

The Compression of Expertise

Flyxion

2026

Abstract

Expertise appears mysterious because experts often arrive at correct judgments without exhibiting the intermediate reasoning a novice would need to reach the same conclusion. This essay argues that expertise is not primarily a stock of knowledge but a compressed repair and intervention history. Drawing on the retention-lattice formalism of *History Before Function*, expertise is characterized as an operator whose generating history has been compressed into a form that supports rapid enactment while reducing visibility into the diagnostic traces that originally produced it. The apparent opacity of expert judgment is therefore not a defect of explanation but a structural consequence of compression. The essay develops this account using the distinction between enactment and representation, interprets deliberate-practice research as trace accumulation prior to compression, and reinterprets classical findings in chess expertise as evidence for retained historical structure rather than superior raw calculation.

This essay is not a theory of skill acquisition, forgetting, or expertise maintenance. It treats the retained trace T as fixed and asks a narrower question: what kind of object is expertise once it exists? The dynamics by which T is accumulated through practice, compressed through experience, maintained through rehearsal, or repaired after degradation are deferred. Those questions belong to a separate theory of expertise as recursive continuation rather than to the present account of expertise as compressed history.

1 The Puzzle of Expertise

Novices show work; experts skip steps. A chess master rejects a losing move before articulating why. A mechanic diagnoses a fault from a sound alone. A physician recognizes a presentation before running the differential a textbook would specify. A senior programmer flags a bug on sight that a code-review checklist would take twenty minutes to surface systematically. The standard explanations on offer — more knowledge, better intuition, faster reasoning — are not wrong so much as they restate the phenomenon in different words without saying what intuition or speed actually consist in.

This essay proposes a specific, structural alternative:

The expert is not performing fewer operations. The operations have been relocated from execution-time to compression-time.

Nothing about the expert's judgment involves less computation than the novice's explicit procedure. What differs is when the computation happened. The novice performs it live, in view, at the moment of the decision. The expert performed an equivalent reduction long before, across many prior cases, and what remains visible now is only the output of that reduction — an operator that runs quickly because its internal search has already been done and discarded.

2 Expertise as a Compressed History

This essay imports its central formal object directly from *History Before Function*: a compression map C taking a history H to an operator F , $C(H) = F$, with F characterized independently of C by its behavior.

Thesis 2.1 (The Expertise Thesis). *An expert operator F is the compressed residue of a large repair and intervention history H : a trajectory of prior cases, mistakes, corrections, and successful interventions. Expertise is not*

$$\text{Knowledge} \longrightarrow \text{Skill},$$

but

$$\text{History} \xrightarrow{C} \text{Compression} \longrightarrow \text{Skill}.$$

Remark. *The distinction matters because “knowledge” suggests a stock of retrievable propositions the expert could, in principle, list. The Expertise Thesis makes no such claim and in fact denies it in general: what has been compressed away in forming F need not be recoverable as an enumerable list of facts at all. It was a trajectory, with an order, with false starts, with corrections applied to earlier errors — and compression is precisely the operation that discards that internal structure while preserving the behavior it produced.*

3 Deliberate Practice as Trace Accumulation

Ericsson’s deliberate-practice literature is standardly read as showing that structured practice increases ability. This essay proposes a reframing that changes what the practice hours are doing, without disputing the empirical findings themselves.

Deliberate practice does not directly increase ability. It accumulates the historical structure H that compression later acts on.

On this reading, the well-established finding that unstructured repetition produces far weaker gains than deliberate, feedback-corrected practice is not a puzzle about motivation or attention; it is a direct consequence of what H needs to contain to compress into something useful. Unstructured repetition produces a long but structurally flat H — many instances of the same uncorrected pattern, offering little for C to fold into a refined F . Deliberate practice, by contrast, deliberately manufactures the repair episodes — attempt, feedback, correction, retry — that give H the kind of internal structure compression can act on. Practice, in this account, is the construction phase; compression is a separate, later operation performed on the material practice supplies. Without accumulated trace there is nothing to compress, however long one merely repeats an action.

4 Enactment Without Representation

History Before Function establishes an asymmetry between two things that can be done with an existing operator: E , enactment, and R , symbolic representation, with E total on the class of genuine operators and R only partially defined:

$$F \in \text{Dom}(E) \not\Rightarrow F \in \text{Dom}(R).$$

This asymmetry is, on the present account, the direct mechanism behind expert opacity. A skilled cyclist enacts balance corrections reliably without possessing, or being able to produce on

demand, a symbolic account of what those corrections are. A physician’s diagnostic recognition, a chess player’s immediate rejection of a losing move, and a reviewer’s sense that a code change is unsafe are all instances of $E(F)$ occurring without an accompanying $R(F)$ — not because the expert is withholding an explanation they possess, but because no such explanation was ever generated in the first place. Compression discarded exactly the intermediate structure that a symbolic account would need to draw on.

Remark. *This reframes a familiar and often frustrating request: explain why you knew that. The request presupposes $R(F)$ exists and asks the expert to produce it. Where only $E(F)$ exists, the expert’s honest answer is unavailable, and what is typically supplied instead — a plausible-sounding chain of reasoning constructed after the fact — is not a report of the process that produced the judgment but a fresh act of representation performed under demand, using whatever partial cues happen to be available at the moment of being asked. Post-hoc rationalization, on this reading, is not dishonesty; it is R being invoked where $R(F)$ was never defined, and something has to be returned regardless.*

5 Expertise on the Retention Lattice

History Before Function’s retention lattice $\mathcal{L}_F = \{T : T_{\min} \subseteq T \subseteq H\}$ and its Anti-Alignment Theorem give this essay’s central taxonomy a precise location rather than a metaphor.

Novice. Retained trace close to H in full. Many intermediate steps remain visible and inspectable. Low repair entropy $S_R(F | T)$: a failure can usually be traced to a specific step. Enactment is slow, since little has been folded into a fast-running F .

Expert. Retained trace closer to T_{\min} . Rapid enactment; the operator runs with little visible intermediate structure. Repair entropy is higher than the novice’s by Lemma 3.3 of *History Before Function* (Trace-Retention Monotonicity), though not yet at its theoretical maximum.

Brittle expert. Retained trace at or near T_{\min} itself. Enactment is fastest here, and by the Anti-Alignment Theorem, repair entropy is at or near its maximum over \mathcal{L}_F : when F fails outside the conditions its compressed history anticipated, there is little retained structure left to localize why, and failure is accordingly difficult to diagnose and prone to being catastrophic rather than gracefully degraded.

Remark. *This section is, in effect, a human-facing reading of a theorem already proved elsewhere rather than a new result. What it adds is the observation that “expertise” as ordinarily used conflates the middle and rightmost categories: a skilled practitioner and a brittle overfit pattern-matcher can look identical under routine conditions, since both enact quickly, and differ only in what happens under distribution shift — precisely where the Anti-Alignment Theorem predicts they should diverge.*

6 Chess as a Compression Phenomenon

De Groot’s early studies and Chase and Simon’s subsequent work on chess perception are usually read as showing that masters calculate faster or further ahead than novices. The evidence does not straightforwardly support this: masters and novices search comparably few moves ahead under time pressure, and the decisive difference instead appears in masters’ ability to reconstruct briefly-viewed, game-plausible board positions from memory — an ability that collapses to novice level

when pieces are arranged randomly rather than according to positions that could arise from actual play.

This is precisely what the compression account predicts and the calculation account does not. What is retained is not a faster search procedure but a compressed operator built from an immense history of prior positions, each previously subjected to search, correction, and evaluation. Confronted with a new position, the master is not searching more; they are pattern-matching against a compressed F built from thousands of previously compressed instances, which is why the advantage vanishes precisely when the position no longer resembles anything that history could have generated. What looks like

$$\text{position} \longrightarrow \text{move}$$

is, on this account,

$$\text{history} \xrightarrow{C} \text{compression} \longrightarrow \text{operator} \longrightarrow \text{move},$$

with the move as the only step a observer standing outside the process ever sees.

7 Relation to *History Before Function*

History Before Function developed the general theory of compressed operators: existence and non-uniqueness of generating histories, the asymmetry between enactment and representation, and the Anti-Alignment Theorem governing retention and repairability. The present essay treats expertise as a single, richly worked special case of that theory, contributing an interpretation of what the abstract apparatus looks like when the operator in question is a human skill, not new formal machinery of its own.

A separate account is required to explain how expertise survives, decays, and regenerates through time — how T is built up through deliberate practice, how it degrades through disuse, and how a lapsed skill is relearned rather than learned anew. That question concerns the dynamics of compressed operators, draws simultaneously on *Recursive Continuation's* treatment of continuation and repair and on *Threshold of Deferral's* treatment of maintenance timing, and belongs to a theory of expertise as recursive continuation rather than to the static account developed here.

Open Problems

1. **Transferability.** Can compressed expertise be transferred directly to another agent, or must the receiving agent rebuild an equivalent F from their own history? If transfer is possible at all, what is actually transferred — is it ever T itself, or only T_{\min} , with the recipient left to accumulate their own surplus independently?
2. **Optimal placement.** What determines where, along \mathcal{L}_F , a given domain's expertise should optimally sit? *Threshold of Deferral's* crossover apparatus may be directly relevant here, with storage/inspection cost playing the role κ_M played there.
3. **Brittleness onset.** At what point does compression cross from productive (expert) into brittle? Is this a sharp threshold on \mathcal{L}_F or a gradual one, and can it be detected before a failure exposes it?
4. **Expertise maintenance as recursive continuation.** How should the deliberate re-exposure to fundamentals that many disciplines require of even advanced practitioners be modeled as a maintenance process on a compressed operator, in the sense of *Threshold of Deferral*?

5. **Expert failure and repair entropy.** When an expert is wrong, is the resulting repair entropy systematically higher than when a novice is wrong, and does this predict differences in how expert versus novice errors are corrected in practice?

References

1. Michael Polanyi. *The Tacit Dimension*. University of Chicago Press, 1966.
2. K. Anders Ericsson, Ralf Th. Krampe, and Clemens Tesch-Römer. "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review*, 100(3), 1993.
3. Adriaan D. de Groot. *Thought and Choice in Chess*. Mouton Publishers, 1965.
4. William G. Chase and Herbert A. Simon. "Perception in Chess." *Cognitive Psychology*, 4(1), 1973.
5. Hubert L. Dreyfus and Stuart E. Dreyfus. *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Free Press, 1986.
6. Gary Klein. *Sources of Power: How People Make Decisions*. MIT Press, 1998.
7. Daniel Kahneman and Gary Klein. "Conditions for Intuitive Expertise: A Failure to Disagree." *American Psychologist*, 64(6), 2009.
8. Flyxion. *History Before Function: Compression as the Origin of Operators*. 2026.
9. Flyxion. *Recursive Continuation: Autopoiesis, Symptoiesis, and the Compression of Self-Modification*. 2026.