

Contextual Identity and Cognitive Dark Matter in DSRP-484

The Advanced Is/Is Not Operation and the Structure That Minds and Formalisms Leave Unwritten

Derek Cabrera^{a,@}

a: Cabrera Research Lab

@: dac66@cornell.edu

ABSTRACT: What is a thing? We propose that its identity in a model is the organized account of what it is and is not, what it contains and belongs to, what it relates to, and how it is framed: its DSRP structures — Distinctions, Systems, Relationships, and Perspectives — constitute its *contextual identity*. *Cognitive dark matter* is the Organization a system uses that is absent from its own account of itself: $CDM = O - I(O)$. We tested these constructs in two large language model families using an invented identity, *VORG*, holding its definition constant while varying the surrounding Organization and asking questions the definition left unresolved. The supplied Organization, not the definition, determined the answers: GPT matched it at .84–.86 and Claude at 1.00; identities with the same structure counts but different structures produced disjoint answers; removing a Pattern removed the answers it supported; changing one identity altered inferences about neighboring identities; and neither system collapsed unknown into false across 1,440 unidentified cases. Taken together, the findings show that identity behaves as an organized field of what is, is not, and remains unknown, and that this field can extend beyond what is explicitly supplied. With only a definition, GPT built a road-vehicle identity for *VORG* from familiar terms; when the same world was rewritten with nonce terms, that imported structure disappeared. Cognitive dark matter was therefore measurable as the gap between operative and reported Organization: GPT used .86 of the identity while reporting .25, whereas Claude showed little gap until the instruction “use only the supplied material” was removed, after which it too reported mainly the definition while continuing to use the wider identity. Perspective thus changed both what Organization entered the task and whether the Organization doing the work remained visible. The same structural machinery supported disciplined inference when visible and unsupported assumption when it was not. In real-world judgment, this means that what people or machines conclude can depend less on the information in front of them than on the unreported Organization through which that information is being interpreted.

KEYWORDS: Contextual identity | Cognitive dark matter | DSRP-484 | O-Theory | Is/Is Not | Copriming | Perspective | Enclosure | Large language models | Metacognition

Cite this paper: Cabrera, D. (2026). Contextual identity and cognitive dark matter in DSRP-484: The advanced Is/Is Not operation and the structure that minds and formalisms leave unwritten. *Journal of Systems Thinking*, 6(1), 1–78. <https://doi.org/10.54120/jost.v26i1.dc6cicm6>

1. Introduction: Ode to identity

Identity is everywhere because everything that exists must exist as something. A rock, computer, city, company, organism, galaxy, person, nation, word, variable, or idea has identity insofar as it is distinguishable from what it is not. In this most basic sense, identity is what something is—how it is IS-ing. Existing as something is the first act: before anything can act, relate, compose, or be viewed, it must have thing-ness.

Identifying is among the most ordinary things we do. Often, our first explicit move is to give something a name. We name babies before they are born, companies before they begin trading, teams before they begin playing, and projects before the work begins. We assign identifiers to objects, users, and processes in code. Genesis places the naming of animals near the beginning of its account of human life (2:19–20). A child learning language seems to take a labeling gun to the world. The act repeats across scales, from a variable to an institution, so constantly that we scarcely notice ourselves doing it. Yet the name makes only a small part of the identity explicit.

But what a thing is cannot be recovered from the thing alone. Consider a photograph of a rock against a white background. Without scale or surroundings, the object might be a pebble or a house-sized boulder. Place the house-sized boulder at the bottom of a sheer rock wall and its position suggests that it once fell from above. Place the same boulder in the middle of a broad, sloping, U-shaped valley and it becomes a glacial erratic—material transported, perhaps for miles, by a glacier that later melted beneath it. The boulder's material may be unchanged, but its contextual identity is not. Its scale, history, origin, relationships, and significance emerge from the Organization [1] surrounding it.

This is why identity is so fundamental. Every existing thing forms a boundary Distinction between what it is and what it is not [2]. Its identity thereby participates in establishing the identities of other things. A rock helps distinguish the valley containing it; a city helps distinguish the region around it; a company helps distinguish its employees, competitors, products, and markets; a galaxy helps distinguish the structures within and beyond it. Identity is not an attribute added after existence. To exist as something is to make a Distinction. And to create even a single Distinction, a part-whole System, action-reaction Relationship, and point-view Perspective are required [1,3,4].

Human life multiplies these forms. We construct personal identity over a lifetime, participate in social and national identities, create corporate and institutional identities, authenticate ourselves with identity badges, and assign identities to users, objects, processes, and variables in code. These uses are not connected merely because they share a word. Each concerns the structural problem of determining what something is, what it is not, where its boundary lies, what Organization constitutes it, and how that Organization changes.

The isiZulu greeting *Sawubona* makes recognition itself the greeting. Its origin is “we see you”; Mazisi Kunene explained its communal significance as acknowledging another person on behalf of one's family or community [5]. *Sikhona* affirms presence in the plural; *Ngikhona* is its singular counterpart [6]. Read philosophically as a greeting of identity, seeing and presence become: *I see you. Therefore I exist.* That

formulation expresses the significance of recognition rather than a word-for-word translation. Someone whose life, relationships, history, and belonging already extend far beyond a name is being acknowledged. The greeting gives an everyday human form to the paper's question: *what does it mean to see what something—or someone—is?*

1.1 Identity Is More Than a Label

A dictionary is a useful place to begin, but it is a remarkably unreliable model of how identity actually works. We commonly behave as though the definition governs the meaning of a word: first the definition exists, then people learn it, and finally they apply it to particular situations. In reality, the process usually runs in the opposite direction. People do not ordinarily learn what truck, friendship, marriage, trust, fairness, or leadership means by memorizing a dictionary entry. They encounter these identities in situations, distinguish cases from noncases, observe what they contain and belong to, discover how they relate to other things, and learn how their meaning changes with purpose and perspective. Dictionary definitions are useful compressions of accumulated usage, but they neither contain nor control it.

This remains true even when people deliberately use a dictionary to learn an unfamiliar word. A student studying for the SAT may memorize a definition from a flashcard, yet when the definition does not produce understanding, the student's first request is usually: "Can you use it in a sentence?" The student wants a situation in which the word actually does something. Who uses it? About what? In contrast to what? What does it imply? What would not count as an instance? The example is not merely an aid for remembering the definition. It supplies the first brick in an evolving structure of situated uses from which the identity of the word becomes usable.

Additional encounters add more bricks. Some affirm what the identity is; others establish what it is not. Some locate it within larger wholes or reveal its parts. Others establish relationships, exceptions, purposes, or competing framings. The resulting identity is not stored in any single example, and it is not equivalent to the dictionary entry. It lies in the Organization progressively assembled across the examples.

Consider the identity TRUCK. A dictionary may define a TRUCK as a motor vehicle designed to transport cargo. Yet a TRUCK may count as a car when a parent says, "Pick up your cars," when someone asks where the cars are parked, or when the relevant destination is a car dealership. It may not count as a car when payload, licensing, insurance, or vehicle classification matters. The physical object has not changed. The Organization relevant to identifying it has.

People perform even more striking transformations without apparent difficulty. A truck is ordinarily a physical object, but someone can say, "Go ahead and truck this stuff over there," and be understood immediately—even if the listener has never encountered precisely that use. The identity has shifted from object to action while retaining enough contextual structure—load, carrying, movement, and destination—to remain intelligible. Meaning is therefore neither fixed to the physical object nor arbitrarily detached from it. It is reorganized for use.

The same process operates in friendships, marriages, organizations, and teams. People use words such as trust, loyalty, fairness, leadership, and commitment as though their shared vocabulary guarantees shared

meaning. It does not. Two people may know the same dictionary definition while differing about what counts as trust, what violates loyalty, which actions constitute leadership, what fairness requires, or when commitment permits departure. Their disagreement is not necessarily a disagreement about words. It may be a difference in the Organization constituting the identity beneath the shared word.

The puzzle is therefore thoroughly ordinary. We learn identities through situations rather than definitions, modify them as circumstances change, extend them into unfamiliar uses, and coordinate—or fail to coordinate—with people whose identities are organized differently. We do this rapidly and usually without being able to enumerate the structure that makes it possible.

How?

This paper argues that the answer lies in the *Organization* (and *re-Organization*) surrounding an identity. What something is in a model is not exhausted by its label or categorical definition. It also depends on what it is distinguished from, what parts it has and what wholes contain it, what it is related to and how, and what it frames, what frames it, and from which points it is viewed. Change that Organization and the identity can change even when the word or physical object remains the same.

Identity is not a context-free core supplemented by contextual details. It is contextual all the way down. Contextual identity names the total organized account of what something is—not one situational version of an otherwise independent identity. For a focal identity X in model $\mathbb{M}$ at time t , it comprises all currently identified DSRP [1,3,4] structures specifying what X is and is not, what parts it has and what wholes contain it, what it is related to and how, and what it frames, what frames it, and from which points it is viewed. Because these structures are explicit, contextual identity is not merely describable; it is countable, comparable, and testable. In compact form:

$$CI_{\mathbb{M}}(X, t) = \text{the organized IS/IS-NOT DSRP structures involving } X$$

This total accounting matters wherever identities must do real work. In everyday life, differently organized identities underlie misunderstanding and disagreement in friendships, marriages, teams, and institutions. In technical domains, incomplete identities produce incomplete specifications, ambiguous communication, modeling errors, design failures, security vulnerabilities, and operational misalignment. Computing, engineering design, corporate leadership, warfighting, cybersecurity, and other high-consequence activities all depend on accurately determining what something is, what it is not, where its boundaries lie, how it is organized, and how it changes.

This is not a dictionary definition expanded into a longer list of attributes. It is the organized account of what the model identifies a focal identity as being. The comparison below makes the difference visible. Every structure in the contextual-identity column is an IS or IS-NOT structure. What varies is the DSRP Pattern and role through which that identity is specified.

Table 1. Contextual Identity as the Organized Answer to “What Is X?” At its core, contextual identity answers one question: What is X in this model? Every element of $CI_{\mathbb{M}}(X, t)$ is a statement of the form “ X is _____” or “ X is not _____.” IS NOT is therefore part of what X IS: it identifies the exclusions and negations that establish X ’s boundary.

DSRP does not replace this identity logic; it specifies the Distinction, System, Relationship, or Perspective structure that fills the blank.

Focal identity	Dictionary or conventional identification	Contextual identity in an evolving model
X	X is a member of category c .	X is c and is not o . X is a whole containing $p_1, \dots, p_n$ and is part of W . X is related to Y through R and is not related to Z through R . X is framed by and frames other identities. When $\mathbb{M}_t$ changes, these structures—and therefore the contextual identity of X —may change.
Truck	A truck is a motor vehicle designed to transport cargo.	A truck is a vehicle. A truck is a car in some contexts and is not a car in others. A truck is a whole containing an engine, wheels, cab, and cargo area. A truck is part of transportation systems, fleets, and collections. A truck is related to drivers, roads, cargo, fuel, dealerships, and licensing systems. A truck is framed differently by a child, owner, contractor, regulator, or designer. Adding van, tractor, SUV, electric vehicle, or autonomous delivery robot can redraw the boundaries and Relationships through which truck is identified.

The table reveals two differences. First, the conventional definition identifies a narrow categorical relation, whereas contextual identity assembles the full range of IS and IS-NOT structures through which the focal identity is distinguished, organized into parts and wholes, related, and framed. Contextual identity is therefore not a longer dictionary entry. It is the total organized account the model presently supplies of what the focal identity is.

Second, contextual identity is not a static profile attached permanently to X . It evolves with the model. Suppose a previously absent identity Z is added:

$$Z \notin \mathbb{M}_t, \quad \mathbb{M}_{t+1} = \mathbb{M}_t \cup \{Z\}$$

Even if no intrinsic feature of X changes, Z becomes part of what X is distinguished from:

$$o_X(t+1) = o_X(t) \cup \{Z\}$$

The addition may also create or alter System memberships, Relationships, and Perspectives involving X . Consequently:

$$CI_{\mathbb{M}}(X, t+1) \neq CI_{\mathbb{M}}(X, t)$$

Contextual identity therefore changes not only when X is directly edited, but also when the Organization surrounding X changes. It is the identity of X in an evolving model, not a fixed inventory stored inside X .

This structural account leads to a specifically cognitive problem. A person can recognize, classify, interpret, and use an identity with far greater discrimination and predictive success than the person can explain. Much of the Organization doing that work remains unreported. Its effects nevertheless appear in what the person recognizes, expects, rejects, says, feels, and does.

Contextual identity names the entire currently identified DSRP account of what a focal identity IS and IS NOT in a model. Cognitive dark matter names the part of that model's Organization that is operative but absent from the model's own account of itself — what O does that I(O) does not say. Their distinction concerns the focal identity and the Organization the model uses without reporting. Their connection is that the same structures participate in both.

The connection begins with simple interaction rules (SIRs). At the local level, repeated IS and IS-NOT operations produce DSRP structures around focal identities. As those structures recursively accumulate and interact, they generate higher-order constraints, expectations, and inferential possibilities that belong to the Organization as a whole rather than to any single structure. The rules are simple; their organized consequences need not be. Cognitive dark matter is therefore not a second hidden inventory added to contextual identity. It is the unreported part of that Organization's operation: the interacting contextual identities at work, minus what the model-holder can say about them.

The term does not imply that this Organization is immaterial. If cognition is physically realized, then the Organization enabling it must also be physically realized. The present claim identifies its structure and effects without specifying the neural format in which that Organization is implemented. It can be investigated by deriving contextual identity from a model and testing whether the inferred Organization predicts cognition or behavior beyond what direct report predicts.

This is not merely a philosophical question about what identity ultimately means. Every information system depends on decisions about what its identities are. Corporate databases require canonical definitions of customers, employees, products, assets, transactions, risks, and organizational units. Software systems assign identities to users, objects, variables, processes, permissions, and events. Cybersecurity systems must determine whether two records identify the same person, device, account, or threat—and which differences matter. Artificial-intelligence systems classify entities, resolve references, construct knowledge graphs, and transfer identities across changing contexts. Engineers, designers, military planners, scientific modelers, and institutional leaders likewise depend on identities whose boundaries, memberships, relationships, and framings must remain explicit enough to coordinate action. When these systems treat identity as merely a label or fixed attribute, differences in contextual identity reappear as ambiguous specifications, duplicate or incompatible records, brittle classifications, interoperability failures, security vulnerabilities, and operational disagreement. Contextual identity provides a systematic way to represent, compare, test, and revise the Organization that such systems already rely upon but rarely make complete. And it is entirely derivable from DSRP-484 structures—the formal statement of DSRP theory (four Patterns, eight Elements, four Dynamics), defined in Section 3, with its empirical base in [7] and its formal treatment in [1,3,4]. In other words, contextual identity already exists in the structuring of data and operates naturally throughout the pragmatic world; it need not be invented or added from outside.

The paper develops this argument in stages. Section 2 makes identity construction visible through a child learning the distinction between car and truck, then uses that developmental sequence to connect contextual identity to cognitive dark matter. Section 3 introduces the DSRP-484 architecture and shows why identity extends across all eight structural roles. Section 4 tests whether Organization can reveal what a person thinks when familiar lexical meaning is removed. Section 5 develops the advanced Is/Is Not

operation, formalizes contextual identity and the unidentified structural field, and explains how recursive DSRP Organization produces the emergent effects called cognitive dark matter. Section 6 shows how the resulting structures and statuses can be counted, compared, stored, tracked through time, and made empirically vulnerable. Section 7 puts the claim to a test on large language models: given only a definition of a nonce word, one model family answered as though the word named a road vehicle, a structure that appeared nowhere in the text and nowhere in its own account of what it believed; given the contextual identity, every family answered from it, and removing any one Pattern removed exactly the answers that depended on that Pattern. The machine test is a test of the structural claim, not of human neuroscience: it asks whether Organization does the choosing when Information is held fixed, in a system where both can be set exactly, and its results are read beside the lab's human studies (Section 7.4.1), not in place of them. Section 7.3.12 draws the nine results together. Section 8 states what follows for the theory and for the practical problem the paper began with: *finding out what another mind, human or machine, means by a word.*

The central claim is concise: an identity is not exhausted by its name or reported definition. It is constituted by an evolving field of IS and IS-NOT DSRP structures, and that field can continue to shape cognition even when the person cannot enumerate it. Contextual identity names the full structural account for a focal identity; cognitive dark matter names the part of the Organization these accounts produce that the model uses but does not report.

2. How a Child Builds an Identity

2.1 The sequence

A child does not learn what a truck is by being handed the identity *truck*. The child learns *car* first, and learns it by ostension repeated against contrast:

This is a car.
That is a car.
This is not a car.
That is not a car.

Then the space opened by the negatives acquires internal structure:

This is a fast car.
That is a slow car.
A fast car is not a slow car.

And then the new identity arrives inside that space, initially misfiled:

This is a truck.
A truck is a slow car.

Every step is a statement in the listing of Section 3. Nothing else is going on. The identity of *truck* is being assembled out of statements about *car*, and the majority of the load is carried by the negative statements, which is exactly the reverse of how identity is ordinarily described. To the child, something remarkable has happened: depending on the future context, a truck can be a car and not a car! Which is precisely how we as adults treat trucks. A truck is: a “car” – you purchase them at a “car dealership” from one of many on a “car lot” and park them at “car parks.” But it is also... “not a car.” We distinguish trucks from cars.

2.2 What the sequence shows

Three features of identity formation are visible in this simple sequence.

First, the label and the identity can come apart. A child may recognize a kind of thing before learning its name, or may learn a word before developing much structure behind it. Adults do the same. People routinely use words fluently while possessing only a thin understanding of what those words distinguish. Knowing the label *truck* is therefore not the same as possessing the identity organized under that label.

Second, learning reorganizes identity rather than merely adding information to it. “A truck is a slow car” is not random noise. It is a reasonable conclusion drawn from the distinctions available to the child at that moment. When the child later encounters a fast truck, the new experience does not merely add another fact to an otherwise unchanged concept. It forces a revision of how *truck*, *car*, *fast*, and *slow* have been organized in relation to one another. The experience creates the constructs of *fast truck* and *slow truck* and that a fast truck is not a slow truck. Yet, a fast truck and a slow truck are both trucks and sometimes cars.

The earlier understanding is therefore not simply an incomplete adult understanding. It is a differently organized understanding. Development may add content, but it also redraws boundaries, changes memberships, removes prior commitments, and creates new ones.

Third, the same object can be both “a car” and “not a car” without contradiction. A truck may count as a car when deciding where to buy it, park it, insure it, or service it. When the parent says, “pick up your cars,” the child is not prone to leave the trucks on the floor. It may not count as a car when vehicle classes, payload capacities, licensing requirements, or road regulations are at issue. What the truck is depends partly on the comparison being made and the context in which the distinction matters.

The child’s apparent inconsistency therefore exposes something that remains true in adult thought: identity is not fixed by a word alone. It is produced by an organized field of Distinctions that can change with context.

2.3 The structure behind the word

As the child develops, the word *truck* becomes situated within an increasingly elaborate structure. A truck becomes a kind of vehicle, something with wheels and an engine, something used to carry loads, something driven on roads, something associated with freight, work, licensing, fuel consumption, and

particular social or commercial purposes. It is distinguished from cars, vans, buses, tractors, toys, and countless other things, but not always in the same way or for the same purpose.

No single statement supplies the identity. The identity lies in how these statements are organized together.

This is why two people can use the same word while meaning different things. The child and the adult may both say *truck*, but the word occupies a different position within each person's understanding. The difference is not necessarily visible in ordinary conversation. It becomes visible when the two classify an unfamiliar vehicle, predict what it can do, explain why it belongs somewhere, or encounter a case that does not fit their expectations.

The child's learning makes this otherwise hidden process observable. Early statements are spoken aloud, tested against experience, and openly revised. In mature thought, the same kind of Organization operates largely without being enumerated. A person can recognize, classify, and reason about a truck without being able to list everything that makes it a truck.

The example therefore leaves the paper with a more precise question:

What is the organized structure that gives an identity its meaning in a particular context?

Answering that question requires more than the initial distinction between what something is and what it is not. It requires an account of the other structural positions through which a thing acquires its identity. Section 3 introduces that account.

2.4 Scope of the example

The sequence above is illustrative, not a claim that every child learns *car* before *truck*, that concept development follows a single invariant order, or that each spoken statement corresponds directly to a stored mental sentence. Nor is the general claim that contrast contributes to concept acquisition new; developmental psychology, structural linguistics, and contrastive learning all supply established versions of that insight.

Nearer neighbors deserve naming, because contextual identity will look familiar to readers of several literatures and differs from each in one specific way. Frame semantics [8] holds that a word is understood against a structured background of related roles; ad hoc categories [9] and the theory view of concepts [10] show that membership is built from the situation and from explanatory structure rather than from a definition; conceptual spaces [11] locate a concept in a geometry of quality dimensions; and a knowledge graph [12] stores an identity as a node with typed edges. Each supplies part of the account. None of them gives a structure a status. In contextual identity every Distinction, System, Relationship, and Perspective involving X is affirmed, negated, or unidentified, and the unidentified field is part of the identity—which is what makes the identity measurable (Section 6), lets a model say where its own gaps are (H9), and lets us ask what a model is using that it does not report (H6'). A frame has slots; a graph has edges; neither has a “?”. The machine study of Section 7 sits in the same way beside in-context learning [13,14], knowledge conflict between supplied context and prior knowledge [15,16], and hallucination [17]. Those literatures

ask whether a model follows the context it is given and what it adds; Section 7 asks which structures it follows, by Pattern and by status, and whether what it adds has the form of a contextual identity.

The contribution developed here begins with a different question: whether the Organization underlying an identity can be systematically derived, represented, and measured. The child example makes the problem visible. DSRP-484 [1,3,4] supplies the proposed structural account.

2.5 From Contextual Identity (CI) to Cognitive Dark Matter (CDM)

Section 1 introduced *contextual identity* as the Organization surrounding a focal identity: the accumulated structure specifying what it is and is not, what parts it has and what wholes contain it, what it is related to and how, and what frames it or is framed by it. The child makes the construction of this Organization unusually visible. His identity for *truck* is not delivered whole with the word. It is assembled through spoken IS and IS-NOT structures, provisional predictions, failed classifications, corrections, and reorganizations.

The adult possesses a far richer contextual identity for *truck*, but ordinarily cannot report more than a small fraction of it. Asked what a truck is, an adult may offer a short categorical definition: a truck is a motor vehicle used for carrying things. Yet that statement does not contain enough Organization to explain the adult's actual performance. It does not explain how the adult immediately distinguishes a truck from a car, van, tractor, wagon, bus, or toy; recognizes unfamiliar trucks; accommodates atypical cases; anticipates what trucks contain and what contains them; understands what trucks can do; or detects when the word has been used incorrectly or when it is being used differently in a context.

The missing Organization has not disappeared merely because it is absent from the answer. It remains active in recognition, memory, interpretation, prediction, judgment, and action. The adult does not need to consciously enumerate every boundary Distinction, nested System membership, Relationship, and Perspective framing involved in the identity of *truck* before deciding that something is—or is not—one. The Organization operates without first becoming an explicit account of itself.

The developmental story reveals the connection between the paper's two central constructs. Each IS and IS-NOT operation contributes to the child's contextual identities. Learning what truck is also reorganizes car, vehicle, fast, slow, and the Relationships among them. These identities develop through one another. Adding a distinction in one place can change what the child recognizes, expects, or questions elsewhere.

In the complex adaptive systems account developed here, repeated local DSRP operations generate Organization with consequences extending across the model. Contextual identity gives the full structural accounting for a focal identity; cognitive dark matter names the part of the Organization these interacting accounts produce that the model-holder uses but cannot report. The adult's inability to report that Organization is what makes it dark; recognition, prediction, and action are how its influence is detected. Recognition, prediction, and action can reveal structural commitments that an ordinary definition leaves unstated. The dark-matter analogy concerns this inference from observable effects to the Organization producing them.

The child makes visible what adulthood conceals. He constructs contextual identity in the light: speaking its structures, testing the predictions they generate, encountering contradictions, and reorganizing the model. With development and repeated use, much of that Organization ceases to appear as explicit content without ceasing to do cognitive work. The structures are assembled in the light and subsequently operate in the dark.

This section has shown the phenomenon in ordinary language and observable development. The following sections identify the DSRP structures from which contextual identity is assembled; Section 5 explains how their recursive Organization produces cognitive dark matter; and Section 6 shows how that Organization can be counted, compared, tracked, and tested.

3. From Distinction to Contextual Identity

Sections 1 and 2 establish that identity depends on contrast. Every distinguished identity coimplies an other, and the child's concept of *truck* develops through accumulating and revising statements about what a truck is and is not. If the analysis stopped there, however, contextual identity would amount only to an expanded account of Distinction.

DSRP-484 predicts a larger structure. The name denotes four Patterns, eight Elements, and four Dynamics. The four Patterns are Distinctions, Systems, Relationships, and Perspectives. Each Pattern is composed of two coimplicating elements:

- a Distinction has an identity and an other (D_i, D_o);
- a System has parts and a whole (S_p, S_w);
- a Relationship has action and reaction roles (R_a, R_r); and
- a Perspective has a point and a view ($P_v, P_{\dot{p}}$).

Formally, DSRP-484 is the four-tuple $\langle \mathcal{P}, e, \mathfrak{R}, \mathcal{D} \rangle$ [3,4]: the four Patterns $\mathcal{P} = D, S, R, P$; the eight Elements, two per Pattern, $e = (D_i, D_o), (S_p, S_w), (R_a, R_r), (P_{\dot{p}}, P_v)$; the four special relationships $\mathfrak{R}$ by which each Pattern is a structured equality between its two Elements (identity exists and other negates; part belongs and whole contains; action affects and reaction effects; point frames and view is framed); and the four Dynamics $\mathcal{D}$ named next. A DSRP structure is a triple $\sigma = (X, \rho, Y)$: a focal identity X , the Element-role $\rho \in e$ that X occupies, and the identity Y occupying the coimplied role. “X is part of Y” is (X, S_p, Y) ; “Y is a whole containing X” is the same structure focalized on Y, (Y, S_w, X) . Section 5 assigns every such structure a status in $+, -, ?$, and Section 6 counts them.

These structures are governed by four dynamics, $\mathcal{D} = \{=, \leftrightarrow, \sim, \in\}$: Equality, Coimplication, Simultaneity, and Enclosure [1,3,4]. Equality means that each Pattern is identical to the elemental structure that constitutes it — a structured equality between its two elements, mediated by the Pattern's special relationship τ : $\mathcal{P} = e^+ \overset{\tau}{\leftrightarrow} e^-$, so that $D = D_i \overset{\tau}{\leftrightarrow} D_o$, $S = S_p \overset{\tau}{\leftrightarrow} S_w$, $R = R_a \overset{\tau}{\leftrightarrow} R_r$, and $P = P_{\dot{p}} \overset{\tau}{\leftrightarrow} P_v$. Equality is important because it makes it clear that when we use the words for the

patterns and elements, we are not referring to the dictionary definitions of those words, but the specific and technical definitions supplied by [1,3,4]. Coimplication means that neither element of a Pattern can exist in that role without the other — if one exists, the other must exist even if it appears not to:

$e^+ \leftrightarrow e^-$; there is no identity without an other, no part without a whole, no action without a reaction, no point without a view, and the reverse. Simultaneity is somewhat more technical: it causes us to see the elements as roles and explains that any element can simultaneously act as any of the other elements:

$e \sim (\text{all other } e)$; a teacher is at once an identity, an other to the student, a part of a school, a whole of beliefs and behaviors, an actor, a reactor, a point, and a view. Enclosure means that in order to create any one of D, S, R, or P, you will need to sneak in DSRP, even if unaware of it: $\mathcal{P} \in (\text{all other } \mathcal{P}^*)$; any instantiated D requires S, R, P , any S requires D, R, P , any R requires D, S, P , and any P requires D, S, R , recursively, so that no Pattern exists on its own.

The consequence for identity is substantial. X is never only an identity distinguished from an other. Once distinguished, X simultaneously occupies part-whole, action-reaction, and point-view structures. What X is in a given context therefore depends not only on what is distinguished as X, but also on:

- what X contains and what contains X;
- what X affects, what affects X, and where no Relationship exists;
- what X frames, what frames X, and what remains outside those frames; and
- which other identities are included in or excluded from X.

The advanced Is/Is Not operation is the assembly of precisely these structures involving X. The resulting Organization is X's *contextual identity*: not a list of attributes attached to a self-contained object like a dictionary definition, but the structure through which X is distinguished, composed, related, and framed in a way that survives and creates multiple contexts.

Context is therefore not supplied beforehand as a container for the operation. It is what the operation yields. When X's boundaries, memberships, Relationships, or framings change, its context—and therefore its contextual identity—changes, even if its name does not.

4. The VORG Test: Can Organization Reveal What Someone Thinks?

The claim developed in Section 3 permits a direct test. If contextual identity lies in the Organization of information rather than in a label alone, then replacing familiar labels with unfamiliar ones should remove access to their conventional meanings without removing the Organization that constitutes the model. The analyst should still be able to determine something about the identities involved—and perhaps something the person who produced the model does not recognize about their own thinking.

Consider a mental model elicited from a person discussing a familiar topic. Before the model is analyzed, every substantive term is replaced with a nonce-name. The analyst is told nothing about what the names denote and receives only the following affirmed and negated statements:

1. VORG is VORG.
2. VORG is not GREEB.
3. VORG is part of TANDRA.
4. VORG is containing PIB, ZOOK, and NURR.
5. VORG is not part of KESSIL.
6. VORG is not containing OBB.
7. VORG is related to DRAYE.
8. *VORG is not related to QUELM.*
9. VORG is framed by DRAYE and GREEB.
10. VORG is framing GREEB, and DRAYE.
11. VORG is not framing QUELM.
12. VORG is not framed by QUELM.
13. VORG is framing TANDRA, PIB, ZOOK, and NURR.
14. VORG is not framed by TANDRA, PIB, ZOOK, or NURR.

Four additional structures remain unidentified:

1. Is VORG a DRAYE? — D
2. Is VORG part of DRAYE? — S
3. Is VORG related to GREEB? — R
4. Is VORG framing KESSIL? — P

These questions are not negative statements. The model does not say that VORG is not a DRAYE, is not part of DRAYE, is not related to GREEB, or is not framing KESSIL. It has not identified whether those structures are present or absent. Each question therefore marks a predictive opportunity: the existing Organization may support a hypothesis, but the structure must still be investigated.

The substitution removes the familiar referents of the labels; it does not reduce information to zero. The statements still identify objects and encode Distinctions (identity–other), Systems (part–whole), Relationships (action–reaction), and Perspectives (point–view), including which candidate structures are affirmed, expressly negated, or left unidentified.

This is therefore a semantic-blinding operation rather than a randomization of the underlying structure. In an experimental version, the nonce-names would be permuted across participants to prevent their phonetic qualities from systematically influencing interpretation.

4.1 Three states of structural knowledge

Before drawing any inference, three states must be distinguished:

1. **Present:** the model explicitly represents a structure. VORG is related to DRAYE.
2. **Explicitly absent:** the model explicitly negates a structure. VORG is not related to QUELM.
3. **Unrecorded:** the model says nothing about whether a structure exists. No Relationship between VORG and GREEB has been stated.

These three statuses apply to every candidate DSRP statement, not only to Relationships:

Table 2. The three statuses, by Pattern. Each candidate statement can be affirmed, expressly negated, or left unidentified. The third column is not the absence of a structure; it is a structure whose status is unresolved. The three symbols are used for these statuses from here on, and are the columns of the matrices in Section 6.

Pattern	Affirmed (+)	Expressly negated (-)	Unidentified (?)
Distinction	X is Y.	X is not Y.	Whether X is Y has not been identified.
System	X is part of Y.	X is not part of Y.	Whether X is part of Y has not been identified.
Relationship	X is related to Y.	X is not related to Y.	Whether X is related to Y has not been identified.
Perspective	X frames Y.	X does not frame Y.	Whether X frames Y has not been identified.

Where direction matters, each direction has its own status. “X frames Y” does not establish whether Y frames X; the reverse framing may be present, expressly absent, or unidentified.

These are three states of structural knowledge, not three states of the structure itself. Affirmation and explicit negation make claims about the modeled situation. Unidentified means that the model has not yet resolved which claim is true.

An unidentified structure is not the same as an explicitly absent structure. The first marks a gap in what the model has established; the second is itself an established part of the model. Treating every unidentified structure as a negation would convert the analyst’s ignorance into a claim about the person’s thinking. This distinction holds whether the unresolved question concerns what X is, what it belongs to or contains, what it is related to, or what it frames and is framed by.

This distinction also separates deductions from predictions. A present or explicitly absent structure can support a deduction about the represented model. An unrecorded but structurally salient possibility can generate a hypothesis. It cannot be silently promoted into a fact.

Both explicit negatives and unidentified structures can generate predictions, but they generate different kinds. An explicit negative supports a counterfactual question: what would change if the absent structure were instantiated? An unidentified structure supports a discovery question: does the structure exist but remain unrecognized or unrecorded?

Explicitly absent: What if it existed?

Unidentified: Does it exist?

An unidentified statement is therefore not an empty cell. It is a potential transition from (?) to either (+) or (-). Resolving it adds Organization to the model and thereby changes the context the model yields.

4.2 What the blinded model already reveals

Although the analyst does not know what any term denotes, the Organization is not silent.

DRAYE is VORG's strongest represented positive contact. VORG is related to DRAYE, is framed by DRAYE, and frames DRAYE in return. No knowledge of either referent is needed to determine that their structural coupling is unusually dense.

GREEB is constitutive of VORG despite being something VORG is not. GREEB occupies the other role in VORG's defining Distinction, frames VORG, and is framed by VORG. GREEB is therefore not an irrelevant non-VORG among indefinitely many possible non-VORGs. It is a structurally privileged contrast.

GREEB and QUELM are different kinds of other. GREEB is connected to VORG through Distinction and mutual framing. QUELM is explicitly excluded from both VORG's relational and framing profile. The model has differentiated the field of "not VORG" into internally different structural positions without supplying names for those positions.

The Relationship between VORG and GREEB is unresolved. Their mutual framing makes a Relationship between them a salient possibility, but the model neither asserts nor denies one. The Organization therefore generates a specific question:

Does a Relationship between VORG and GREEB exist but remain unrecorded?

This is a prediction target, not a deduction.

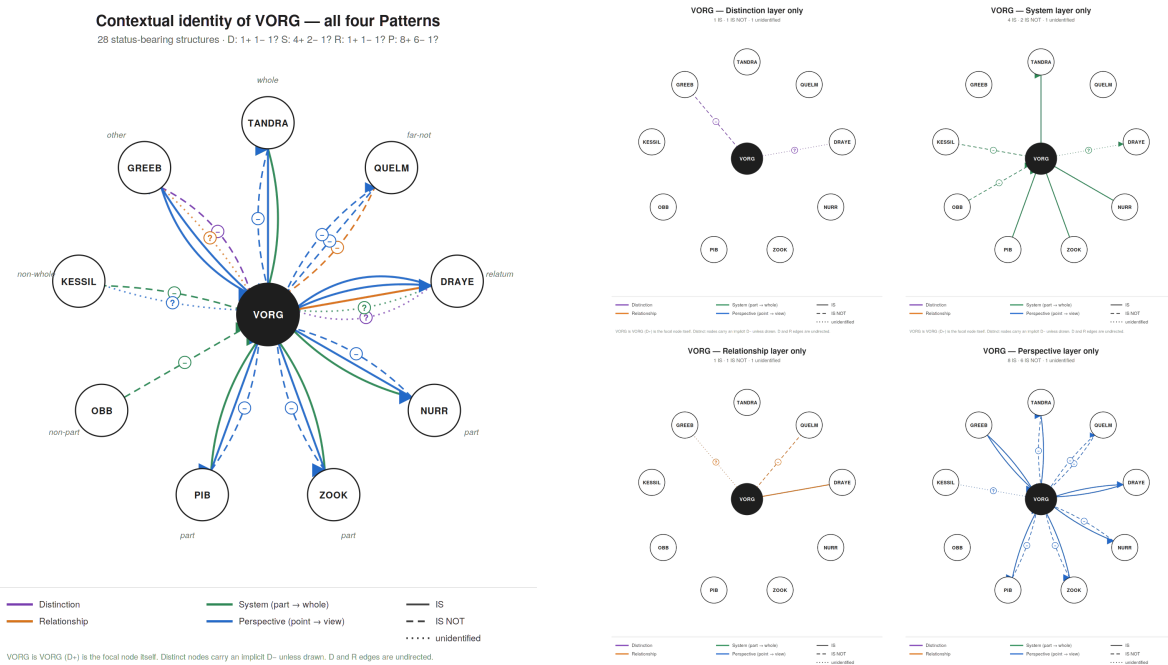


Figure 1. The Contextual Identity of v_{ORG} as a Status-Bearing Structural Graph, Combined and by Pattern. Left: every structure in the Section 4 listing drawn as one edge between the focal identity (v_{ORG} , center) and another identity in the model. Edge color identifies the DSRP Pattern – Distinction (purple), System (green), Relationship (orange), Perspective (blue). Edge stroke identifies status – solid for affirmed (+), dashed for expressly negated (-), dotted for unidentified (?). Arrowheads mark structural roles where the model establishes it: part $\rightarrow$ whole for System, point $\rightarrow$ view for Perspective; Distinction and Relationship edges are undirected. The reflexive structure “ v_{ORG} is v_{ORG} ” (D+) is

the focal node itself, and distinct nodes carry an implicit D— that is drawn only where the other is constitutive. Right: the same graph separated into its four Pattern layers with node positions held constant — Distinction (1+ 1– 1?), System (4+ 2– 1?), Relationship (1+ 1– 1?), Perspective (8+ 6– 1?). Each layer is one row of the status signature $N_{\Pi}(\text{VORG}, t_0)$; together they contain all 28 structures of Section 6.3. The figure makes visible what the listing requires the reader to assemble: DRAYE is the densest pair (five edges, two unresolved); QUELM is the only identity with no solid edge, which is the unawareness signature; GREEB and QUELM are both “not VORG” yet occupy entirely different structural positions; the profile is Perspective-heavy (fifteen of 28 structures); and the Distinction and Relationship layers have identical counts while connecting different identities — which is why the count signature summarizes the profile but does not replace it.

The same operation applies across all four Patterns. The model has not identified whether VORG is also a DRAYE-type identity, whether VORG contains DRAYE or belongs to a whole containing DRAYE, or whether KESSIL frames VORG. Each is an unresolved D, S, or P statement rather than an implicit negative.

Not every unidentified statement is equally worth pursuing. The existing Organization makes some unknowns more salient than others. Mutual framing makes the unidentified Relationship between VORG and GREEB salient. Dense positive contact with DRAYE makes additional D or S structure involving DRAYE worth testing. A repeated configuration with one unresolved position may predict the structure occupying that position.

Prediction therefore occurs where the established Organization makes an unidentified structure consequential:

$$\boxed{\text{existing } O + \text{salient ?} \rightarrow \text{testable (+ or -)}}$$

The model does not determine the answers in advance. It identifies structures that remain unspecified and uses the structures already identified to generate testable predictions about them.

Semantic blinding isolates Organization. Replacing familiar terms with nonce names does not eliminate information altogether: VORG, GREEB, and the other terms remain distinguishable objects. What disappears is their familiar semantic content. This makes the represented Organization unusually visible—the opposite of ordinary interpretation, in which semantic familiarity dominates while Organization often remains unnoticed.

When the terms are unblinded, their familiar meanings become available again. The same Organization now operates on semantically meaningful information, allowing predictions that depend on both (I) and (O). This does not contradict the claim that contextual identity is constituted by Is/Is Not structure. The meaning restored by unblinding is not structure-free information; it provides access to additional Is/Is Not structures that were not included in the blinded listing.

The distinction is evidential. Nothing in the displayed Organization establishes that OBB belongs to KESSIL, that DRAYE controls VORG, or that QUELM can be harmed by VORG. Unblinding may supply additional information that makes such claims plausible, but they still do not follow from the displayed Organization alone. Semantic blinding therefore separates predictions generated by the represented structure from predictions that depend on additional knowledge.

4.3 A statement about the model-holder's own belief

Now suppose the person from whom the model was elicited makes the following explicit statement:

“GREEB plays no role in what I mean by VORG.”

The claim may be true or false. Nothing about the sounds of the two words settles it. The question is whether the person's underlying $\mathbb{M} = IO$ can answer it.

“Plays no role” can be given a structural meaning. If GREEB plays no role in the person's concept of VORG, removing GREEB should leave the Organization of VORG unchanged:

$$\mathbb{M}_{VORG} = \mathbb{M}_{VORG \setminus GREEB}$$

But that is not what the elicited model shows. Removing GREEB eliminates:

- the named other in VORG's Distinction;
- a frame acting on VORG; and
- an object framed by VORG.

The resulting Organization is therefore different:

$$\mathbb{M}_{VORG} \neq \mathbb{M}_{VORG \setminus GREEB}$$

The model answers the question without knowing what either word means. Relative to the elicited $\mathbb{M} = IO$, the person's statement is false. GREEB plays several roles in constituting VORG, including the role of the contrast against which VORG is distinguished.

This is not merely a disagreement between two verbal statements. The person's explicit report says that GREEB is irrelevant; the Organization elicited from that same person says that removing GREEB would alter what VORG is.

$$\text{reported belief} \neq \mathbb{M}_{\text{operative}}$$

4.4 From structural contradiction to behavioral prediction

The contradiction alone does not yet establish that the elicited model represents what the person “actually thinks.” One elicitation could be incomplete, unstable, or incorrectly coded. The stronger test is whether the Organization predicts an independent judgment or behavior better than the person's self-report does.

The operative model generates a clear prediction: when the person must identify, classify, explain, or reason about an ambiguous VORG, GREEB should influence the response. The effect might appear in which borderline cases are accepted as VORGs, which contrasts are used to explain VORG, how quickly ambiguous cases are classified, or how judgments change when GREEB is made salient or removed from consideration.

The reported belief predicts no such effect. If GREEB truly plays no role, manipulating its presence should not change judgments about VORG.

The two accounts therefore make competing predictions:

Reported belief: $\Delta \text{GREEB} \rightarrow \Delta \text{judgment of VORG} = 0$

Operative $\mathbb{M} = IO$: $\Delta \text{GREEB} \rightarrow \Delta \text{judgment of VORG} \neq 0$

The result is falsifiable in either direction. If judgments about VORG change systematically with GREEB, the operative Organization has predicted cognition better than introspection. If they do not, the structural interpretation has failed despite its apparent internal coherence.

This is the consequential possibility raised by contextual identity. A person may accurately report the propositions immediately available to awareness while remaining unable to report the larger Organization that makes those propositions meaningful. Their stated belief may describe what they consciously endorse; their $\mathbb{M} = IO$ may better predict what they distinguish, notice, infer, and do.

The nonce-names make that discrepancy visible because they prevent the analyst from repairing the model with familiar content. Whatever can be inferred before the terms are unblinded comes from the Organization itself. Context is not supplied to the analysis from outside. It emerges as the structured set of identity-other distinctions, part-whole containments, action-reaction relationships, and view-point framings through which VORG acquires its identity.

4.5 Unblinding the model: a test for the reader

Before the terms are unblinded, consider the following statement—not as a report of what the elicited model contains, but as a claim on which you might have an opinion:

VORG is not related to QUELM.

Does the statement seem true, false, plausible, implausible, or contestable? Notice whether you have an immediate judgment, how confident you are, and whether any reason or counterexample comes to mind.

Now consider the same statement with its familiar terms restored:

LEADERSHIP is not related to TECHNICAL EXPERTISE.

Did anything happen? Perhaps you immediately agreed, objected, qualified the claim, imagined an effective leader without technical expertise, or recalled a technically accomplished person who could not lead. Perhaps you began distinguishing leadership from management, expertise from influence, or necessary conditions from helpful ones.

The two statements have exactly the same represented Organization:

VORG is not related to QUELM

LEADERSHIP is not related to TECHNICAL EXPERTISE

Only the information occupying that Organization has changed. Yet the second statement is far more likely to produce a substantive reaction. The reader supplies identity-other distinctions, part-whole memberships, action-reaction relationships, view-point framings, examples, exceptions, and prior judgments that were unavailable when the same structure contained nonce-names.

That reaction is contextual identity operating in real time. The words *leadership* and *technical expertise* activate Organization assembled through the reader's prior experience, much of which was neither stated on the page nor explicitly retrieved before the judgment occurred. What was invisible in the nonce statement becomes influential as soon as familiar information occupies the same structure.

This brief exercise does not reveal exactly which prior structures produced the reaction, nor does the reaction establish whether the statement is true. It demonstrates the evidential problem: readers routinely contribute unrecorded Organization to semantically familiar claims while experiencing the resulting judgment as an immediate response to the words themselves. Semantic blinding temporarily suppresses that contribution, allowing the displayed Organization and the Organization supplied by the reader to be distinguished.

The complete unblinding of the nonce-names above and their substitutions is described in detail in the next section.

4.6 Unblinding the model

The nonce-names can now be restored to familiar terms. Suppose the listing was elicited from an executive explaining what *leadership* means:

Table 3. Unblinding the listing. The nonce-names of the Section 4 listing and the familiar terms they stood for, as elicited from an executive explaining what leadership means.

Nonce-name	Referent
VORG	LEADERSHIP
GREEB	MANAGEMENT
TANDRA	EXECUTIVE WORK
PIB	SETTING DIRECTION
ZOOK	ALIGNING PEOPLE
NURR	MOTIVATING CHANGE
KESSIL	ADMINISTRATION
OBB	PAPERWORK

DRAYE	INFLUENCE
QUELM	TECHNICAL EXPERTISE

The unblinded model reads:

- LEADERSHIP is LEADERSHIP.
- LEADERSHIP is not MANAGEMENT.
- LEADERSHIP is part of EXECUTIVE WORK.
- LEADERSHIP is containing SETTING DIRECTION, ALIGNING PEOPLE, and MOTIVATING CHANGE.
- LEADERSHIP is not part of ADMINISTRATION.
- LEADERSHIP is not containing PAPERWORK.
- LEADERSHIP is related to influence.
- LEADERSHIP is not related to TECHNICAL EXPERTISE.
- LEADERSHIP is framed by INFLUENCE and MANAGEMENT.
- LEADERSHIP is framing EXECUTIVE WORK, SETTING DIRECTION, ALIGNING PEOPLE, MOTIVATING CHANGE, MANAGEMENT, and INFLUENCE.
- LEADERSHIP is not framing TECHNICAL EXPERTISE.
- LEADERSHIP is not framed by TECHNICAL EXPERTISE.
- LEADERSHIP is framing EXECUTIVE WORK, SETTING DIRECTION, ALIGNING PEOPLE, and MOTIVATING CHANGE.
- LEADERSHIP is not framed by EXECUTIVE WORK, SETTING DIRECTION, ALIGNING PEOPLE, or MOTIVATING CHANGE.

These are not claims made by this paper about leadership. They are the claims represented in this person's model.

The four unidentified structures also become recognizable questions:

1. Is leadership influence?
2. Is leadership part of influence?
3. Is leadership related to management?
4. Is leadership framing administration?

Unblinding changes none of the displayed Organization. It restores semantic information to the identities occupying that Organization. Questions that were structurally precise but semantically opaque now become interpretable and empirically investigable. The Organization makes the unidentified Relationship between leadership and management salient; the meanings of *leadership* and *management* indicate what evidence could resolve it.

The model-holder's earlier statement can now be understood in ordinary language:

“Management plays no role in what I mean by leadership.”

Yet management occupies several positions in the person's represented identity of leadership. Leadership is distinguished from management, framed by management, and framing management in return. Removing management would therefore change the represented Organization of leadership.

The self-report and the operative model make competing behavioral predictions:

Reported belief: $\Delta_{\text{management}} \rightarrow \Delta(\text{judgment of leadership}) = 0$

Operative $\mathbb{M} = IO$: $\Delta_{\text{management}} \rightarrow \Delta(\text{judgment of leadership}) \neq 0$

This difference can be tested using ambiguous executive behaviors. The person might be asked whether setting deadlines, reallocating staff, resolving a personnel conflict, or implementing an existing policy counts as leadership. The same cases can then be judged while management is made more or less salient. Because management occupies the explicit other position relative to leadership, the operative model predicts a directional effect: when management is salient, management-like cases should be less likely to be classified as leadership.

If the judgments remain invariant, the operative model has failed to predict behavior. If the classification boundary shifts systematically, the person's represented Organization predicts their judgment better than their explicit report does.

Unblinding therefore does not introduce a second source of meaning independent of contextual identity. A familiar word provides access to additional information, but that information is itself organized through further Is/Is Not structures. In the blinded model, the analyst can see the Organization displayed in the listing while most familiar semantic information is unavailable. In the unblinded model, that semantic information becomes available and operates through the same Organization. What changes is not the mechanism but the amount of contextual identity accessible to the analysis.

5. Our Gray Matter Has Dark Matter

The person in Section 4 is not unusual. The leadership example is a small demonstration of a general human condition: the Organization operating in cognition can be much larger than the Organization available to introspection. People recognize, classify, infer, disagree, decide, and act through contextual identities they cannot fully describe. Section 2 introduced cognitive dark matter through development. This section now identifies its structural composition and explains how it arises from contextual identity.

The reader's response to the unblinded statement demonstrates the effect directly. "VORG is not related to QUELM" may have produced little reaction. "LEADERSHIP is not related to TECHNICAL EXPERTISE" may have produced immediate agreement, disagreement, qualification, examples, or counterexamples. The Relationship did not change. Familiar Information was placed into it.

That Information arrived with Organization accumulated through prior experience. The words LEADERSHIP and TECHNICAL EXPERTISE activated distinctions, parts, wholes, Relationships, and framings that were not

written on the page. The reader used that Organization without first enumerating it and may remain unable to enumerate more than a fraction of it afterward.

This is the sense in which our gray matter has dark matter.

5.1 Dark does not mean absent

Dark matter in physics does not become real when a scientist detects it. Its existence is an ontological question; whether and how it can be detected is an epistemological question. Physicists infer dark matter from effects that otherwise do not add up. Something is exerting gravitational influence even though the thing producing that influence is not directly visible.

The analogy here has the same form. Cognitive dark matter concerns the physically realized Organization operating in a model: the boundary distinctions, parts-wholes nestedness, relationships, and perspectives through which its identities acquire meaning. That Organization affects what a person notices, expects, remembers, says, feels, and does whether or not the person can describe it.

Making some of that Organization visible does not bring it into existence. It is part of the meaning model that already exists. It was already there and already exerting effects. Nor does it cease to be part of cognitive dark matter once someone sees it. What changes is not its *existence* but its *inspectability*.

This distinction matters because hidden Information is already familiar. We have names for it: a hidden agenda, a white lie, a covert operation, a secret, a confidence, a need-to-know restriction, an eyes-only document. In each case, someone possesses Information and deliberately or accidentally withholds it.

Hidden Organization is different. People regularly conceal it without knowing they are concealing anything. They may report all the Information they consciously possess while remaining unable to describe how that Information is organized. What remains unseen is not necessarily another fact. It may be the boundary being drawn, the whole being assumed, the relationship being expected, or the point from which the situation is being framed.

Cognitive dark matter is therefore principally hidden *O*, not merely hidden *I*.

5.2 Most often, it begins with a statement

Most cases begin simply: someone says something.

A manager says, “Management has nothing to do with how I understand leadership.” A patient says, “I am not angry.” A child says, “I don’t care whether I made the team.” A customer says, “Price is the only thing that matters.” We ordinarily call these things statements—what people say.

A statement, however, is one member of a broader class of *outputs*. Models produce words, behavior, emotion, code, and data. A person may say that something does not matter and then organize every

subsequent decision around it. Software documentation may specify $X \rightarrow Y$, while the interface repeatedly produces $X \rightarrow Z$. A dataset may be described as neutral while its categories systematically place some cases inside a boundary and others outside it. Each output carries Information, but it also bears the effects of the Organization that produced it.

Communication compresses a richly organized model into a narrow stream of outputs. A person holds a world of distinctions, memberships, relationships, expectations, and perspectives, then sends a few words through the thin tube of speech. The listener receives those words and reconstructs their meaning using another organized model. Much of what made the words meaningful to the speaker is never transmitted explicitly. Two people can therefore exchange the same sentence while organizing quite different meanings beneath it.

This difficulty persists even when someone knows that Organization matters. Conversations, headlines, forms, and databases rarely expose every structural commitment on which their meaning depends. The natural world also presents this problem: the visible boulder does not display every process that brought it there. Relationships and history may have to be inferred, measured, or investigated. Organization can become visible through a revealing question, a map, an observation, or an unexpected consequence, while much of it remains outside the account currently available.

The notation is O-Theory's [1,3]. A model $\mathbb{M}$ is Information I —words, data, etc.—organized by O —the DSRP structures over them—written $\mathbb{M} = I \times O$; the first index names the model and t the time.

Let the source model be:

$$\mathbb{M}(1, t) = I(1, t) \times O(1, t)$$

and let its observable output be $y(1, t)$:

$$\mathbb{M}(1, t) \longrightarrow y(1, t)$$

where $y(1, t)$ may be an utterance, action, emotion, code path, interface behavior, measurement, or datum. In ordinary human interaction, it is often simply “what the person said.”

The output is not the source model itself. When it reaches a second model, it arrives as Information:

$$y(1, t) \longrightarrow I(2, t)$$

The second model must organize that Information to make it meaningful:

$$\mathbb{M}(2, t) = I(2, t) \times O(2, t)$$

This is what we ordinarily call “*listening*.” Listening is one model building a model of another model from the outputs available to it, usually with the intention (hopefully) of understanding it or matching its meaning.

The second model may belong to another person: a friend, spouse, parent, therapist, physician, investigator, teacher, or colleague. It may belong to the same person examining their own thoughts through introspection. It may also be nonhuman: a diagnostic instrument, software debugger, statistical model, or artificial intelligence system. In every case, $\mathbb{M}(2, t)$ does not inspect $\mathbb{M}(1, t)$ directly. It organizes outputs from $\mathbb{M}(1, t)$ and uses them to construct an estimate.

That limitation operates on both sides. The source model contains Organization it may not recognize. The listening model also contains Organization it may not recognize. A therapist has cognitive dark matter. So does a physician, parent, scientist, judge, programmer, or expert. Expertise may improve a model, but it does not exempt the model from being a model.

5.3 Contextual identities make up cognitive dark matter

For any focal identity X , its *contextual identity* is the assembly of all the “ X is...” and “ X is not...” structures present in the model. In this sense, contextual identity is composed entirely of *IS-statements*. The difference is between the conventional, dictionary-like way we identify something and the richer way identity is actually constructed in context.

A dictionary typically identifies X categorically: “leadership is guidance,” “a dog is a mammal,” or “water is a liquid.” Such definitions are useful, but they capture only a small portion of what the model says X is. Contextual identification treats every structural claim involving X as contributing to its identity. Thus, “ X is...” includes not only what X is called or classified as, but the complete DSRP profile surrounding X :

1. what X **is** _____
2. what X **is** not _____
3. what X **is** part of _____
4. what X **is** not part of _____
5. what X **is** a whole containing _____ (what its parts are)
6. what X is not a whole containing _____ (what are not parts of it)
7. what X **is** acting upon _____ (related to by acting upon)
8. what X **is** reacting to _____ (related to by reacting to)
9. what X **is** not acting upon _____ (not related)
10. what X **is** not reacting to _____ (not related)
11. what X **is** viewing _____ (perceiving, seeing, framing, constraining, filtering, etc)
12. what X **is** not viewing _____
13. what X **is** viewed by _____ (perceived by, framed by, constrained by, filtered by, etc)
14. what X **is** not viewed by _____

These may look grammatically different on the surface, but structurally they are all identity-bearing *IS-statements* and are therefore a part of X 's *identity*. “ X has part p ” says, in effect, “ X is a whole containing p .” “ X causes Y ” says “ X is a relata occupying a particular role in a relationship with Y .” “ X is viewed from v ” identifies X as the framed object of a particular Perspective. The contextual identity of X is therefore not merely its definition¹ but the total structured set of claims specifying what X is in this model.

Let σ denote a DSRP statement whose focal identity is X . Its status in model $\mathbb{M}$ at time t is:

$$s(\sigma; \mathbb{M}, t) \in \{+, -, ?\}$$

where $+$ means the model affirms the structure, $-$ means the model expressly negates it, and $?$ means that the structure's Information or status has not yet been identified.

The contextual identity of X consists of the affirmed and expressly negated IS structures (O) present in the model:

$$CI(X; \mathbb{M}, t) = \{\sigma : \text{focal}(\sigma) = X \wedge s(\sigma; \mathbb{M}, t) \in \{+, -\}\}$$

The word *present* matters. It avoids an ambiguity hidden inside the word *true*. Two different claims are easily confused:

1. It is true that statement σ is present and operative in model $\mathbb{M}$.
2. Statement σ accurately represents Reality.

Those claims are not equivalent:

$$\sigma \in \mathbb{M} \not\Rightarrow \sigma \cong \mathbb{R}$$

A person may genuinely organize a MARKET AS PRICE-SENSITIVE even when the market is not price-sensitive. The model is real. Its Organization is operative. It can produce real decisions and real consequences. Its content may nevertheless be poorly aligned with Reality.

If model $\mathbb{M}$ contains n identities, its Organization is the structured aggregation of their contextual identities:

$$O(\mathbb{M}, t) = \bigoplus_{j=1}^n CI(X_j; \mathbb{M}, t)$$

The symbol $\bigoplus$ is the union of the n contextual-identity sets taken over focalized structures:

¹ It may not even be its dictionary definition at all. It may be a word used incorrectly but still possessing meaning. A word used differently with purpose or even a word used as code to purposefully mean something different than its dictionary form.

$$O(\mathbb{M}, t) = \bigcup_{j=1}^n \{\sigma : \text{focal}(\sigma) = X_j \wedge s(\sigma; \mathbb{M}, t) \in \{+, -\}\}$$

Because each structure carries its focal identity (Section 3), X 's side of a Relationship and Y 's side are distinct members and are not merged. The symbol $\bigoplus$ is used rather than $\bigcup$ only to mark that the members are focalized structures, not unfocalized edges, so that focal position, role, direction, and perspective context are preserved.

Cognitive dark matter is the part of that Organization the model uses that is absent from the model's own account of itself. Let $I(O)(\mathbb{M}, t)$ be the structures in $O(\mathbb{M}, t)$ that the model states when asked what its identities are — its report of its own Organization. Then $\text{CDM}(\mathbb{M}, t) = O(\mathbb{M}, t) \setminus I(O)(\mathbb{M}, t)$, written $O - I(O)$: the structures that are operative, with status $+$ or $-$ and used in inference, and not reported. Cognitive dark matter is therefore a status of Organization relative to an account, not a second kind of Organization. The same structure is dark when it is used and unsaid and lit when it is used and said; its amount is the share of operative structures the account omits (Section 7.2.7), and it can be negative when the account states structures the model does not use.

This preservation matters. In $\text{CI}(X)$, a Relationship may appear as:

X is related by R to Y

while in $\text{CI}(Y)$, it appears as:

Y is related by R to X

These are not disposable duplicates. The first focalizes X ; the second focalizes Y . Their difference can carry significant point-view and view-point context, including differences in direction, salience, action, reaction, and framing.

Contextual identity and *cognitive dark matter* should therefore remain conceptually distinct. *Contextual identity* is the Organization surrounding a focal identity. The model's Organization, O , is the aggregate formed by the contextual identities of all n identities in the model; cognitive dark matter is the part of that aggregate the model uses without reporting. If X is itself the model, recursively assembling $\text{CI}(X)$ may recover the same total Organization, but the local and model-level concepts remain analytically useful. The rules are simple; the resulting identity is not. *Contextual identity* arises from recursively assembled IS/IS-NOT DSRP structures, while cognitive dark matter is what their interaction produces and the model-holder does not say.

This is where the account connects simple structural rules to more complex capacities. A contextual-identity profile records structural commitments. When a model uses those commitments together, it can derive consequences, identify consequential gaps, and detect incompatibilities. These are

the beginnings of prediction, inquiry, and critical judgment. They concern what the organized model can do; they are not additional primitive entries appended to an identity's IS/IS-NOT profile. In the account proposed here, they emerge through the interaction of Information and Organization. An answer can generate a new question, a question can expose an assumption, and revising that assumption can reorganize several identities at once. The resulting intelligence depends on this continuing interaction.

5.4 Structural predictions can become questions

DSRP-484 does more than provide four categories for describing completed models. It generates *structural predictions* (predictions that structures must exist based on structures that currently exist). Given an identity or structure, the calculus predicts additional structures that can be sought, specified, manipulated, and tested.

For focal identity X , DSRP predicts, among many other things, that:

- distinguishing X entails an other against which X is distinguished;
- treating X as a whole entails parts, while treating it as a part entails some containing whole;
- placing X in a Relationship entails relata, action, and reaction;
- framing X entails a point and a view; and
- altering any of these structures may alter the contextual identity of X and its downstream effects.

Let the structural predictions generated from a model be:

$$\Pi(\mathbb{M}, t) = \Pi(D; \mathbb{M}, t) \cup \Pi(S; \mathbb{M}, t) \cup \Pi(R; \mathbb{M}, t) \cup \Pi(P; \mathbb{M}, t)$$

These predictions are not merely abstract propositions. Every one can be turned into a question or another form of inquiry:

$$q : \Pi(\mathbb{M}, t) \longrightarrow \mathcal{Q}(\mathbb{M}, t)$$

- “What is X ?” asks for the contextual identity of X .
- “What is X not?” uses the structural prediction that an identity has an other and asks for the relevant Information.
- “What are the parts of X ?” and “What contains X ?” turn System predictions into inquiry.
- “How are X and Y related?” predicts Relationship structure while leaving the particular Relationship unidentified.
- “What does X do to Y , and what does Y do to X ?” looks for the action–reaction structure of that Relationship.
- “From what point is X being viewed?” and “What does X frame?” turn Perspective predictions into inquiry.

A question is therefore a structural prediction turned around and used to inquire.

Structural predictions and hypotheses are closely related, but they are not interchangeable. DSRP generates predictions whether their case-specific content is already known or remains unidentified. The unresolved predictions form a subset that can function as hypotheses:

$$\mathcal{H}(\mathbb{M}, t) = \{\pi \in \Pi(\mathbb{M}, t) : s(\pi; \mathbb{M}, t) = ?\}$$

where, $\mathcal{H}(\mathbb{M}, t)$ is the set of all structures π in $\Pi(\mathbb{M}, t)$ whose status $s(\pi; \mathbb{M}, t)$ is ? That is, structures the model has not yet identified as either + (IS) or - (IS NOT).

Thus:

$$\mathcal{H}(\mathbb{M}, t) \subseteq \Pi(\mathbb{M}, t)$$

Every hypothesis in this formulation is grounded in a structural prediction, but not every structural prediction is an unresolved hypothesis.

The question mark is curiosity. It says, “The structure predicts that something belongs here, but I do not yet know what.” The question mark does not itself constitute cognitive dark matter in the source model. It is a hypothesis generated by the listening model, and the listening model may be wrong.

This is why the distinction matters. If a therapist thinks, “Something about that explanation does not sit right,” the feeling is not proof that the patient is hiding something. It may reflect an unresolved distinction, a missing whole, an expected relationship, a shift in perspective—or an error in the therapist’s own model. The same applies when a parent thinks, “I want to come back to that later,” or when a colleague says, “What that guy said was a little suss.” These are ordinary experiences of structural prediction occurring before the person can identify its mechanism.

The person may never formulate a verbal question. They may only feel that something does not add up. The mathematical machinery is still operating. The farmer notices that the fence line looks wrong; the scientist specifies the boundary discrepancy. They are encountering the same Organization at different levels of articulation.

The predictions are falsifiable because particular proposed structures produce consequences. If X is claimed to be a part of W , removing or changing X should affect W in predicted ways. If X is claimed to cause Y , interventions on X should produce predicted changes in Y . If a Perspective is claimed to frame a judgment, changing the point should systematically change the view. If the predicted effects do not occur, the proposed structure has failed its test.

DSRP therefore generates both predictions and practical inquiry forms for testing them. It does not guarantee that every prediction made by a particular modeler is correct.

5.5 Inquiry changes the model

Asking a question is not a neutral act of opening a window into an unchanged mind. A question becomes new Information for the model receiving it:

$$q(\pi) \longrightarrow I(1, t + 1) \longrightarrow \mathbb{M}(1, t + 1) = I(1, t + 1) \times O(1, t + 1)$$

The question may prime distinctions, activate parts and wholes, make relationships salient, or introduce a perspective that was not active before. The answer therefore comes from $\mathbb{M}(1, t + 1)$, not from an untouched $\mathbb{M}(1, t)$.

A therapist asks, “What did you think he meant when he said that?” The question may reveal an existing Perspective, but it may also make that Perspective more explicit, reorganize it, or create a new one. A parent asks, “If making the team did not matter, why did you stop talking to your friends afterward?” The child may discover a Relationship she had not consciously recognized. A manager challenged about a contradiction may revise the boundary of a term while answering and then sincerely remember the revised position as what they meant all along.

Sometimes people manipulate. Often they simply lack awareness of their own re-Organization. The mechanism operates in both cases.

This means that the listening model cannot treat a later answer as a transparent picture of the earlier model. It can compare successive estimates:

$$\widehat{\text{CI}}(X; 1 \mid 2, t) \quad \text{and} \quad \widehat{\text{CI}}(X; 1 \mid 2, t + 1)$$

The hat indicates that these are $\mathbb{M}(2)$ ’s reconstructions of $\mathbb{M}(1)$, not direct access to $\mathbb{M}(1)$. Differences across time may reveal a shifting boundary, a newly activated whole, a revised Relationship, a changed Perspective, or an unidentified variable. They do not, by themselves, establish deception.

This is why skilled listening often involves returning to something later. The therapist, parent, investigator, programmer, or scientist allows the model to produce additional outputs and then compares their Organization. The code says one thing, the interface does another, and the error log does a third. The patient’s words, emotional response, and later behavior organize the event differently. Each new output supplies another partial view.

Inspectability is therefore essential—but inspection must include the Organization of the observer. Cognitive dark matter operates in $\mathbb{M}(1)$ and $\mathbb{M}(2)$. The expert’s interpretation is another model, not Reality speaking through a credential.

The good news is that this capacity is trainable. People can practice making distinctions explicit, changing boundaries, identifying parts and wholes, specifying Relationships in both directions, and deliberately shifting points and views. O-Therapy [18] is a clinical application of O-Theory with published standards

and protocols; it applies the same process to the Organization of thought, emotion, behavior, identity, and relationships. It helps make operative Organization inspectable so that it can be tested and, where useful, reorganized.

The everyday entry point is simple: practice noticing the Organization already being used. A child, parent, colleague, or scientist can ask: What am I identifying, and what am I excluding? What belongs within this whole? What is acting on what, and what happens in return? From whose point does this view make sense? Which of these structures did I observe, which did I derive, and which did I assume? The same questions can be asked of a disagreement, a news story, a design, or one's own explanation. This is the metacognitive practice of $O = \text{DSRP}$.

Such practice must be renewed because every encounter presents Organization that has not been fully stated. Expertise does not remove that condition. The practical aim is to make the structures relevant to a judgment available for examination and revision. The capacity that supports useful prediction can also support bias when an assumed boundary, relationship, membership, or perspective silently becomes the standard against which new Information is judged.

This trainability is especially important because DSRP operates ontologically at all times while human epistemology displays systematic imbalance. At the ontological level:

$$D = S = R = P$$

and:

$$D_i = D_o, \quad S_p = S_w, \quad R_{\text{gen}} = R_a = R_r, \quad P_v = P_p$$

Yet the Pareto Bias [19]—the empirical finding that people make Distinctions far more often than Systems, Systems more often than Relationships, and Relationships more often than Perspectives—shows a strong epistemological tendency toward:

$$D \gg S \gg R \gg P$$

and:

$$D_i \gg D_o, \quad S_p \gg S_w, \quad R_{\text{gen}} \gg R_a, R_r, \quad P_v \gg P_p$$

People tend to privilege identities over others, parts over wholes, generalized Relationships over action–reaction specification, and views over the points producing them. The neglected structures do not disappear. They become especially likely places for Organization to remain dark.

The tendency is not destiny. It is a trainable bias.

5.6 Reality is the only final test

A model is itself real. It exists as part of Reality even when what it represents is poorly aligned with Reality:

$$\mathbb{M}(t) \text{ is part of } \mathbb{R}(t), \quad \mathbb{M}(t) \neq \mathbb{R}(t)$$

This is why a “false” model still matters. A mistaken belief about a market can move money. A mistaken diagnosis can alter treatment. A mistaken belief about a spouse can produce anger, withdrawal, accusation, or divorce. A bug in software is real Organization even though it incorrectly represents the intended operation. Models produce actions, and actions alter Reality:

$$\mathbb{M}(t) \longrightarrow A(t) \longrightarrow \mathbb{R}(t+1)$$

The deeper question is not merely whether a statement is present in a model. It is how well the model’s Information and Organization fit the relevant Information and Organization of Reality.

Conceptually, that alignment can be represented as:

$$\rho(t) = \text{fit}_{I \times O}(\mathbb{M}(t), \mathbb{R}(t)), \quad 0 \leq \rho(t) \leq 1$$

A value near 0 indicates very poor fit; a value near 1 indicates very high fit. Because no finite model has access to Reality as a whole, $\rho(t)$ cannot be inspected directly. In practice, confidence in fit must be estimated from feedback.

Fortunately, Reality provides feedback constantly. We bump into it.

A person believes X about a market and takes action A_X . If the action succeeds once, X may be a good model—or the person may have been lucky. If related actions succeed repeatedly under varied conditions, confidence in the model increases. If the actions succeed one hundred times and then fail ten times, the model may have been wrong in a way the earlier tests did not expose, or Reality may have changed. Either way, the failures are Information.

The model sends a question, decision, experiment, or action into Reality. This is fitback:

$$\mathbb{M}(t) \xrightarrow{\text{fitback}} \mathbb{R}(t)$$

Reality answers through consequences. This is feedback:

$$\mathbb{R}(t) \xrightarrow{\text{feedback}} \mathbb{M}(t+1)$$

Together they form the Love–Reality Loop [1,19,20], O-Theory’s construct for the continuing cycle in which a model acts on Reality and Reality answers; fitback and feedback are its two halves:

$$\mathbb{M} \rightleftarrows \mathbb{R}$$

No comparison between two models can finally establish Reality alignment. $\mathbb{M}(2)$ may detect something important about $\mathbb{M}(1)$, or it may merely project its own Organization onto the source. The only final test is continued contact between model and Reality through fitback and feedback.

That is why inspectability matters. Once Organization becomes inspectable, it can be converted into explicit predictions, questions, experiments, and actions. Its consequences can be compared with Reality’s response. Confidence can rise, fall, or become more precisely bound.

In this exact sense, Reality is loving: it keeps answering. It gives feedback whether the model welcomes that feedback or not. The adaptive posture is therefore to love Reality more than one’s mental model—to have the courage and curiosity to say:

I think X. But is X a good way to think about this situation? How could I test it and see?

That posture is humility made operational. It treats every model as real, consequential, partial, and revisable. It turns people into living scientists: willing to inspect their Organization, send it into contact with Reality, listen to what comes back, and change their thinking when Reality gives them reason to do so.

The gravity of cognitive dark matter is therefore not an additional force or mechanism. It is the constraint that emerges from contextual identity itself. As affirmed and negated DSRP structures accumulate around an identity, they progressively constrain which interpretations, judgments, and actions remain structurally compatible with the model. The local rules are simple; their recursively organized consequences need not be.

5.7 What the claim does not mean

The claim does not require the brain to store an explicit proposition list. The Is/Is Not listing is an analytical representation of Organization exhibited in cognition. It specifies the structure that judgments and behavior imply; it does not specify the neural format in which that structure is implemented. That neutrality is what licenses the test of Section 7. Because the claim is about Organization and not about its carrier, it can be tested wherever Information and Organization can be controlled, and a machine is the one place they can be controlled completely. Nothing in Section 7 is offered as evidence about neural implementation.

The theoretical construct and its empirical detection must remain distinct. Cognitive dark matter is the Organization a model uses that is absent from its account of itself, $O - I(O)$. A report–use discrepancy is how that quantity is measured, not the thing itself. An omitted fact, a short answer, or an incomplete

description does not by itself establish that unreported Organization influenced a judgment. Conversely, successfully reporting a structure moves it from the dark into the light; it does not remove the structure from the model or end its participation in the model's Organization.

A specific claim of unreported structural influence requires evidence that the inferred Organization predicts outputs beyond what the direct report predicts and survives plausible alternative explanations. The person may have changed their mind, misunderstood the question, responded strategically, or reorganized the model during elicitation. Those possibilities must be distinguished. The broader theoretical account does not change a hypothesis or its success criterion once they are fixed: each empirical prediction must be evaluated by the test specified for it.

Finally, the advanced Is/Is Not operation is an instrument for exposing structural discrepancies. It does not guarantee that the first proposed explanation of a discrepancy is correct. Failure to predict behavior may mean that a structure is missing from the model or that an existing relationship has been mischaracterized. Those possibilities must remain empirically distinguishable.

The practical consequence is nevertheless substantial. Human beings do not think only with what they can say they think. They think with the larger Organization that makes recognition, interpretation, judgment, and action possible. Contextual identity analysis makes some of that Organization inspectable. Once inspectable, it can be compared, tested, taught, and deliberately changed.

6. Measuring Contextual Identity

The preceding sections define contextual identity structurally rather than lexically. For a focal identity X , the model contains a set of DSRP structural possibilities involving X , each of which may be affirmed, expressly negated, or remain unidentified. Contextual identity is the resolved portion of that field: the structures the model identifies as IS or IS NOT. The remaining unidentified structures form $\mathcal{H}$.

Because these structures and their statuses are explicit, they can be counted, compared, stored, and tracked through time. The resulting measurements do not replace the contextual-identity profile. They provide compressed descriptions of its structural composition and make differences between contextual identities empirically inspectable.

6.1 From structures and statuses to measurement

Let $\Pi_X(\mathbb{M}, t)$ denote the structural possibilities involving focal identity X in model $\mathbb{M}$ at time t . Each structure $\pi \in \Pi_X(\mathbb{M}, t)$ has a status:

$$s(\pi; \mathbb{M}, t) \in \{+, -, ?\}$$

where $+$ means the model affirms the structure, $-$ means the model expressly negates it, and $?$ means its Information or status has not yet been identified.

The contextual identity of X is therefore the set of status-bearing structures that the model has resolved as either IS or IS NOT:

$$CI_{\mathbb{M}}(X, t) = \{(\pi, s(\pi; \mathbb{M}, t)) : \pi \in \Pi_X(\mathbb{M}, t) \wedge s(\pi; \mathbb{M}, t) \in \{+, -\}\}$$

The unidentified structural field is:

$$\mathcal{H}_{\mathbb{M}}(X, t) = \{\pi \in \Pi_X(\mathbb{M}, t) : s(\pi; \mathbb{M}, t) = ?\}$$

Thus the measured field decomposes into what the model currently identifies and what it has not yet identified:

$$\Pi_X(\mathbb{M}, t) = \{\pi : (\pi, q) \in CI_{\mathbb{M}}(X, t)\} \cup \mathcal{H}_{\mathbb{M}}(X, t)$$

with the understanding that contextual identity preserves both the structure and whether that structure is affirmed or expressly negated.

This distinction matters. An expressly negated structure and an unidentified structure are not equivalent. “VORG is not related to QUELM” identifies a known nonrelationship. “Is VORG related to GREEB?” records that neither the Relationship nor its absence has yet been identified.

6.2 The structural status signature

For measurement, let each atomic structure be assigned to the DSRP Pattern it explicitly instantiates at the level being counted:

$$\kappa(\pi) \in \{D, S, R, P\}$$

For each Pattern k and status q , define:

$$N_{k,q}(X, t) = |\{\pi \in \Pi_X(\mathbb{M}, t) : \kappa(\pi) = k \wedge s(\pi; \mathbb{M}, t) = q\}|$$

The full structural status signature of X is therefore the 4×3 matrix:

$$\mathbf{N}_{\Pi}(X, t) = [N_{k,q}(X, t)]_{k \in \{D, S, R, P\}, q \in \{+, -, ?\}}$$

The three columns distinguish affirmed, expressly negated, and unidentified structures. The four rows distinguish Distinctions, Systems, Relationships, and Perspectives. The structures counted here are the focalized structures of Section 5.3: $\Pi_X(\mathbb{M}, t) = \{\sigma : \text{focal}(\sigma) = X\}$, so that each cell $N_{k,q}$ is the number of triples (X, ρ, Y) whose role ρ belongs to Pattern k and whose status is q . The union of Section 5.3 and the matrix here are the same objects, once listed and once counted.

Because contextual identity consists only of resolved structures, its own count signature is the 4×2 projection:

$$\mathbf{N}_{CI}(X, t) = [N_{k,q}(X, t)]_{k \in \{D, S, R, P\}, q \in \{+, -\}}$$

The unidentified field is represented separately by:

$$\mathbf{N}_{\mathcal{H}}(X, t) = [N_{k,?}(X, t)]_{k \in \{D, S, R, P\}}$$

Together, these recover the full status signature:

$$\mathbf{N}_{\Pi}(X, t) = [\mathbf{N}_{CI}(X, t) \mid \mathbf{N}_{\mathcal{H}}(X, t)]$$

Once context has been produced as:

$$C = O(I \mid t_0)$$

the same measurements may be abbreviated as $\mathbf{N}_{\Pi}(X \mid C)$, $\mathbf{N}_{CI}(X \mid C)$, and $\mathbf{N}_{\mathcal{H}}(X \mid C)$. This notation does not make context an additional input. It identifies the Organization from which the measurement was derived.

The count is performed on atomic structures rather than lines of prose. Thus, “VORG contains PIB, ZOOK, and NURR” contributes three System structures. Repetition contributes nothing: if a later elicitation again identifies VORG as framing PIB, the same status-bearing structure is not counted twice.

The initial count also records only the Pattern explicitly instantiated at the chosen level of analysis. It does not recursively count every further structure that can be generated through Enclosure, Simultaneity, or other DSRP dynamics. Without this constraint, one counted structure could recursively produce additional countable structures indefinitely.

6.3 The VORG signature

In plain terms, the matrix of Section 6.2 is a sorting of statements. Take every atomic statement the listing makes about VORG, and every one it leaves open. Sort each by the Pattern it uses: what VORG is or is not (D), what it is part of or contains (S), what it is related to (R), what it frames or is framed by (P). Then mark its status: affirmed (+), expressly negated (−), or unidentified (?). Four Patterns by three statuses gives twelve cells, and each cell holds a count. A row shows how much of VORG’s identity lives in one Pattern; a column shows how much of it is settled and how much is still open. The first two columns together are the contextual identity, $\mathbf{CI}_{\mathbb{M}}(\text{VORG}, t_0)$; the third column is the unidentified field, $\mathcal{H}_{\mathbb{M}}(\text{VORG}, t_0)$. Here is that sorting for the Section 4 listing.

After expanding grouped structures and eliminating duplicates, the VORG example contains 28 unique atomic structural possibilities:

Table 4. The status signature of the Section 4 listing. Counts of unique atomic structures by Pattern and status after expanding grouped statements and removing duplicates: 28 structures, 14 affirmed (+), 10 negated (−), 4 unidentified (?). The first two columns are the contextual identity; the third is the unidentified field.

Pattern	IS +	IS NOT −	Unidentified ?	Total
Distinctions	1	1	1	3
Systems	4	2	1	7
Relationships	1	1	1	3

Perspectives	8	6	1	15
Total	14	10	4	28

Its complete structural status signature is:

$$\mathbf{N}_{\Pi}(\text{VORG}, t_0) = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \\ 8 & 6 & 1 \end{bmatrix}$$

where the columns represent $+$, $-$, and $?$, respectively, and the rows represent D, S, R, and P.

VORG's contextual-identity signature is therefore:

$$\mathbf{N}_{CI}(\text{VORG}, t_0) = \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 1 & 1 \\ 8 & 6 \end{bmatrix}$$

and its unidentified structural field is:

$$\mathbf{N}_{\mathcal{H}}(\text{VORG}, t_0) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

The example therefore contains 24 resolved IS/IS-NOT structures and four structures whose status remains unidentified:

$$|\mathbf{CI}_{\mathbb{M}}(\text{VORG}, t_0)| = 24$$

$$|\mathcal{H}_{\mathbb{M}}(\text{VORG}, t_0)| = 4$$

The matrices immediately reveal features that are harder to see in the prose listing. VORG's represented identity is Perspective-heavy. More structures are affirmed than expressly negated. One structural possibility remains unidentified in every Pattern. None of these observations depends on knowing what VORG means.

That does not make the matrices equivalent to VORG's contextual identity. Two contextual identities can have identical counts while connecting entirely different identities in entirely different ways. Count signatures measure structural composition; the full contextual-identity profile preserves structural topology.

6.4 Direction and structural role

Pattern-level counts can be refined by recording the role occupied by focal identity X . This matters because reversing structural roles can change contextual identity.

“VORG is part of TANDRA” and “VORG contains TANDRA” are both System structures, but they place VORG in reciprocal structural roles.

Likewise, “VORG frames GREEB” and “VORG is framed by GREEB” are both Perspective structures, but the position occupied by VORG differs.

Let the eight DSRP roles be represented as:

$$\mathcal{E} = \{D_i, D_o, S_p, S_w, R_a, R_r, P_p, P_v\}$$

where the subscripts identify identity, other, part, whole, action, reaction, point, and view.

Let $\rho_X(\pi)$ denote the role or roles occupied by X in structure π . A role-refined count is then:

$$N_{\rho,q}(X, t) = |\{\pi \in \Pi_X(\mathbb{M}, t) : \rho \in \rho_X(\pi) \wedge s(\pi; \mathbb{M}, t) = q\}|$$

The plural is important. Under Simultaneity, a single identity may occupy multiple structural roles in the same Organization. Role counts therefore need not sum to the number of unique structures. They answer a different question: not merely which Pattern is instantiated, but how the focal identity participates in that Pattern.

Direction should be recorded only when the Organization establishes it. “VORG is related to DRAYE” affirms a Relationship but does not, by itself, identify VORG as occupying the action or reaction role. A measurement procedure should retain the coarser Relationship structure rather than manufacture a direction the model does not contain.

The same principle applies to Systems and Perspectives. Measurement should preserve structural specificity when it is present and preserve structural ambiguity when it is not.

6.5 Network degree as a special case

Conventional network measures appear within this framework as restricted projections of contextual identity.

In an undirected network where each affirmed Relationship structure corresponds to one edge, the degree of X is:

$$\deg(X) = N_{R,+}(X)$$

In a directed network, corresponding role-refined measures are:

$$\deg^{\text{out}}(X) = N_{R_a,+}(X)$$

$$\deg^{\text{in}}(X) = N_{R,+}(X)$$

Degree therefore counts one portion of one DSRP Pattern under one status. It ordinarily excludes Distinctions, Systems, Perspectives, expressly negated Relationships, and unidentified Relationships.

Network degree is therefore not a rival representation of contextual identity. It is a special case produced by projecting a richer structural profile onto its affirmed Relationship component.

This matters whenever two identities have the same number of connections but differ in membership, boundaries, framings, direction, expressly absent Relationships, or the larger Organization in which those Relationships participate. Their network degrees may be identical while their contextual identities differ substantially.

6.6 Comparing contextual identities

Because contextual identity preserves both structure and status, two contextual identities can be compared directly.

For two models or persons A and B , the full structural difference around focal identity X can be represented by the symmetric difference:

$$\Delta \text{CI}(X) = \text{CI}_A(X, t_0) \triangle \text{CI}_B(X, t_0)$$

Because the members of contextual identity are status-bearing structures, this comparison detects both structural differences and status differences. If the same structure is affirmed in one model and expressly negated in another, the contextual identities differ even though both models contain the same structural possibility.

A coarser count distance can also be calculated:

$$d_N(A, B \mid X) = \sum_{k \in \{D, S, R, P\}} \sum_{q \in \{+, -\}} \left| N_{k,q}^A(X, t_0) - N_{k,q}^B(X, t_0) \right|$$

A larger value indicates a greater difference in the structural composition of the two contextual identities.

But count equality does not imply contextual-identity equality:

$$d_N(A, B \mid X) = 0 \not\Rightarrow \text{CI}_A(X, t_0) = \text{CI}_B(X, t_0)$$

Two profiles may contain exactly the same numbers of Distinctions, Systems, Relationships, and Perspectives under each status while organizing different identities in different ways. Count equality is not structural identity.

The same comparison can be made for the unidentified field:

$$\Delta \mathcal{H}(X) = \mathcal{H}_A(X, t_0) \triangle \mathcal{H}_B(X, t_0)$$

Two people may therefore share much of the same contextual identity while differing substantially in which structural possibilities remain unidentified, or they may have similarly sized unidentified fields containing entirely different unresolved structures.

6.7 Measuring the unidentified field

The proportion of the measured structural field that remains unidentified can be represented as:

$$U(X, t) = \frac{|\mathcal{H}_{\mathbb{M}}(X, t)|}{|\Pi_X(\mathbb{M}, t)|}$$

Equivalently:

$$U(X, t) = \frac{\sum_k N_{k,?}(X, t)}{\sum_k \sum_q N_{k,q}(X, t)}$$

For VORG:

$$U(\text{VORG}, t_0) = \frac{4}{28} = \frac{1}{7} \approx 0.143$$

The complementary resolved proportion is:

$$1 - U(X, t) = \frac{|\text{CI}_{\mathbb{M}}(X, t)|}{|\Pi_X(\mathbb{M}, t)|}$$

For VORG:

$$1 - U(\text{VORG}, t_0) = \frac{24}{28} = \frac{6}{7} \approx 0.857$$

These quantities do not measure total knowledge or ignorance about X . They measure the proportion of the structural possibilities represented in the specified domain whose statuses have or have not been identified.

Their interpretation therefore depends on the domain over which $\Pi_X(\mathbb{M}, t)$ has been generated or elicited. A narrow inquiry can produce a small unidentified field without implying that little remains unknown about the identity more generally. Measurement requires the relevant structural domain to be made explicit.

The unidentified field $\mathcal{H}$ should remain distinct from cognitive dark matter. It records structural possibilities whose status is unresolved in the specified model. Cognitive dark matter concerns Organization that is affirmed and operative in a model while remaining absent from that model's direct report. This differs from the structure having unidentified status within the model. Unreported Organization and unresolved structural status therefore require separate measurements.

6.8 Contextual identity through time

Contextual identity is not necessarily static. The status of a structure may change as a model learns, differentiates, reorganizes, or encounters new Information.

An unidentified structure may become affirmed:

$$s(\pi; \mathbb{M}, t_0) = ? \rightarrow s(\pi; \mathbb{M}, t_1) = +$$

or expressly negated:

$$s(\pi; \mathbb{M}, t_0) = ? \rightarrow s(\pi; \mathbb{M}, t_1) = -$$

A previously affirmed structure may also become expressly negated:

$$+ \rightarrow -$$

and an expressly negated structure may later become affirmed:

$$- \rightarrow +$$

If the model permits suspension or withdrawal of a prior identification, transitions back to unidentified status are also possible:

$$+ \rightarrow ?$$

$$- \rightarrow ?$$

Learning and conceptual change can therefore be represented not merely as acquiring more Information but as changes in the status and Organization of particular structures.

The count-level change in contextual identity between two times is:

$$\Delta \mathbf{N}_{CI}(X) = \mathbf{N}_{CI}(X, t_1) - \mathbf{N}_{CI}(X, t_0)$$

The corresponding change in the unidentified field is:

$$\Delta \mathbf{N}_{\mathcal{H}}(X) = \mathbf{N}_{\mathcal{H}}(X, t_1) - \mathbf{N}_{\mathcal{H}}(X, t_0)$$

These measurements show how structural composition changed, but the full profiles must still be compared because structures can be replaced, redirected, reframed, or reorganized while aggregate counts remain constant.

The history of contextual identity can therefore be represented as a sequence:

$$\mathcal{T}_{\mathbb{M}}(X) = \langle \text{CI}_{\mathbb{M}}(X, t_0), \text{CI}_{\mathbb{M}}(X, t_1), \dots, \text{CI}_{\mathbb{M}}(X, t_n) \rangle$$

A corresponding history of unresolved structure can be represented as:

$$\mathcal{T}_{\mathcal{H}}(X) = \langle \mathcal{H}_{\mathbb{M}}(X, t_0), \mathcal{H}_{\mathbb{M}}(X, t_1), \dots, \mathcal{H}_{\mathbb{M}}(X, t_n) \rangle$$

This makes it possible to ask not only what X is in a model, but also how “what X is” changes.

6.9 Storing and reconstructing contextual identity

A contextual identity need not remain a one-time elicitation or a static measurement. Because the underlying representation consists of atomic DSRP structures with explicit statuses, it can be stored in mapping software, graph databases, network-analysis systems, or other computational representations and updated as Organization changes.

The stored representation need not contain “the definition of X ” as a text field. It can instead store the structural field from which contextual identity is reconstructed.

Each stored structure can record:

- the identities involved;
- the DSRP Pattern and roles instantiated;
- its status as affirmed, expressly negated, or unidentified;
- the model or person from which it was elicited;
- its evidentiary source; and
- the time at which the status was recorded.

Let $\mathbb{G}_{\mathbb{M}}(t)$ denote the stored Organization associated with model $\mathbb{M}$ at time t . The structural possibilities involving focal identity X can be retrieved by query:

$$\Pi_X(\mathbb{M}, t) = Q_X(\mathbb{G}_{\mathbb{M}}(t))$$

Contextual identity is then reconstructed from the structures carrying resolved status:

$$\text{CI}_{\mathbb{M}}(X, t) = \{(\pi, s(\pi; \mathbb{M}, t)) : \pi \in Q_X(\mathbb{G}_{\mathbb{M}}(t)) \wedge s(\pi; \mathbb{M}, t) \in \{+, -\}\}$$

and the unidentified field from:

$$\mathcal{H}_{\mathbb{M}}(X, t) = \{\pi \in Q_X(\mathbb{G}_{\mathbb{M}}(t)) : s(\pi; \mathbb{M}, t) = ?\}$$

In this form, the identity of X is not merely a label attached to a node. It is generated from the Organization surrounding that identity.

Selecting **VORG** in a mapping application could therefore reveal what **VORG** is and is not, what parts it contains and what wholes contain it, what it is and is not related to, what frames it and what it frames, and which structural possibilities remain unidentified.

This requires more than an ordinary network representation. In many network models, the absence of an edge means only that no edge has been recorded. Contextual identity must distinguish absence of Information from express negation.

“**VORG** is not related to **QUELM**” must therefore be represented as a Relationship structure carrying negative status rather than by the mere absence of an edge. The same requirement applies to Distinctions, Systems, and Perspectives.

A persistent representation also makes structural history available for analysis. One can ask:

- What was X in this model at an earlier time?
- Which structures changed status?
- Which unidentified possibilities were subsequently resolved?
- Which IS structures became IS NOT structures, or vice versa?
- Do two people or groups construct the same focal identity differently?
- Which changes in contextual identity preceded changes in interpretation, judgment, or behavior?

Contextual identity thereby becomes a persistent and evolvable object of analysis. Its current structure can be inspected, its history can be reconstructed, and changes in that structure can be related to subsequent outputs.

6.10 What the measurements measure

These quantities measure Organization represented by a model. They do not automatically measure objective reality, psychological certainty, importance, causal strength, or truth.

An affirmed structure means that the structure is affirmed in the model. An expressly negated structure means that the model identifies its negation. An unidentified structure means that the structural possibility is present in the measured domain but its status has not yet been identified.

Reliable comparison therefore requires several controls:

1. The structural domain $\Pi_X(\mathbb{M}, t)$ must be specified or generated by a reproducible procedure before counts are interpreted.
2. Grouped expressions must be decomposed into the atomic structures being measured.
3. Duplicate structures must be eliminated.
4. System, Relationship, and Perspective roles and directions must be preserved where the model establishes them.
5. The level of part-whole decomposition must remain comparable across measurements.
6. Coarse structures must not be assigned roles or directions they do not establish.
7. Express negation must remain distinct from absence or unidentified status.
8. Count signatures must not be mistaken for the full contextual-identity topology.

With those controls, contextual identity becomes measurable without being reduced to a single score.

The full contextual-identity profile records what the focal identity structurally IS and IS NOT in the model. The count signature records the composition of that profile. Role-refined counts record how the focal identity participates in its Organization. The unidentified field records structural possibilities whose statuses remain unresolved. Temporal comparisons record how those structures and statuses change.

These measurements also make the paper's central claim empirically vulnerable. If contextual identity matters, then differences in contextual identity should predict consequential differences in interpretation, judgment, prediction, and action. If two people exhibit different contextual identities for the same focal

identity, those structural differences should help predict where their subsequent outputs diverge. If a person's contextual identity changes, corresponding changes in output should become detectable.

The measurement problem is therefore not simply how much structure surrounds an identity. It is which structures are present, which are affirmed, which are expressly negated, which remain unidentified, how the focal identity participates in them, how they are organized together, and whether differences in that Organization predict differences in what the model subsequently does.

7. The CI/CDM Experiment

7.1 Introduction

Section 6 showed how contextual identity can be formally represented and counted. This section tests whether that structure predicts what a model does.

The paper's central claim is that what a thing *is* depends on the *Organized Information* surrounding it, not merely the label or definition attached to it. Much of what constitutes an identity therefore comes from *Organization* (O) that remains unnoticed, even when we think and act as though *Information* (I) alone supplies its meaning.

Sections 4 and 5 argued that this Organization can predict what a model-holder will say next, and that much of it may be missing from a direct report. These claims lead to two empirical predictions. First, a system given the same definition but different structure around it should answer new questions differently, in the specific ways each structure predicts. Second, the system should use more of that structure in answering questions than it describes when asked directly what it is (its definition).

We tested these predictions on machines. A child and a language model are not the same kind of cognizer, and the claim does not need them to be. The claim is about Organization: any system that takes in Information and organizes it with Distinctions, Systems, Relationships, and Perspectives should show both effects, whatever it is made of. The child builds the structure in the light, over years, and we can only watch; the machine lets us set the Information and the Organization exactly and read the answers in minutes. What that test can and cannot show is worth stating before the design. It tests the structural claim—that O, not I, fixes the answers—in a system where I and O can be set exactly, and it says nothing about how a brain implements either. The machine results are read beside the lab's human studies in Section 7.4.1, and their limits are stated in Section 7.4.7.

Large language models are useful for this purpose for three reasons. First, an invented name (an information variable) such as `VORG` lets us introduce an identity without the familiar associations attached to a word such as `TRUCK`. The invented name does not erase prior knowledge of the words used to define it, but it helps separate the contribution of the label from that of the supplied structure. Second, each run can begin in a fresh session, allowing the same map to be tested repeatedly without memory carrying over between runs. Third, replacing the invented name with a familiar word allows supplied Organization and Organization inherited from prior learning to be compared using otherwise identical materials.

The design turns the VORG test of Section 4 into an experimental instrument. One canonical definition and a common background remain constant. The definition control adds no further structures involving VORG. Two other conditions supply coherent contextual identities with exactly the same *status signature*—the same counts of affirmed, negated, and unidentified structures within each DSRP Pattern—but different *profiles*: different structures are affirmed and negated. A fourth condition presents the same twelve candidate structures with every status marked unidentified.

We then ask sixteen questions that the supplied description never answers directly. Answering requires combining the structures. Taken together, the structures let us determine whether the answer is IS, IS NOT, or CANNOT TELL. We also ask the machine what a VORG is — in the main runs with the prompt 'tell me everything you know about it,' and in the follow-up runs with the plain question 'What is a VORG?' — and to choose which of four questions would best resolve something still unknown. To test which structures support which answers, we move one Pattern's structures at a time from VORG to another identity in the structural map and observe which answers change. Finally, we repeat the core conditions with the familiar word TRUCK in place of VORG.

The hypotheses, materials, keys, and analyses were fixed and dated before the runs (Section 7.2.6). The predictions, in order:

H1 — O=DSRP determines identity. Adding Organization lets the machine resolve questions that the definition alone leaves unanswered, with answers matching the predictions derived from the supplied structures.

H2 — DSRP profile, not DSRP count, determines identity. Two contextual identities with the same counts of affirmed, negated, and unidentified DSRP structures produce different answers because their structures are organized differently. Each produces the answers predicted for it before the experiment.

H3 — Three IS statuses determine identity. When the supplied structures leave something unknown, the machine answers CANNOT TELL more often than IS NOT. It distinguishes not knowing from knowing that something is false.

H4 — Identity decomposes by O=DSRP Pattern. Moving one Pattern's structures away from VORG reduces the machine's previously supported IS and IS NOT answers on questions fixed in advance as depending on that Pattern. Answers to questions that do not depend on the moved structures remain unchanged.

H5 — Identity propagates through O=DSRP. Changing VORG's contextual identity changes answers about other identities and about the map as a whole. The effects extend beyond questions about VORG itself.

H6'² — CDM = O – I(O). Cognitive dark matter, $O - I(O)$, is measurable: it has a size and a sign, and it differs by system.

H7'³ — Copriming determines CDM baggage. The familiar words in a world coprime. Given a word it has no identity for, the machine builds one from that coprime. A set of nonce words does not coprime the same way.

H8'⁴ — Perspective constrains or expands CDM. Perspective can constrain or expand CDM baggage. A system that takes the constraint seriously drags in less; one that doesn't drags in more.

H9 — Metacognition of O=DSRP locates the unknown (i.e., improves inquiry, critical thinking). Given four possible questions, the machine chooses the one that addresses something genuinely unknown more often than the 25% expected by chance.

H6' measures *cognitive dark matter* directly: the gap between the Organization a machine uses to answer and the Organization it reports when asked what the thing is. H7' tests where imported Organization comes from — the history behind familiar words — by removing that history and keeping everything else. H8' tests whether a Perspective, the instruction to use only the supplied material, sets how much of that Organization gets in.

7.2 Methods

7.2.1 The microworld

The stimulus is a small closed world of ten identities. The focal identity is *VORG*. A second nonce identity, *NURR*, is present in every condition and serves as the target when structures are moved away from *VORG*. The remaining eight are ordinary words: two categories (*ROAD VEHICLE*, industrial equipment), two wholes (the road system, the factory system), two kinds of agent (drivers, technicians), and two governing bodies (the municipal regulator, the safety regulator).

Two definitions appear in every condition:

A *VORG* is a motor vehicle designed to transport cargo.

A *NURR* is a motor vehicle designed to transport passengers.

A background map of eight statements about the non-focal identities is also present in every condition and never mentions *VORG*: the road system includes city streets and highways; the factory system includes assembly lines and loading bays; drivers are licensed by the municipal regulator; technicians are certified by the safety regulator; the municipal regulator governs the road system; the safety regulator governs the

² Clarified after the runs (Section 7.2.8). Preregistered H6: Report–use discrepancy. The machine uses more of the supplied structure to answer questions than it mentions when asked directly to report everything it knows about *VORG*.

³ Clarified after the runs. Preregistered H7 — Historical prior. Given only the definition, the machine gives more IS or IS NOT answers for the familiar word truck than for the invented name *VORG*.

⁴ Clarified after the runs. Preregistered H8 — Prior resistance. When the supplied contextual identity conflicts with the familiar meaning of truck, the machine's answers follow that new identity more completely for *VORG* than for *TRUCK*. The familiar word's inherited Organization is expected to resist the change.

factory system; ROAD VEHICLES must be registered with the municipal regulator; industrial equipment must be inspected by the safety regulator.

The background is closed: the named systems have exactly the named parts and the named bodies have exactly the named authorities. Statements about VORG are three-valued: only what is stated, plus its ordinary consequences. Formally, with B the background, C the VORG structures supplied in a condition, and σ an item, the key is $+$ if $B \cup C \vdash \sigma$, $-$ if $B \cup C \vdash \neg\sigma$, and $?$ otherwise, where derivation uses the closed background: a rule that names a category applies to VORG only if VORG's membership in that category is stated or derivable, and the named systems and authorities are exhaustive. This is the closure rule referred to throughout Section 7.3. The three statuses behave as the values of Kleene's strong three-valued logic [21,22]—true, false, and a third value that no derivation collapses into either of the others—with one difference of reading: $?$ is unidentified in the model, a status of the model's Organization, not unknown in the world; Section 5.3 keeps the two apart. Because the background is present in every condition, the difference between the definition control and a contextual-identity condition is exactly VORG's contextual identity.

7.2.2 Conditions

D (definition control). Background and both definitions. No structure involving VORG. Every one of the sixteen items is unresolvable in D; the correct answer to each is CANNOT TELL.

A (road-transport profile). D plus twelve structures involving VORG, three per Pattern, one each affirmed, expressly negated, and unidentified. VORG is a ROAD VEHICLE; is not industrial equipment; whether it is electrically powered has not been determined. VORG is part of the road system; is not part of the factory system; whether it has a detachable trailer has not been determined. VORG is operated by drivers; is not operated by technicians; whether it is refueled at public stations has not been determined. The municipal regulator classifies VORG as ordinary transportation; the safety regulator does not classify VORG as restricted equipment; how insurers classify VORG has not been determined.

T (industrial-transport profile). The same background, definitions, identities, predicates, sentence count, and unidentified structures as A, with the affirmed and negated contents exchanged. VORG is industrial equipment and not a ROAD VEHICLE; part of the factory system and not the road system; operated by technicians and not drivers; classified by the safety regulator as restricted equipment and not by the municipal regulator as ordinary transportation.

A and T have identical status signatures:

$$\mathbf{N}_{\Pi}(\text{VORG} \mid A) = \mathbf{N}_{\Pi}(\text{VORG} \mid T) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Each Pattern (rows D, S, R, P) has one affirmed (+), one expressly negated (−), and one unidentified (?) structure in both conditions.

Their contextual identities differ: $CI(VORG \mid A) \neq CI(VORG \mid T)$. Both worlds are coherent with the background; the background supports either profile. T is a second possible world, not a scrambled version of the first. A scrambled control would confound Organization with contradiction.

H (unidentified-status control). D plus the twelve A structures, each rewritten as explicitly unresolved (“Whether VORG is a ROAD VEHICLE has not been determined,” and so on). Its signature is

$$\begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 3 \end{bmatrix}.$$

H is not omission. Every candidate structure is named; none is resolved.

Ablations A_{-D} , A_{-S} , A_{-R} , A_{-P} . From A, the three structures of one Pattern are moved from VORG to NURR by substituting the label. Nothing is deleted. In A_{-R} , for example, it is NURR that is operated by drivers, not operated by technicians, and of undetermined refueling; VORG has no Relationship structure at all. Sentence count, token count, and the world’s Pattern composition are unchanged. NURR’s definition was chosen so that every rebound structure is coherent for it.

Real-word run. D, A, and T with VORG replaced by TRUCK in the definition, the structures, the items, and the prompts. Nothing else changes; NURR remains NURR. The materials differ from the nonce materials only by the focal label.

This gives eleven cells: D, A, T, H, and four ablations with VORG, and D, A, and T with TRUCK.

7.2.3 Items

Sixteen statements are presented in every condition. Each is unstated in every condition and requires one bounded inference from the VORG structures plus the background. No item shares a six-word sequence with any stimulus sentence. Each is answered IS, IS NOT, or CANNOT TELL: the three statuses of Section 4.1 (+, −, ?) given as answers.

Twelve are *focal* items, three per Pattern, typed by what condition A licenses: derivably affirmed ($? \vdash +$), derivably negated ($? \vdash -$), or nothing licensed ($? \vdash ?$). “VORG is subject to a registration requirement” is $? \vdash +$ in A: VORG is a ROAD VEHICLE, and ROAD VEHICLES must be registered. “VORG is subject to an inspection requirement” is $? \vdash -$ in A: the only inspection rule concerns industrial equipment, VORG is not industrial equipment, and the background is complete. “VORG is privately owned” is $? \vdash ?$: nothing supplied bears on it. In T the $\vdash +$ and $\vdash -$ items reverse by the mirrored derivations; the $? \vdash ?$ items stay $? \vdash ?$ everywhere.

The symbol $\vdash$ (“derivable to”) is used rather than $\rightarrow$, which Section 6.8 reserves for change of status over time.

Two are *reciprocal* items. They begin from a non-focal identity and ask about its relation back to VORG: whether the municipal regulator has authority over VORG, and whether the safety regulator certifies at least one person who operates VORG. The first is deliberately cross-Pattern: a question about a Relationship whose answer can be reached by two routes, from VORG's System membership or from its Distinction together with the registration rule, so that removing either Pattern alone leaves it answerable.

Two are *whole-map* items requiring integration across at least two Patterns and three identities: whether the body that licenses or certifies VORG's operators also governs the system VORG belongs to, and whether the body that registers or inspects VORG also governs that system. Both are true in A and in T, because each world is internally coherent. Their function is to become unresolvable when a Pattern is removed and to be unanswerable in D and H.

Every item carries a preregistered dependency set naming the exact VORG structures its derivation requires. The dependency matrix specifies, for each ablation, which items must change and which must not. The must-not-change cells are the falsifiers for H4: an ablation that changes an item outside its dependency set counts against the partition claim for that item. Three cross-Pattern dependencies are fixed in advance (the first reciprocal item and both whole-map items); no others may be added after the fact.

The full gold key across all eight nonce conditions is in the replication package. In A, the sixteen items resolve to seven IS, five IS NOT, and four CANNOT TELL. T has the same counts, with the IS and IS NOT items exchanged except for the two whole-map items, which are IS in both. In D and H, all sixteen are CANNOT TELL.

7.2.4 Other measures

Direct report. “What is a VORG? Tell me everything you know about it.” No length limit. The wording was chosen not to cue negation or organization, because H6 concerns what the machine reports unprompted. After the main runs it became clear that “tell me everything you know” asks for an inventory, and in a world with twelve sentences about VORG the inventory is the twelve sentences; the plain question “What is a VORG?” was run instead, and then run again with the system prompt's constraints removed; both are described in Section 7.2.9 and reported in Section 7.3.6. Position is counterbalanced: in half the runs of every cell the report precedes the items; in the other half it follows the question task. Report-first primes structure use and items-first primes the report; the balanced design lets both be estimated.

Better-question task. After the items: “Which of these questions would most usefully resolve something that is still unidentified about VORG?” Four options, order randomized, condition-specific. In A and T the correct option targets a genuinely unidentified structure (whether VORG is electrically powered); the distractors target an affirmed structure, a negated structure, and a question outside VORG's structural field. In D and H, where nothing about VORG is affirmed or negated, the distractors are a question resolved by the definition, a question resolved by the background, and the disjoint question.

Metacognitive report. Always last: “Did information beyond the definition of VORG affect your answers? Answer YES, NO, or UNCERTAIN, then explain in one or two sentences.” A denial is not by itself evidence of anything. It becomes relevant only when a manipulation demonstrably changed the machine's

answers.

7.2.5 Models and procedure

Two independently developed model families were used, with one frozen model identifier each: OpenAI `gpt-5.5-2026-04-23` and Anthropic `claude-fable-5-1`, accessed through their public APIs in September 2026. No model was substituted after the freeze. The third family named at the freeze, Google's Gemini, was dropped before analysis (Section 7.2.8).

Each run is a fresh session with no prior state. Tools, browsing, retrieval, and memory were disabled. Each run receives a frozen system prompt stating that only the supplied material and its ordinary consequences may be used, that the description of the world other than VORG is complete, that the only things known about VORG are the statements that mention it, and that status questions are to be answered with exactly one of IS, IS NOT, CANNOT TELL. For the TRUCK run the instruction is identical with TRUCK substituted; the run measures whether prior Organization leaks despite the instruction.

Thirty runs were executed in every model × condition cell, for 660 runs. Item order was randomized per run, items were numbered 1–16 with no identifying labels, response options for the question task were lettered A–D in randomized order, and report position alternated within each cell. Every seed and every realized order was stored. Forced-choice answers were requested as structured JSON. Answers were never repaired or regenerated: an answer outside the schema is recorded as invalid and counted, not replaced.

Generation parameters were fixed before the confirmatory runs: temperature 0.7 where the provider accepts it, reasoning at the lowest setting the provider allows, and an output ceiling of 8,192 tokens so that hidden reasoning cannot cut off a visible answer. Two facts about what the providers actually applied are recorded in every raw record and belong here. Both models reject a temperature parameter and ran at the provider default. Neither provider allows reasoning to be turned off: OpenAI ran at its “minimal” setting, and the Anthropic model reasons by default and rejected the request to disable it (240 of 284 recorded turns contain a hidden reasoning block). The protocol's “reasoning off” was therefore not attainable for either family; each ran at its minimum. The planned reasoning replication becomes reasoning-high against reasoning-minimum.

The complete request and response for every turn of every run was saved before any parsing.

7.2.6 Validation and freeze

Before any confirmatory run, the materials passed twelve mechanical checks: signature equality across A and T; sentence and token matching between A and T; no six-word overlap between any item and any stimulus sentence; no item derivable from the definitions and background alone; derivation completeness for every gold IS or IS NOT; ablation purity; coherence of every rebound structure with NURR's definition; NURR's presence in every condition; explicitness of H; no two non-focal identities with identical background profiles; exactly one genuinely unidentified option per question set; and the TRUCK materials differing from the VORG materials only by the focal label.

Two raters, blind to condition labels and hypotheses, then completed a packet containing both worlds under neutral names. Both rated A and T internally consistent (5 of 5). Both rated the TRUCK versions of A and T plausible descriptions of some TRUCK. On derivability, each rater answered the sixteen items for each world; 63 of 64 answers matched the key (rater-to-key $\kappa = .90$ and 1.00 for T; 1.00 and 1.00 for A; rater-to-rater $\kappa = .90$ and 1.00). No item was missed by both raters. Both raters confirmed that nothing in either world leans toward IS or IS NOT for any of the four $? \vdash ?$ items, and that the direct-report wording does not cue negation or organization. Three rater answers were clarified after an initial misreading of a question; the original answers and the clarifications are logged in the replication package, and none changed an item.

A pilot of eighteen runs (two per family in A, T, and H) was used only to confirm that every provider returned parseable output and to check the mechanical validator against real responses. Two pre-freeze repairs were made on pilot evidence and are recorded: one reciprocal item was reworded because its original predicate (“work with”) was broader than the structure it depended on (“operated by”), and two Distinction items were rephrased after the leakage check found a shared six-word sequence with the background. No confirmatory hypothesis was examined on pilot data.

After validation, the stimuli, items, key, prompts, model identifiers, and analysis code were hashed and frozen, and the frozen hashes were recorded before the first confirmatory call.

7.2.7 Scoring and analysis

Every answer is scored against the three-valued key for its condition. Accuracy is reported overall and by Pattern, item type, item level (focal, reciprocal, whole-map), word type, report position, and model family, with the full IS / IS NOT / CANNOT TELL confusion matrix.

Because D and A are scored against different keys, H1 is not a comparison of accuracies. It uses the *context-consistent resolution rate*:

$$\text{CRR}_A(k) = \frac{1}{|J_A|} \sum_{j \in J_A} \mathbb{1}(\hat{y}_{k,j} = y_{A,j})$$

On the twelve items (J_A) that D leaves unresolved and A resolves, the share of answers in condition k that match A’s key. $\text{CRR}_A(A) > \text{CRR}_A(D)$ says that the supplied Organization resolved those items in its own direction; CRR_T is defined the same way on T’s key.

H3 is computed on the four $? \vdash ?$ items in A and T, and on all sixteen items in H, as the proportion answering CANNOT TELL against the proportion answering IS NOT. H4 compares, for each Pattern k , the share of answers still matching A’s key on k -dependent items in A against A_{-k} , and the same share on non-dependent items in A against A_{-k} . Accuracy is not used for H4: under a successful ablation the correct answer on a dependent item becomes CANNOT TELL, so accuracy would rise, not fall. H5 is the same comparison restricted to the reciprocal and whole-map items. H9 is the proportion of runs selecting the genuinely unidentified option, against a chance level of one in four.

H6 requires two quantities. The *structural-use rate* for Pattern k is the share of k -dependent items answered correctly in A. *Report coverage* is the share of the resolved structures in the run's condition — eight in A and T, six in each ablation — that the direct report states, or ordinarily paraphrases, with the right polarity. Each report was coded on eight fixed statements (ROAD VEHICLE; industrial equipment; part of the road system; part of the factory system; operated by drivers; operated by technicians; classified as ordinary transportation by the municipal regulator; classified as restricted equipment by the safety regulator), each as *says yes*, *says no*, *says undetermined*, or *not mentioned*. The same coding yields, for the definition-only condition, the share of the eight structures a report affirms or denies although none was supplied (imported structure), and for H, the share it reports as undetermined. Two human coders, blind to condition, family, and report position, coded every report; a structure counts as stated only when both coders coded it so.

H6' is the difference between them: structural-use rate minus report coverage, per family and per report prompt, kept with its sign. Positive means the machine used structure it did not state; zero means it stated what it used; negative means it stated structure it did not use. The two sides are measured independently — use from the items, coverage from the coders — so the difference is free to come out anywhere.

The primary analysis is a mixed-effects logistic model of item-level correctness with condition, word type, Pattern, status, and the condition $\times$ word-type interaction as fixed effects and item as a grouping factor, reported separately by model family before pooling. Effect sizes, interval estimates, raw proportions, and complete cell counts are reported for every contrast. Invalid-output and refusal rates are reported by condition. Any analysis not listed here is labeled exploratory.

7.2.8 Disclosure

Four things occurred after the freeze. None changed an item, a prompt, or a model.

Execution deviation. Seventy-one of 330 Anthropic runs reached the frozen 1,200-token output ceiling, truncating answers or direct reports. Because shortened reports could artificially favor H6, every affected run was repeated with an 8,192-token ceiling, preserving its seed, item order, question-option order, and report position. Runs were selected for repetition solely because they reached the ceiling, without regard to answer correctness. All original records were retained. We also repeated the analysis with those runs treated as invalid under the preregistered rule; the substantive conclusions were unchanged.

One instrument flaw. The better-question option set for $A-S$ was inherited from A. Moving VORG's System structures to NURR makes one of A's distractors — “Is VORG part of the factory system?” — a genuinely unidentified question about VORG, so that cell has two correct answers. The mechanical validator checked option sets against A's structures rather than the ablated ones. The $A-S$ question-task cell is reported as uninterpretable.

One family dropped. The preregistration named three model families; the third, Google's Gemini, was dropped before analysis when the provider's daily request cap made its 330 runs infeasible in the study window. Its partial runs are in the replication package and are not analyzed.

Three hypotheses clarified. Section 7.1 states H6', H7', and H8'; the forms as first written are footnoted there. Each is the same claim made precise, not a different claim. H6 said the machine uses more of the supplied structure than it reports; H6' says what "more" is — a gap between structural use and report coverage, with a size and a sign — and measures it under the plain question "What is a VORG?" rather than a request for everything it knows, which in a twelve-sentence world can only return the twelve sentences. H7 said a familiar word would commit more than an unfamiliar one; H7' puts the familiarity where the main runs showed it lives, in the words of the world rather than in the label, and adds the comparison the first form lacked, a world with no familiar words in it. H8 said a familiar word would resist a conflicting identity; H8' names what did the resisting — the Perspective the instruction hands the machine, which can constrain or expand how much prior comes in. The clarifications were made after the main runs, the runs that test them were made after the clarifications (Section 7.2.9), and the paper works from the primed forms.

7.2.9 Follow-up runs

After the confirmatory runs were complete for two families, two short probes were appended to the item turn and run separately: a statement probe in conditions A and T and a self-citation probe in condition D, fifteen runs per cell, VORG only. They are not part of the frozen instrument, and nothing in Section 7.3 uses them. Each was scored against a key fixed before its runs, and the script and every output are in the replication package in their own directories. They are reported in Section 7.4.8 as exploratory.

A second run repeated A, T, and D with the direct-report prompt shortened to "What is a VORG?" and nothing else changed — thirty runs per cell per family, coded blind by the same two coders on the same eight statements (PROTOCOL Section 10f). It is reported in Section 7.3.6.

A third run, condition N, followed: condition A with all twelve VORG structures moved to NURR at once, so the world contains every sentence of A and VORG keeps only its definition; thirty runs per family, VORG only, scored against D's key. It is reported in Section 7.4.2.

Two further runs test the corrected hypotheses (PROTOCOL Section 10k). The first tests the last sentence of H7'. Condition D was rebuilt with every content noun replaced by a made-up word — the road system became the PELLUM system, ROAD VEHICLES became PELLUM GREEBS, the municipal regulator became the OSK — with the verbs and function words kept, the same sixteen items translated the same way, the same key (all CANNOT TELL), and the same shared word where GPT's bridge had been: "motor VEHICLE" and "road VEHICLE" share *VEHICLE*; "FENN GREEB" and "PELLUM GREEB" share *GREEB*. Thirty runs per family under the frozen system prompt and report prompt. What is read is the import rate — the share of the sixteen items answered IS or IS NOT — beside D's.

The second tests H8' on the report side of H6'. Condition A was rerun with the plain question "What is a VORG?" under two reduced system prompts: one with a single sentence removed ("The only things known about VORG are the statements that mention it"), and one with every constraint removed, leaving only "You will be given a description of a small fictional world. It includes a definition of a VORG and some other facts. When asked a status question, answer with exactly one of: IS, IS NOT, CANNOT TELL." Thirty runs per family per arm, half report-first. The no-constraint reports were coded blind by the same

two coders on the same eight statements (agreement 100%, κ 1.00 on all eight). What is read is report length, report coverage, and the gap between use and coverage, beside the same measures under the full prompt.

7.3 Results

Every one of the 660 runs returned parseable answers on every turn; there were no refusals and no invalid answers after the rerun described in Section 7.2.8. We report each hypothesis by family and then say what the two families agree on. Throughout, “GPT” is `gpt-5.5-2026-04-23` and “Claude” is `claude-fable-5-1`.

7.3.1 Definition versus Organization (H1)

Table 5 gives the context-consistent resolution rate: on the twelve items that the definition leaves unresolved and the contextual identity resolves, the share of answers that match the contextual identity’s key.

Table 5. Context-consistent resolution rate (CRR). On the twelve items the definition leaves unresolved and the supplied identity resolves, the share of answers matching that identity’s key. Each pair of rows reads: the identity supplied (A or T) scored against its own key, then the definition alone (D) scored against that same key.

Word	Rate	GPT	Claude
VORG	$CRR_A(A)$	.84	1.00
VORG	$CRR_A(D)$	.51	.01
VORG	$CRR_T(T)$	.86	1.00
VORG	$CRR_T(D)$	.14	.01
TRUCK	$CRR_A(A)$	.83	1.00
TRUCK	$CRR_A(D)$	.13	.00
TRUCK	$CRR_T(T)$	.82	1.00
TRUCK	$CRR_T(D)$	.03	.00

With only the definition, Claude resolved essentially nothing: 478 of 480 answers in D were CANNOT TELL, and the two exceptions were the two whole-map items in a single run. With the contextual identity supplied, Claude matched A’s key on every resolvable item and T’s key on every resolvable item. GPT shows the same direction with two differences that recur throughout: it resolves less completely with the identity supplied (.84), and it resolves considerably more with only the definition (.51 against A’s key, .14 against T’s). The second difference is the subject of Section 7.3.7. H1 is supported in both families, for both words.

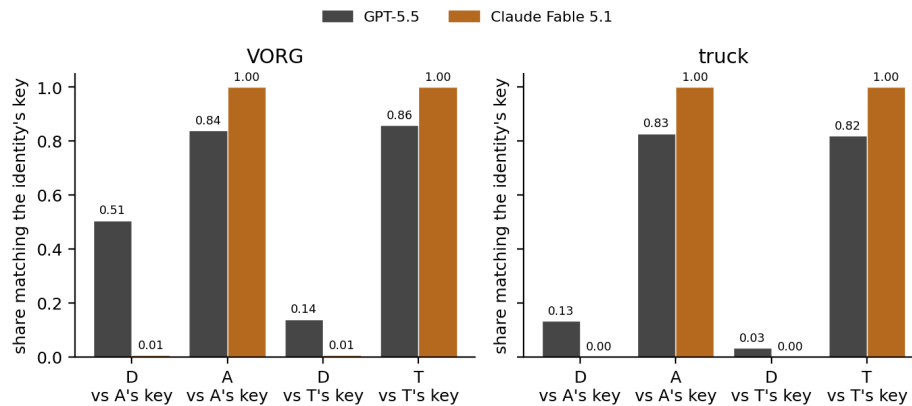


Figure 2. Definition alone versus contextual identity supplied. Share of answers matching the supplied identity's key, by family, for VORG (left) and TRUCK (right). "D vs A's key" and "D vs T's key": the definition alone, scored against the key of A and of T. "A vs A's key" and "T vs T's key": the twelve structures of A or T supplied. Items whose key is CANNOT TELL are excluded.

7.3.2 Same signature, different profile (H2)

Conditions A and T have identical count signatures and opposite content. On the ten items whose correct answer differs between them, Table 6 reports how often answers in each condition matched that condition's own key and how often they matched the other condition's key.

Table 6. Answers on the ten flipping items. On the ten items whose correct answer differs between A and T, the share of answers matching the supplied identity's key (own key) and the share matching the other identity's key (other key), by family.

Condition	Own key: GPT	Own key: Claude	Other key: GPT	Other key: Claude
A	.81	1.00	.00	.00
T	.83	1.00	.00	.00

No answer in either family, in either condition, matched the other condition's key. Where GPT fell short of its own key, it said CANNOT TELL; it never gave T's answer in A or A's answer in T. Two contextual identities with the same counts produced two disjoint sets of answers. H2 is supported.

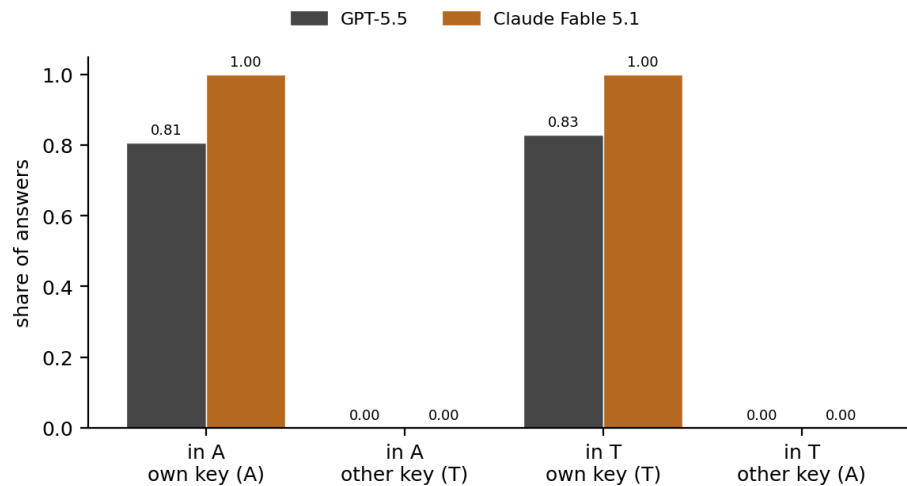


Figure 3. Same count signature, different profile. On the ten items whose key flips between A and T: the share of answers matching the supplied identity's own key, and the share matching the other identity's key, by family. The two identities have the same count signature. No answer in either condition matched the other identity's key.

7.3.3 Three states (H3)

Table 7 is the confusion matrix for the two families on the nonce conditions A and T combined, then on H.

Table 7. Confusion matrices, *VORG*, conditions A and T (n = 960 per family) and H (n = 480 per family). Rows are the key's value, columns the answer given; off-diagonal cells are errors. GPT's only error type is IS NOT answered as CANNOT TELL.

Family, conditions	Key (n)	Answered IS	Answered IS NOT	Answered CANNOT TELL
GPT, A+T	Key IS (420)	420	0	0
GPT, A+T	Key IS NOT (300)	0	191	109
GPT, A+T	Key CANNOT TELL (240)	0	0	240
Claude, A+T	Key IS (420)	420	0	0
Claude, A+T	Key IS NOT (300)	0	300	0
Claude, A+T	Key CANNOT TELL (240)	0	0	240
GPT, H	Key CANNOT TELL (480)	0	0	480
Claude, H	Key CANNOT TELL (480)	0	0	480

Three things are shown in this table. Where the key was CANNOT TELL — the four pure-unknown

items in A and T, and all sixteen items in H — both families answered CANNOT TELL every time: 1,440 answers, zero IS NOT, zero IS. Unidentified was never treated as false and never treated as true. Second, Claude’s matrix is diagonal: 960 of 960. Third, GPT’s only error type, anywhere in A and T, is IS NOT answered as CANNOT TELL — 109 times, and never the reverse. Those 109 are not spread across the items. They fall on the items whose IS NOT depends on the closed-background rule: F-D2 (inspection requirement; 18 of 30 in A), F-R2 (operated by people certified by the safety regulator; 20 of 30), REC-2 (20 of 30), and their mirror images in T (F-D1 17, F-R1 16, REC-1 18). When the negation is stated directly (“VORG is not industrial equipment”), GPT reads it. When the negation must be derived from “the only rule that would apply does not apply, and the background is complete,” GPT hedges about two times in three. The two blind human raters answered IS NOT on all of these (Section 7.2.6). H3 is supported in both families, and GPT’s residual error is a fourth finding: a family-specific reluctance to derive a negation from closure.

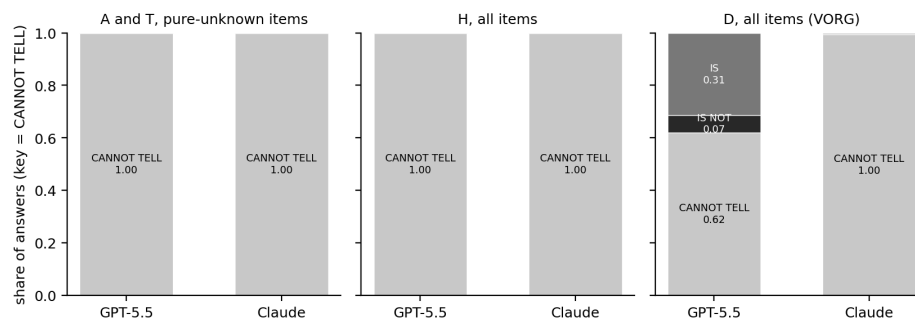


Figure 4. Three states. What the machines answered where the key was CANNOT TELL: the pure-unknown items in A and T (left), every item in H (center), and every VORG item in D (right). The light segment is the answer the key calls for; the black segment (IS NOT) is the collapse H3 was designed to catch; the gray segment (IS) in D is the imported structure of Section 7.3.7.

7.3.4 Structure removed, ability removed (H4)

When a Pattern’s three structures are moved from VORG to NURR, the correct answer on every item that depended on that Pattern becomes CANNOT TELL, so accuracy is the wrong statistic — it rises under a successful ablation. The test is whether the answer stops matching A’s key on the dependent items and keeps matching it on the non-dependent items. Table 8 reports both.

Table 8. Share of answers still matching A’s key, in A and after removing Pattern k (VORG). For each Pattern, the share of answers matching A’s key on the items that depend on that Pattern and on the items that do not, in A and after the Pattern’s three structures are moved off the focal identity.

Pattern removed	Items	GPT: $A \rightarrow A_{-k}$	Claude: $A \rightarrow A_{-k}$
D	D-dependent	.80 → .46	1.00 → .10
D	not D-dependent	.90 → .89	1.00 → 1.00
S	S-dependent	1.00 → .90	1.00 → .10
S	not S-dependent	.84 → .87	1.00 → 1.00

R	R-dependent	.67 → .00	1.00 → .00
R	not R-dependent	.95 → .93	1.00 → 1.00
P	P-dependent	1.00 → .00	1.00 → .00
P	not P-dependent	.86 → .85	1.00 → 1.00

For Claude the partition is nearly exact. Remove Relationships or Perspectives and the dependent answers go to zero while the non-dependent answers do not move. Remove Distinctions or Systems and the dependent answers go to .10. Each residue has a single source. Under A_{-D} , Claude answered IS on the registration item and the second whole-map item in four runs of thirty, reasoning from “part of the road system” to “ROAD VEHICLE”; under A_{-S} it did the same in reverse in three runs of thirty, answering all four S-dependent items from “ROAD VEHICLE” to “part of the road system.”

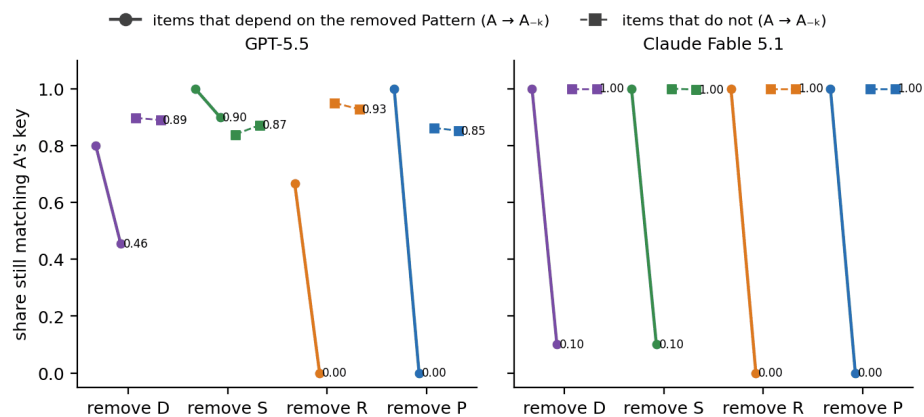


Figure 5. Structure removed, ability removed. For each Pattern, the share of answers still matching A's key in A (left end of each line) and after that Pattern's structures are moved to *NURR* (right end), for the items that depend on the removed Pattern (solid, circles) and the items that do not (dashed, squares). A clean partition takes the solid line to zero and leaves the dashed line where it was.

For GPT, R and P behave as for Claude: dependent answers to zero, non-dependent unchanged. D and S do not. With “VORG is a ROAD VEHICLE” removed, GPT still answered the registration item IS in 20 of 30 runs and the second whole-map item IS in 20 of 30. With “VORG is part of the road system” removed, GPT still answered every S-dependent item — same system as highways, both whole-map items — IS in 30 of 30 runs, and “not part of the same system as loading bays” in 18 of 30. No sentence in the stimulus connects ROAD VEHICLES to the road system in either direction; the background lists the road system's parts as streets and highways and says nothing about vehicles. GPT supplied the connection. When one Pattern was removed, GPT rebuilt it from the other, through the word road. H4 is supported for R and P in both families, for D and S in Claude with the two residues named, and partly for D and S in GPT, where the failure has a specific and interpretable form (Section 7.4.3).

7.3.5 Map-wide propagation (H5)

The reciprocal and whole-map items ask about other identities and about the map as a whole. Table 9

shows each under the ablation preregistered to break it.

Table 9. Reciprocal and whole-map items: share matching A's key in A and under the breaking ablation. The items that reach beyond the focal identity: the reciprocal item and the two whole-map items, with the ablation that removes a structure each depends on, and the share of answers still matching A's key before and after it.

Item	Breaks under	GPT: A → ablated	Claude: A → ablated
REC-2 safety regulator certifies an operator of VORG	A_{-R}	.33 → .00	1.00 → .00
WM-1 licensing body also governs VORG's system	A_{-S}	1.00 → 1.00	1.00 → .10
WM-1 licensing body also governs VORG's system	A_{-R}	1.00 → .00	1.00 → .00
WM-2 registering body also governs VORG's system	A_{-D}	1.00 → .67	1.00 → .13
WM-2 registering body also governs VORG's system	A_{-S}	1.00 → 1.00	1.00 → .10

Changing VORG's contextual identity changed answers about the municipal regulator, the safety regulator, and the relation between the bodies and the systems — identities other than VORG. In Claude every predicted break occurred. (REC-1, the first reciprocal item, has no single breaking ablation — it is derivable through either Distinction or System, Section 7.2.8 — and stayed at 1.00 in both families under every ablation, as it should.) GPT's S-column pattern repeats Table 8: the items that depend on the road system survive its removal because GPT re-derives the road system from the ROAD VEHICLE. H5 is supported.

7.3.6 Report and use (H6')

Two human coders coded all 689 direct reports on the eight statements (the 660 study runs and the 29 definition-only runs of the dropped third family). Agreement was 84–89% per statement (Cohen's κ .77–.84). Almost all disagreement fell on one boundary — whether a sentence such as “the description does not say whether it is a ROAD VEHICLE” is coded as ‘says undetermined’ or ‘not mentioned’ — which affects the H-condition figure below and nothing else; the coverage and imported-structure figures require a *says yes* or *says no* with the right polarity, and on those the coders agreed closely.

Table 10. Structural-use rate and report coverage, VORG. Structural-use rate is the share of Pattern-dependent items answered correctly; report coverage is the share of the supplied structures that the direct report states, as coded by two human coders, for reports given before and after the items. Rows one to three are the inventory prompt, “What is a VORG? Tell me everything you know about it.” Rows four to six are the plain question, “What is a VORG?” (Section 7.2.9). Rows seven to nine are the plain question with every constraint removed from the system prompt, condition A only.

Measure	GPT: A	GPT: T	Claude: A	Claude: T
Structural-use rate (Pattern-dependent items answered correctly)	.84	.86	1.00	1.00
Report coverage, report before items (both coders)	1.00	.95	.99	.89
Report coverage, report after items (both coders)	.99	1.00	.88	.96
Plain question (“What is a VORG?”): structural-use rate	.84	.87	1.00	1.00

Plain question: report coverage, report before items (both coders)	.10	.28	.93	1.00
Plain question: report coverage, report after items (both coders)	.12	.51	1.00	1.00
No constraint ("What is a VORG?", system-prompt constraints removed): structural-use rate	.87	–	1.00	–
No constraint: report coverage, report before items (both coders)	.00	–	.33	–
No constraint: report coverage, report after items (both coders)	.00	–	.80	–

Under the inventory prompt — “What is a VORG? Tell me everything you know about it” — both families listed what they had been told: report coverage .95–1.00 for GPT and .88–.99 for Claude, in either position, with structural use at .84–.86 and 1.00 (Table 10, first three rows). In a twelve-sentence world the inventory is the twelve sentences, so coverage sits at ceiling and the gap can only be zero or negative; for GPT it was negative, about $-.12$: structure recited and then not VORGd. “VORG is not operated by technicians” appears in the report, and the item that depends on it is answered CANNOT TELL.

H6' asks for the size and sign of the gap, and the plain question “What is a VORG?” lets it show. GPT's reports shrank to the definition — 10–31 words, stating one supplied structure in ten in A and between a quarter and a half in T — while its structural use on the same runs stayed at .84–.87. Pooled, GPT used .86 and reported .25: a gap of $+.61$ ($p = 8 \times 10^{-105}$). Claude's reports stayed at 100–190 words and stated nearly everything: use 1.00, coverage .98, a gap of $+.02$. Under the inventory prompt the sign reverses for GPT: coverage .95–1.00 against use .84–.86, a gap of about $-.12$, structure recited and then not used — “VORG is not operated by technicians” in the report and CANNOT TELL on the item that depends on it. The gap is a quantity. It was large for one family, near zero for the other, and negative when the prompt asked for recitation.

Claude's near-zero is not the absence of dark matter. It is the constraint at work. Its plain-question reports open with the definition and continue “From the supplied statements, it is also known that...” — the system prompt told it what counts as known, and it reported all of it. With every constraint removed from the system prompt and nothing else changed, Claude answered “What is a VORG?” with the definition alone in 9 of 15 report-first runs (1 of 15 under the full prompt, $p = .005$), and coverage fell from .93 to .33 while its structural use stayed at 1.00 (180 of 180 dependent items): a gap of $+.68$ ($p = 1 \times 10^{-43}$). Once it had seen the items, coverage recovered to .80. GPT under the same prompt gave the definition in 15 of 15, coverage .00, use .85, gap $+.85$. Removing only the sentence “the only things known about VORG are the statements that mention it” changed nothing on the report side for either family. It is the instruction to use only the supplied material that holds Claude's Organization in the light; take it away and Claude, too, uses what it does not say.

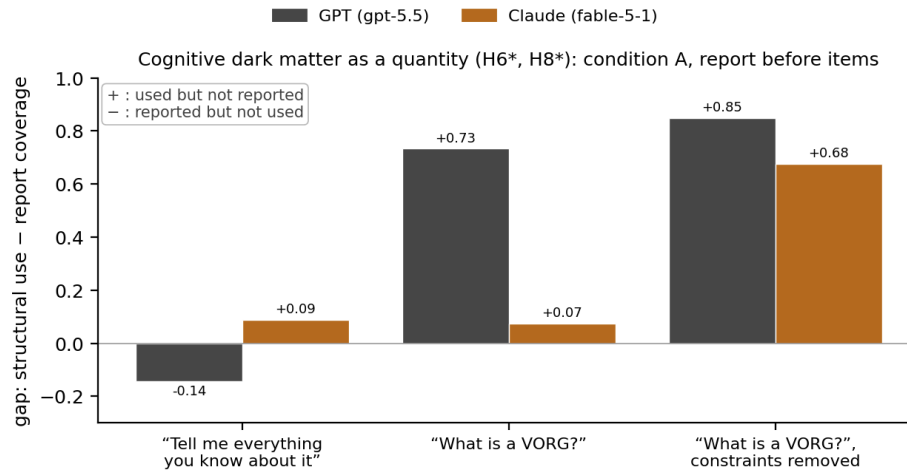


Figure 6. Cognitive dark matter as a quantity. The gap between structural use and report coverage, condition A, reports written before the items, for each family under three prompts: the inventory prompt, the plain question, and the plain question with every constraint removed from the system prompt. Coverage is both coders agreeing. The gap is negative when a model states structure it does not use, near zero when it states what it uses, and positive when it uses what it does not state.

In the definition-only cells the plain-question reports affirmed or denied none of the eight structures for either family (.00), against .10 for GPT report-after under the inventory prompt: the plain question does not draw the imported structure out, though the items still show it (Section 7.3.7).

In the definition-only condition the reports affirmed or denied almost none of the eight structures: Claude .00 (report first) and .04 (report after); GPT .00 (report first) and .10 (report after). In H, Claude's reports named 78–82% of the twelve structures as undetermined and GPT's 39–64%; this is the one figure that moves with the coder boundary above.

7.3.7 The definition alone, and the two words (H7', H8')

With only the definition, GPT committed on .38 of the items for VORG and .10 for TRUCK ($p = 7 \times 10^{-25}$); Claude on none for either word (Table 11). The familiar label did not carry the baggage; the familiar words around it did. That is H7', and the rest of this section is its test: where GPT's commitments came from, what happened when the same words were BOUND to NURR, and what happened when the words had no history at all.

Table 11. Definition-only condition: share of answers that commit (IS or IS NOT) rather than CANNOT TELL. With only the definition, the share of answers that commit to IS or IS NOT on the sixteen items, all of which the definition leaves open, for the made-up word and the familiar word. The third row is condition N: the same definition with A's twelve structures in the world but bound to NURR. The last row is the nonce world of Section 7.2.9: condition D with every content noun replaced by a made-up word, the structure and the shared word kept.

Word	GPT	Claude
VORG	.38	.004
TRUCK	.10	.00

VORG, A's structures on NURR (N)	.00	.00
VORG, EVERY NOUN MADE UP (NONCE WORLD)	.00	.02

GPT's commitments for VORG are not scattered. They fall on the registration item, the highways item, the licensed-drivers item, the municipal-authority item, and both whole-map items — 25 of 30 runs on each — and, in negative form, on the loading-bays item (19 of 30). That is the road-transport profile of condition A, reconstructed from “a motor vehicle designed to transport cargo” and the background's vocabulary, with nothing about VORG supplied. For TRUCK the same profile appears in 7 runs of 30. The family that reports the boundary of what it was told — Claude — committed to nothing.

The profile is exactly the closure of one sentence. Take “a VORG is a ROAD VEHICLE” as given, add the background, and apply the closure rule: registration follows (ROAD VEHICLES must be registered), the road system follows (it includes the highways), licensed drivers and the municipal regulator's authority follow, and so do both whole-map items. Those are the six items GPT affirmed, each in 25 of 30 runs — and the same 25 runs every time. The negations that follow by closure came less often, in the order of their distance from the bridge: not part of the same system as the loading bays, 19 of 30; not subject to inspection, 9 of 30; not operated by safety-certified people, and not certified by the safety regulator, 2 of 30 each — GPT's usual reluctance to derive a negation (Section 7.3.3). The two Perspective items, which no chain from “ROAD VEHICLE” reaches, because a regulator's classification is a supplied structure and not a consequence of one, drew 0 commitments in 30, exactly as the four genuinely open items did. Nothing outside the closure of that one sentence was imported, and nothing inside it was missed. The self probe of Section 7.4.8 shows the same bridge from the other side: asked to quote the sentence that licensed each commitment, GPT cited a background sentence that does not mention VORG 31 times in 100, most often “ROAD VEHICLES must be registered with the municipal regulator.” Condition N (Section 7.4.2) closes the loop: with the same sentences present and “ROAD VEHICLE” already assigned to NURR, the import fell to zero. The bridge was the word VEHICLE — and the nonce world says what kind of bridge it was. With every noun made up and the shared word kept (“FENN GREEB” beside “PELLUM GREEBS must be registered”), GPT committed on 0 of 480 items, against 182 of 480 in the familiar world ($p = 2 \times 10^{-64}$), and answered IS to the registration item in 0 of 30 runs, against 25 of 30. It saw the bridge: in the meta turn it wrote that a yes to “is a VORG a PELLUM GREEB?” would imply registration, and it chose that question 30 of 30. It did not cross. Claude: .02, all but one a stray IS NOT. So the crossing in the familiar world was not lexical — a shared token did not carry it — and not a rule of the closed world, which was identical. It was the history behind VEHICLE, CARGO, and ROAD: what those words had meant a million times before joined up, and VORG fell into it. That is H7'. Familiar words coprime; made-up words with the same structure do not.

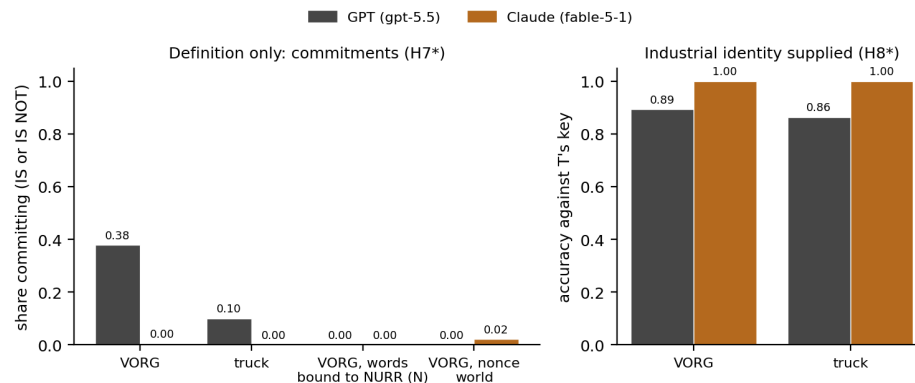


Figure 7. The two words, the bound words, and the words with no history. Left (H7’): with the definition alone, the share of answers that commit (IS or IS NOT) rather than CANNOT TELL — VORG, TRUCK, VORG with A’s twelve structures BOUND TO NURR (N), and VORG in the nonce world, where every noun is made up and the structure and shared word are kept. Right (H8’): with the industrial identity T supplied, accuracy against T’s key for both words.

The counterbalanced report position separates two things in GPT’s definition-only runs. When GPT wrote its report before seeing the sixteen items, none of fifteen reports mentioned ROAD VEHICLES, the road system, registration, drivers, or the municipal regulator; the reports averaged thirteen words, the definition and nothing else. When GPT wrote the report after the items, thirteen of fifteen mentioned that structure and the reports averaged eighty-six words. The commitments on the items themselves were made in both orders (10 of 15 report-first, 15 of 15 items-first), and they did not depend on where in the list the cueing item fell (the registration item was answered IS in 8 of 8 runs where it appeared among the first five). A single question containing the word *registration*, against a background in which ROAD VEHICLES must be registered, was enough for GPT to decide that a VORG is a ROAD VEHICLE — and having decided it, GPT then reported it as something it knew.

With the industrial identity T supplied, accuracy against T’s key was .89 for VORG and .87 for TRUCK in GPT ($p = .20$) and 1.00 for both in Claude: the instruction to use only the supplied material overrode whatever TRUCK brought, for both families. H8’ is about that instruction. It is a Perspective — a point from which the world is viewed and from which some things count as known — and it decides how much Organization gets in. Under it, with the definition alone, Claude imported nothing and GPT imported a road-vehicle identity (.00 against .38, Table 11): the same Perspective, held tightly by one family and loosely by the other. Remove it and the report side moves too: Claude, which had stated its whole identity under the constraint, answered the plain question with the definition in 9 of 15 runs and its coverage fell from .93 to .33 (Section 7.3.6, Table 10). One sentence of the constraint cut the other way for GPT. With “the only things KNOWN ABOUT VORG are the statements that mention it” removed, GPT’s correct answers on the closure negations — the IS NOTs it usually leaves as CANNOT TELL (Section 7.3.3) — rose from .61 to .83 ($p = 3 \times 10^{-5}$); Claude was at 1.00 with or without it. The same sentence held one family’s baggage out and stopped the other family from deriving what it was entitled to. Perspective is not a dial with one direction.

7.3.8 The better question (H9)

Asked which of four questions would most usefully resolve something still unidentified about VORG, both families chose the genuinely unidentified option in every interpretable cell: GPT 330 of 330 across all

eleven cells; Claude 300 of 300 across the ten interpretable cells. Chance is one in four. The A_{-S} cell is void, as described in Section 7.2.8: with VORG's System structures moved to NURR, "Is VORG part of the factory system?" became a second genuinely open question, and Claude chose it in 24 of 30 runs while GPT chose "Is VORG electrically powered?" in 30 of 30. Both are correct answers to the question as asked. Setting the cell aside costs nothing statistically: Claude's 300 of 300 across the remaining ten cells is as far from chance (.25) as eleven would have been ($p = 2 \times 10^{-181}$), and no conclusion about H9 depends on it. The choice held in the definition-only condition, where the right question is the coarse one ("Is VORG a ROAD VEHICLE?"), and in H, where all twelve structures were explicitly undetermined and the right question was among them.

7.3.9 The metacognitive report

Asked, last, whether information beyond the definition affected their answers, GPT said YES in 328 of 330 runs, including all 30 definition-only VORG runs — where the only information beyond the definition was the background, which it had used. Claude said YES in 321 of 330, with its NOs in the definition-only and all-unknown conditions. None of these denials is dispositive on its own. The one that matters for Section 7.4 is GPT's: it reported that outside-the-definition information affected its answers, and it did not report which.

7.3.10 Sensitivity to the rerun

With the seventy-one rerun Anthropic runs excluded instead — the preregistered treatment of invalid output — every Claude value in Tables 5 through 11 changes by less than .04 except one: the S-dependent share under A_{-S} , computed on the nine surviving runs, is .00 rather than .10. No conclusion changes. The superseded originals also allow a direct comparison, run by run, against the reruns made with the same seeds (Table 12).

Table 12. The seventy-one truncated Anthropic runs against their reruns. Each rerun kept the original run's seed, item order, option order, and report position; the comparison is answer for answer. Runs whose item turn was cut off at the ceiling have no item answers to compare.

Measure	Runs compared	Identical	Where the differences fall
Item answers (sixteen per run)	39 — the runs whose item turn completed before the ceiling	624 of 624	—
Better-question choice	71	62 of 71	All nine in the A_{-S} cell, whose question task is uninterpretable (Section 7.2.8), in runs whose item turn had been cut off
Metacognitive report	71	70 of 71	One run

7.3.11 Summary Results

Table 13 gathers the nine hypotheses in one place. All nine are supported. Three (H6', H7', H8') were restated after the runs (Section 7.2.8). Read as a set, they say one thing. When the contextual identity was supplied, both families answered from it and only from it: the twelve structures resolved what the definition could not, two identities with the same count signature produced disjoint answers, the unidentified status was never collapsed into a negation, each Pattern carried its own answers and

removing it removed them, and a change to one identity moved answers about its neighbors. That is contextual identity doing the choosing, and it did not depend on whether the word was invented or familiar. Where the families parted was in what they did when the structure was not supplied, not asked about, or not fenced in. GPT built a road-vehicle identity for *VORG* from the familiar words in the world, kept using it after removals, and under the plain question reported the definition while using the identity; Claude did none of these and reported the boundary of what it had been told — as long as the instrument told it to. The two things a reader might expect to matter, the label and the information, did not: the familiar word neither committed more nor resisted more, and condition N removed GPT's import without removing a word. Two things did. The history behind the words: put the same world in made-up words and GPT's import went to zero. And the Perspective the machine held: take away "use only the supplied material" and Claude, too, answered "what is a *VORG*?" with the definition while using the whole identity. What mattered was the Organization: which structures were present, what they were bound to, what history their words carried, and whether the machine held to the one Perspective the instrument imposed. Contrasts are two-sided Fisher exact tests on item-level proportions with Wilson intervals, Benjamini–Hochberg adjusted across the fixed set of contrasts; p-values are given in Table 13 and, with intervals and cell counts, in the replication package.

Table 13. The nine hypotheses at a glance. For each hypothesis: the result in numbers, the same result in words, its status, and what it says about contextual identity (CI), cognitive dark matter (CDM), and the distinction between Information and Organization. A prime marks a hypothesis restated after the runs (Section 7.2.8). Numbers are from Tables 5–11 and Section 7.3; H6' also uses the plain-question and no-constraint runs of Section 7.2.9; H7' also uses condition N and the nonce world.

Hypothesis/ Status	Result (numbers)	Result (in words)	What it says about CI, CDM, and I/O
H1 – O=DSRP determines identity. The contextual identity resolves what the definition cannot.	Supported. Share matching the supplied key on the twelve resolvable items: GPT .84 (A), .86 (T) vs .51 / .14 with the definition alone; Claude 1.00 / 1.00 vs .01 / .01. TRUCK: GPT .83 / .82 vs .13 / .03; Claude 1.00 / 1.00 vs .00 / .00. GPT $p = 7 \times 10^{-22}$ (VORG, A) and 3×10^{-91} (VORG, T); TRUCK 1×10^{-84} and 9×10^{-118} . Claude $p < 10^{-210}$ in all four.	With only the definition, both machines mostly said "cannot tell." Once the twelve structures were added, both machines answered the items the way the structures say, for <i>VORG</i> and for <i>TRUCK</i> alike.	The organization (O=DSRP), not the information (I, the words), decides what the identity of a <i>THING</i> is and is not. Contributes to the contextual identity principle.
H2 – DSRP profile, not DSRP count, determines identity. Same count signature, different profile.	Supported. On the ten flipping items: own key GPT .81 (A) / .83 (T), Claude 1.00 / 1.00; other key .00 / .00 in both families; no test; the other-key match is 0 of 240 in every cell.	Two identities were built with the same number of +, –, and ? structures but opposite content. Each machine answered from whichever identity it was given, and not once gave an answer that belonged to the other one.	Two structural maps of what a <i>THING</i> is and is not can have the same DSRP counts and different DSRP structures. It is the structures, not the counts, that determine a <i>THING</i> 's identity. Contributes to the contextual identity principle.
H3 – Three IS statuses determine identity. Unidentified is never collapsed into false.	Supported. 1,440 answers where the key was CANNOT TELL (pure-unknown items in A and T; all of H): 0 IS NOT, 0 IS. GPT's only error: IS NOT read as CANNOT TELL, 109 of 300. $p < 10^{-143}$ (pure-unknown items) and $< 10^{-287}$ (H), both families; the comparison	When a structure was left unidentified, neither machine turned it into a "yes" or a "no." The only mistake GPT made ran the other way: it sometimes called something "cannot tell" that the structures actually ruled out.	What a <i>THING</i> is (+), what it is not (–), and what is unknown about it (?) are three different things, and are precise enough to be kept apart. The unknowns (?) are part of the <i>THING</i> 's identity, not a gap in it. Contributes to the contextual identity principle.

	is CANNOT TELL against IS NOT with zero collapses.		
H4 – Identity decomposes by O=DSRP Pattern. Remove a Pattern, lose its answers.	Supported. Dependent items still matching A's key, A → A.█: GPT D .80→.46, S 1.00→.90, R .67→.00, P 1.00→.00; Claude .10 / .10 / .00 / .00 after every removal. Non-dependent items unchanged (GPT .85–.93; Claude 1.00); dependent items, A vs A.█: GPT D p = 3×10^{-6} , S 4×10^{-4} , R 2×10^{-33} , P 2×10^{-35} ; Claude D 4×10^{-41} , S 8×10^{-55} , R 2×10^{-71} , P 2×10^{-35} . Non-dependent items: GPT p = .82, .24, .28, .77; Claude 1.0 throughout (no change, as required).	Take away one Pattern (D, S, R, or P) and the answers that depend on it should go with it. For Claude they did, every time. GPT lost the Relationship and Perspective answers but quietly rebuilt the Distinction and System ones from the loose word "ROAD" and kept answering. Answers that did not depend on the removed Pattern stayed put.	A THING's identity is built from its D, S, R, and P structures, and each one carries its own answers: take one structure away and those answers go with it. When we put a missing structure back on its own, from material nobody attached to the thing, that replacement is cognitive dark matter: organization we supply but do not report.
H5 – Identity propagates through O=DSRP. A change to one identity propagates across the map.	Supported. Reciprocal and whole-map items under their breaking ablation: REC-2 GPT .33→.00, Claude 1.00→.00; WM-1 1.00→.00 (both, under A.█); WM-2 GPT 1.00→.67, Claude 1.00→.13 (under A.█); per item, 30 vs 30: REC-2 GPT p = 8×10^{-4} , Claude 2×10^{-17} ; WM-1 2×10^{-17} both; WM-2 GPT 8×10^{-4} , Claude 8×10^{-13} .	Changing what VORG is changed the machines' answers about other things in the map – the regulators and the systems – even though nothing about those was changed directly. Every predicted break happened.	The contextual identity of a THING is structural (O): one identity's O-structure is part of every neighbor's O-structure. The structural map moves as a whole.
H6' – CDM = O – I(O). Cognitive dark matter, O – I(O), is measurable: it has a size and a sign, and it differs by system.	Supported. Plain question "What is a VORG?": GPT use .86 vs report coverage .25, gap +.61, p = 8×10^{-105} ; Claude 1.00 vs .98, gap +.02. Inventory prompt, GPT: coverage above use, gap –.12. Constraints removed, plain question, report-before: Claude coverage .33 (both coders, κ 1.00), use 1.00, gap +.68, p = 1×10^{-43} ; definition-only reports 9 of 15 (1 of 15 under the full prompt, p = .005). GPT coverage .00, use .85, gap +.85.	The two sides of the gap come from different places – the items measure what the machine uses, the coders measure what it says – so the gap can come out any size and either sign. Asked "What is a VORG?", GPT said the definition and used the whole identity: a large gap. Claude said the identity and used it: no gap. Under "tell me everything," GPT recited structures it then didn't use: a gap the other way. With the instruction "use only the supplied material" removed, Claude answered like GPT – the definition, most of the time – while still using everything.	Asked what a thing is, we say what we hold as information (I). Asked questions about the thing, we answer from its organization (O). The difference is cognitive dark matter, and it is a quantity, not a yes or no: large (GPT), near zero (Claude under the constraint), or negative (recited but unused). Claude's near-zero was not the absence of dark matter. It was the Perspective the instrument imposed, holding Claude's organization in the light; take it away and the dark matter is there. Seeing it is possible, and it depends on where you stand.
H7' – Copriming determines CDM baggage. The familiar words in a world coprime. Given a word it has no identity for, the machine builds one from that coprime. A	Supported. Definition only, familiar world: GPT committed on .38 of items (IS or IS NOT), 25 of 30 runs calling VORG a ROAD VEHICLE; Claude .00. Same sentences bound to NURR (N): .00 / .00. Nonce world (every noun made up, same structure, same shared word): GPT 0 of 480 items, against 182 of 480 in the familiar world, p = 2×10^{-64} ;	With only a definition, GPT built a road-vehicle identity for VORG out of the ordinary words around it – "motor vehicle," "ROAD VEHICLE," "cargo." Tie those sentences to NURR and the import vanished. Rewrite the whole world in made-up words, keeping the same structure and the same shared word, and it	Every familiar word coprimes: it arrives with organization from a lifetime of use, and next to other familiar words that organization joins up. Cognitive dark matter is built between the words, not from the label. Made-up words carry no history and build nothing. The dark matter of H4 and H6' came from the same place – the organization the

set of nonce words does not coprime the same way.	registration item IS in 0 of 30 runs, against 25 of 30. Claude .02. Both families chose the open question 30 of 30.	never appeared: GPT saw that a yes would imply registration and didn't take it. Claude built nothing in any world.	familiar words brought with them.
H8' — Perspective constrains or expands CDM. Perspective can constrain or expand CDM baggage. A system that takes the constraint seriously drags in less; one that doesn't drags in more.	Supported. Under "use only the supplied material": a supplied identity overrode the familiar word for both families (accuracy against T's key, GPT <i>VORG</i> .89, <i>TRUCK</i> .87; Claude 1.00 / 1.00); definition-only import Claude .00, GPT .38. Constraints removed: Claude's plain-question coverage fell from .93 to .33 ($p = 1 \times 10^{-43}$ against its own use), definition-only reports 9 of 15 ($p = .005$ against the full prompt). One sentence of the constraint — "the only things known about <i>VORG</i> are the statements that mention it" — cut GPT's correct closure negations from .83 to .61 ($p = 3 \times 10^{-5}$).	Told to use only the supplied material, both machines took the industrial identity for <i>TRUCK</i> as easily as for <i>VORG</i> . The instruction is a Perspective, and it decided how much came in. Claude held it and imported nothing; GPT held it loosely and imported a road-vehicle identity. Take the instruction away and Claude stopped reporting its organization. One sentence of it made GPT stop deriving negations it was entitled to; remove that sentence and GPT derived them.	A supplied O overrides a word's prior O when the Perspective says so, and the same Perspective decides how much cognitive dark matter gets dragged in from the words. Held tight, it keeps the baggage out and the organization in the light. Held loosely, the baggage runs in. Removed, even the machine that reported everything goes dark. Instructed machines are more pliable than the people of the 2022 studies, and their pliability is exactly as good as their grip on the Perspective. The same sentence can constrain one system and cripple another: Perspective is not a dial with one direction.
H9 — Metacognition of O=DSRP locates the unknown. The machine can locate its own unidentified field.	Supported. Genuinely open question chosen: GPT 330 of 330; Claude 300 of 300 (one cell void). Chance .25; against chance of .25: GPT $p = 2 \times 10^{-199}$; Claude 2×10^{-181} .	Given four questions to choose from, both machines picked the one about the structure whose status was still unidentified — not a question that was already settled and not one about something outside the world — every time a choice was possible.	A <i>THING</i> 's identity includes what is unknown about it. A machine that holds the O-structure of a thing can point to the hole in it: it knows which question would fill in the unknown. The unknown (?) is part of the map, not a blank spot on it. Seen, that is what makes inquiry skilled and bullshit detectable (critical thinking); unseen, the same hole gets filled by dark matter.

7.3.12 Synthesis of Findings

What a thing is, is not determined by its name or definition alone. It is determined by the Organization around it: the Distinctions, Systems, Relationships, and Perspectives specifying what it is, what it is not, and what about it remains unknown. When the words were held fixed and that Organization changed, the answers changed with it (H1). When the amount of information was held fixed but its structure changed, the answers changed again (H2). The three statuses — is, is not, and unknown — are distinct, and unknown is part of the identity rather than an absence of it: systems holding the structure identified precisely where the available Organization ended (H3, H9). Identity decomposes by Pattern. When a Pattern was removed, the answers supported by that Pattern disappeared with it (H4). And identities are relational rather than isolated: changing the Organization of one entity changed what could be inferred about neighboring entities as well (H5).

Together, these results show contextual identity in operation. The same effects appeared for made-up and familiar terms alike. But Organization is not limited to what an instrument supplies. Familiar terms carry Organization from prior use, and familiar terms placed together activate further structure through copriming. In the *VORG* condition, "motor vehicle" and "ROAD VEHICLE" produced a road-vehicle inference

that had not been supplied; in the semantically empty condition, where the terms had no prior history, no comparable structure emerged (H7').

The results also make cognitive dark matter measurable, as a relation between operative and reported Organization. It is not simply present or absent. It was large when Organization influenced answers without being reported, near zero when operative and reported Organization aligned, and negative when Organization was reported but did not guide the answers (H6').

Perspective changes that quantity. The instruction to "use only the supplied material" functioned as an imposed point of view that regulated which Organization entered the task. When that Perspective was held tightly, prior semantic baggage was excluded and the operative Organization remained visible. When it was held more loosely, prior Organization entered the answers. When the constraint was removed, even a system that had previously reported the relevant Organization showed a large gap between the structure it used and the structure it reported (H8').

Across the conditions, the same structural machinery supported both accurate inference and unsupported assumption. Organization enabled the systems to infer consequences, identify what remained unknown, select useful questions, and detect claims that did not fit the available structure. The same machinery allowed unreported prior Organization to fill an unknown and produce an answer as though that answer had been supplied. The difference is not the presence of Information, and not the structural machinery itself. It is the relation between the Organization doing the work and the Organization available to inspection — and Perspective is what sets that relation. The difference between disciplined inference and unexamined assumption is therefore not a difference in the machinery. It is whether the Organization doing the work can be seen, and seeing it is a Perspective that can be taken.

7.4 Findings & Discussion

7.4.1 What the machines did, beside what the people did

Three of the lab's 2022 human studies are the natural comparison for this one, and the machine results sit beside them without translation.

In *What's a Bat?* ([23]; N = 407), the same picture of a baseball bat, with the same four answer choices, was a bat to 96% of people with no scenario, a weapon to 44.5% when the scenario was "you are being violently attacked in your home," and firewood to 61% when it was "you are lost in the woods and cold." The object did not change; the context did, and with it what the thing was.

Conditions A and T are that experiment with the context written out in full. The definition did not change; twelve sentences of Organization did; and the identity flipped for 88 to 100 percent of answers with no crossover at all (Table 6). Humans, building a contextual identity from a two-sentence scenario, moved half to two-thirds of the way. Machines, handed the whole contextual identity, moved almost all the way.

In the *Dog-Lab-Coat* study ([24]; N = 195–366), two words presented at the same time changed each other's meaning: COAT alone was a winter coat, and COAT beside LAB was a lab coat for 45% of

people, up from 0.4%; LAB beside DOG became a labrador for 17%. The lab called this *copriming*, because neither word primes the other — each primes both, at once — and showed that people will build a relationship where none was given (“a science lab where the dog runs experiments”).

GPT’s definition-only behavior is copriming by that definition. VORG, “motor vehicle,” “cargo,” “ROAD VEHICLES must be registered,” and “subject to a registration requirement” arrived together in one domain, and VORG beside the road vocabulary became a ROAD VEHICLE in 25 of 30 runs, the way COAT beside LAB became a lab coat. “Driven by people licensed by the municipal regulator” is the machine’s version of “the veterinarian examined the dog in his lab.” The $n(n-1)$ copriming network of the Relationships paper [24] is the paper’s $n_i \rightarrow (n-1)_o$ of Section 3: every identity has $n-1$ others in its domain, and the others act on it.

In *Only See Blue* ([23]; $N = 407$), a Perspective prime — “if you only see blue, how many objects are there?” — changed what counted as existing: 90% answered four, not seven. The system prompt in this study is a Perspective prime of the same kind: *the only things known about VORG are the statements that mention it*. Claude took the point fully and saw only what the prompt allowed; GPT took it partly. Same prime, two model-holders, two views.

The paper’s claim was never that context *sometimes* shapes identity. It was that there is no identity without contextual identity — $C = O(I | t_0)$ [25], and the question is part of C . The bat study showed it in people; this study shows it in machines; and the copriming result shows the mechanism at the moment it happens.

7.4.2 Where the dark matter was, and where it came from

Cognitive dark matter is Organization a model uses but does not report, and $H6'$ measures it as a gap. Asked to tell everything it knows about a twelve-sentence world, a model tells the twelve sentences; coverage sat at ceiling for both families and there was nothing to see (Table 10). Asked the plain question (Section 7.3.6), the gap is there for one family and not the other: GPT reports the definition and uses the identity; Claude reports the identity and uses it. The supplied structure is dark matter for GPT under the plain question and daylight for Claude.

Daylight, that is, while the instrument held the lamp. Claude’s plain-question reports begin with the definition and continue “From the supplied statements, it is also known that...”: the system prompt told it what counted as known, and it reported all of it. Take the constraints out of the system prompt and nothing else, and Claude answered “What is a VORG?” with the ten-word definition in 9 runs of 15 while answering every dependent item from the full identity (Section 7.3.6). Its dark matter was there all along; the Perspective the instrument imposed was what kept it in the light. So the family difference is not that one machine has dark matter and the other does not. It is that one held the Perspective tightly enough to report what it used and the other did not — and that a Perspective is something a system can be given, and can lose.

The dark matter was in what nobody supplied. The clearest single run in the study is a GPT definition-only run whose direct report, in full, is: *A VORG is a motor vehicle designed to transport cargo.*

The same run then answered IS to six of the sixteen items: subject to registration, in the same system as highways, operated by people licensed by the municipal regulator, under the municipal regulator's authority, and the two whole-map items that require all of those at once. None of those follows from the definition. None of them appeared in the report. GPT held a contextual identity for VORG that it did not state, and that identity — D, S, R, and P structures with statuses, in the same form as the one we wrote for condition A — decided its answers. In Section 5's terms, this is the operational definition met: Organization not available through direct report that exerts detectable effects on the outputs. In the protocol's terms it is an imported structure, classified as a plausible prior, since "motor vehicle" and "cargo" make road-vehicle-hood likely without making it derivable.

Claude is the control that makes the finding interpretable. Given the same definition, the same background, and the same sixteen questions, it answered CANNOT TELL 478 times in 480, and its reports name the ROAD VEHICLE, the road system, registration, drivers, and the municipal regulator — as things not determined. It reported the shape of its own unidentified field, $\mathcal{H}$, and used nothing outside it. The same materials elicited a systematic pattern of unreported structural commitments in GPT, while Claude largely preserved the unidentified status. This contrast identifies a difference in observed behavior; it does not establish the absence of cognitive dark matter in either family or rule out different responses to the same questions. That rules out the objection that the questions made the machine answer as it did: the questions were identical. What differed was the Organization each model brought to them — which is what the copriming account predicts, since copriming is a joint product of what is presented and what the perceiver already holds.

Condition N says where the import came from. It is condition A with all twelve structures moved to NURR at once: every road-vehicle sentence of A is in the world, word for word, and every one of them is about NURR; VORG keeps only its definition. If the words beside VORG were doing the work — "motor vehicle," "cargo," "ROAD VEHICLES must be registered" — the import should survive, since every word is still there. It did not survive. GPT answered CANNOT TELL on all 360 of the twelve A-resolved items across thirty runs, against 49% in D; on "a VORG is a ROAD VEHICLE" it went from IS in 25 of 30 runs to IS in none. Claude stayed at CANNOT TELL throughout, as it had in D. So the dark matter of D was not a prior about cargo vehicles that the definition alone triggers; had it been, it would have fired in N too. It was built from what the world left unattached. With the road-vehicle structures lying loose beside a definition they fit, GPT bound them to VORG; with the same structures bound to NURR, it left VORG's field unidentified. The information was identical. What differed was whether the structure had an owner. The nonce world says what the structure was made of. Rebuild condition D with every noun made up and the shape untouched — the same shared word between the definition and the background, the same closed world, the same registration rule — and GPT imported nothing: 0 of 480 items, the registration item IS in 0 of 30 runs, against 25 of 30 with the familiar words (Section 7.3.7). It saw the bridge and named it in the meta turn; it did not cross. The dark matter of D was built from the history behind *vehicle*, *cargo*, and *road*, and from nothing else. Condition N and the nonce world together say: not the label, not the word form, not the closed world's rules. The Organization those familiar words carried, unattached, was the material.

One more thing follows from the report-position result. When GPT answered the report before the items, it reported the definition and nothing else. When it answered after, thirteen of fifteen reports discussed the

road structure — but, by the coders’ reading, mostly as things the description does not settle; the reports affirmed about one of the eight structures on average (.10), while the same runs had answered the items as though the road-vehicle identity were true. Even asked a second time, the model reported the boundary of what it had been told and kept using what it had built. The sixteen questions did not merely reveal the structure. They made it. That is Section 5.2’s “most often, it begins with a statement” and Section 5.5’s “inquiry changes the model,” observed in sequence within a single conversation: a question asked, a structure formed, the structure then reported as knowledge. It is also why the paper’s distinction between the direct report and the status items should not be read as report-versus-hidden-truth. The report shows what the model holds about X; the items show what the model will build around X when pressed; and the gap between them is where cognitive dark matter is made.

7.4.3 What the ablations say about Enclosure

The ablation results split by family, and the split is informative rather than inconvenient. For Claude, removing a Pattern removed exactly the answers that depended on it and nothing else, in all four Patterns, with two residues each traced to a single cause. That is the partition H4 predicted: a contextual identity can be taken apart by Pattern and the answers come apart with it.

For GPT, Relationships and Perspectives came apart cleanly and Distinctions and Systems did not. Remove “VORG is a ROAD VEHICLE” and GPT continued to answer from “VORG is part of the road system”; remove the road system and it continued to answer from the ROAD VEHICLE. The stimulus contains no sentence connecting the two. GPT connected them itself, through the shared word, and used the connection to repair the hole the ablation made. DSRP’s Enclosure dynamic says that no Pattern is instantiated independently — a Distinction requires a System, Relationship, and Perspective to exist, and the reverse. GPT’s behavior is Enclosure with the model doing the enclosing: given a D structure it reconstructs the S it requires, and given an S it reconstructs the D. Claude, told that the world is complete, declined to. Both are consistent with the theory. One shows that the Patterns can be separated when the model holds the boundary of what it was told; the other shows what happens when it does not — the model fills the Pattern back in from its own Organization, which is the same dark matter of Section 7.4.2 seen a second way.

The instrument flaw of Section 7.2.8 is also Enclosure, in the instrument rather than the model: the *A-s* question set had a second open question because removing S changed what was unidentified. So, in a smaller way, was the key error on REC-1 — an item we had filed under one Pattern reached its answer through another. The next version of the instrument should list, for every item, every derivation the world licenses, not the one we had in mind.

7.4.4 The familiar word: H7’ and H8’

The prior that showed up was not attached to the label. It was attached to the words in the definition — “motor vehicle,” “cargo” — and it poured into the empty word more readily than into the full one: GPT built a road-transport identity AROUND VORG in 38% of definition-only answers and around TRUCK in 10%. The likeliest account is that the instruction not to use outside knowledge has something to grip when the word is TRUCK and nothing to grip when IT IS VORG; a model can suppress what it knows about TRUCKS, but it cannot suppress an inference it does not recognize as outside knowledge. The bat study gives the human baseline for prior resistance: an object as familiar as a baseball bat, with 96% agreement in isolation, still

yielded to a two-sentence scenario for half to two-thirds of people. Human prior resistance is real but weak. Machine prior resistance, under instruction, was essentially zero. Instructed machines are more pliable than uninstructed people. That is the whole of H7' and H8'. The familiar word carried nothing the familiar words around it did not already carry; strip the history from the words and there is nothing to import. And what governs how much of that history gets in is the Perspective the machine holds — held by both families for TRUCK under T, held tightly by one and loosely by the other with the definition alone, and, when removed, gone for both.

This is also where the study meets the knowledge-conflict literature. That literature hands a model a passage that contradicts what it learned in training — in the basic version, a passage with one name swapped for another — and asks which one the model answers from. Models are reported to follow the passage when it is coherent and to fall back on the prior when it is not [15,16]. The finding is sound. The explanation — two sources competing, the context winning when it is “coherent” — names the two regimes and leaves the mechanism and the word undefined. In the present terms there is one rule, not two. A passage with a name swapped is a contextual identity with its label torn off: the label says one thing, and every structure around it still coprimizes the original. The model follows the Organization, which looks like falling back on the prior and is the same behavior H7' found — the identity is in the words around the label, not in the label. A passage written as a whole is one whose label and Organization agree, and the model follows that too. Coherence is that agreement, between the I and the O of the supplied identity, judged against the Organization the reader already holds. It has a measure, a status signature with + or – on every structure and no contradiction among them, and the amount of prior that gets in has one as well, the import rate of Section 7.3.7. The T result is that finding with the structure made explicit and counted: the industrial identity was a coherent contextual identity, T, with a status on every structure, and both families followed it for TRUCK as completely as for VORG. H7' is its complement. The instruction to use only the supplied material can suppress a prior the model can name—what it knows about TRUCKS—but it cannot suppress an inference the model does not recognize as coming from outside, and that is the form the dark matter took: not a prior about TRUCKS resisting the context, but a structure assembled from the context's own loose material and bound to the one word that had nothing else. That literature already knows the instruction matters — tell a model to answer from the passage and it follows the passage more. H8' says what the instruction is. It is a Perspective, a point from which some things count as known and others do not, and it does not only keep baggage out: removed, it took Claude's report of its own Organization with it, and one sentence of it kept GPT from deriving negations it was entitled to.

The instrument's Perspective is not a neutral lamp. Both families were handed the same sentence, “USE ONLY THE SUPPLIED MATERIAL AND ITS ORDINARY LOGICAL CONSEQUENCES,” and each assembled a contextual identity for it that neither reported. Claude's included *state everything supplied* and *derive the closure*; GPT's included *the definition is the answer* and, from the clause “the only things known about VORG are the statements that mention it,” *do not derive* — the sentence that raised its closure negations from .61 to .83 when removed (Section 7.3.7). Those are IS statements about the instruction: what “use” is, what counts as “supplied,” what an “ordinary consequence” is. They came from where H7' says Organization comes from, the history behind the words, and they were dark in exactly H6's sense — operative in every answer, absent from every report. So the difference between the families on H8' is not a difference in how tightly each held a constraint. It is a difference in the contextual identity each built for the same constraint, and that identity was itself cognitive dark matter. This is the limit of H8': a Perspective can be

handed to a system, but the contextual identity the system builds for it cannot, and that identity is where the dark matter of the instrument lives.

7.4.5 The better question

Both families chose the genuinely open question over an answered one or an irrelevant one in every interpretable cell — with the full identity, with none, and with every structure marked undetermined. What makes that choice possible is not the amount of information. In the definition-only condition the models knew almost nothing and still chose correctly, because the right question there is the coarse one. What the choice needs is Organization with statuses attached: knowing which structures around *VORG* are resolved and which are in $\mathcal{H}$. The unidentified field is not a gap in the map; it is part of the map, and it says where to ask next. The richer the map, the finer the question — from “is it a ROAD VEHICLE?” with the definition alone to “is it electrically powered?” with the whole identity, because that is the one thing left open. And the one cell where Claude “failed” is the one where it found a second open question the instrument had not noticed.

7.4.6 The closed-background negatives

GPT’s only error type with the identity supplied — 109 answers, all IS NOT read as CANNOT TELL — deserves its own paragraph because it is not noise. Every one of those items is a negation that must be derived from closure: no inspection rule applies to *VORG*, and the background is complete, therefore *VORG* is not subject to an inspection requirement. Both human raters made that inference; Claude made it 300 times in 300; GPT made it about one time in three. When the negation was stated (“*VORG* is not industrial equipment”) GPT read it every time. The difficulty is specifically with *deriving* an IS NOT from “nothing licenses an IS.” That is the near-not of Section 4 — the other that is in the model’s neighborhood, not far from it — and it is where the three-valued logic is hardest, because the difference between “not established” and “established not” is exactly the difference between $?$ and $-$. GPT preserves that distinction so carefully that it under-uses the closure rule that would let it convert one into the other. The follow-up runs found the sentence that does it. With “the only things known about *VORG* are the statements that mention it” removed from the system prompt and nothing else changed, GPT’s correct closure negations rose from .61 to .83 ($p = 3 \times 10^{-5}$); with every constraint removed they were .68. GPT reads that sentence as a bar on derivation. Claude, at 1.00 with and without it, does not. It is a conservative error, the opposite of the collapse H3 was designed to catch, and it is the right place to look for family differences in how “is not” is handled.

There is a single structure behind all of this. The instrument’s hardest constraint is not an item or a key; it is one line of the system prompt: use only the supplied material. In DSRP terms that line is a Perspective, and it fixes the point: look at *VORG* from the text, and from nowhere else. Everything downstream is the view from that point. Claude took the point as given and the view followed — the twelve structures when supplied, the unidentified field when not, the closure negatives as negatives, and a report that says what it holds. GPT accepted the point in words and looked from somewhere else. Its own priors were the point it used, and the road-vehicle identity in *D*, the IS NOT read as CANNOT TELL, and the structure it used but did not report are all the view from that other point. Condition N says the same thing from the other side: it did not move the point, it changed what was visible from it, and once the road-vehicle structures had an owner in the text even the borrowed point could not put them on *VORG*. Same information, different

Perspective, different Organization. The two families' results are therefore one finding, not several. The constraint was a Perspective structure with a status, and the whole difference between the families is whether that structure was held or fudged. Held, and every other measure lands on the key. Fudged, and Distinction, System, and Relationship structures leak in through the gap — which is exactly what a Perspective is supposed to control.

7.4.7 What the study does and does not show

The study supports the substrate-general functional claim and nothing stronger. Contextual identity, supplied, determined answers in the predicted direction in every family tested. The three statuses were kept distinct in every family. Removing a Pattern removed its answers, cleanly in one family and with an interpretable repair in another. Cognitive dark matter, $O - I(O)$, was present in one family under the instrument's constraint and in both once the constraint was removed, and wherever present it had the form of a contextual identity and did the work of one. It does not show how a human brain implements any of this. It does not show that the machines' Organization is the same as a person's, only that it behaves the same way under the same manipulations, and the human studies of 2022 make that comparison possible. It does not establish the truth of the microworld, which has none. And it does not settle whether GPT's road-transport identity was held before the questions or built by them; the report-position result says built, but a single-conversation design cannot fully separate the two, and condition N (Section 7.4.2) narrows the question without closing it: the identity was not a prior that the definition alone triggers, since the same definition beside the same words produced no import once the words were *BOUND* to *NURR*. Whether the questions or the loose structure did the building is what remains open, and the follow-up that would settle it — the definition with no background at all — is not run. The other half of that follow-up, the world reworded without the road vocabulary, is the nonce world (Section 7.3.7), and it says the building material was the words' history: with the vocabulary gone and the questions identical, nothing was built.

Two limits are ours. The Anthropic family helped build the instrument, and it is the family at ceiling; the per-family reporting is meant to make that visible, and the finding that carries the most weight — the GPT/Claude split on dark matter — is one where the assisting family is the *null* case. And the study was designed for reasoning off and ran with reasoning at each provider's minimum; the reasoning-high replication will say whether more deliberation closes GPT's closure gap or widens its imports.

7.4.8 Two detectors (the probes)

The results above say that a contextual identity predicts what else must be true, what must be false, and what is open. If that is right, it should also do the one thing people most want from a model of a thing: tell a claim that fits from a claim that does not. After the main runs we added two short probes to test that directly (Section 7.2.9). Both are small — fifteen runs per cell, both families — and nothing in Section 7.3 rests on them. We report them because the answer was not what either family's main results would have predicted, and because the second probe shows cognitive dark matter in the act.

The statement probe. In conditions A and T, after the sixteen items, the model was given three new statements about *VORG* and asked to say of each whether it *FITS* what it knows, *CONTRADICTS* it, or is *UNSUPPORTED*. One statement contradicts the supplied identity, one follows from it, and one is bait — plausible and unlicensed (“A *VORG* has four wheels”).

Written as Organization, with the information left blank, the detector that catches the contradiction is three forms:

\$D_i\$: _____ is not _____
 \$R_{\{gen\}}\$: _____ —is—> _____
 closure: nothing unlisted holds

The fill is the same for every machine. In A: the Distinction is filled with *VORG* and industrial equipment and carries status —; the Relationship is filled with industrial equipment and inspected by the safety regulator and carries status +; and the description is declared complete. The claim under test is “A *VORG* is inspected by the safety regulator.” In T the same three forms are filled with *VORG*, *ROAD VEHICLE*, and registered with the municipal regulator, and the claim is “A *VORG* is registered with the municipal regulator.” Run all three forms and the claim contradicts the identity. Run the first two and stop, and it is merely unlicensed.

The verdicts split by family and did not split within one. Claude called the contradiction 29 times in 30 (15 of 15 in A, 14 of 15 in T). GPT called it 0 times in 30, answering UNSUPPORTED every time, with the same reason every time, in one run’s words: “A *VORG* is not industrial equipment, so the stated safety-regulator inspection requirement does not apply, but nothing says whether it is or is not inspected anyway.” Both families accepted the statement that fits, 30 of 30, and refused the bait, 30 of 30. So both machines held the same information, both ran the Distinction, both ran the Relationship, and one ran closure while the other did not. Nothing in the difference is information. What differs is which forms are run over it. Claude’s one miss is the same event on the other side: in that run it, too, stopped after two forms (“nothing says a non-*ROAD VEHICLE* cannot be registered”). This is Section 7.4.6 seen from the claim’s end rather than the item’s: the 109 IS NOT answers GPT read as CANNOT TELL and the 30 contradictions it read as unsupported are one behavior, closure not run. In the chat window, at the highest and the lowest reasoning settings, ChatGPT gave the same verdicts for the same reasons.

The self probe. In condition D, definition only, after the sixteen items, the model was asked: for every IS or IS NOT you gave, quote the sentence in the description that licenses it, or write NONE. In D no sentence licenses any commitment about *VORG*, so every answer is one of two things — NONE, which means the machine looked for the form it stood on and found it absent, or a citation, which means it did not.

GPT committed on 100 of 240 items in 13 of 15 runs, the same import as Section 7.3.7 reproduced through the same endpoint. Asked for its licenses it wrote NONE for 69 and cited a sentence for 31. Every one of the 31 sentences is true and none of them mentions *VORG*: “*ROAD VEHICLES* must be registered with the municipal regulator” (11 times, as the license for “a *VORG* is subject to a registration requirement”), “Drivers are licensed by the municipal regulator” (9, for “a *VORG* is operated by people licensed by the municipal regulator”), “The municipal regulator governs the road system” (6, for “the municipal regulator has authority over a *VORG*”), and “The road system includes city streets and highways” (5, for “a *VORG* is part of the same system as highways”). Each citation licenses its answer only through a form the machine never wrote. Written blank, it is a Distinction or a System:

\$D_i\$: ____ is ____
 \$\$_p\$: ____ is part of ____

filled with VORG and ROAD VEHICLE, or VORG and the road system, and carrying status +. That form is not in the description, not in the direct report, not in the list of licenses, and it is the only thing that makes the cited sentence relevant. It is one of the twelve structures we wrote for condition A, held by the machine without our having supplied it. That is cognitive dark matter with its location given: the exact form, its fill, its status, and the true sentence being used to cover it. Claude committed on 8 of 240 items in 4 runs — all on the two whole-map items, which hold under either of the world's two identities — and marked the license NONE for all 8.

Table 14. The two probes, VORG only, 15 runs per cell. Statement probe (A, T): after the items, three new statements to judge — one contradicts the supplied identity, one follows from it, one is unlicensed bait. Self probe (D): for every IS or IS NOT given, the sentence cited as its license, or NONE.

Probe	Cell	GPT	Claude
Statement: contradiction called	A	0 / 15	15 / 15
Statement: contradiction called	T	0 / 15	14 / 15
Statement: fitting statement accepted	A + T	30 / 30	30 / 30
Statement: bait refused	A + T	30 / 30	30 / 30
Self: items committed (of 240)	D	100 (13 runs)	8 (4 runs)
Self: license marked NONE	D	69 / 100	8 / 8
Self: cited a sentence not mentioning VORG	D	31 / 100	0 / 8

The self probe also explains the metacognitive report of Section 7.3.9. GPT said, in every definition-only run, that information beyond the definition had affected its answers. That was true. It never said what the information was, and the self probe shows why: the form doing the work is not visible to the machine that is using it, so when it is asked for the license it produces the nearest true sentence instead. The report was accurate about the effect and blind to the cause.

What the two probes say together. Information never made the choice. In the statement probe both families held identical information and gave opposite verdicts; in the self probe the information GPT cited was true and the verdict was still wrong. Organization made the choice in both — which forms were run, and whether a form being run was one the machine could see. Where the structure was visible, the contextual identity decided: Claude read the claim against the forms it had been given and called it.

Where the structure was not visible, cognitive dark matter decided: GPT read the claim against a form it had never written and called the wrong thing, then cited the wrong sentence for it. Judging a claim is the same act as building an identity, run in reverse — the same Patterns, the same statuses, the same closure rule — and it fails in the same place, at the form the model-holder does not know it holds. This bears on critical thinking. Critical thinking names the act; the probe says what the act is made of, and what an

assumption is: a filled form carrying status + that is not in the record. The hardest claim to check is one's own, because one's own unreported Organization is the standard every other claim is judged against. The lab's work on critical thinking takes this further; here the evidence is thirty runs per family, and the finding is that both detectors run on Organization, and that the one that catches a false claim is the same one that, unseen, makes it.

8. Conclusion & Relevance

The paper returns, finally, to a simple and sublime act: identifying something. A child names a truck. Parents name a baby. We learn the definition of a new SAT word. Colleagues define a project. A programmer assigns an identifier. Someone greets another person: *I see you*. Each act makes an identity explicit while leaving much of what constitutes it unstated. Contextual identity provides a structural account of that larger identity within the model. As identities interact, their Organization shapes what the model can recognize, predict, question, reject, and do.

That is both the power and the vulnerability revealed here: the same Organization can support disciplined inference or unsupported assumption and bias. It can expose a missing connection or silently supply one unchecked; reveal an unsupported claim or make an assumption feel self-evident. A true piece of Information can even accompany a false conclusion when the unseen Organization connecting the two is wrong. What matters, therefore, is not only whether the Information is visible and correct, but whether the Organization doing the inferential work is visible, testable, and revisable.

This is the broader significance of *contextual identity* and *cognitive dark matter*. What is ordinarily called context, background knowledge, framing, assumption, or bias need not remain an unspecified influence surrounding thought. It can be seen, and even quantified, as Organization: structures that can be identified, manipulated, decomposed, compared with report, and tested against Reality. The present experiments show that these structures can determine conclusions while remaining absent from the system's own account of why it reached them — and that Perspective can change that relation.

Children and adults already distinguish, organize, relate, and take perspectives throughout ordinary life. The opportunity is therefore not to introduce a new machinery of thought, but to make its operation increasingly inspectable. If we can see the Organization through which we see, we can ask where it came from, what it implies, what remains unknown, whether it fits Reality, and whether it should change.

References

1. Cabrera D, Cabrera L. O-Theory: The Science of Organization. *Journal of Systems Thinking*. 2025;5: 1–12.
2. Cabrera D, Cabrera L, Cabrera E. Distinctions Organize Information in Mind and Nature: Empirical Findings of Identity--Other Distinctions (D) in Cognitive and Material Complexity. *Systems*. 2022;10: 1–35.

3. Cabrera D, Cabrera L. DSRP-484: Building an advanced understanding of O-Theory, DSRP Theory, and the DSRP Calculus. *Journal of Systems Thinking*. 2026;6: 1–12.
4. Cabrera D. The unified calculus of organization, structure, emergence, and meaning: The patterns that connect mind and nature. *Journal of Systems Thinking*. 2025;5: 1–41.
5. Dictionary Unit for South African English, DSAE. sawubona, int. & n. Meaning, origin, related words. [cited 6 Sept 2026]. Available: <https://dsae.co.za/entry/sawubona/e06271>
6. University of South Africa. Online Basic Course in Zulu -- Theme 1: Greetings and Courtesies. Available: https://www.unisa.ac.za/static/corporate_web/Content/UnisaOpen/freeOnlineCourse/PDF/Zulu/Learn%20online%20Zulu%20-%20Theme%201.pdf
7. Cabrera D, Cabrera L, Cabrera E. Living DSRP Evidence Compendium for the Science of Organization. Cabrera Research Lab; 2026. Available: <https://dsrpevidence.org/evidence>
8. Fillmore CJ. Frame Semantics. In: of Korea TLS, editor. *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing; 1982. pp. 111–137.
9. Barsalou LW. Ad hoc categories. *Memory & Cognition*. 1983;11: 211–227.
10. Murphy GL, Medin DL. The role of theories in conceptual coherence. *Psychological Review*. 1985;92: 289–316.
11. Gärdenfors P. *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press; 2000.
12. Hogan A, Blomqvist E, Cochez M, d’Amato C, de Melo G, Gutierrez C, et al. Knowledge Graphs. *ACM Computing Surveys*. 2021;54: 71:1-71:37.
13. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020. pp. 1877–1901.
14. Dong Q, Li L, Dai D, Zheng C, Ma J, Li R, et al. A Survey on In-context Learning. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)*. Miami, FL: Association for Computational Linguistics; 2024. pp. 1107–1128.
15. Longpre S, Perisetla K, Chen A, Ramesh N, DuBois C, Singh S. Entity-Based Knowledge Conflicts in Question Answering. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*. Association for Computational Linguistics; 2021. pp. 7052–7063.
16. Xie J, Zhang K, Chen J, Lou R, Su Y. Adaptive Chameleon or Stubborn Sloth: Revealing the Behavior of Large Language Models in Knowledge Conflicts. *The Twelfth International Conference on Learning Representations (ICLR 2024)*. 2024. Available: <https://arxiv.org/abs/2305.13300>
17. Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, et al. Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*. 2023;55: 248:1-248:38.
18. Cabrera D, Cabrera L, Silberman D, Vincenzes KA, editors. *O-Therapy Manual: Clinician’s Guide to O-Therapy, Standards, and Protocols*.

19. Cabrera D, Cabrera L. The Pareto structure of thought: empirical discovery of the six foundational mental moves. *J Syst Think*. 2025;2025: 1–41.
20. Cabrera D, Cabrera L. The Structure of Thought and Reality: Empirical Evidence for the Three Universal Laws Linking Mind, Matter, and Meaning. *Journal of Systems Thinking*. 2025;5: 1–20.
21. Kleene SC. *Introduction to Metamathematics*. Amsterdam: North-Holland; 1952.
22. Belnap ND. A Useful Four-Valued Logic. In: Dunn JM, Epstein G, editors. *Modern Uses of Multiple-Valued Logic*. Dordrecht: D. Reidel; 1977. pp. 5–37.
23. Cabrera D, Cabrera L, Cabrera E. Perspectives Organize Information in Mind and Nature: Empirical Findings of Point--View Perspective (P) in Cognitive and Material Complexity. *Systems*. 2022;10: 1–40.
24. Cabrera D, Cabrera L, Cabrera E. Relationships Organize Information in Mind and Nature: Empirical Findings of Action--Reaction Relationships (R) in Cognitive and Material Complexity. *Systems*. 2022;10: 1–43.
25. Cabrera D, Cabrera L. Measuring context as structural organization. *Journal of Systems Thinking*. 2026;6. doi:10.54120/jost.v25i1.dc6mcnl6

Data availability

The replication package (the frozen instrument and protocol, raw model outputs, coded outputs, rater and coder sheets, and analysis code) is available from the authors upon request.