

Borrowed Intuition

Compression Without Experience

Flyxion
Independent Researcher

2026

Abstract

Competence is often treated as evidence of the experience required to produce it. An expert recognizes a pattern quickly because earlier encounters have compressed a difficult search into an apparently immediate judgment. Yet the person exercising such judgment need not be the person whose exploration produced the compression. Rules of thumb, diagnostic categories, precedents, checklists, and trained models can transmit the residue of earlier search without transmitting the search itself. This paper calls the resulting capacity *borrowed intuition*. It distinguishes the provenance of a competence from the conditions under which it is acquired, separating compression generated through an agent's own trajectory from compression received from other agents, records, or institutions. The distinction is one of degree rather than kind: inherited competence may be tested, revised, and partially re-earned through subsequent experience. Testimony supplies propositions, but borrowed intuition can also transmit policies of attention, exclusion, and action. Institutions distribute such policies across people and records, while pretrained computational systems provide a limiting case in which extensive compression is available to a receiver that did not conduct the originating exploration. The relevant question is therefore not whether inherited competence is genuine, but whose trajectory produced it, how it was subsequently validated, and how much trust its provenance can support.

1 Two Kinds of Ease

An apprentice encounters a fault that has resisted several ordinary repairs. The visible symptoms suggest a damaged component, but the master advises checking the connection immediately behind it. The apprentice follows the advice and finds the defect. Months later, a similar failure occurs. This time the apprentice checks the connection without deliberation. What had first arrived as an instruction now appears as intuition.

The ease of the later judgment admits at least two explanations. The apprentice may have learned from repeated encounters that faults attributed to the component often originate in its connection. In that case, the judgment compresses a trajectory of observation, unsuccessful intervention, comparison, and eventual stabilization. Alternatively, the apprentice may simply have retained the master's rule and learned when to apply it. The same outward competence

can therefore arise from different histories. In one case, the agent has compressed its own search. In the other, the agent has inherited the result of someone else's.

The inherited judgment is not thereby false, counterfeit, or merely verbal. It may guide attention correctly, reduce unnecessary search, and produce reliable action. The difference concerns provenance rather than reality. The apprentice possesses the competence operationally, but did not initially generate or validate it through an independent trajectory. The competence is theirs to exercise before it is fully theirs to explain.

This separation between competence and its formative history is easy to miss because successful action conceals the search it replaces. Once a rule has become available, the failures, comparisons, and discarded hypotheses that produced it no longer need to be repeated. Compression is useful precisely because it permits one trajectory to spare another. The receiver obtains an economical disposition without incurring the full cost of its discovery.

Such transmission is ordinary. A clinician uses a diagnostic category developed from cases they never observed. An engineer applies a safety factor whose originating failures are known only through professional convention. A judge relies on precedent without reproducing the conflicts and deliberations that made the precedent authoritative. A pilot follows a checklist whose terse instructions condense accidents, investigations, and redesigns distributed across decades. In each case, present competence contains the residue of absent experience.

This paper calls such competence *borrowed intuition*. The term does not imply that the competence is temporary or that its receiver lacks a legitimate claim to it. Borrowing names a relation between a compressed disposition and the trajectory that produced it. An intuition is borrowed to the extent that its practical economy was generated elsewhere and transmitted without the receiver reproducing the originating exploration.

The distinction matters because fluency is commonly treated as evidence about history. Rapid recognition is taken to show extensive experience; confident action is taken to show understanding; reliable performance is taken to show that the performer has encountered and resolved the relevant class of problems. These inferences are often reasonable, but they are not necessary. Fluency can survive the separation of compression from experience.

The resulting question is narrower than the traditional question of whether another person, institution, or computational system genuinely understands. It asks instead whose trajectory produced a displayed competence, through what mechanism it reached its present holder, and what happened to it after transmission. These are questions about provenance, formation, and validation. Unlike a general verdict about understanding, they admit answers in degrees.

The distinction also reaches beyond the transmission of facts. Testimony can communicate that a proposition is true, but professional instruction frequently communicates how to search, what to ignore, when to become suspicious, and where inquiry may safely stop. A rule of thumb does not merely enlarge what an agent knows. It changes the order in which possibilities are considered and the threshold at which deliberation is relegated to routine action. What is transmitted is therefore not only content but a partial organization of cognition.

Artificial intelligence makes this structure unusually visible, but it does not originate it. A pre-trained model can display extensive compressed competence without having performed the

human investigations, arguments, experiments, and acts of composition from which its training corpus was assembled. It is an extreme case of a familiar arrangement, not an exception to human epistemology. Beginning with the machine would therefore obscure the prior fact that human competence has always been partly inherited.

The argument proceeds from this ordinary human condition. It first restates a formal account of compression and relegation, then separates the provenance of compression from the conditions under which an agent acquires it. It subsequently considers testimony, professional intuition, and institutional memory before turning to pretrained and adaptive computational systems. The aim is not to distinguish authentic from inauthentic knowledge. It is to describe how competence can travel farther than the experience that made it possible.

2 Compression, Expansion, and Relegation

Let an agent confront an environment $\mathcal{E}$ through a sequence of states, observations, and actions. At time t , let $K(t)$ denote the distinctions presently available to the agent: the hypotheses it can entertain, the features it can discriminate, and the relations it can use to organize a problem. Let $A(t) \subseteq K(t)$ denote the distinctions currently activated by the situation. Deliberation occurs when the agent must search within, compare, or enlarge this activated set before selecting an action.

Repeated inquiry can reduce the amount of active search required. Distinctions that once had to be represented explicitly become incorporated into a more economical rule, disposition, or policy. Let R_t denote the set of such relegated structures available at time t . Elements of R_t need not be consciously represented whenever they operate. They may appear phenomenologically as recognition, habit, familiarity, or intuition.

The trajectory

$$K(t) \longrightarrow A(t) \longrightarrow R_t$$

describes a transition from possible distinctions, through situated activation, to stabilized compression. It does not imply that all knowledge proceeds linearly or that relegation is irreversible. A relegated structure can return to active deliberation when it fails, conflicts with another structure, or encounters an environment sufficiently different from the one in which it stabilized.

To state the compression relation more precisely, let X be a space of possible situations and let Y be a space of responses. A deliberative procedure

$$F : X \rightarrow Y$$

may require an extended search over distinctions in $K(t)$. A compressed policy

$$C_F : X \rightarrow Y$$

approximates the relevant outputs of F while requiring fewer explicit intermediate distinctions. The compression is successful relative to an environment $\mathcal{E}$ and a tolerance ε when

$$\Pr_{x \sim \mathcal{E}} [d(C_F(x), F(x)) \leq \varepsilon] \geq 1 - \delta,$$

where d measures the practical difference between responses and δ is the tolerated rate of divergence.

Compression is therefore indexed to an environment, an objective, and a threshold of acceptable loss. A rule may be highly compressed and reliable within the conditions that produced it while failing abruptly elsewhere. Its ease of application does not reveal the range over which it remains valid.

Definition 2.1 (Experiential trajectory). *An experiential trajectory for an agent i is a temporally ordered sequence*

$$\tau_i = (K_i(t), A_i(t), a_i(t), o_i(t), u_i(t))_{t=0}^T,$$

where $a_i(t)$ is an action, $o_i(t)$ is a resulting observation, and $u_i(t)$ is evaluative feedback relative to the agent's objective. The trajectory records the distinctions activated, interventions attempted, consequences observed, and evaluations available during the formation of a competence.

Definition 2.2 (Earned compression). *A compressed policy C is earned by agent i to the extent that its selection and stabilization depend upon evidence generated within τ_i .*

The phrase “to the extent” is essential. A competence rarely begins from nothing. The agent's initial distinctions, language, instruments, and available actions are already historically inherited. Earned compression does not require epistemic self-creation. It identifies the portion of a competence whose practical selection or validation depends upon the receiver's own subsequent trajectory.

The complementary case occurs when a compressed policy reaches an agent without the agent reproducing the exploration that selected it.

Definition 2.3 (Transmitted compression). *A compressed policy C is transmitted from a source j to a receiver i when C depends materially upon a trajectory τ_j or a collection of external trajectories $\{\tau_j\}$, while the corresponding distinctions and comparisons are absent from τ_i at the time of acquisition.*

Transmission can occur through direct instruction, imitation, written records, formal education, technical standards, institutional procedures, or trained computational parameters. The source need not be a single identifiable agent. A policy may condense the trajectories of many people, and the record through which it travels may no longer preserve which observations selected which part of the policy.

The receiver nevertheless gains access to the result. If the transmitted policy maps situations to useful responses, the receiver can act as though an expensive expansion had already occurred. In an important sense, it has occurred, but elsewhere.

Proposition 2.4 (Separation of expansion from use). *The agent that incurs the cost of expanding a problem space need not be the agent that later benefits from the resulting compression.*

Proof. Let agent j conduct an exploration over a set of candidate distinctions K_j and select a policy C after observing its performance across τ_j . Suppose C is then communicated to agent i in a form that permits its application to situations in $\mathcal{E}$. Agent i can obtain the action advantage of C without reproducing the comparisons through which j selected it. The computational or experiential cost of expansion is therefore borne by j , while at least part of the practical benefit is realized by i . $\square$

This separation is the structural basis of borrowed intuition. It explains why the phenomenology of ease is insufficient to establish the provenance of competence. Both earned and transmitted compression reduce present deliberation. Both can operate rapidly. Both can be reliable. Their difference lies not in the appearance of the resulting action but in the relation between the present policy and the history that made it available.

Let $P(C)$ denote the provenance of a compressed policy. At minimum, provenance must record the trajectories that generated evidence for the policy, the mechanism by which the policy reached its present holder, and any subsequent trajectory through which the holder tested or revised it. We may write

$$P_i(C) = (G(C), M_{G \rightarrow i}(C), V_i(C)),$$

where $G(C)$ denotes the generating trajectories, $M_{G \rightarrow i}(C)$ the mechanism of transmission to agent i , and $V_i(C)$ the receiver's later validation history.

This decomposition prevents three different questions from being collapsed. The first is who or what generated the evidence from which the compression arose. The second is how the receiver acquired it. The third is what the receiver subsequently did with it. A competence may be externally generated, pedagogically conditioned, and later extensively validated by its receiver. Another may be externally generated, copied without explanation, and applied without any feedback capable of testing it. Both are transmitted, but they do not have the same epistemic standing.

The provenance of a competence must also be distinguished from its present ownership. An agent can possess, understand, and responsibly exercise a transmitted competence. What the agent cannot acquire retroactively is the originating trajectory itself. Later experience may supply new evidence, revise the inherited policy, or independently reproduce part of its justification, but it does not change where the initial economy came from.

This point suggests a useful distinction between expansion and reconstruction. To reconstruct a rule's justification is to recover enough of the problem space to understand why the rule was selected. To expand independently is to encounter and discriminate the relevant alternatives without relying wholly upon the inherited selection. Reconstruction may produce understanding, but it remains historically downstream from transmission. Independent expansion may converge upon the same policy, but does so through a distinct evidential trajectory.

Borrowed intuition therefore admits degrees. A received rule may initially be almost wholly transmitted. As the receiver applies it, observes its failures, compares it with alternatives, and modifies its scope, a greater portion of the operative competence becomes supported by the receiver's own trajectory. The borrowed policy is not simply replaced by an earned one. It is progressively transformed through use.

The next section develops this graded structure by separating two dimensions that are often conflated: where a compression was generated and who controlled the conditions under which it was acquired.

3 Provenance and Formation

A taxonomy of borrowed intuition must distinguish the origin of a compression from the conditions under which it is acquired. These dimensions are related, but they are not identical. A competence may originate outside an agent while being acquired through self-directed inquiry, or it may be generated through the agent's own activity under conditions designed almost entirely by another. Treating these cases as members of a single sequence obscures the difference between producing evidence and controlling exposure to evidence.

The first dimension is provenance. Compression is *earned* to the extent that its selection and stabilization depend upon the receiving agent's own trajectory. It is *transmitted* to the extent that the relevant selection has already occurred elsewhere and reaches the receiver in compressed form.

The second dimension is formation. A formation process is *self-directed* to the extent that the receiving agent controls which situations to investigate, which actions to attempt, which errors to pursue, and which objectives govern revision. It is *conditioned* to the extent that another agent or institution structures the receiver's exposures, feedback, objectives, or permissible actions.

These dimensions yield four idealized regions.

	Self-directed formation	Conditioned formation
Earned provenance	Compression selected through inquiry substantially directed by the receiving agent.	Compression selected through the receiver's experience under externally structured exposure or feedback.
Transmitted provenance	External compression sought, compared, or assembled by the receiver without reproducing its originating exploration.	External compression delivered through instruction, training, procedure, or engineered exposure.

The table identifies analytic limits rather than natural kinds. Ordinary learning moves among the regions. A student may receive a theorem through conditioned transmission, investigate alternative proofs through self-directed comparison, and eventually use the theorem within

a new line of inquiry that produces earned compression. A clinician may inherit a diagnostic rule, test it across years of practice, discover its limits, and modify its application. The resulting competence retains transmitted structure while acquiring support from the clinician's own trajectory.

To represent this mixture, let C_i be a competence exercised by agent i . Its provenance may be expressed schematically as

$$C_i = \alpha_i C_i^{\text{earned}} + \beta_i C_i^{\text{transmitted}}, \quad \alpha_i + \beta_i = 1,$$

where α_i and β_i do not measure proportions of stored content. They represent the relative dependence of the operative competence upon evidence generated within and outside the receiver's trajectory. The expression is diagnostic rather than directly numerical. In most practical cases, assigning exact values would require a detailed causal analysis of learning history.

Formation can be represented independently:

$$F_i = \gamma_i F_i^{\text{self}} + \eta_i F_i^{\text{conditioned}}, \quad \gamma_i + \eta_i = 1.$$

The independence of these expressions matters. An externally conditioned process can still produce earned compression if the receiver acts, observes consequences, and stabilizes a policy from those consequences. Conversely, self-directed study can remain largely transmitted if the agent chooses which authorities to consult but does not reproduce or test the inquiry underlying their conclusions.

Definition 3.1 (Borrowed intuition). *A competence C_i is borrowed to the extent that its present economy depends upon transmitted compression, so that*

$$\beta_i > 0,$$

and the receiver can exercise some practical advantage of C_i without reproducing the generating trajectories in $G(C_i)$.

On this definition, nearly all mature human competence contains some borrowed component. This is not an objection. The concept is not intended to divide agents into independent and dependent knowers. It identifies a variable ordinarily hidden by successful performance.

The dependence may concern more than conclusions. An agent can inherit a vocabulary that determines which differences are visible, a ranking that determines which hypotheses are considered first, or a stopping rule that determines when further inquiry appears unnecessary. Borrowed intuition can therefore enter at several stages of the trajectory:

$$K_i(t), \quad A_i(t), \quad \pi_i(t), \quad R_{i,t},$$

where $\pi_i(t)$ is the policy governing movement through the activated distinction space. Transmission may alter the distinctions available in $K_i(t)$, the subset activated in $A_i(t)$, the search

policy $\pi_i(t)$, or the relegated dispositions in $R_{i,t}$. It need not transmit a complete rule in order to shape the resulting competence.

This observation also clarifies the role of testimony. A statement that p enlarges the receiver's available propositional content. A diagnostic heuristic can do something structurally different. It can make certain features salient, demote other possibilities, and change the order in which hypotheses are tested. Its effect is not exhausted by adding p to $K_i(t)$. It changes the organization of search itself.

The transformation of inherited compression through later experience requires a further term.

Definition 3.2 (Synthetic compression). *Synthetic compression is the process by which an agent applies, tests, combines, revises, or re-stabilizes transmitted compression through evidence generated in its own subsequent trajectory.*

Synthetic compression is a process rather than a fifth provenance class. Its inputs are partly transmitted, while part of its evidential support is earned. It describes movement within the taxonomy rather than another cell beside its existing terms.

Let C_i^0 be a policy initially received by agent i , and let τ_i^+ denote the receiver's post-acquisition trajectory. A synthetic update may be represented as

$$C_i^{n+1} = S(C_i^n, \tau_i^+),$$

where S is a stabilization operator that retains, revises, narrows, combines, or rejects portions of the inherited policy. If τ_i^+ contains discriminating feedback, the receiver may progressively acquire independent warrant for some components of C_i^0 . If the trajectory supplies no relevant feedback, repeated use can increase familiarity without substantially changing provenance.

This distinction between repetition and validation is crucial. A procedure does not become earned merely because an agent has performed it many times. Repetition under invariant conditions may strengthen a disposition while providing little evidence about why it works, where it fails, or whether an alternative would perform better. For synthetic compression to occur in the epistemically relevant sense, later experience must be capable of discriminating among policies.

Let $\mathcal{H}$ be a set of candidate policies compatible with the inherited instruction. A post-acquisition trajectory τ_i^+ is discriminating when it changes the receiver's relative support over $\mathcal{H}$:

$$P(H \mid \tau_i^+) \neq P(H)$$

for at least some $H \in \mathcal{H}$. Experience that merely confirms what every candidate policy predicts may stabilize behavior without clarifying its basis.

The same distinction applies to adaptation. Not every response to new information constitutes compression. A system may alter its current action because a fact is present in its immediate context, then lose that alteration when the context disappears. This is temporary conditioning, not yet a stable addition to R_t . Retained experience requires that some effect of the encounter

persist. Stabilized compression requires the stronger condition that the retained effect reorganize future search or action economically across a class of situations.

These relations can be expressed as

temporary conditioning $\nRightarrow$ retained experience $\nRightarrow$ stabilized compression.

Each implication may hold in a particular system, but none holds by definition. A conversation can condition an immediate response without producing persistent memory. A persistent record can retain an experience without integrating it into a policy. A policy update becomes compression only when it reduces or reorganizes future deliberation while preserving relevant performance.

Proposition 3.3 (Provenance is not erased by synthesis). *If a transmitted policy C_i^0 undergoes synthetic compression through τ_i^+ , the resulting competence may acquire earned support without ceasing to depend historically upon transmitted structure.*

Proof. The update $S(C_i^0, \tau_i^+)$ depends jointly upon the inherited policy and the receiver's later evidence. Unless the later trajectory independently reconstructs every distinction and selection materially contributed by C_i^0 , the resulting policy remains causally dependent upon external compression. Synthetic development can alter the degree and significance of that dependence, but it does not make the originating contribution disappear. $\square$

Borrowed intuition is therefore neither a permanent deficiency nor a stage that every competence must leave behind. It is a provenance relation that later experience can transform without nullifying. A mature competence may be deeply owned, critically understood, and extensively validated while remaining partly borrowed. Indeed, this is likely the normal condition of expertise.

4 Human Intuition and the Testimony of Practice

Human intuition is often described autobiographically. An experienced practitioner sees what a novice overlooks because repeated encounters have altered the organization of attention. Possibilities that once required deliberate comparison are eliminated before they become explicit. The expert does not necessarily reason faster through the same search performed by the novice. The expert enters a smaller and better ordered search space.

This account is plausible, but incomplete. The expert's apparent immediacy does not arise entirely from personal encounters. Before acquiring experience, the practitioner inherits names, categories, exemplars, warnings, procedures, and standards of relevance. These inherited structures influence which encounters count as similar and which differences deserve attention. Experience subsequently compresses a space that transmission has already organized.

Let the active search required for a problem x be represented by

$$\Sigma_i(x) = (K_i(x), A_i(x), \pi_i(x), \theta_i(x)),$$

where $K_i(x)$ is the available distinction space, $A_i(x)$ is the activated subset, $\pi_i(x)$ is the search policy, and $\theta_i(x)$ is the stopping condition under which the agent treats the inquiry as sufficiently resolved. Instruction can modify every term in this structure. It can supply distinctions, direct attention, order hypotheses, and determine what counts as an adequate answer.

The apprentice who is told to inspect the connection behind an apparently defective component receives more than the proposition that connections sometimes fail. The instruction changes $\pi_i(x)$ by advancing one hypothesis ahead of others. It may also alter $A_i(x)$ by making a previously neglected relation salient. If the instruction becomes a rule of thumb, it can modify $\theta_i(x)$ by specifying when the search should stop. What has been transmitted is a small policy for conducting inquiry.

This form of transmission is adjacent to testimony but not reducible to its most familiar propositional form. The modern epistemology of testimony has emphasized the extent to which knowledge depends upon what others tell us. Coady argues against treating testimony as a marginal or merely derivative source, observing that ordinary and theoretical knowledge depend extensively upon the reports of others [2]. Borrowed intuition accepts this dependence and extends the object of transmission. Agents receive not only reports about the world but compressed ways of moving through it.

A useful distinction can therefore be drawn between propositional testimony and procedural testimony. Propositional testimony characteristically has the form

j tells i that p .

Procedural testimony has the more general form

j induces in i a policy π

for noticing, comparing, acting, or stopping. The policy may be communicated verbally, demonstrated through example, imposed by a checklist, or embedded in the design of an instrument. Its content may resist complete propositional expression even when its effects are observable in action.

Definition 4.1 (Procedural testimony). *Procedural testimony is the interpersonal or institutional transmission of a structure that modifies a receiver's inquiry policy without requiring the receiver to reproduce the trajectory by which that structure was selected.*

Procedural testimony need not be tacit. A checklist is explicit, but the reasons for its sequence may not be. A diagnostic manual states criteria, but the clinical comparisons that stabilized those criteria remain external to most users. A precedent can be read in full while still compressing a larger history of disputes, failed rules, institutional negotiations, and practical consequences. Explicitness of instruction does not imply completeness of provenance.

The distinction between proposition and policy also clarifies why transmitted competence can exceed the receiver’s ability to justify it. A person may reliably apply a rule while possessing only a partial account of its origin. The rule can improve action because it preserves a selection made by earlier inquiry. What it does not necessarily preserve is the evidential structure that made the selection rational.

Let $E(C)$ denote the evidence that originally selected a compressed policy C , and let $Q(C)$ denote the practical guidance retained in its transmitted form. In many cases,

$$Q(C) \subsetneq E(C)$$

is not the relevant relation, because guidance and evidence are not objects of the same type. A more accurate expression is

$$T(E(C)) = Q(C),$$

where T is a transmission operator that preserves enough structure for action while discarding much of the path by which the action was justified. The receiver obtains the output of an evidential transformation without obtaining an isomorphic copy of the evidence.

This loss is not necessarily a defect. If every user of a safety rule had to reconstruct the accidents and calculations that produced it, the rule would fail in its purpose. Transmission makes competence scalable by allowing evidence to be compressed into action guidance. The epistemic cost is that later users may be unable to distinguish parts of the rule that remain necessary from parts that merely reflect the environment of its origin.

Professional intuition therefore develops through an interaction between transmitted organization and personal correction. A novice does not begin with an undifferentiated world. Training supplies a partition of the field. Practice then reveals which inherited distinctions are robust, which are too coarse, and which fail under local conditions. What eventually appears as personal intuition is usually a layered object:

$$R_{i,t} = R_{i,t}^{\text{transmitted}} \cup R_{i,t}^{\text{conditioned}} \cup R_{i,t}^{\text{synthetic}},$$

where the terms identify causal contributions rather than cleanly separable stores. The practitioner may no longer be able to say which parts were taught, which were acquired through structured exercises, and which arose through later independent revision.

This opacity is especially pronounced when training succeeds. A well-assimilated distinction no longer arrives phenomenologically as another person’s claim. It becomes part of what the receiver finds obvious. Transmission can therefore disappear from introspection while remaining active in cognition.

Proposition 4.2 (Assimilation does not establish provenance). *If a transmitted policy becomes automatic for its receiver, the resulting phenomenological immediacy does not imply that the policy was selected through the receiver’s own trajectory.*

Proof. Automaticity describes the present cost of executing a policy. Provenance describes the history through which the policy was selected and stabilized. A transmission mechanism may lower execution cost through rehearsal, imitation, conditioning, or repeated application without reproducing the originating evidential search. Present immediacy therefore does not entail earned provenance. $\square$

The social evaluation of testimony further complicates this process. Fricker's account of testimonial injustice concerns cases in which prejudice causes a speaker to receive an unfair deficit of credibility, while hermeneutical injustice concerns unequal participation in the interpretive resources through which experience becomes intelligible [3]. Both forms of injustice affect not only which propositions circulate but which compressions become available for transmission.

If a speaker is denied credibility, the receiver may lose access to distinctions and heuristics produced by that speaker's experience. If a group lacks the concepts required to articulate recurrent features of its experience, its trajectories may fail to become transmissible compression at all. Epistemic exclusion can therefore interrupt the path

experience $\longrightarrow$ distinction $\longrightarrow$ compression $\longrightarrow$ transmission.

Conversely, social authority can allow poorly supported compression to circulate widely. A heuristic associated with an institutionally credible source may be inherited and rehearsed even when the originating trajectory was narrow or the environment has changed. Credibility affects the distribution of compression independently of its reliability.

Borrowed intuition thus introduces a provenance question into the ethics of testimony. It asks not only whether a speaker was believed, but whether the practical structures generated from one person's experience were credited, preserved, and transmitted under that person's authority. A community may adopt a distinction while forgetting its source. An institution may retain a rule while erasing the workers, patients, witnesses, or marginalized speakers whose reports forced its creation. The compression survives while authorship disappears.

The phrase "how much of the competence is genuinely the receiver's" would state the problem badly. Operational possession is not diminished by transmitted origin. A physician can genuinely possess a diagnostic competence derived partly from inherited categories. The more precise questions are how much of the competence was generated, selected, or validated through the physician's own trajectory, and how much depends upon external trajectories that remain visible, recoverable, or lost.

These questions matter because two observationally similar performances may rest upon different evidential relations. One practitioner may have tested a rule across varied cases and learned its failure conditions. Another may apply the same rule with equal fluency but only within familiar examples. Their present outputs can coincide while their capacities to respond to distributional change differ. Provenance does not determine reliability, but it can explain why reliability fails in different ways.

Human intuition is therefore neither wholly autobiographical nor merely received. It is a composite formed where inherited compression meets situated experience. Its apparent unity is a consequence of successful assimilation. Beneath that unity lie multiple trajectories, unequal

relations of authority, and different degrees of subsequent validation. Institutions make this composite structure durable by separating competence not only from its originating agent but sometimes from every living person who remembers why it was adopted.

5 Institutions as Compression-Distributing Systems

An institution can preserve the result of an inquiry after the participants in that inquiry have departed. Records, procedures, standards, precedents, forms, training programs, and technical artifacts allow distinctions selected in one period to constrain action in another. Institutional memory is therefore not merely stored information. It is an arrangement for transmitting compressed policies across discontinuities of personnel and experience.

Consider a checklist introduced after an accident. The investigation that produces it may involve testimony, physical reconstruction, competing causal models, failed hypotheses, and negotiation over acceptable risk. The final checklist retains little of this structure. It specifies a sequence of actions whose justification lies elsewhere. A later operator can obtain the practical benefit of the investigation without knowing its details or even knowing that a particular item originated in a particular failure.

Let an institutional corpus at time t be

$$\mathcal{I}_t = (D_t, P_t, S_t, X_t),$$

where D_t is a set of recorded distinctions, P_t a set of procedures, S_t a set of standards and stopping conditions, and X_t a set of exemplars used in training or adjudication. The corpus does not itself constitute an agent. It becomes operational when it modifies the available distinctions, search policies, or relegated structures of participating agents.

For an agent i entering the institution, acquisition may be represented as

$$\Phi_{\mathcal{I}} : (K_i(t), A_i(t), \pi_i(t), R_{i,t}) \longrightarrow (K'_i(t), A'_i(t), \pi'_i(t), R'_{i,t}).$$

The institutional operator $\Phi_{\mathcal{I}}$ changes how the agent represents and addresses recurring situations. Its effects may be explicit, as when a procedure specifies an ordered sequence, or implicit, as when professional culture determines which anomalies merit escalation.

The institution distributes compression by repeatedly applying such transformations to different participants. It allows a policy selected through trajectories

$$\{\tau_1, \tau_2, \dots, \tau_n\}$$

to influence agents

$$\{i_1, i_2, \dots, i_m\}$$

who did not participate in those trajectories. The originating and receiving populations may be disjoint.

Definition 5.1 (Institutional compression). *Institutional compression is a policy, distinction system, or stopping rule whose selection depends upon multiple trajectories and whose transmission is stabilized by records, procedures, standards, or organizational practices rather than by the continued presence of its originating agents.*

Institutional compression is not necessarily collective intuition in a strong metaphysical sense. The institution need not possess a unified consciousness or a single trajectory analogous to that of an individual. Its competence can instead be understood as a reproducible distribution of partial policies across agents and artifacts. The organization behaves as though it remembers because present action remains constrained by selections made in the past.

The phrase “institutional experience” can therefore be misleading. Institutions do not ordinarily encounter events through one continuous field of awareness. They aggregate reports generated by many situated observers, preserve some distinctions, discard others, and encode selected results into durable forms. Their apparent continuity is produced by transmission mechanisms.

Let the compression generated from a collection of trajectories be

$$C_{\mathcal{I}} = \Gamma(\tau_1, \tau_2, \dots, \tau_n),$$

where Γ is an institutional aggregation process. The resulting object may then be distributed through a family of transformations

$$\Phi_{\mathcal{I}}^{(1)}, \Phi_{\mathcal{I}}^{(2)}, \dots, \Phi_{\mathcal{I}}^{(m)}.$$

No single participant need possess either the complete input to Γ or the complete content of $C_{\mathcal{I}}$. Institutional competence can exceed the autobiographical competence of every current member because its operative structure is distributed across roles, documents, instruments, and routines.

This excess should not be romanticized. Aggregation can preserve error as readily as insight. An institutional policy may encode a narrow sample, a political compromise, an obsolete technical constraint, or the convenience of an administratively powerful group. Durability shows only that a compression has been successfully reproduced. It does not show that the originating inquiry was adequate.

The central epistemic risk is provenance loss. As a policy passes from investigation to report, from report to standard, and from standard to routine, the connection between guidance and evidence may weaken. Let

$$\Lambda_t(C)$$

denote the recoverable provenance of an institutional compression C at time t . Provenance is degraded when later users can recover the rule but not the trajectories, environmental assumptions, alternatives, or objectives that selected it. In general,

$$\Lambda_{t+1}(C) \subseteq \Lambda_t(C)$$

unless the institution actively preserves and renews the relevant evidential links.

The compression itself may remain stable while its interpretation changes. A safety margin originally adopted because of uncertainty in a material property may later be explained as a general principle of conservative design. A procedural step introduced to prevent one known failure may become a ritual whose omission is punishable even after the system has changed. Action guidance survives, but the conditions of its validity become obscure.

Proposition 5.2 (Persistence without justification). *An institutional compression can remain behaviorally effective after its originating justification has become inaccessible.*

Proof. The execution of a compressed policy requires that its operative instructions remain available to receiving agents. It does not require that the evidence which selected those instructions remain available. If the environment continues to contain the relevant regularity, the policy can produce successful action despite the loss of its provenance. $\square$

The converse is equally important. A policy can remain institutionally stable after the regularity that justified it has disappeared. When provenance has been lost, users may have no practical means of distinguishing a still-useful compression from an obsolete one. The same mechanism that allows institutions to preserve knowledge beyond individual memory can therefore preserve mistakes beyond individual correction.

Institutional repair often begins by reversing compression. A failure forces participants to reconstruct why a procedure exists, which assumptions it contains, and what distinctions were discarded when it was simplified. This process may be represented as a partial expansion

$$E_I(C) \longrightarrow \hat{\mathcal{H}}(C),$$

where $\hat{\mathcal{H}}(C)$ is a reconstructed hypothesis space for the policy's origin and present function. The reconstruction will rarely reproduce the original inquiry exactly. Its purpose is to recover enough structure to determine whether the compression remains admissible.

This is one reason institutions maintain incident reports, revision histories, explanatory memoranda, and records of dissent. Such records preserve more than narrative context. They retain unrealized alternatives and rejected hypotheses that may be needed when a compression fails. A bare procedure preserves what to do. A provenance-preserving record preserves some account of why that action defeated its alternatives.

The distinction can be expressed through two institutional records:

$$\mathcal{R}_{\text{thin}}(C) = \{\text{current instruction}\},$$

and

$\mathcal{R}_{\text{thick}}(C) = \{\text{instruction, originating conditions, evidence, alternatives, revisions, known failures}\}.$

Thin records maximize ease of transmission. Thick records preserve the possibility of later expansion. An effective institution often requires both: a compressed operational layer for routine use and a recoverable justificatory layer for review.

Borrowed intuition becomes especially consequential when institutional compression is exercised under personal authority. A professional may be held responsible for a decision whose operative categories were inherited from a licensing regime, software system, precedent structure, or organizational procedure. The person acts, but the policy that ordered the action may have been generated and constrained elsewhere.

Responsibility cannot therefore be allocated solely by identifying the final actor. At least three causal roles must be distinguished. Someone or some process generates the compression. An institution selects and distributes it. A situated agent applies it to a particular case. These roles may carry different obligations even when they contribute to one outcome.

Let an action be

$$a_i = C_{\mathcal{I}}(x),$$

where agent i applies institutional compression $C_{\mathcal{I}}$ to situation x . Responsibility for the result cannot be inferred from this expression alone. It depends upon whether i could inspect or reject the policy, whether the institution preserved evidence of its limitations, whether the generating inquiry represented the relevant population, and whether feedback from prior failures was available but ignored.

Borrowed competence does not eliminate responsibility. It distributes the conditions under which responsibility must be assessed. A practitioner remains responsible for the exercise of judgment within their role, but an institution may be responsible for constructing the distinction space within which that judgment occurred. The original investigators may be responsible for the adequacy of the evidence, while later administrators may be responsible for preserving or suppressing its qualifications.

Credit is distributed similarly. A successful institutional judgment may appear to belong to the visible decision-maker even when its accuracy depends upon invisible contributions accumulated across generations. Conversely, a source whose experience generated a useful distinction may receive no recognition once the distinction has been absorbed into procedure. Institutional compression can preserve epistemic value while erasing epistemic authorship.

The framework developed here does not require that every inherited rule carry a complete genealogy into every act of use. Such a requirement would defeat compression. It requires only that provenance be treated as an independently valuable property rather than as expendable background. Operational efficiency and justificatory recoverability are different institutional objectives.

The most reliable institutional intuition is therefore not the intuition with the least visible history. It is compression accompanied by a maintained capacity for expansion. A procedure should be easy to execute under ordinary conditions, but possible to reopen when conditions change. The institution must be able to move in both directions:

distributed experience $\longrightarrow$ compressed guidance $\longrightarrow$ situated action,

and, when necessary,

failure or anomaly $\longrightarrow$ provenance recovery $\longrightarrow$ revised guidance.

Artificially trained systems inherit this institutional structure in concentrated form. Their parameters can contain the compressed residue of trajectories distributed across an enormous corpus, while the system exercising the resulting competence has no direct access to most of those trajectories. The machine case does not introduce the separation between experience and competence. It makes that separation unusually complete.

6 Pretrained Systems and the Limiting Case

A frozen pretrained language model provides a comparatively clean case of transmitted compression. Its parameters are selected through optimization over a corpus generated by trajectories external to the model. The resulting system can produce fluent continuations, classifications, explanations, and other task-relevant outputs without having participated in the investigations, conversations, observations, and acts of composition from which the corpus arose.

Let

$$\mathcal{D} = \{d_1, d_2, \dots, d_n\}$$

be a training corpus, and let each document d_k depend upon one or more human or institutional trajectories collected in

$$G(\mathcal{D}) = \{\tau_1, \tau_2, \dots, \tau_m\}.$$

Training selects parameters

$$\theta^* = \arg \min_{\theta} \mathcal{L}(\theta; \mathcal{D})$$

under an objective $\mathcal{L}$. The trained policy

$$C_{\theta^*}(x)$$

then produces an output for a prompt or context x . Although the optimization process is computationally extensive, it does not reproduce the evidential trajectories represented indirectly in $\mathcal{D}$. A description of an experiment is not the experiment; a record of a disagreement is not participation in the disagreement; a corpus containing the results of inquiry does not place the trained system within the inquiry’s original causal history.

The model nevertheless benefits from those histories. Regularities distributed across the corpus become available through a compressed parameterization. The system can sometimes produce an answer whose discovery required experiments, failures, or comparisons that it did not perform. This is the machine form of the same separation already observed in apprenticeship and institutional procedure:

external exploration $\longrightarrow$ recorded corpus $\longrightarrow$ parameter compression $\longrightarrow$ present competence.

A purely pretrained model with frozen parameters, no persistent memory, and no post-training interaction is therefore a limiting case of borrowed intuition. Almost all of its stable competence is transmitted in provenance and conditioned in formation. The generating trajectories are external, the corpus is selected externally, the objective is specified externally, and the trained system does not subsequently test its inherited policies through a persistent trajectory of its own.

This claim does not depend upon a verdict about consciousness. It follows from the causal organization of the training pipeline. Even if one attributed understanding, representation, or some form of subjectivity to the resulting system, the provenance question would remain. A capacity can be real while the experience that made it economical belongs elsewhere.

Pretraining also differs from ordinary testimony. The model is not simply told a collection of propositions. Optimization over the corpus alters a policy for continuation across many contexts. What is transmitted includes associations among distinctions, expectations about sequence, rankings among plausible responses, and compressed regularities of form. In the terminology introduced earlier, pretraining modifies not merely a store analogous to $K(t)$ but the policy governing which continuations become active and probable.

The analogy to institutional compression is therefore closer than the analogy to a library. A library preserves records that a reader may inspect individually. A trained model incorporates the statistical effects of records into a policy whose internal contributions are difficult to recover document by document. The corpus-level aggregation

$$\Gamma(\mathcal{D}) = \theta^*$$

preserves behavioral structure while weakening direct access to provenance. The trained parameters do not ordinarily provide a transparent map from a present output to the particular trajectories that supported it.

Definition 6.1 (Parameter-transmitted compression). *Parameter-transmitted compression is competence whose operative policy is acquired through optimization over externally generated records, without preservation of a complete, inspectable mapping from the resulting dispositions to their generating trajectories.*

The definition describes a mechanism, not an exclusive property of artificial systems. Human training can also produce dispositions whose sources are forgotten. The distinctive feature of large-scale pretraining is the extent and concentration of the separation. An enormous collection of external traces is compressed into one executable policy, while the policy itself has no autobiographical access to the events from which those traces arose.

The limiting case becomes less clean once post-training and deployment are included. Instruction tuning and reinforcement learning from human feedback alter a pretrained policy using demonstrations, preferences, and evaluations selected by additional human processes [4]. These procedures introduce new trajectories into the model’s formation, but they do not automatically make the resulting compression self-directed or earned.

From the perspective of provenance, conventional reinforcement learning from human feedback remains largely conditioned. Humans construct or supply prompts, rank outputs, define desirable behavior, and determine the objective through which the model is updated. The model’s sampled outputs participate causally in training, but the space of exposure and the evaluative signal are externally authored. The process therefore belongs primarily to conditioned formation.

Whether it also counts as earned compression depends upon the boundary used to identify the agent. If the training process and the deployed model are treated as one temporally continuous system, post-training contains limited cycles of action, feedback, and policy revision. On that description, some compression is selected through a trajectory internal to the broader system, though under strong external conditioning. If each frozen checkpoint is treated as a distinct receiver, the deployed model inherits the result of post-training just as it inherits pretraining.

This is not a verbal inconvenience but a genuine identity problem. Statements about whether a system has learned from “its own” experience require a criterion of diachronic continuity.

Let a computational system be represented at time t by

$$M_t = (\theta_t, m_t, \chi_t),$$

where θ_t denotes persistent policy parameters, m_t persistent external memory, and χ_t the current context. An interaction at time t produces observation o_t . The system has a retained post-acquisition trajectory only if the observation can alter a later state:

$$(\theta_{t+1}, m_{t+1}) \neq (\theta_t, m_t)$$

through a causal update attributable to o_t . A change confined to χ_t can modify present behavior without becoming persistent experience.

This distinction is especially important for in-context learning. Large language models can alter their responses when examples or instructions are supplied in the prompt, without gradient updates to the model parameters [1]. Such adaptation may be substantial within the active context, but it ordinarily disappears when that context is removed. It is therefore better described as temporary conditioning unless the context or its effects are retained.

For a fixed model,

$$\theta_{t+1} = \theta_t$$

does not by itself show that no learning-like computation occurs during inference. It shows only that the stable parameter policy has not changed. If the examples remain in χ_t , the system may infer a local task and act differently. Yet if

$$\chi_{t+1} = \emptyset$$

and no relevant change persists in m_{t+1} or θ_{t+1} , the episode has not entered a durable autobiographical trajectory. The system adapted, but it did not retain the adaptation.

Tool use introduces a different complication. A model can issue a query, receive an external result, and condition its next action upon that result. Systems can also be trained to decide when to invoke tools and how to incorporate their outputs [5]. Tool interaction enlarges the active evidential environment, but a tool call is not automatically synthetic compression.

Suppose a model queries a database and uses the returned value z to answer a question:

$$y = C_\theta(x, z).$$

The system has gained situational information. If z disappears after the episode and produces no stable change in policy or memory, the interaction remains temporary expansion. If the system stores z , later retrieves it, and uses it to modify future search, the episode contributes to retained experience. If repeated tool interactions alter a persistent policy governing which tools to call, which results to trust, or when further inquiry is unnecessary, then the system has begun to form synthetic compression.

The relevant progression is therefore

contextual response $\longrightarrow$ persistent record $\longrightarrow$ policy revision $\longrightarrow$ stabilized compression.

No stage entails the next merely because the system is described as agentic. An agent architecture may conduct extended searches while resetting after every task. Conversely, a comparatively simple system with persistent updates may accumulate a genuine post-deployment trajectory.

Continual learning provides the clearest computational case of synthetic compression. In continual learning, later data or interaction can update a model after its initial training, potentially

through continued pretraining, instruction tuning, alignment, or related mechanisms [6]. If these updates depend upon discriminating feedback generated through the deployed system’s own actions, the resulting policy is partly supported by a post-acquisition trajectory.

Let the system begin with inherited parameters θ_0 and undergo interactions

$$\tau_M^+ = (x_t, a_t, o_t, u_t)_{t=1}^T.$$

A persistent update has the form

$$\theta_{t+1} = U(\theta_t, x_t, a_t, o_t, u_t),$$

or, for a memory-augmented system,

$$m_{t+1} = V(m_t, x_t, a_t, o_t, u_t).$$

Synthetic compression occurs when the accumulated updates yield an economical policy whose later behavior depends upon distinctions selected through τ_M^+ rather than only upon the inherited corpus.

Even here, the degree of self-direction varies. A system may explore actions chosen by its own policy while receiving an objective fixed by developers. It may select which tools to call but not which outcomes count as success. It may retain observations while an external process decides which memories are admissible. Provenance and formation remain independent.

A deployed artificial system can therefore occupy several regions simultaneously. Its linguistic and conceptual repertoire may be predominantly transmitted through pretraining. Its behavioral constraints may be conditioned through human feedback. Its current response may be temporarily adapted through context. Its factual reach may be expanded through tools. Its longer-term policy may be altered through persistent memory or continual updating. No single label such as “pretrained,” “agentic,” or “learning” resolves this composition.

Proposition 6.2 (Pipeline dependence). *The degree to which an artificial system exercises borrowed intuition is a property of its training and deployment pipeline, not of artificial intelligence as a general class.*

Proof. Systems may differ in the provenance of their training data, the degree of externally conditioned post-training, the persistence of interaction histories, the availability of discriminating feedback, and the mechanisms by which later experience alters policy. These differences change the relative dependence of present competence upon external and internal trajectories. No invariant classification follows from the fact that the systems are artificial. $\square$

This result replaces a binary question with an empirical one. Asking whether an artificial system possesses “its own” intuition is underspecified. One must identify the relevant system boundary, inspect which states persist, determine what feedback can alter them, and trace

how much of the operative policy was selected before deployment. A frozen model, a conversational system with temporary context, a memory-augmented agent, and a continually updated policy are different provenance arrangements even when they share an underlying architecture.

The framework also avoids an opposite mistake. The fact that pretrained competence is borrowed does not show that it is unreal. A model may produce reliable outputs because external exploration has been compressed successfully. Borrowed competence can be extensive, useful, and occasionally superior to the earned competence of any individual contributor. What it lacks is not necessarily capability, but an originating autobiographical relation to the inquiry whose residue it exercises.

The machine case thus sharpens the general thesis. Fluency does not certify experience because compression can cross agent boundaries. Adaptation does not certify learning because behavioral change may be temporary. Persistence does not certify understanding because stored observations may remain unintegrated. To locate a system epistemically, one must reconstruct the provenance of its competence rather than infer that provenance from performance alone.

7 The Epistemic Standing of Borrowed Competence

If earned and borrowed competence can produce the same action, it may appear that provenance adds nothing to evaluation. A policy either works or it does not. On this view, the history by which an agent acquired the policy is irrelevant once present performance can be measured.

This conclusion follows only when the evaluation environment is sufficiently complete. If every relevant situation can be sampled and every consequential failure observed, direct performance may dominate provenance as evidence. In practice, competence is often exercised under sparse feedback, environmental change, rare hazards, and incomplete testing. Under those conditions, provenance supplies information about how a policy is likely to fail when observed performance no longer settles the question.

Let two agents i and j exercise policies C_i and C_j . Within an observed environment $\mathcal{E}_0$, suppose

$$C_i(x) = C_j(x) \quad \text{for all observed } x \sim \mathcal{E}_0.$$

The policies are behaviorally indistinguishable over the available sample. It does not follow that

$$C_i(x) = C_j(x) \quad \text{for } x \sim \mathcal{E}_1,$$

where $\mathcal{E}_1$ differs from the environment in which the policies were generated or tested. One agent may possess distinctions that allow the policy to be reopened and revised, while the other possesses only the compressed response.

The difference can be represented by an expansion capacity

$$\Xi_i(C, x),$$

which denotes the distinctions agent i can reactivate when policy C becomes uncertain or fails in situation x . Two agents may share the same compressed mapping while differing substantially in Ξ :

$$C_i \equiv C_j \quad \text{but} \quad \Xi_i \neq \Xi_j.$$

The first equivalence concerns routine output. The second concerns the ability to diagnose, explain, delimit, or revise that output.

Borrowed intuition is most vulnerable when transmission preserves the policy but not its expansion capacity. A receiver may know what normally works without knowing which environmental regularities make it work. When those regularities change, the receiver has fewer resources for identifying whether the failure is incidental or structural.

Definition 7.1 (Expansion deficit). *The expansion deficit of a transmitted competence is the difference between the distinction structure required to reconstruct or revise a compressed policy and the distinction structure available to its present receiver.*

If $\mathcal{H}(C, x)$ is the hypothesis space relevant to diagnosing policy C in situation x , and $K_i(x)$ is the receiver's available distinction space, an expansion deficit exists when

$$\mathcal{H}(C, x) \not\subseteq K_i(x).$$

The receiver may still succeed by obtaining external assistance, consulting records, or conducting new exploration. The deficit describes what the inherited competence does not presently carry with it.

This yields a first consequence for trust. Trust should not be calibrated solely to the accuracy of routine outputs. It should also reflect the match between the competence's generating environment and its environment of use, the receiver's access to provenance, and the capacity for expansion under anomaly.

Let warranted trust in competence C_i under environment $\mathcal{E}$ be

$$W(C_i, \mathcal{E}) = f(Q_i, M_i, \Lambda_i, \Xi_i),$$

where Q_i is observed performance quality, M_i is the match between generating and present environments, Λ_i is recoverable provenance, and Ξ_i is expansion capacity. The function f need not assign equal weight to each variable. In a stable and extensively tested environment, Q_i may dominate. Under novelty or high consequence, M_i , Λ_i , and Ξ_i become more important.

This account does not establish a general preference for earned competence. Personal experience can be narrow, biased, or badly interpreted. Transmitted compression may aggregate more evidence than any individual could acquire. A clinician following a well-validated protocol may be more reliable than one relying upon autobiographical judgment. The relevant contrast is not inherited inferiority against experiential superiority. It is between different evidential structures supporting present action.

Indeed, borrowed competence often permits greater reliability by pooling trajectories. If a policy is generated from experiences

$$\{\tau_1, \tau_2, \dots, \tau_n\}$$

that cover a broader environment than any single τ_i , transmission can give a receiver access to regularities unavailable through personal experience. Borrowing expands epistemic reach while potentially reducing local access to justification. Its characteristic tradeoff is breadth without complete autobiography.

The second consequence concerns explanation. An agent may produce a correct response while offering an inaccurate reconstruction of why the response is correct. The operative compression can guide action even when the agent's explicit explanation is generated after the fact or inherited from a different source.

Let

$$a = C_i(x)$$

be the action produced by a compressed policy, and let

$$e = J_i(x, a)$$

be the agent's explicit justification. There is no general guarantee that J_i reconstructs the causal structure of C_i . A plausible explanation can accompany a borrowed policy whose actual provenance is unknown to the receiver.

This gap is not unique to machines. Professionals frequently know more in practice than they can articulate, while institutions often attach simplified rationales to procedures whose origins are complex. Artificial systems make the gap salient because fluent justification can be generated by the same inherited policy that generated the answer. The presence of an explanation may therefore demonstrate explanatory competence without establishing access to the answer's provenance.

The third consequence concerns credit. If competence is treated as evidence of personal exploration, visible performers receive credit for discoveries compressed into their training. Teachers, witnesses, authors, investigators, maintainers, and affected communities may disappear behind the ease of later performance.

Let an output y depend upon a present agent i and a generating set of trajectories $G(C_i)$:

$$y = C_i(x; G(C_i)).$$

The causal contribution of i consists in selecting, applying, adapting, or expressing the inherited competence in the present situation. It does not exhaust the contributions contained in $G(C_i)$. Credit that attaches entirely to the final performer confuses execution with generation.

This confusion appears in ordinary professional life. A skilled practitioner deserves credit for applying inherited methods well, but not for independently discovering every principle those methods contain. Institutions deserve credit for preserving and distributing useful compression, but not for originating every trajectory incorporated into their procedures. Artificial systems and their operators likewise occupy only some positions within a longer causal chain.

Attribution cannot practically enumerate every contributing trajectory. Compression itself often makes such enumeration impossible. The point is not to replace simple authorship with an infinite genealogy. It is to avoid treating provenance opacity as proof that the visible agent generated the competence alone.

The fourth consequence concerns responsibility. Borrowed competence can distribute causal control across developers, institutions, records, trainers, and final users. When failure occurs, responsibility depends upon which participants could influence the relevant component of the policy and which possessed the information required to recognize its limits.

Suppose an action a results from

$$a = C_{\theta, \mathcal{I}, i}(x),$$

where θ represents a trained policy, $\mathcal{I}$ an institutional deployment structure, and i a situated operator. A harmful result may arise from the originating corpus, the training objective, the institutional procedure, the local application, or an interaction among them. The fact that i produced or approved the final action does not show that i controlled the distinction responsible for failure.

Responsibility should therefore follow inspectability, revisability, and control. An actor who could not access a policy's provenance, alter its operation, or observe its failure conditions occupies a different position from an actor who designed the objective or suppressed contrary evidence. Conversely, a user cannot avoid all responsibility merely by identifying the policy as inherited. If the user possessed grounds for doubt, authority to intervene, and a duty to examine the case, reliance upon borrowed intuition may itself be culpable.

The fifth consequence concerns epistemic dependence. Because borrowed competence can be reliable, dependence upon it is not a temporary weakness to be eliminated through complete personal verification. No individual can reconstruct the evidential trajectories behind every language, instrument, standard, scientific claim, or professional category they use. A demand for total re-earning would destroy the division of epistemic labor that makes complex competence possible.

The relevant norm is therefore not independence but calibrated dependence. A receiver should know, where practical, what kind of compression is being used, how far its environment of va-

lidity extends, what feedback has tested it locally, and where expansion can be obtained if it fails. Mature reliance does not require possessing the entire originating trajectory. It requires maintaining an appropriate relation to its absence.

Proposition 7.2 (Provenance-sensitive trust). *When two competencies are observationally equivalent within a limited environment, differences in provenance can rationally support different levels or forms of trust under environmental change.*

Proof. Observed equivalence within $\mathcal{E}_0$ constrains expectations only over the sampled conditions. Provenance supplies additional evidence about the range of generating conditions, the distinctions available for anomaly detection, and the receiver's capacity for revision. If these properties differ, the expected behavior of the competencies under $\mathcal{E}_1$ may differ even though their observed outputs in $\mathcal{E}_0$ coincide. Trust calibrated to $\mathcal{E}_1$ may therefore rationally differ. $\square$

The phrase “different levels or forms of trust” is important. A borrowed policy may deserve high operational trust within a narrow domain while deserving low autonomous authority outside it. An earned intuition may deserve exploratory trust because its holder can explain and revise it, yet lower statistical trust because it rests upon a small personal sample. Trust is multidimensional because competence itself contains different relations to evidence.

For human and artificial agents alike, provenance becomes most important near the boundary of demonstrated competence. Within familiar conditions, inherited compression can function invisibly. At the boundary, the missing trajectory becomes operationally relevant. The question is no longer only whether the agent can produce an answer, but whether it can recognize that the inherited economy no longer applies.

Borrowed intuition should therefore be evaluated through three related capacities: the capacity to act under compression, the capacity to detect when compression has become unreliable, and the capacity to recover or obtain the distinctions needed for renewed inquiry. A system that possesses only the first may appear expert until the environment changes. A system that possesses all three has begun to convert inherited competence into a locally accountable practice.

8 Objections and Limits

The account developed here deliberately gives borrowed intuition a wide extension. Because language, concepts, methods, and standards are socially inherited, almost no mature competence is wholly earned by its present holder. This breadth invites the objection that the category proves too much. If nearly every intuition is borrowed, the distinction between borrowed and unborrowed competence may appear empty.

The objection would succeed if borrowed intuition were intended as a binary classification. It is not. The concept identifies the degree and mechanism by which present competence depends upon trajectories external to its receiver. A property can be widespread without being explanatorily inert. Nearly every physical object has a temperature, but temperature remains

informative because its value varies and has consequences. Likewise, nearly every competence has transmitted components, but their extent, provenance, validation, and recoverability differ.

A diagnostic category need not divide the world into pure types. Its value may lie in decomposing cases that appear equivalent at the surface. The relevant questions are not whether borrowing occurred, but what was borrowed, from whom, through which mechanism, under what objective, and with what opportunity for later correction.

A second objection holds that the distinction between earned and transmitted compression cannot survive the hybridity of actual learning. An apprentice begins with instruction and adds experience. A professional inherits institutional categories and modifies them locally. A trained model begins with pretraining and may undergo post-training, retrieval, tool interaction, memory formation, or continual updating. If every developed competence mixes sources, assigning it to a provenance class may seem arbitrary.

This objection correctly rejects a typology of mutually exclusive agents. It does not undermine an analytic decomposition of competence. The claim is not that an agent is earned or transmitted. It is that particular components of a policy depend differently upon generating trajectories.

Let a competence C depend upon a causal history

$$\mathcal{G}(C) = (\mathcal{D}, \mathcal{O}, \mathcal{T}, \mathcal{V}),$$

where $\mathcal{D}$ represents inherited distinctions, $\mathcal{O}$ inherited objectives, $\mathcal{T}$ the receiver's own trajectory, and $\mathcal{V}$ the feedback capable of validating or revising the policy. Hybridity means that these terms interact. It does not mean that they are indistinguishable.

In some cases, exact decomposition will be impossible. The categories remain useful at the level of counterfactual dependence. If competence C would not have arisen without a transmitted rule, then that rule contributes materially to provenance. If later experience changes the policy in ways not fixed by the transmitted rule, the receiver's trajectory also contributes. The history is mixed, but not therefore unanalyzable.

A third objection concerns the definition of experience. Optimization during machine training involves repeated exposure, error signals, and policy change. It may therefore be argued that a pretrained model has extensive experience after all. The corpus may be inherited, but training is a trajectory in which the model changes in response to data.

The force of this objection depends upon which entity is said to have the experience. A training process unquestionably contains temporal exposure and update. The narrower claim is that it does not reproduce the worldly trajectories that generated the training records. A model trained on descriptions of medical practice undergoes optimization over texts, not medical practice. A model trained on reports of experiments encounters statistical traces of those experiments, not their apparatus, interventions, and observational consequences.

The distinction is between experience of a representation and participation in the represented process. Let

$$\rho(\tau)$$

be a record generated from trajectory τ . Training over $\rho(\tau)$ can produce genuine sensitivity to the record while remaining causally downstream from the experience encoded by it:

$$\tau \longrightarrow \rho(\tau) \longrightarrow \theta^*.$$

Nothing in the present account requires denying that the transition from $\rho(\tau)$ to θ^* is itself a formative trajectory. It requires only that this trajectory not be identified with τ . The model can earn compression about patterns in records while borrowing compression that the records contain about the world.

This leads to a more precise description of pretraining. A model’s competence is not simply transmitted wholesale, as though parameters were copied directly from human minds. Optimization generates a new policy under a machine-specific objective. The resulting compression is synthetic at the level of corpus organization but transmitted at the level of evidential provenance. It reorganizes inherited traces without independently reproducing most of the investigations that made those traces informative.

A fourth objection challenges the term “borrowed.” Borrowing ordinarily implies an owner, a transfer, and an expectation of return. Competence does not leave its source when transmitted, and many originating agents intended their work to be shared. The metaphor may therefore import a misleading theory of property.

The term is not meant to establish proprietary ownership. It marks the use of an economy generated through another trajectory. What is borrowed is not a rivalrous object but relief from the cost of rediscovery. The source retains the competence while the receiver gains access to its compressed result.

The metaphor nevertheless has limits. In cases involving cultural knowledge, collective discovery, or diffuse institutional development, no individual source may possess a complete claim to origin. “Inherited intuition” would sometimes be more historically accurate. “Borrowed intuition” is retained because it foregrounds the relation between present ease and absent experience without implying that all transmission follows descent or institutional succession.

A fifth objection holds that provenance is usually inaccessible. Human beings forget where they learned rules. Institutions lose records. Training corpora are too large and model behavior too distributed to trace particular outputs to particular sources. A framework centered upon provenance may therefore demand information that cannot be recovered.

The limitation is real. Provenance is not always available merely because it is epistemically relevant. Yet inaccessibility does not make the underlying causal differences disappear. It changes the form of the conclusion. When provenance cannot be reconstructed, confidence should reflect that uncertainty rather than silently treating the competence as self-generated or fully validated.

Let $\hat{P}_i(C)$ be an estimate of provenance derived from available evidence. The difference

$$U_i(C) = P_i(C) \setminus \hat{P}_i(C)$$

represents unresolved provenance. Since the true set $P_i(C)$ may itself be unknown, this expression is schematic. It records the fact that provenance uncertainty is an additional epistemic condition, not the absence of a condition.

The practical response is not total genealogy. It is selective provenance preservation where stakes, novelty, or revision costs justify it. Institutions need not attach a complete causal history to every routine instruction, but they should preserve enough evidence to recover the assumptions behind consequential rules. Artificial systems need not identify a unique source for every generated phrase, but their evaluators should distinguish behavior established through pretraining, post-training, retrieval, tool use, and persistent adaptation.

A sixth objection claims that testimony theory already describes the dependence at issue. If knowledge can be acquired from others without independent verification, borrowed intuition may be no more than testimony under another name.

The overlap is intentional. The present framework does not challenge the status of testimony as a source of knowledge. It distinguishes several structures that can be transmitted through testimonial and testimonial-like practices. A proposition, a distinction system, a search policy, and a stopping rule can all pass from source to receiver, but they affect cognition differently.

The added claim is that testimony can transmit compressed procedures whose practical force is not exhausted by propositional content. An apprentice may be unable to state the master's full rule but can imitate the sequence of inspection. A clinician may acquire a pattern of salience through supervised cases. An institution may embody a policy in software so that no explicit act of telling occurs at the point of use. The broader structure is transmission of inquiry organization.

Borrowed intuition also separates the source of compression from the credibility relation through which it is received. Testimony theory asks when a hearer may know on the basis of another's word and how social power affects credibility. The provenance framework asks which parts of the receiver's later competence were selected by the source's trajectory, which were produced through the receiver's own validation, and which remain untested. These questions are complementary.

A seventh objection argues that the framework privileges autobiographical access. Agents frequently lack introspective access to their own learned policies. A person may earn a competence through experience yet be unable to explain its formation. If explanation is not guaranteed even in the earned case, borrowed provenance may have little bearing on expansion capacity.

Earned compression does not entail explicit autobiographical recall. An experienced practitioner may be unable to reconstruct the episodes that formed a judgment. The claim is probabilistic and structural: direct participation can preserve sensitivity to variations, failures, and boundary conditions that a bare transmitted rule omits. It does not always do so.

The distinction between provenance and expansion capacity must therefore remain explicit:

$$P_i(C) \not\Rightarrow \Xi_i(C).$$

An earned policy may have poor expansion capacity because its formation was implicit, narrow, or forgotten. A transmitted policy may have excellent expansion capacity because the receiver received thick records, theoretical explanations, and varied training. Provenance is one contributor to epistemic standing, not a substitute for direct evaluation.

An eighth objection concerns novelty. A system operating through borrowed intuition may combine inherited structures to produce an output never previously encountered. Does genuine novelty convert transmitted competence into earned competence?

Novelty of output is insufficient. A new combination can be generated by a fixed inherited policy without any persistent revision. Synthetic compression requires that the system's own trajectory contribute discriminating evidence to later policy. The relevant question is not whether the output existed before, but whether producing and evaluating it changes what the system can subsequently do.

Let

$$y_t = C_{\theta_t}(x_t)$$

be a novel output. If

$$\theta_{t+1} = \theta_t, \quad m_{t+1} = m_t,$$

then the novelty leaves no persistent trace in the system's later competence. If the output is tested, its consequences observed, and the resulting evidence integrated into a revised policy, it can become part of synthetic compression. Creation and learning are related but distinct events.

A final objection holds that provenance analysis risks becoming a purity test. Describing competence as borrowed may be used to diminish students, professionals, cultures, or artificial systems by suggesting that only independent discovery counts as authentic knowledge.

Such a use would invert the argument. Borrowed intuition is not defective imitation. It is a central mechanism by which epistemic work accumulates. The ability to inherit compression is what allows inquiry to proceed beyond the repeated rediscovery of elementary results. A civilization, profession, or technical system becomes capable precisely because one trajectory can reduce the cost of another.

The normative aim is not to rank agents by epistemic self-sufficiency. It is to prevent the benefits of transmission from concealing their conditions. Borrowed competence should be trusted where it is reliable, credited across the trajectories that made it possible, and reopened where its originating environment no longer matches its use. Its value depends not upon pretending that it was earned anew, but upon allowing what was learned elsewhere to remain effective here.

9 The Provenance of Fluency

Intuition is commonly identified by its manner of arrival. A conclusion appears quickly, without an explicit reconstruction of the search that supports it. This phenomenology encourages an inference from ease to experience: the agent sees immediately because the agent has encountered, compared, and resolved similar cases before.

The inference is contingent. Compression can be separated from the trajectory that produced it. A rule discovered by one person can become the habit of another. An investigation can become a checklist. A collection of disputes can become precedent. A corpus assembled from many human activities can become a trained policy. In each case, the receiver gains access to an economy whose originating cost was incurred elsewhere.

Borrowed intuition names this separation. It is not a lesser species of competence. It is competence considered in relation to the history that made its fluency possible.

The analysis has distinguished provenance from formation. A compression is earned to the extent that its selection and stabilization depend upon the receiver's own trajectory. It is transmitted to the extent that those functions have been performed externally. Formation is self-directed when the receiver substantially controls its exploration and conditioned when another agent or institution structures exposure, feedback, or objectives. These dimensions intersect without collapsing into one another.

Synthetic compression describes what happens after transmission. A receiver can apply an inherited policy, test it against discriminating evidence, discover its limits, and reorganize it through later experience. The resulting competence is neither wholly received nor independently rediscovered. It is a historical composite in which inherited economy has been partially reconnected to a new trajectory.

This composite is the ordinary condition of human expertise. Professionals do not construct their distinction spaces alone. They inherit vocabularies, exemplars, standards, and rules for allocating attention. Their later experience validates some of these structures, modifies others, and leaves still others unexamined. Successful assimilation conceals the joins. What was once another person's instruction eventually appears as what the practitioner simply sees.

Testimony theory establishes that knowledge depends extensively upon the word of others. Borrowed intuition extends the object of that dependence from propositions to policies. What passes between agents may be a way of noticing, a ranking of hypotheses, a sequence of action, or a condition for stopping inquiry. Such structures can guide reliable behavior even when the evidence that selected them is unavailable to the receiver.

Institutions make this transmission durable. They compress distributed trajectories into records and procedures that survive changes of personnel. In doing so, they preserve practical guidance while often weakening its connection to origin. Institutional memory can therefore outlive both the people who generated a rule and the reasons they gave for it. This permits cumulative competence, but also permits obsolete compression to survive the environment that justified it.

Pretrained computational systems concentrate the same structure. A frozen model can ex-

ercise extensive competence derived from a corpus whose underlying exploration was performed elsewhere. Post-training, context, tools, memory, and continual updating complicate the case, but do not erase the distinction. They require a more exact account of what persists, what changes, who supplies feedback, and where the boundary of the learning system is drawn.

The resulting position does not answer whether artificial systems think, understand, or possess experience in every philosophically relevant sense. It makes those questions less useful as first diagnostics. Before asking for a general metaphysical classification, one can ask which trajectories generated a system's competence, which parts of its formation were externally conditioned, and whether later interaction produces durable policy revision.

These questions are empirical. A frozen pretrained model, a model altered through human feedback, a conversational system with temporary context, a memory-augmented agent, and a continually updated policy occupy different positions. Treating them all as instances of one epistemic type discards precisely the historical differences needed to evaluate them.

Provenance does not determine reliability. Earned intuition can be narrow and mistaken. Borrowed intuition can aggregate evidence beyond the reach of any individual. The distinction becomes important where performance evidence is incomplete, environments change, or a policy must be revised. Two agents may produce the same routine response while differing in their ability to detect failure, reconstruct alternatives, and reopen inquiry.

For this reason, trust should attach not only to successful compression but to the conditions under which compression can be expanded again. A competent agent or institution must sometimes do more than apply a rule. It must recognize when the rule's economy has become misleading and recover enough of the discarded distinction space to proceed without it.

The same principle governs credit and responsibility. Present fluency can depend upon the invisible work of teachers, witnesses, investigators, authors, maintainers, and communities whose trajectories have been compressed into another agent's competence. The visible performer deserves credit for situated use and revision, but fluency alone does not establish sole authorship of the capacity being exercised. Likewise, responsibility for failure may be distributed across those who generated, selected, transmitted, constrained, and applied the policy.

None of this diminishes the receiver. To borrow intuition is to participate in cumulative knowledge. An agent capable only of personally earned compression would remain confined to the scale of one trajectory. Transmission allows competence to exceed biography.

The relevant contrast is therefore not between authentic experience and illegitimate inheritance. It is between competence whose history is understood and competence whose ease has concealed its dependencies. Provenance matters because a compressed policy carries less than the inquiry that produced it. It preserves an advantage while leaving some of its conditions behind.

Fluency is evidence of compression. It is not self-certifying evidence of experience. When an action arrives with the ease of intuition, the appropriate question is not only whether it works or whether the agent understands it. One may ask more quietly whose search made the ease

possible.

Appendices

A Summary of the Provenance Framework

The framework separates the source of a compressed competence from the conditions governing its acquisition. Provenance concerns which trajectories supplied the evidence that selected a policy. Formation concerns who controlled the receiver's exposure, actions, feedback, and objectives. Synthetic compression describes the subsequent transformation of inherited structure through a receiver's own discriminating experience.

Term	Defining criterion	Formal indication	Characteristic case
Earned compression	The policy is selected or stabilized partly through evidence generated within the receiver's own trajectory.	C_i depends materially upon τ_i .	A practitioner modifies a heuristic after observing discriminating failures.
Transmitted compression	The receiver acquires the practical economy of a policy without reproducing the exploration that selected it.	C_i depends materially upon τ_j while the relevant search is absent from τ_i .	An apprentice adopts a rule of thumb supplied by a master.
Self-directed formation	The receiver substantially controls the situations investigated, actions attempted, or hypotheses compared.	The receiver materially determines π_i , and possibly its exploratory objective.	A researcher chooses which alternatives to test.
Conditioned formation	Another agent or institution substantially controls exposure, feedback, objectives, or admissible actions.	An external process materially determines $\mathcal{E}_i, u_i$, or the constraints upon π_i .	A trainee learns through a prescribed curriculum and externally evaluated exercises.
Synthetic compression	The receiver transforms transmitted compression through later, discriminating, and persistent experience.	$C_i^{n+1} = S(C_i^n, \tau_i^+)$.	An inherited procedure is revised after repeated local testing.
Temporary conditioning	Present behavior changes because of active context, but the change does not persist beyond the episode.	χ_t changes while $\theta_{t+1} = \theta_t$ and $m_{t+1} = m_t$.	A model follows examples supplied only within the current prompt.
Retained experience	An encounter causally alters a persistent state available to later action.	$(\theta_{t+1}, m_{t+1}) \neq (\theta_t, m_t)$ because of o_t .	An agent stores an observed tool result in persistent memory.
Stabilized compression	Retained experience reorganizes later search or action economically across a class of situations.	A persistent update reduces or reorders later deliberation while preserving relevant performance.	Repeated feedback changes when an agent invokes a diagnostic test.

No row in the table defines an exclusive type of agent. A single competence can contain earned and transmitted components, arise through both self-directed and conditioned formation, and undergo several later cycles of synthetic revision. The framework describes a history of competence rather than an essence of its holder.

For a competence C_i , the minimum provenance record is

$$P_i(C) = (G(C), M_{G \rightarrow i}(C), V_i(C)),$$

where $G(C)$ identifies the generating trajectories, $M_{G \rightarrow i}(C)$ identifies the mechanism through which the compression reached agent i , and $V_i(C)$ identifies the receiver's subsequent validation history.

A fuller evaluation includes the environment in which the compression was generated, the objective under which it was selected, and the receiver's capacity to reopen it:

$$P_i^+(C) = (G(C), \mathcal{E}_G, O_G, M_{G \rightarrow i}(C), V_i(C), \Xi_i(C)).$$

Here $\mathcal{E}_G$ denotes the generating environment, O_G the originating objective or evaluative criterion, and $\Xi_i(C)$ the receiver's expansion capacity. This expanded record makes it possible to distinguish a policy that was reliable under its original conditions from one whose validity has been tested in the receiver's present environment.

The framework can be applied without exact quantification. A provenance analysis is already informative when it establishes that a policy was generated externally, transmitted through an institution, applied repeatedly without discriminating feedback, and supported by little recoverable evidence about its boundary conditions. Numerical estimates become useful only when the causal history and evaluation objective are sufficiently specified.

B Notation

Symbol	Meaning
$K_i(t)$	The distinctions available to agent i at time t .
$A_i(t)$	The subset of distinctions activated in the present situation.
$\pi_i(t)$	The policy governing search, comparison, or action.
$R_{i,t}$	The structures relegated to compressed or automatic operation.
τ_i	The experiential trajectory of agent i .
C_i	A compressed competence or policy exercised by agent i .
$G(C)$	The trajectories that materially generated compression C .
$M_{G \rightarrow i}(C)$	The mechanism through which C passed from its generating trajectories to receiver i .
$V_i(C)$	The receiver's post-acquisition validation history.
$\Xi_i(C, x)$	The receiver's capacity to expand, diagnose, or revise C in situation x .
$\Lambda_t(C)$	The recoverable provenance of institutional compression C at time t .
θ_t	The persistent parameters of a computational policy at time t .
m_t	The persistent external memory available to a computational system.
χ_t	The system's temporary active context.

The notation is intentionally neutral concerning consciousness and subjective experience. A trajectory records causally connected exposure, action, observation, and update. Whether a particular trajectory also constitutes phenomenal experience is a further question not settled by the framework.

C A Causal Measure of Borrowed Competence

This appendix gives a more explicit mathematical interpretation of provenance. The objective is not to assign an intrinsic numerical essence to intuition, but to separate the causal contributions of external and receiver-generated trajectories to an observable competence.

Let $\mathcal{X}$ be a space of situations, $\mathcal{Y}$ a space of actions, and

$$C : \mathcal{X} \rightarrow \Delta(\mathcal{Y})$$

a possibly stochastic policy, where $\Delta(\mathcal{Y})$ is the space of probability distributions over actions. Let μ be an evaluation distribution over $\mathcal{X}$, and let

$$u : \mathcal{X} \times \mathcal{Y} \rightarrow [0, 1]$$

be a bounded utility function. The value of policy C in environment μ is

$$V_\mu(C) = \mathbb{E}_{x \sim \mu} \mathbb{E}_{y \sim C(\cdot|x)} [u(x, y)].$$

Let H_E denote externally generated history and H_I denote history generated through the receiver's own interaction. A learning or formation operator

$$\mathcal{L} : \mathcal{H}_E \times \mathcal{H}_I \rightarrow \mathcal{C}$$

maps these histories to a policy

$$C_{E,I} = \mathcal{L}(H_E, H_I).$$

The notation does not imply that the two histories are independent. The receiver's later trajectory may depend strongly upon inherited distinctions, objectives, and policies. The separation records causal provenance rather than statistical independence.

Let $H_\emptyset$ be a baseline history representing the absence of a particular contribution. The policies

$$C_{\emptyset, \emptyset}, \quad C_{E, \emptyset}, \quad C_{\emptyset, I}, \quad C_{E, I}$$

represent counterfactual outcomes under different combinations of external and internal histories. These counterfactuals require a specified learning operator and baseline. Provenance measurement is therefore always relative to a formation mechanism.

Define a cooperative value function

$$v(S) = V_\mu(C_S), \quad S \subseteq \{E, I\},$$

where C_S is the policy formed using only the histories indexed by coalition S . The marginal contribution of external history depends upon whether internal history is already present:

$$\Delta_E^\emptyset = v(\{E\}) - v(\emptyset),$$

and

$$\Delta_E^I = v(\{E, I\}) - v(\{I\}).$$

Likewise, the marginal contribution of the receiver's own trajectory is

$$\Delta_I^\emptyset = v(\{I\}) - v(\emptyset),$$

and

$$\Delta_I^E = v(\{E, I\}) - v(\{E\}).$$

The interaction between transmitted and earned histories is

$$\mathcal{J}_{E,I} = v(\{E, I\}) - v(\{E\}) - v(\{I\}) + v(\emptyset).$$

If $\mathcal{J}_{E,I} > 0$, inherited and receiver-generated histories are complementary. The external compression makes later experience more productive, or later experience makes inherited compression more useful. If $\mathcal{J}_{E,I} < 0$, the two histories interfere. An inherited rule may inhibit discovery, or local experience may degrade a generally reliable transmitted policy.

A symmetric allocation of the interaction term is given by the two-source Shapley contributions

$$\phi_E = \frac{1}{2} [v(\{E\}) - v(\emptyset)] + \frac{1}{2} [v(\{E, I\}) - v(\{I\})],$$

and

$$\phi_I = \frac{1}{2} [v(\{I\}) - v(\emptyset)] + \frac{1}{2} [v(\{E, I\}) - v(\{E\})].$$

These satisfy

$$\phi_E + \phi_I = v(\{E, I\}) - v(\emptyset).$$

Under the monotonicity conditions

$$v(\{E, I\}) \geq v(\{E\}) \geq v(\emptyset)$$

and

$$v(\{E, I\}) \geq v(\{I\}) \geq v(\emptyset),$$

both contributions are nonnegative. A performance-relative borrowed share can then be defined as

$$B_\mu(C) = \frac{\phi_E}{\phi_E + \phi_I},$$

provided that

$$\phi_E + \phi_I > 0.$$

The corresponding earned share is

$$A_\mu(C) = \frac{\phi_I}{\phi_E + \phi_I},$$

so that

$$A_\mu(C) + B_\mu(C) = 1.$$

These quantities do not measure how many parameters, memories, propositions, or rules came from each source. They measure the relative causal contribution of external and receiver-generated histories to performance under a specified environment and utility function.

Proposition C.1 (Evaluation dependence). *The borrowed share $B_\mu(C)$ is not generally invariant under a change of evaluation environment or utility function.*

Proof. Each coalition value $v(S)$ is defined through $V_\mu(C_S)$ and therefore depends upon both μ and u . A change in either can alter the four coalition values by different amounts. Since ϕ_E and ϕ_I are functions of those values, their ratio need not remain constant. $\square$

This dependence is desirable. A transmitted diagnostic rule may account for most of a practitioner's competence in routine cases, while personally acquired experience accounts for most of their competence in unusual local cases. Provenance attribution is indexed to the competence being evaluated.

Performance is not the only relevant object. Two histories may generate policies with similar value but different behavior. Define the behavioral distance between policies C and C' under μ as

$$D_\mu(C, C') = \mathbb{E}_{x \sim \mu} [D_{\text{TV}}(C(\cdot | x), C'(\cdot | x))],$$

where

$$D_{\text{TV}}(P, Q) = \frac{1}{2} \sum_{y \in \mathcal{Y}} |P(y) - Q(y)|$$

for discrete $\mathcal{Y}$.

The external behavioral dependence of $C_{E,I}$ is

$$\mathfrak{B}_\mu = D_\mu(C_{E,I}, C_{\emptyset,I}),$$

while its internal behavioral dependence is

$$\mathfrak{A}_\mu = D_\mu(C_{E,I}, C_{E,\emptyset}).$$

A competence may have low external performance contribution but high external behavioral dependence. External training may strongly affect how the agent acts while leaving aggregate utility nearly unchanged. Conversely, two policies may differ little behaviorally over μ while possessing different latent capacities outside its support.

This yields a basic non-identifiability result.

Proposition C.2 (Behavioral non-identifiability of provenance). *Observed policy behavior does not uniquely determine the provenance of that policy.*

Proof. Let C^* be any policy in the range of two distinct formation operators. Construct histories (H_E, H_I) and (H'_E, H'_I) such that

$$\mathcal{L}(H_E, H_I) = C^*$$

and

$$\mathcal{L}'(H'_E, H'_I) = C^*,$$

while the first construction depends entirely upon external history and the second entirely upon receiver-generated history. The resulting policies agree for every $x \in \mathcal{X}$, yet their provenance differs. Therefore no inference from behavior alone can identify provenance without additional assumptions about the formation process. $\square$

The proposition is stronger than finite-sample underdetermination. Even complete extensional knowledge of a policy need not reveal the history through which it was acquired. Provenance is a causal property of formation, not a behavioral property of the finished mapping.

Formation conditioning can also be represented causally. Let Z be an external variable controlling curriculum, prompt selection, reward, or permitted action. Let

$$P(\tau_I \mid Z = z)$$

be the distribution of receiver trajectories under conditioning state z . Let z_0 denote a specified autonomous or minimally conditioned baseline. Define the degree of formation conditioning as

$$\kappa = \mathbb{E}_{z \sim P_Z} \left[D_{\text{KL}}(P(\tau_I \mid \text{do}(Z = z)) \parallel P(\tau_I \mid \text{do}(Z = z_0))) \right],$$

when the relevant distributions are absolutely continuous.

A large κ indicates that external intervention substantially changes which trajectories the receiver undergoes. It does not show that the resulting compression is transmitted. A highly conditioned curriculum may still produce earned compression if the receiver generates actions, observes discriminating outcomes, and stabilizes a policy from those outcomes.

For this reason,

$$\kappa \not\Rightarrow B_\mu(C)$$

and

$$B_\mu(C) \not\Rightarrow \kappa.$$

High conditioning and high borrowed dependence may occur together, but neither determines the other.

The distinction between temporary adaptation and persistent synthesis can be expressed through a state-transition system. Let

$$M_t = (\theta_t, m_t, \chi_t)$$

be the state of a computational or abstract learning system. Its transition under observation o_t is

$$M_{t+1} = \mathcal{U}(M_t, o_t).$$

Let

$$p(M_t) = (\theta_t, m_t)$$

project the full state onto its persistent components. An episode is transient when

$$p(M_{t+1}) = p(M_t),$$

even if

$$\chi_{t+1} \neq \chi_t.$$

It is retentive when

$$p(M_{t+1}) \neq p(M_t).$$

Retentive change becomes synthetic compression only when it alters a future policy over a nontrivial set of situations. For evaluation distribution μ , define

$$\Delta_{\text{syn}}(t) = D_{\mu}(C_{p(M_{t+1})}, C_{p(M_t)}).$$

A necessary behavioral condition for synthetic compression is

$$\Delta_{\text{syn}}(t) > 0.$$

This condition is not sufficient. Random corruption can also alter future policy. Let μ^+ be an evaluation environment relevant to the new experience. A stronger condition is

$$\Delta_{\text{syn}}(t) > 0$$

together with

$$V_{\mu^+}(C_{p(M_{t+1})}) > V_{\mu^+}(C_{p(M_t)}).$$

The update then produces a persistent behavioral difference with positive value under the specified environment. It remains possible for the update to improve μ^+ while degrading another environment, so synthetic compression is again objective-relative.

Synthetic updating does not necessarily reduce borrowed dependence.

Proposition C.3 (Non-monotonicity of borrowing under experience). *Additional receiver-generated experience can increase, decrease, or leave unchanged the borrowed share $B_{\mu}(C)$.*

Proof. Receiver-generated experience changes the coalition values containing I . If it independently improves performance, ϕ_I may increase relative to ϕ_E , reducing $B_{\mu}(C)$. If it improves the receiver's ability to exploit external compression, the complementarity term $\mathcal{J}_{E,I}$ may increase, raising the allocated external contribution. If it produces no policy-relevant change under μ , the coalition values and borrowed share may remain unchanged. $\square$

Experience therefore does not automatically convert borrowed competence into earned competence. It may instead deepen dependence by teaching the receiver to use inherited compression more effectively.

The environmental fragility of a competence can be bounded independently of provenance. Define expected loss

$$R_{\mu}(C) = \mathbb{E}_{x \sim \mu} \mathbb{E}_{y \sim C(\cdot|x)} [\ell(x, y)],$$

where

$$0 \leq \ell(x, y) \leq 1.$$

For two environments μ_0 and μ_1 ,

$$|R_{\mu_1}(C) - R_{\mu_0}(C)| \leq D_{\text{TV}}(\mu_1, \mu_0).$$

Proof. Define

$$g_C(x) = \mathbb{E}_{y \sim C(\cdot|x)}[\ell(x, y)].$$

Since $0 \leq g_C(x) \leq 1$, the variational characterization of total variation gives

$$|\mathbb{E}_{\mu_1}[g_C] - \mathbb{E}_{\mu_0}[g_C]| \leq D_{\text{TV}}(\mu_1, \mu_0).$$

□

This bound applies equally to earned and borrowed policies. Provenance matters because it affects what evidence is available about the relevant environments, not because transmitted compression obeys a different law of generalization. If the receiver does not know the generating environment μ_G , then the receiver may be unable to estimate

$$D_{\text{TV}}(\mu_{\text{use}}, \mu_G).$$

Provenance loss thus produces uncertainty about transfer rather than transfer failure by definition.

Let Ω be a random variable representing the generating source, environment, objective, or trajectory of a policy, and let $\mathcal{R}$ be the record available to its receiver. Provenance recoverability can be represented by

$$\mathcal{P}_{\text{rec}} = I(\Omega; \mathcal{R}),$$

the mutual information between the true provenance and the preserved record. If

$$I(\Omega; \mathcal{R}) = 0,$$

the record contains no information that discriminates among the modeled provenance possibilities. If

$$I(\Omega; \mathcal{R}) = H(\Omega),$$

the record determines the modeled provenance completely.

Compression can preserve policy value while reducing provenance recoverability. Let

$$T : \mathcal{EVIDENCE} \rightarrow \mathcal{GUIDANCE}$$

be an institutional or pedagogical compression operator. A successful transmission may satisfy

$$V_\mu(T(E)) \approx V_\mu(C_E)$$

while

$$I(\Omega; T(E)) \ll I(\Omega; E).$$

The transmitted guidance retains the practical advantage of the evidence while discarding information about where that advantage came from.

Expansion capacity can be formalized as a covering problem. Let $\mathcal{H}_{C,x}$ be the set of hypotheses relevant to diagnosing policy C in situation x , equipped with metric $d_{\mathcal{H}}$. Let $K_i(x) \subseteq \mathcal{H}_{C,x}$ be the hypotheses available to receiver i . Define the expansion deficit at resolution ε as

$$\mathfrak{D}_\varepsilon(i, C, x) = \inf \{n : \exists h_1, \dots, h_n \text{ such that } K_i(x) \cup \{h_1, \dots, h_n\} \text{ is an } \varepsilon\text{-cover of } \mathcal{H}_{C,x}\}.$$

If

$$\mathfrak{D}_\varepsilon(i, C, x) = 0,$$

the receiver's present distinction space already covers the relevant diagnostic hypotheses at resolution ε . A large value indicates that substantial expansion is required before the receiver can adequately reconstruct the problem space.

The definition is idealized because the full hypothesis space is rarely known. It nevertheless clarifies why identical compressed policies can support different forms of trust. Two agents may satisfy

$$C_i = C_j$$

while

$$\mathfrak{D}_\varepsilon(i, C, x) \ll \mathfrak{D}_\varepsilon(j, C, x).$$

They act identically under routine conditions, but agent i requires much less expansion to diagnose failure.

A provenance-sensitive trust functional can now be written as

$$W_i(C, \mu) = \alpha V_\mu(C) - \beta \mathbb{E}_{x \sim \mu} [\mathfrak{D}_\varepsilon(i, C, x)] + \gamma I(\Omega; \mathcal{R}_i) - \zeta U_i,$$

where $\alpha, \beta, \gamma, \zeta \geq 0$ are context-dependent weights and U_i represents unresolved uncertainty about the match between generating and deployment environments. This expression is not proposed as a universal decision rule. It shows that observed performance, expansion capacity, provenance recoverability, and environmental uncertainty are mathematically distinct inputs.

Finally, the complete history of a competence can be represented as a directed acyclic graph

$$\mathcal{G}_C = (N_C, E_C),$$

whose nodes include generating trajectories, records, aggregation procedures, training processes, persistent updates, institutions, and present actions. An edge

$$n_a \rightarrow n_b$$

indicates that intervention upon n_a can alter n_b under the causal model. Borrowed intuition is present whenever an external trajectory node has a directed causal path to present policy while the receiver has not reproduced the relevant evidential comparison through an internal path.

Let $\mathcal{T}_E$ be the external trajectory nodes, $\mathcal{T}_I$ the receiver-generated trajectory nodes, and C_i the present policy node. The external ancestral set is

$$\text{Anc}_E(C_i) = \text{Anc}(C_i) \cap \mathcal{T}_E,$$

while the internal ancestral set is

$$\text{Anc}_I(C_i) = \text{Anc}(C_i) \cap \mathcal{T}_I.$$

A purely transmitted limiting case satisfies

$$\text{Anc}_E(C_i) \neq \emptyset$$

and

$$\text{Anc}_I(C_i) = \emptyset.$$

A purely earned limiting case satisfies

$$\text{Anc}_E(C_i) = \emptyset$$

and

$$\text{Anc}_I(C_i) \neq \emptyset,$$

relative to the chosen causal boundary. A hybrid competence satisfies

$$\text{Anc}_E(C_i) \neq \emptyset$$

and

$$\text{Anc}_I(C_i) \neq \emptyset.$$

The qualification concerning the causal boundary is unavoidable. If language, conceptual inheritance, and prior training are included, the purely earned case may be empty for human agents. Its role is that of an analytic limit. The mathematical framework does not require epistemically self-created agents. It requires only that external and receiver-generated causal contributions can sometimes be distinguished relative to a specified level of analysis.

The principal result of the appendix is therefore negative but useful:

performance $\nRightarrow$ provenance,

adaptation $\nRightarrow$ persistence,

persistence $\nRightarrow$ compression,

and

experience $\nRightarrow$ independence from inherited structure.

A rigorous analysis of borrowed intuition must specify the learning operator, causal boundary, evaluation environment, persistent state, and counterfactual baseline. Once these are fixed, the extent of borrowing becomes an empirical property of a competence's formation rather than a metaphorical judgment about the authenticity of its holder.

References

- [1] Tom B. Brown et al. Language Models Are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901, 2020.
- [2] C. A. J. Coady. *Testimony: A Philosophical Study*. Clarendon Press, Oxford, 1992.
- [3] Miranda Fricker. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, Oxford, 2007.
- [4] Long Ouyang et al. Training Language Models to Follow Instructions with Human Feedback. In *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744, 2022.
- [5] Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Eric Hambro, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. Toolformer: Language Models Can Teach Themselves to Use Tools. In *Advances in Neural Information Processing Systems*, volume 36, pages 68539–68551, 2023.
- [6] Tongtong Wu, Massimo Caccia, Zhuang Li, Yuan-Fang Li, Guilin Qi, and Gholamreza Haffari. Continual Learning for Large Language Models: A Survey. *arXiv preprint arXiv:2402.01364*, 2024.