Crossing the Threshold: A Falsifiable Definition of Symbolic Stabilization

Anthony Vondoom 0009-0003-4953-1427
Independent Researcher, Cognitive Archaeology

Abstract
Cognitive archaeology has repeatedly gestured toward the durable, carrier-independent fixity of shared symbolic meaning — under labels like external symbolic storage, behavioral-modernity trait lists, and material engagement — without isolating it as a definable property in its own right. This paper gives that property, symbolic stabilization, a definition precise enough to survive independent testing: a symbolic system S is stabilized when it jointly satisfies intergenerational repetition, spatial constraint, cross-media redundancy, and persistence under stress — a threshold condition, not a matter of degree. The definition is distinguished from three constructs it is routinely confused with: Donald's external symbolic storage, Donald's mimetic-mythic-theoretic staging, and Clark and Chalmers' extended mind, none of which specify the semantic-fixity condition stabilization requires. An unsupported assertion in prior work — that pre-stabilization symbolic cognition was more fluid — is replaced with a mechanism: symbolic convergence is driven by environmental unpredictability acting through information-exchange networks, not by a universal cognitive capacity, which is why it is contingent rather than inevitable, and why it differs in kind from the tool-standardization proxy, which has not held up as a modernity marker under direct testing. This mechanism, and the Upper Palaeolithic/Middle Stone Age contrast offered in its support, are proposed and motivated rather than established; the paper's definitional and differentiating claims do not depend on them. The paper draws no conclusions about writing, institutions, or population thresholds; those are separate problems, treated in Vondoom (2026), and this paper does not conflate them with the one problem it solves.

Keywords: symbolic stabilization; cognitive archaeology; external symbolic storage; extended mind; information exchange; Deep Symbolic Systems Model

1. Introduction

Symbolic stabilization is a term that has circulated in cognitive archaeology without a precise, freestanding definition. This paper proposes one: symbolic stabilization as a definable, falsifiable structural condition, distinct from the constructs it most resembles, with a stated mechanism for why some symbolic systems reach it and others do not.

The paper is organized around three commitments.

First, it addresses a single question. It does not ask why writing emerged, why formal institutions appeared, or what population size triggers organizational complexity. Those are downstream, separable questions that the author takes up elsewhere; each tracks a different kind of change, with its own evidentiary requirements, and treating them as facets of one problem would leave it unclear, for any single claim in this paper, what evidence would count for or against it. This paper does not attempt those broader questions here.

Second, it differentiates stabilization from the constructs most likely to be mistaken for it: Merlin Donald's external symbolic storage, Donald's mimetic-mythic-theoretic developmental staging, and Andy Clark and David Chalmers' extended mind thesis. Each shares surface features with stabilization. None specifies the property stabilization is actually about: whether a symbolic system's meaning has become fixed and transmissible independent of any individual carrier.

Third, it grounds its central empirical claim — that pre-stabilization symbolic systems are measurably more fluid than post-stabilization ones — in a stated mechanism rather than an assumption. That mechanism is environmental unpredictability acting through information-exchange networks. It is presented here as a testable prediction with a named dataset against which it can be run, not as a settled finding. Where the underlying comparison has not yet been quantified, this paper says so directly, rather than implying more certainty than the evidence currently supports.

This distinction matters beyond terminology. Cognitive archaeology currently lacks a shared, checkable criterion for when a symbolic practice has become genuinely stabilized, which means claims about symbolic revolutions or behavioral modernity are routinely made and contested without a common yardstick for what would count as evidence either way. A definition precise enough to be applied consistently by different researchers to different regional records — rather than argued about case by case — makes systematic cross-regional and cross-period comparison possible in a way it currently is not.

2. Formal Definition of Symbolic Stabilization

2.1 How the Field Has Talked About Stabilization Without Naming It

Cognitive archaeology has repeatedly circled the property this paper isolates — durable, carrier-independent symbolic fixity — under several different names, without treating it as a separable, gradeable structural property in its own right. McBrearty and Brooks' influential critique of 'behavioral modernity' framed the concept substantially as a trait list, in which symbolic

durability appears as one item to be checked off alongside others, rather than as a property with its own criteria and its own conditions of failure (McBrearty and Brooks 2000). On that kind of account, a symbolic system either counts as evidence of modernity or does not, with no mechanism offered for why some symbolic material crosses that line while other, contemporaneous material does not. Donald's external symbolic storage (1991) comes closer, describing a shift from individually carried representation to externally storable symbol systems, but as Section 3.1 argues in detail, ESS is a claim about storage capacity, not about whether the meaning stored has stabilized into a fixed, shared reference independent of its carriers. The debate over stone-tool standardization as a proxy for a 'clearer mental template' (Monnier and McNulty 2010, discussed at length in Section 4.1, and refuted on empirical grounds by Marks, Hietala and Williams 2001) is itself a prior instance of the field reaching for something like a stabilization signature and settling on the wrong operationalization: testing for a fixed physical target rather than a fixed shared meaning, and finding that the proxy did not discriminate between the two industries. Material engagement theory (Malafouris 2013) supplies a rich account of how meaning becomes inseparable from material practice at the point of use, but as Section 3.4 shows, it does not by itself address whether that constitutive relationship persists, spreads, or survives its originator across generations. Across these three literatures, then, something resembling stabilization has repeatedly been gestured toward — as a trait to check off, as a storage capacity, as a constitutive engagement at the point of use — without ever being isolated, defined, and given criteria of its own. This paper names and formally specifies that property.

2.2 The property in question

Existing accounts of early symbolic behavior document that meaning-bearing marks, objects, and practices exist. They do less well at specifying when a symbolic practice has crossed from something a given individual or small group does, to something that persists as a shared, fixed reference independent of any one carrier. That crossing is what stabilization names. It is not a claim about the presence of symbols — Middle Stone Age ochre engravings and Aurignacian ivory carvings are both, uncontroversially, symbolic. It is a claim about a specific structural property some symbolic systems acquire and others do not: independence from the individuals who produced them.

2.3 Definition

A symbolic system S is stabilized if and only if it jointly satisfies:

Criterion 1 — Intergenerational Repetition: S reproduces reliably across at least three generations, independent of any specific individual carrier.

Criterion 2 — Spatial Constraint: S imposes recurrent, non-arbitrary structure on space, movement, or behavior.

Criterion 3 — Cross-Media Redundancy: S is expressed in at least two distinct media (material, ritual, landscape, linguistic).

Criterion 4 — Persistence Under Stress: S retains recognizable form across ecological disruption, demographic pressure, or political reorganization.

These four criteria are not an arbitrary checklist; each targets a distinct way a symbolic system can remain dependent on its originating individuals rather than becoming carrier-independent, and the four together are intended to close off the main routes by which that dependence could persist undetected. A system could be carried by a single lineage rather than a population (addressed by Criterion 1); it could be behaviorally optional, imposing no real constraint on conduct, and therefore trivially easy to abandon (addressed by Criterion 2); it could exist in only one medium, such that the loss of a single practice, object type, or context erases it entirely (addressed by Criterion 3); or it could be a fair-weather convention that survives only because it has never been tested by disruption (addressed by Criterion 4). A system that fails any one of these is dependent on some contingency — a lineage, a single behavioral setting, a single medium, or an undisturbed environment — that stabilization is specifically meant to rule out. This is a design rationale, not a proof that exactly four criteria are necessary and sufficient; a reviewer who identifies a fifth route to carrier-dependence not covered by Criteria 1–4 would identify a genuine gap in the definition, and that possibility is left open rather than foreclosed.

These four criteria are conjunctive, not additive. A symbolic system that satisfies three of four is not seventy-five percent stabilized — it has not stabilized. This is deliberate. Stabilization is modeled here as a phase transition, structurally closer to a percolation threshold than to a running average: below the threshold, a symbolic system remains dependent on the specific individuals and contexts that sustain it, and can be lost, reinterpreted, or replaced without resistance; above it, the system persists and constrains behavior regardless of which individuals currently carry it. The conjunctive requirement is what makes the criteria falsifiable rather than merely descriptive — a system can fail stabilization cleanly, on a named criterion, rather than occupying an ambiguous middle ground that absorbs any counterexample.

2.4 Variables

Analysis under this definition operates on four variables: the symbolic system S under examination; the substrate M in which S is encoded (material, oral, ritual, spatial); the temporal depth T over which continuity is documented; and the stress events E across which S has been tested. A stabilization score can be reported as a 0–4 tally of criteria met, but the score is a descriptive summary of the analysis, not the analysis itself — 4/4 is the only value that licenses the term stabilized.

2.5 What stabilization is not

Stabilization is not synonymous with cultural elaboration, artistic skill, or symbolic density. A single, highly skilled engraving with no successors satisfies none of the four criteria and is not stabilized, regardless of its technical sophistication. Conversely, a comparatively simple mark repeated across generations, across sites, and across media, surviving demographic collapse, is stabilized. The definition intentionally decouples technical or aesthetic quality from structural persistence, because the two are answers to different questions.

2.6 Relationship to Prior Published Use of This Term

A related, continuous measure of symbolic stabilization was applied in an earlier study of early Egyptian symbolic material (Vondoom 2026), which scored stabilization using declining coefficient of variation in artifact typologies across stratigraphic phases, with a threshold of $CV < 0.15$ taken to indicate completion. That usage and the definition given here are related but not identical, and the relationship should be stated plainly rather than left for a reader to reconcile. The CV-based measure is an evidentiary proxy for one component of stabilization as defined in this paper: declining typological variation is a plausible material signature of Criterion 1 (intergenerational repetition) and, where the declining CV is expressed across more than one artifact class or medium, of Criterion 3 (cross-media redundancy). It is not, by itself, a full test of stabilization in the sense defined here, because a single continuous measure cannot on its own establish spatial constraint or persistence under stress. That earlier, continuous measure should accordingly be read as reporting partial evidence toward the full four-criteria threshold defined in this paper, not as having already applied that threshold. Future work applying this framework should score all four criteria explicitly rather than treating the CV-based observable as a stand-in for stabilization as a whole.

3. Differentiation from Existing Constructs

3.1 External symbolic storage (Donald 1991)

Donald's external symbolic storage (ESS) describes the offloading of memory onto artifacts external to the brain — notches, tallies, images — that extend the storage capacity of a cognitive system beyond what any individual's memory can hold (Donald 1991). ESS is a claim about storage capacity. It is silent on whether the meaning attached to what is stored is fixed. A tally stick is external symbolic storage whether or not the specific meaning of its notches is agreed upon by anyone beyond its maker; a stored symbol can be freely reinterpreted by each subsequent user without ceasing to be an instance of ESS. Stabilization, by contrast, is a claim about semantic fixity — whether the meaning itself, not merely the physical trace, persists independent of any one interpreter. A system can exhibit ESS without ever stabilizing, and it is common in the archaeological record to see externalized marks, including the Middle Stone Age engraved ochre corpus, that show no evidence of a shared, cross-site interpretive convention. ESS is therefore a precondition stabilization can build on, not a synonym for it.

3.2 Mimetic-mythic-theoretic staging (Donald 1991)

Donald's broader account proposes a sequence of cultural stages — mimetic, mythic, theoretic — through which a culture as a whole develops increasingly abstract representational capacities. This is a diachronic, culture-level model: a population moves through the stages more or less as a unit, and the stages are ordered. Stabilization, as defined here, operates at a different grain. It is a synchronic property of individual symbolic systems, not of cultures as wholes. A single population can carry multiple symbolic systems simultaneously that sit at different stabilization states — a burial practice that is fully stabilized alongside an ornamentation style that is not. Donald's staging cannot represent this within-culture heterogeneity, because it assigns a single developmental stage to the culture as a whole. Stabilization is a finer-grained instrument, applicable system by system, and it makes no developmental-sequence claim: a symbolic system can stabilize or fail to stabilize in a population regardless of where that population would sit on Donald's mimetic-mythic-theoretic scale. Donald's staging model is nonetheless the source of the developmental intuition stabilization refines: that meaning-bearing practices become progressively less tied to individual carriers over evolutionary time. Stabilization takes that intuition and converts it from a stage a culture occupies into a property individual systems either do or do not possess, testable independently of any claim about where a population sits developmentally.

3.3 Extended mind (Clark and Chalmers 1998)

Clark and Chalmers' extended mind thesis holds that an object counts as part of a cognitive system if it plays the right functional role — if it is poised for use in the way an internal memory would be, regardless of its physical location (Clark and Chalmers 1998). This is a parity criterion: functional equivalence to an internal process is sufficient for cognitive extension. A shopping list, by this criterion, is genuinely part of its user's cognitive system while it is being consulted. Stabilization adds a requirement extended mind does not need: durability and fidelity across carriers, generations, and disruption. A shopping list satisfies extended mind's parity criterion completely and satisfies none of stabilization's four criteria — it does not survive its author, does not recur across generations, and is not expressed cross-media. The two frameworks are not competitors; they answer different questions. Extended mind asks whether an external object is functioning as cognition right now, for a given user. Stabilization asks whether a symbolic system has become structurally independent of any given user at all. A system can

be extended-mind-active without ever stabilizing, and a stabilized system typically was extended-mind-active for some individual before it stabilized. The relationship is genetic, not definitional — one is not a special case of the other. This diagnosis converges with an independent critique of extended mind from within philosophy of cognitive science: Sterelny (2010) argues that most genuine extended-mind cases are better understood as limiting instances of a broader class of environmental scaffolding, precisely because they typically lack the durability, generational transmission, and individual-independence that distinguish a fully scaffolded cognitive resource from a merely extended one. Stabilization's four criteria can be read as naming, for the specific case of shared symbolic systems, the durability dimension Sterelny identifies as doing the real explanatory work that the parity criterion alone cannot supply.

3.4 Material engagement theory (Malafouris)

Material engagement theory (MET) holds that the relationship between material culture and cognition is constitutive rather than representational: objects do not merely encode pre-formed symbolic content but actively participate in and transform the cognitive processes that use them (Malafouris 2013). MET is the closest of the four constructs treated in this paper to stabilization as defined here, because both reject a picture in which meaning is fixed privately and then merely expressed externally. MET's own literature has explicitly worked through its relationship to extended mind (Malafouris and Renfrew 2010), which strengthens rather than undermines the case for triangulating carefully between the three frameworks here rather than treating any pair as interchangeable. The distinction is that MET is a claim about the ontological status of the object-cognition relationship at the point of use — what kind of thing a material engagement is — while stabilization is a claim about the population-level persistence of that engagement once established. MET can be fully satisfied by a single individual's single act of constitutive engagement with an object and says nothing about whether that engagement recurs, spreads, or survives its originator. Stabilization presupposes something MET-like has occurred (a symbolic system must be genuinely constitutive of practice for the four criteria to apply meaningfully at all) but adds the independent, testable question of whether that constitutive relationship has become structurally durable across carriers. The two are compatible and largely non-competing: MET describes the mechanism by which a symbolic system engages cognition at all; stabilization describes what happens to that engagement over generational and demographic time. This paper does not treat MET as a rival explanation requiring refutation; the four criteria proposed here are a downstream addition to MET rather than a replacement for it, extending an approach the author has applied elsewhere to early Egyptian symbolic material (Vondoom 2026).

4. The Mechanism: Why Symbolic Systems Converge

4.1 A prior proxy: tool-form standardization

The most obvious candidate mechanism for detecting a fluid-to-fixed transition is convergence — a measurable narrowing of variation in symbolic form over time and across sites. This is precisely the logic that was applied to stone tool morphology under the 'clearer mental template' hypothesis: the claim that anatomically modern humans, unlike Neanderthals, produced more standardized tool forms because they held a more fixed mental template of the intended shape (Monnier and McNulty 2010). Direct testing of this claim on Upper Palaeolithic burin assemblages against Middle Palaeolithic comparanda found no significant standardization advantage for the supposedly more modern industries; both were comparably variable once raw material, reduction sequence, and function were controlled for (Marks, Hietala and Williams 2001). The standardization proxy, applied to stone tools, was not supported by quantitative testing.

This is not a minor caveat. It means convergence toward a fixed form cannot be adopted as a stabilization signature without independently answering the question that already sank it once: convergence toward what, and driven by what mechanism? Stone tools converge, when they do, because they are built against a physical target — hafting geometry, cutting angle, haptic fit to the hand — that is the same target for every knapper regardless of culture, belief, or diet. Convergence under a shared physical constraint is expected and uninformative about symbolic cognition specifically; it would occur even absent any symbolic system at all.

4.2 The mechanism proposed here: environmental unpredictability and information exchange

Symbolic convergence, when it happens, cannot be explained the same way, because there is no equivalent universal physical target for what a shared sign, motif, or practice should look like. There is no haptic or biomechanical reason for two unrelated groups to converge on the same mark. Where convergence in symbolic material is observed, the archaeological literature on hunter-gatherer social networks offers a specific, independently motivated mechanism: information exchange as a risk-buffering strategy against environmental unpredictability (Wobst 1977; Gamble 1982; Whallon 2006; Fitzhugh, Phillips, and Gjesfjeld 2011). On this account, populations facing unpredictable, patchy, or seasonally volatile resource conditions maintain long-range social networks — and the shared, legible symbolic material that sustains those networks — because the networks function as insurance against local resource failure. Populations in ecologically stable, low-variance environments face weaker selective pressure to build or maintain such networks, because there is less to insure against.

This gives stabilization a mechanism that is the mirror image of the tool-standardization proxy, not a restatement of it. Tool form converges because the physical target is universal and the pressure to converge is constant across all populations, which

is exactly why it is a poor discriminator. Symbolic convergence, on this account, is contingent on a variable — environmental unpredictability — that differs sharply across populations, which is exactly what makes its presence or absence a potentially useful discriminator rather than a foregone conclusion. A population under stable, low-variance conditions is predicted, on this account, to show persistent local idiosyncrasy in its symbolic material even over long durations and even at large population sizes, because it has not faced the adaptive pressure that produces cross-group symbolic convergence. This is the sense in which pre-stabilization symbolic cognition is proposed to be fluid: not fluid in the individual mind, but fluid at the population level, in the absence of the specific pressure that fixes a shared form across carriers.

4.3 Rival Mechanisms

A different account of geometric sign convergence must be addressed directly, the first of three considered in this section, because it proposes a second universal target of exactly the kind Section 4.1 argues symbols lack. The neuropsychological model developed by Lewis-Williams and Dowson (1988) holds that a recurring set of geometric forms — grids, dots, zigzags, catenary curves, and related patterns — are generated by the structure of the human visual nervous system itself during altered states of consciousness, independent of culture or environment, and that their recurrence across unrelated rock art traditions reflects shared neurology rather than shared information networks. If correct, this would undercut the argument of this section: sign convergence would have a universal biological target after all, exactly parallel to the haptic target that makes tool convergence uninformative, and its presence would say nothing about environmental unpredictability or information exchange.

Three considerations limit how much weight this model can bear against the argument advanced here, without requiring its outright rejection. First, the entoptic model predicts recurrence of a small set of general shape categories — six broad classes — wherever altered-state imagery is recorded, essentially independent of period or region. It does not, on its own, predict the specific pattern this paper relies on: convergence on a stable, bounded repertoire of approximately thirty specific sign types, shared across an entire continent for thirty thousand years, with documented persistence and disappearance of individual signs at specific times and places (Section 5.1). Shared neurology could explain why grids and zigzags appear repeatedly across unrelated traditions worldwide; it does not by itself explain why the same thirty-odd signs, rather than some other sampling from the neurologically available set, persist together as a bounded regional-to-continental repertoire for a specific thirty-thousand-year window. Second, the entoptic model has drawn sustained methodological criticism within rock art research itself (Bahn 1988; Helvenston, Bahn, Bradshaw, and Chippindale 2003), including the objection that it lacks demonstrated universality across the traditions it is meant to explain and that the conditions under which Palaeolithic populations would have reliably entered the relevant altered states are not independently established; a parallel critique of the shamanic framework within which the entoptic model is typically embedded argues that its supporting ethnographic analogies have not been adequately tested against the archaeological record they are applied to (McCall 2007). The specific empirical prediction the entoptic model would need to satisfy against the Upper Palaeolithic corpus itself — that geometric-sign frequency should track known entoptic categories — has begun to be tested directly, rather than assumed (von Petzinger 2017a). Third, and most importantly, the two mechanisms are not mutually exclusive and this paper does not need to choose between them to make its case: entoptic phenomena could supply a raw perceptual repertoire of candidate forms while information-exchange pressure determines which subset of that repertoire becomes a shared, stabilized, cross-site convention in a given population under given ecological conditions. On that reading, entoptic imagery explains where candidate sign forms come from; information exchange explains why some populations converge on a shared, bounded selection from that pool and others do not. This paper's claim is specifically about the latter question, and it is stated here as an explicit scope limitation rather than folded silently into Section 4.2: the mechanism proposed in this paper explains differential convergence across populations, not the ultimate origin of geometric form as such.

A second alternative mechanism must be addressed for the same reason: it too could produce sign convergence without invoking environmental unpredictability. Conformist transmission — the well-established finding that individuals disproportionately copy the majority variant present in their group, rather than sampling variants in proportion to their frequency (Henrich and Boyd 1998) — could in principle generate a shared, bounded repertoire purely through within-group copying bias, independent of any adaptive pressure tied to resource volatility. Three considerations limit how much weight this alternative can bear against the mechanism proposed here. First, conformist transmission's own foundational literature shows it favoring the maintenance of between-group differences, not their erosion (Boyd and Richerson 1985; Henrich and Boyd 1998): because conformist bias operates by amplifying whichever variant is already locally common, its predicted signature is persistent regional distinctiveness sustained by within-group homogenization, not the specific pattern this paper explains — convergence of the same repertoire across unrelated, geographically dispersed populations. Conformist bias, on its own terms, predicts close to the opposite large-scale pattern from the one documented in Section 5.1. This prediction has since been confirmed experimentally: when conformist and non-conformist learning strategies are directly compared, conformist transmission measurably increases within-group homogeneity and between-group divergence, rather than producing convergence across unrelated populations (Efferson et al. 2008). Second, the conditions under which conformist transmission itself evolves, and its strength relative to other learning strategies, have been extensively debated within cultural evolution research, and its effects are demonstrably sensitive to migration rate, group structure, and the frequency of environmental

change — none of which are held constant across the populations compared here. Third, the two mechanisms are not mutually exclusive: conformist transmission could operate within a population once a convention has begun to spread, stabilizing a variant that information-exchange pressure first established across population boundaries. On that reading, conformist bias is a plausible within-group amplifier of a pattern that environmental unpredictability initiates at the between-group scale, not a competing explanation for the phenomenon this paper addresses.

A third alternative, cultural attractor theory, holds that some representations are inherently more stable and transmissible than others because they align with pre-existing cognitive biases — being easier to remember, reconstruct, or process — and that repeated transmission pulls variants toward these cognitively privileged points regardless of the social or ecological context in which transmission occurs (Sperber 1996). If a subset of the roughly thirty Upper Palaeolithic sign types is simply more cognitively attractive than the alternatives, convergence on that subset would require no appeal to environmental unpredictability at all. Three considerations apply here as well. First, cultural attractor theory predicts convergence on cognitively privileged forms wherever transmission occurs with sufficient fidelity, essentially independent of population-level ecological variation; it does not, on its own, predict why the same bounded repertoire should be confined to a specific thirty-thousand-year Upper Palaeolithic window rather than recurring wherever human transmission has since occurred. Second, cultural attractor theory's own proponents have acknowledged that transmission fidelity itself varies with social and demographic context (Claidière and Sperber 2007), which reintroduces exactly the population-level variables this paper's mechanism addresses rather than eliminating the need for them. Third, the two accounts are compatible rather than competing: cognitive attractors could determine which candidate forms are available for adoption at all, while information-exchange pressure under environmental unpredictability determines whether a population's social network is dense and extensive enough for that attraction to result in cross-site convergence rather than isolated, locally recurring use. This paper's claim concerns the latter question and does not require resolving whether cognitive attraction also plays a role in the former.

4.4 Operationalizing environmental unpredictability

The mechanism proposed here is only as testable as its central variable is measurable, and this must be stated explicitly rather than left implicit. Environmental unpredictability, for the purposes of this framework, refers to interannual and decadal variance in the availability of key resources — not average resource abundance, but the unpredictability of its timing and location. Three proxy classes are available for estimating it in deep-time contexts, none of them novel to this paper: paleoclimate proxies (isotopic and sedimentary records of rainfall and temperature variance, rather than mean values, over the relevant period; cf. Roberts et al. 2016 for an application to the Middle Stone Age itself); faunal and botanical assemblage diversity and turnover, which register the practical consequences of resource patchiness in the record a population actually depended on; and, where dietary isotope data exist, individual-level dietary breadth and its variance across a population, which indexes how much a population's subsistence base fluctuated within a lifetime rather than only across geological time. None of these proxies is available at the resolution or coverage needed to score the Middle Stone Age and Upper Palaeolithic comparison in Section 5 at present, and this paper does not claim otherwise. Proposing a mechanism without specifying how it would be measured would repeat, in a new variable, the same operationalization gap this paper's Section 6.3 already acknowledges for other parts of the framework; naming the candidate proxies here, and naming their current insufficiency, is intended to prevent that gap from being inherited silently.

4.5 What this does and does not claim

This mechanism generates a directional prediction — populations under higher environmental unpredictability should show earlier and more extensive symbolic convergence, independent of population size — but it does not, by itself, establish that any specific archaeological transition instantiates it. Testing that prediction directly would require measuring environmental unpredictability itself, which, as the previous subsection makes clear, is not yet possible at the resolution this comparison needs. What the next section provides is the other half of the correlation: independent documentation of where symbolic convergence itself is, and is not, observed. That falls short of a full test of the mechanism, and this paper says so rather than presenting it as one.

5. A Motivating Case: Convergence Patterns and Their Current Evidentiary Limits

This section is not a completed empirical test of the mechanism proposed in Section 4, and it should not be read as one. It presents the best currently published evidence bearing on the predicted contrast between the Upper Palaeolithic and Middle Stone Age records, using that evidence to illustrate what the mechanism predicts and to motivate a specific future test. As Sections 5.3 and 6.3 make explicit, two conditions required for a completed test are not yet met: the Middle Stone Age record cannot currently be scored on two of the four stabilization criteria (Table 1), and the paleoclimate proxy data needed to measure environmental unpredictability directly at these sites do not yet exist at the necessary resolution (Section 4.4). The paper's contribution does not depend on this section succeeding as a test; it depends on the definition (Section 2), the differentiation from adjacent constructs (Section 3), and the falsifiability criteria (Section 6.2) standing on their own. What follows should be read accordingly, as a worked illustration with a named path to becoming a real test, not as that test itself.

5.1 The convergence case: Upper Palaeolithic sign systems

The strongest documented case of cross-site symbolic convergence in the Pleistocene record is the geometric sign repertoire of the European Upper Palaeolithic. The SignBase database, compiled from over five hundred Aurignacian-period mobile objects across Europe and the Near East, identifies approximately thirty recurring sign types accounting for the great majority of non-figurative marks in that corpus (Dutkiewicz et al. 2020). Independently, von Petzinger's fieldwork documenting parietal signs across fifty-two cave sites in France, Spain, Portugal, and Italy reports a comparably small, repeating set of thirty-two forms in use across the continent for roughly thirty thousand years, with most sign types remaining in continuous use across that entire span (von Petzinger 2017b). Two independently constructed corpora — mobile objects and cave walls — converge on the same finding: a small, shared repertoire, repeated across enormous geographic and temporal distance. On the criteria set out in Section 2, this record is a strong candidate for stabilization: cross-media redundancy, spatial constraint, persistence under stress across multiple glacial-interglacial cycles and documented population turnovers, and intergenerational repetition across thirty thousand years are all independently documented.

The sign counts themselves are not a single analyst's unchecked judgment call, which matters because a bounded repertoire is only informative if the boundary is reliable. Dutkiewicz and colleagues tested inter-coder reliability directly, having independent coders classify a subsample of SignBase entries and reporting agreement rates and Cohen's Kappa scores against the original classifications, rather than presenting the thirty-type scheme as self-evidently correct (Dutkiewicz et al. 2020). Von Petzinger's fieldwork points in a similar direction from the opposite side: her survey documented previously unrecorded instances of already-catalogued sign types at roughly three-quarters of the sites she visited, which, if it moves the count at all, points toward local under-cataloguing rather than toward an inflated or cherry-picked repertoire. Neither fact proves the thirty-odd-sign boundary is exactly right, and this paper does not claim that precision; it claims that the boundary is checked rather than assumed, which is what Criterion 1 requires of a bounded, recurring set.

This convergence is corroborated by an independent artefact class within the same tradition. A recent continent-wide catalogue of Upper Palaeolithic hand stencils records 775 individual representations across 59 caves and shelters in France, Spain (including one site in Gibraltar), Italy, and the United Kingdom, concentrated in two areas — the Cantabrian coast and the Dordogne — but present in smaller numbers well beyond either (Fernández-Navarro, Spaey, and Garate 2026). Direct Uranium-Thorium dating places individual hand stencils across a span exceeding thirty thousand years, from Gravettian-period examples at Cosquer Cave (approximately 23,000–28,000 years before present) to a minimum age exceeding 33,000 years at La Garma and, subject to ongoing methodological debate over the dating technique, potentially exceeding 37,000 years at El Castillo, with the corpus as a whole persisting through the Last Glacial Maximum. Within this corpus, statistical analysis of finger-combination patterns identifies regionally patterned but recurring configurations — a small number of dominant combinations account for the great majority of hands within each region, with the specific dominant combination differing by region — a pattern the study's authors describe as consistent with, though not proof of, a shared but geographically differentiated visual convention (Fernández-Navarro, Spaey, and Garate 2026). As a distinct medium co-occurring with the geometric sign repertoire discussed above, this corpus offers independent support for Criterion 3 within the same tradition, and its multi-millennial persistence across a documented climatic disruption offers a second, independently sourced line of evidence for Criterion 4.

5.2 The comparison case: Middle Stone Age engravings

The African Middle Stone Age (MSA) engraved ochre and bone corpus — Blombos Cave, Diepkloof, Klasies River — dated from roughly 100,000 to 70,000 years before present, is the primary comparison case (d'Errico, Henshilwood, and Nilssen 2001; Henshilwood et al. 2002; Henshilwood, d'Errico, and Watts 2009). This record demonstrates deliberate, controlled, repeated engraving activity at a single site over a documented span of at least twenty-five thousand years (Henshilwood and d'Errico 2011). Published descriptions of the Blombos material characterize individual pieces as variations on a small number of related motifs rather than as instances of a repertoire explicitly documented as replicated across unrelated sites at the scale reported for the Aurignacian.

That characterization deserves to be unpacked rather than taken on faith, since a specialist reader would rightly want to know what "a small number of related motifs" is actually doing here. The published catalogue of Blombos engravings itself distinguishes at least four pattern categories — cross-hatched designs, dendritic shapes, parallel lines, and right-angled juxtapositions (Henshilwood et al. 2009) — and whether those four count as variations on one underlying motif or as several distinct ones is a classification judgment, not a fact read directly off the material. This paper follows the excavators' own characterization rather than proposing an independent one. That is a defensible choice, but it is a choice: a reader who weighs the four categories as more distinct than related would score this record even more conservatively than Table 1 already does, which would strengthen rather than undermine the contrast this paper draws between the two records.

This comparison must be stated carefully, because the two records as currently published are not measuring the same thing. The Upper Palaeolithic evidence in Section 5.1 is explicitly cross-site: the same sign types recur across dozens of caves and hundreds of mobile objects spread over a continent. The Middle Stone Age evidence available here is substantially within-site: it establishes long-duration, controlled production at Blombos specifically, with comparison to a small number of other MSA localities, not a systematic multi-site survey of the scale SignBase and von Petzinger's cave survey represent for the

Aurignacian. The absence of documented cross-site MSA convergence is therefore not yet evidence of absence — it may simply reflect that the equivalent multi-site survey has not been done at comparable scale, a point developed further in Section 5.3.

Table 1. Scoring of the four stabilization criteria (Section 2.3) against the two records discussed in Section 5. The Middle Stone Age column is not scoreable as 4/4 or as failing outright; two criteria remain untested rather than unmet, which is precisely the evidentiary gap Section 5.3 identifies and Section 5.3 proposes to close via a direct SignBase-based comparison.

Criterion	Upper Palaeolithic sign repertoire	Middle Stone Age engravings
1. Intergenerational repetition	Documented: majority of sign types in continuous use across ~30,000 years (von Petzinger 2017b; Dutkiewicz et al. 2020)	Documented at Blombos alone: repeated motif production over ~25,000 years (Henshilwood & d'Errico 2011)
2. Spatial constraint	Documented: recurrent placement patterns across cave and mobile-object contexts (von Petzinger 2017b)	Not systematically assessed at comparable multi-site scale (Section 5.3)
3. Cross-media redundancy	Documented: same repertoire independently confirmed on mobile objects (Dutkiewicz et al. 2020) and cave walls (von Petzinger 2017b)	Documented across ochre and bone media at Blombos (d'Errico et al. 2001; Henshilwood et al. 2009)
4. Persistence under stress	Documented: repertoire persists across multiple glacial–interglacial cycles and population turnovers (Section 5.1)	Not yet assessed against a documented stress event at comparable resolution
Stabilization score	4/4 — stabilized	2/4 confirmed, 2/4 untested — not yet scoreable

5.3 What has and has not been established

The contrast between these two records is consistent with the mechanism proposed in Section 4, and it is the contrast this paper's central claim rests on. It is important to state precisely what has and has not been shown. What has been shown, in the existing published literature, is sustained, repeated, controlled engraving activity at Blombos across roughly twenty-five thousand years, without published evidence of the same motifs recurring across unrelated Middle Stone Age sites at a scale comparable to the Aurignacian sign repertoire; and, separately, a small, shared, cross-site sign repertoire in the Upper Palaeolithic, independently documented by two different research programs. What has not been shown, because no one has run the direct comparison, is a formal quantitative measure of inter-site variability or repertoire convergence computed identically across both corpora.

A second, equally important gap must be stated alongside the first: an apparent difference in cross-site convergence between the two records could reflect a genuine difference in past behavior, or it could reflect a difference in how much each record has been looked for. European Upper Palaeolithic cave and mobile-object art has been the subject of over a century of intensive, well-funded, systematically comparative survey across dozens of sites; the African Middle Stone Age record, while the subject of increasingly serious attention including SignBase's own inclusion of MSA material, has not received a comparable density of cross-site systematic survey until relatively recently. A lower documented rate of MSA cross-site convergence is therefore consistent with two very different underlying realities — lower actual convergence, or under-sampled convergence — and the published record as it currently stands cannot fully distinguish between them. This is a limitation of the evidence base, not a detail to be smoothed over, and it is treated as a live constraint on the argument in Section 6.3 rather than resolved here. SignBase already includes Middle Stone Age material alongside its Aurignacian holdings, which means the dataset needed to run the direct, sampling-controlled comparison exists; the comparison itself has not yet been published. This paper treats the Middle Stone Age/Upper Palaeolithic contrast as a strongly motivated, testable prediction of the stabilization mechanism proposed here, and as the specific next step in this research program — not as a result already established in the literature, and not yet as a result that has ruled out a sampling explanation for the pattern it reports.

6. Scope, Limitations, and Falsifiability

This paper makes two different kinds of claim, and they should not be weighted the same way. The definition in Section 2, the differentiation from external symbolic storage, mimetic-mythic-theoretic staging, and extended mind in Section 3, and the

falsifiability criteria in Section 6.2 are established here: they are argued directly, and a reader can accept or reject them on the reasoning given. The mechanism in Section 4 and the Upper Palaeolithic/Middle Stone Age contrast in Section 5 are proposed rather than established: they are a motivated hypothesis and a worked illustration consistent with it, not a completed empirical test, for the reasons given in Sections 4.4, 5.3, and 6.3. The paper's central contribution — a falsifiable, four-criterion definition of symbolic stabilization, differentiated from the constructs it is most often confused with — does not depend on the mechanism or the case-study contrast being correct. Readers who are unpersuaded by Sections 4 and 5 are not thereby given reason to reject Sections 2 and 3, and this paper does not want that distinction lost in the reading.

6.1 What this paper does not claim

This paper makes no claim about the origin of writing, about the emergence of formal political institutions, or about population thresholds for symbolic or organizational complexity. These are related but separable problems, each requiring its own evidentiary standard and its own paper. Treating stabilization, writing, and institutional formation as a single undifferentiated explanatory target was a common failure mode in syntheses of this kind. This paper instead addresses one construct, defined, differentiated, and given a mechanism, with adjacent questions left to other work.

6.2 Falsifiability

The framework advanced here is weakened or falsified by any of the following: a symbolic system meeting all four stabilization criteria in a population with no independent evidence of the information-exchange pressure proposed to drive it; a population under sustained, documented high environmental unpredictability that fails to develop cross-site symbolic convergence despite adequate population size and duration; or a direct quantitative comparison of the Middle Stone Age and Upper Palaeolithic corpora (Section 5.3) that finds no meaningful difference in repertoire convergence, which would remove the empirical basis for the fluidity claim advanced here.

6.3 Limitations

The mechanism proposed in Section 4 is drawn from a body of ethnographic and archaeological theory developed primarily for hunter-gatherer social networks in temperate and glacial contexts; its applicability to earlier, African Middle Stone Age populations under different ecological regimes is an extension, not a direct transfer, and should be treated as such until tested. The comparison in Section 5 draws on published site-level descriptions rather than a re-analysis of primary data, and a formal quantitative test is identified as necessary future work rather than claimed here. That comparison is also, at present, confounded by unequal research intensity between the two regions (Section 5.3): a lower documented rate of Middle Stone Age cross-site convergence cannot currently be fully separated from a lower rate of comparably scaled multi-site survey, and the direct SignBase-based comparison proposed in Section 5.3 would need to either control for this or report it as an explicit caveat on its result.

6.4 Implications for Comparative Method in Deep-Time Symbolic Research

The unequal research intensity flagged in Section 6.3 is not a problem unique to the Middle Stone Age/Upper Palaeolithic comparison advanced here; it is a structural feature of how the field compares deep-time symbolic behavior across regions. European Upper Palaeolithic cave and mobile-object art has been the beneficiary of over a century of dedicated, well-funded, systematically comparative survey, producing dense, cross-site catalogues like SignBase and von Petzinger's cave inventory. The African Middle Stone Age record, by contrast, has historically received a fraction of the comparable multi-site systematic attention, despite growing recognition of its significance (Section 5.3). This general problem — test implications for behavioral or symbolic complexity being derived from the richer, better-surveyed European record and then applied, often unfavorably, to the African Middle Stone Age — has been identified as a structural weakness of the wider behavioral-modernity literature (Henshilwood and Marean 2003). This asymmetry means that any comparison built on the existing published record risks mistaking a sampling gradient for a behavioral one — treating the relative absence of documented cross-site convergence in the MSA as evidence of a genuine behavioral difference, when it may equally reflect where scholarly attention and funding have historically concentrated. This is not a problem specific to this paper's argument: it is a live methodological hazard for any claim, in either direction, about relative complexity or modernity between the African MSA and the European Upper Palaeolithic, and a comparable concern about differential archaeological visibility has been raised independently in recent work on Middle Stone Age technology and behavioral modernity more broadly (Sterelny and Hiscock 2026). Future comparative work in this area, including the direct test proposed in Section 5.3, would need to measure and control for survey intensity as a variable, rather than treat it as an incidental caveat.

References

Bahn, P.G. (1988). Comment on "The Signs of All Times: Entoptic Phenomena in Upper Palaeolithic Art." Current Anthropology, 29(2), 217–218.

Boyd, R., & Richerson, P.J. (1985). Culture and the Evolutionary Process. University of Chicago Press.

Claidière, N., & Sperber, D. (2007). The role of attraction in cultural evolution. Journal of Cognition and Culture, 7(1–2), 89–111.
Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19.
d'Errico, F., Henshilwood, C.S., & Nilssen, P. (2001). An engraved bone fragment from ca. 75,000-year-old Middle Stone Age levels at Blombos Cave, South Africa: implications for the origin of symbolism and language. Antiquity, 75(288), 309–318.
Donald, M. (1991). Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition. Harvard University Press.
Dutkiewicz, E., Russo, G., Lee, S., & Bentz, C. (2020). SignBase, a collection of geometric signs on mobile objects in the Paleolithic. Scientific Data, 7, 364.
Efferson, C., Lalive, R., Richerson, P.J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: The empirics of frequency-dependent cultural transmission. Evolution and Human Behavior, 29(1), 56–64.
Fernández-Navarro, V., Spaey, O., & Garate, D. (2026). Reevaluating hand stencil phenomena in cave art: A step forward towards the characterization of symbolic patterns during the Upper Palaeolithic in Europe. Cambridge Archaeological Journal, 36. Advance online publication. https://doi.org/10.1017/S0959774325100322
Fitzhugh, B., Phillips, S.C., & Gjesfjeld, E. (2011). Modeling hunter-gatherer information networks: An archaeological case study from the Kuril Islands. In R. Whallon, W.A. Lovis, & R.K. Hitchcock (Eds.), Information and Its Role in Hunter-Gatherer Bands (pp. 85–115). Cotsen Institute of Archaeology Press.
Gamble, C. (1982). Interaction and alliance in Palaeolithic society. Man, 17(1), 92–107.
Helvenston, P.A., Bahn, P.G., Bradshaw, J.L., & Chippindale, C. (2003). Testing the "Three Stages of Trance" model. Cambridge Archaeological Journal, 13(2), 213–224.
Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. Evolution and Human Behavior, 19(4), 215–241.
Henshilwood, C.S., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G.A.T., Mercier, N., Sealy, J.C., Valladas, H., Watts, I., & Wintle, A.G. (2002). Emergence of modern human behavior: Middle Stone Age engravings from South Africa. Science, 295(5558), 1278–1280.
Henshilwood, C.S., d'Errico, F., & Watts, I. (2009). Engraved ochres from the Middle Stone Age levels at Blombos Cave, South Africa. Journal of Human Evolution, 57(1), 27–47.
Henshilwood, C.S., & d'Errico, F. (2011). Middle Stone Age engravings and their significance to the debate on the emergence of symbolic material culture. In C.S. Henshilwood & F. d'Errico (Eds.), Homo Symbolicus: The Dawn of Language, Imagination and Spirituality (pp. 75–96). John Benjamins.
Henshilwood, C.S., & Marean, C.W. (2003). The origin of modern human behavior: A review and critique of models and test implications. Current Anthropology, 44(5), 627–651.
Lewis-Williams, J.D., & Dowson, T.A. (1988). The signs of all times: Entoptic phenomena in Upper Palaeolithic art. Current Anthropology, 29(2), 201–245.
Malafouris, L. (2013). How Things Shape the Mind: A Theory of Material Engagement. MIT Press.
Malafouris, L., & Renfrew, C. (Eds.). (2010). The Cognitive Life of Things: Recasting the Boundaries of the Mind. Cambridge: McDonald Institute for Archaeological Research.
Marks, A.E., Hietala, H.J., & Williams, J.K. (2001). Tool standardization in the Middle and Upper Palaeolithic: a closer look. Cambridge Archaeological Journal, 11(1), 17\u201344.
McBrearty, S., & Brooks, A.S. (2000). The revolution that wasn't: A new interpretation of the origin of modern human behavior. Journal of Human Evolution, 39(5), 453–563.
McCall, G.S. (2007). Add shamans and stir? A critical review of the shamanism model of forager rock art production. Journal of Anthropological Archaeology, 26(2), 224–233.
Monnier, G.F., & McNulty, K.P. (2010). Questioning the link between stone tool standardization and behavioral modernity. In S.J. Lycett & P.R. Chauhan (Eds.), New Perspectives on Old Stones: Analytical Approaches to Palaeolithic Technologies (pp. 61–81). New York: Springer.
Roberts, P., Henshilwood, C.S., van Niekerk, K.L., Keene, P., Gledhill, A., Reynard, J., Badenhorst, S., & Lee-Thorp, J. (2016). Climate, environment and early human innovation: Stable isotope and faunal proxy evidence from archaeological sites (98–59ka) in the southern Cape, South Africa. PLoS ONE, 11(7), e0157408.
Sperber, D. (1996). Explaining Culture: A Naturalistic Approach. Oxford: Blackwell.
Sterelny, K. (2010). Minds: Extended or scaffolded? Phenomenology and the Cognitive Sciences, 9(4), 465–481.
Sterelny, K., & Hiscock, P.D. (2026). Farewell to behavioural modernity? Homo sapiens in the Middle Stone Age. Cambridge Archaeological Journal, 36, 259–271.
von Petzinger, G. (2017a, March). The shaman in the cave? Testing for entoptic imagery in Upper Paleolithic geometric rock art. Paper presented at the 81st Annual Meeting of the Society for American Archaeology, Vancouver, British Columbia.
von Petzinger, G. (2017b). The First Signs: Unlocking the Mysteries of the World's Oldest Symbols. New York: Atria.
Vondoom, A. (2026). Cognitive Foundations of Early Egyptian Civilization: Insights from the Deep Symbolic Systems Model (DSSM). IWNW, 1–46. https://doi.org/10.21608/iwnw.2026.464676.1075
Whallon, R. (2006). Social networks and information: Non-"utilitarian" mobility among hunter-gatherers. Journal of Anthropological Archaeology, 25(2), 259–270.
Wobst, H.M. (1977). Stylistic behavior and information exchange. In C.E. Cleland (Ed.), For the Director: Research Essays in Honor of James B. Griffin (pp. 317–342). University of Michigan Museum of Anthropology.

Anthony Vondoom is an independent researcher in cognitive archaeology, systems theory and cognitive neuroscience. He is the originator of the Deep Symbolic Systems Model (DSSM), a framework for the structural conditions under which symbolic systems become fixed and transmissible across generations.

AI Use Disclosure

The author used Claude (Anthropic) as a writing and editorial aid in the preparation of this manuscript, including drafting and revising prose, formatting, and reference verification. All conceptual content, arguments, definitions, and conclusions are the author's own; all AI-assisted output was reviewed, edited, and approved by the author, who takes full responsibility for the accuracy and integrity of the final text.