, and thresholds as ambient surface. The audience sees Stage 1 (what the machine is reading), Stage 2 (the rules by which it chooses silence), and Stage 3 (the rendered output) as three distinguishable layers. The work is a demonstration that the failure critiqued elsewhere — machines dressed as judges — is avoidable by design. The gallery argument and the pipeline argument are identical.

## Animal ethics

The politics of animal rights is a domain where SDC cuts deep. The language of *welfare* (probabilistic, managerial, contextual — how the animal was treated) is different in kind from the language of *rights* (deterministic, categorical — whether the treatment is acceptable at all). "Humane meat" is welfare language smuggled into rights territory, producing a judgement-shaped object that cannot be audited against a rights framework because the framework has been replaced by an adjective. See

[variants/animal-rights-and-veganism.md](#) for the full variant.

## Writing and philosophy

The descriptive–normative distinction is older than SDC. Hume named it; Moore formalised it. What SDC adds is architectural enforceability: the observation that the distinction does not maintain itself in prose and must be held up by form. *Museum of Vestigial Desire* was built so description and evaluation could not collapse. Its founding text ( [symptoms-of-a-museum-part-1](#) , 17 November 2012) states the move directly:

*"I choose to appropriate the word. The word 'museum' now just means what I read into it, it just 'means' what I pronounce. So it becomes a specification more than a description arrived at through careful observation and analysis. Pronouncements, specifications, commandments, hard-knocks."*

This is the Stage-2 claim made before SDC existed as a term. The Museum declares that it *pronounces* rather than *describes*, and marks the declaration explicitly so that description-work and verdict-work stay distinguishable in everything that follows. A further 2012 formulation, in the columnist bio at

[offices/columnists/prayas-abhinav](#) , puts the same move from the other side: *"he reviews only the narrative of the event and not the event itself."* Stage 2 is named as operating on language outputs, not on the underlying world — the split is architectural, not editorial. An essay that fuses description and evaluation moves the reader without letting the reader disagree. An essay that separates them gives the reader the data for their own judgement. The distinction is visible in the prose, not only in the philosophical commitment. [movd-as-precursor.md](#) collects ten such passages from the archive, annotated by the SDC role each performs.

## The practice itself

Koher as a whole is an SDC move at the scale of a life. The Level 1 / Level 2 position ( `positions-source/level-1-living-practice.md` ) names it: the lived practice is generative (Level 1); the shared practice is evaluative (Level 2). Measuring the lived practice by Level 2 outcomes is the collapse. Keeping them structurally unlike is the refusal. The canonical claim recurses at three scales — inside the pipeline, inside a medium, inside a life.

## The structural connection to worth

---

The argument in `worth-is-not-hierarchical.md` (Prayas Abhinav, 11 April 2026) is the deepest philosophical underpinning of the SDC claim, and it is worth surfacing here because the two arguments are structurally the same move applied at different scales.

That argument shows that hierarchical rankings of worth across persons fail at the level of the common unit. Different domains of human excellence (physics, moral perception, aesthetic judgement, craft, relational capacity) do not share a unit in which a global weighted sum could be computed. Einstein's depth in physics and his blindness to the patriarchal arrangement he lived inside are not two readings on the same scale; they are readings on different scales whose conversion factor does not exist. A global hierarchy of worth cannot be constructed without smuggling in an arbitrary weighting, and the judge qualified to construct such a hierarchy would, in the earning of that qualification, have discovered that the hierarchy is structurally impossible.

The SDC claim is a restricted form of the same observation. Where the worth argument says *domains of excellence are incommensurable at the unit level*, SDC says *language work and judgement work are incommensurable at the criterion level*. The larger argument licenses the smaller. If the yardsticks of different domains do not reduce to a common unit, then two such domains cannot be combined into a single act of assessment without the combination being arbitrary in the strict sense. Any act that pretends to combine them (a grade, a verdict, a ranking) is importing a weighting that has no principled ground.

The ethical consequence of the larger argument — that hierarchical worth across persons is incoherent, and by extension across species — grounds the animal-rights variant. But

the structural consequence of the larger argument also grounds the SDC architecture itself. Both are expressions of the refusal to collapse unlike domains into a common unit.

For this reason, `worth-is-not-hierarchical.md` should be cross-referenced by any serious articulation of SDC. It belongs in the intellectual lineage, and it is referenced explicitly in the animal-rights variant where its ethical application is operative.

## What SDC predicts (and how it could be wrong)

---

A claim that cannot be falsified is not empirical. SDC makes predictions that are falsifiable in principle.

1. **Prediction on audits.** Practices that separate language and judgement into distinct artefacts (rubric, description, narration) will produce higher inter-rater agreement on the judgement layer than practices that do not. If explicit rubrics do not improve agreement over silent ones, the mechanism is not what SDC claims it is.
2. **Prediction on teachability.** Domains with explicit rubrics will produce more rapid calibration in newcomers than domains without. If students, authors, candidates, or artists exposed to explicit criteria do not calibrate faster than those exposed only to verdicts, the teachability claim fails.
3. **Prediction on drift.** Practices without explicit rubrics will drift over time in the direction of whichever stakeholder's preferences are most continuously present; practices with explicit rubrics will require visible acts to drift. If drift is similar across both kinds of practice, the co-optation-resistance claim fails.
4. **Prediction on sycophancy.** Language-model systems that perform language work and judgement work in a single prompt will show higher agreement with whatever the user appears to want than systems that separate the stages. If the separation does not reduce sycophancy in comparable settings, the architectural claim about model behaviour fails.
5. **Prediction on legibility.** Outputs from split-domain pipelines will be more successfully re-entered by non-specialists (they can follow the argument,

locate the disagreement, propose a change) than outputs from single-channel processes. If legibility does not improve, the trust-reconstruction claim fails.

These predictions are open. Koher's tools have provided early support for (1), (4), and (5) through their deployment behaviour. (2) and (3) require longitudinal evidence that the practice has not yet collected. The principle does not depend on these predictions being proven for every case; it depends on them being the right predictions to test, and on the tests being possible.

## What the idea is not

---

**Not a framework for all cognition.** SDC addresses a specific failure — the collapse of unlike domains in a single call. Activities that involve only one domain (pure making, pure calculation, pure performance) do not need the architecture.

**Not a claim against intuition.** Expert judgement is not devalued; it is repositioned. The expert's intuition moves into the rubric, where it can be examined, taught, and revised. The expertise is not removed from the practice; it is moved to a layer where it can do more than act silently.

**Not a claim that code is superior to humans.** Code is mentioned often because software forces explicitness, and explicitness is the practice that SDC names. Human practices that are already explicit (a good editorial policy, a well-written rubric, a curated thesis held publicly) are SDC in action without any software present.

**Not a guarantee of good verdicts.** An explicit bad rubric produces bad verdicts with perfect auditability. SDC does not fix bad judgement; it makes judgement inspectable. Whether the judgement is good is a separate question, one the architecture invites but does not answer.

**Not a neutrality claim.** Making the rubric explicit makes it contestable. SDC does not say the rubric is neutral. It says the rubric is legible enough to argue with.

**Not a finished idea.** Each domain the principle is applied to reveals something the principle did not already know. The variants in `variants/` are the ongoing test. The principle is refined by its applications.

## Implications for what Koher builds

---

Three consequences shape Koher's decisions about new work.

First, when evaluating a proposal for a new tool, collaboration, or artwork, the question is not "is this an AI project?" but "is there a risk of collapse between language work and judgement work in this domain, and does the practitioner have genuine stake in preventing the collapse?" If yes, SDC applies and Koher's architectural commitment engages. If no, Koher may still be the right home for the project, but the principle's specific work is elsewhere.

Second, every Koher tool must make its own architecture visible to its users. The Behind the Curtain toggle, the layer cards with structural accents, the explicit display of Stage 1 / Stage 2 / Stage 3 — these are not design decorations. They are the empirical claim made legible. A Koher tool that hid its architecture would contradict its own argument.

Third, when explaining Koher to someone outside AI, the AI pipeline is the best *illustration* of SDC but the worst *definition* of it. A fuller explanation includes the studio example, the curatorial example, the ethics example, the Level 1 / Level 2 example. These are typically more accessible to non-technical audiences and demonstrate that SDC is a practice rather than a technical trick.

## Closing

---

Split-Domain Cognition names a structural move that recurs across domains when cognition is at risk of collapsing between language work and judgement work. It is empirical in that it identifies observable symptoms of the collapse and observable consequences of the separation. It is structural in that it describes an architecture rather than a style. It is narrow in that it addresses a specific failure mode and not cognition as a whole. It is portable in that the same architecture recurs across software, pedagogy, curation, ethics, writing, and the shape of the practice that articulates it.

The principle is the same at every scale. When two kinds of thinking meet, build the work so the kinds stay distinct. Let the language be language. Let the judgement be judgement. Keep unlike domains unlike, because the unlikeness is not a limitation — it is the source of whatever power the architecture has.

---

Version 1.0 — 16 April 2026. Pair-references: `principle-not-pattern.md` ,  
`process.md` , `methodology-funding.md` , `variants/README.md` ,  
`variants/animal-rights-and-veganism.md` ,  
`/home/prayas/Dropbox/personal_projects/gabor.mate/worth-  
is-not-hierarchical.md` .