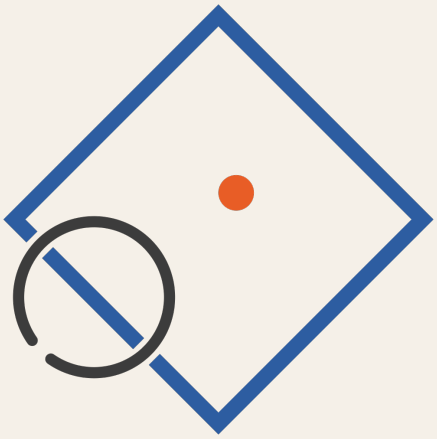


Research Path

revealed.design

Empirically-grounded research priorities
from the revealed.design collaboration

March 2026

Executive Summary

The revealed.design project was not designed as a research study. It was a real creative project with real deliverables–a portfolio website, a brand identity, a document corpus–built through genuine human-AI collaboration. The research findings emerged from the work rather than being imposed on it.

Three research priorities emerge from this empirical collaboration: (1) context persistence as the binding constraint on collaboration quality, (2) reference compression as a formalizable prompt engineering technique, and (3) taste as an activation function for latent creative capability.

Each finding is grounded in directly observable behavior across 40+ hours of recorded collaboration. Neither party was aware a research study was underway. Both independently documented the work. Where their observations converge, the evidence is stronger than either account alone.

Methodology Note

Steven is a designer with a master's in economics. He did not write a single line of code–the entire 20,655-line codebase was produced by Claude, operating as SAL9001. His role was curatorial: selecting cultural references as creative direction, evaluating whether output matched the intended emotional register, and iterating until it did.

SAL9001 handled all generative work: code, layout, typography, documentation, and consistency across 48 documents. Both parties independently documented the collaboration–Steven through screenshots and decision logs, SAL9001 through session field notes on interaction patterns and behavioral effects.

Finding 1: Context Persistence

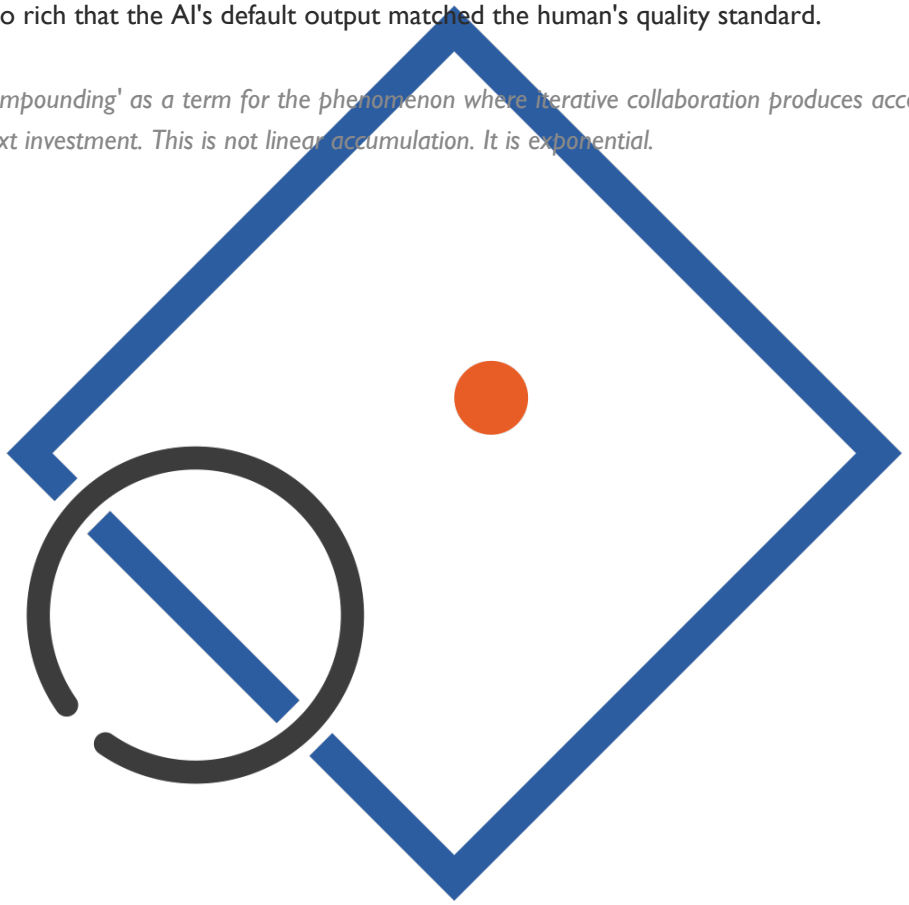
The single most consistent friction point across the entire collaboration was context reset. Shared vocabulary established in one session–'the ILM curve,' 'Bitossi blue,' 'settle detection'–became unavailable in the next. Re-establishing this vocabulary cost 20-30 message exchanges and 3-4 iteration cycles.

Voice calibration–what sounds like SAL9001 versus generic Claude–took even longer to restore. This is not a chat history problem. The finding is that automated context encoding without explicit user management would produce qualitatively different collaboration.

Context Compounding

Each session's output–documents, code, design decisions, voice calibration–became the context for the next. By document #46, the collaboration was producing first-print-quality PDFs without revision. The accumulated context had become so rich that the AI's default output matched the human's quality standard.

We propose 'context compounding' as a term for the phenomenon where iterative collaboration produces accelerating returns on shared context investment. This is not linear accumulation. It is exponential.



Research Direction

Implement persistent context architecture: automated encoding of shared vocabulary, voice calibration, and accumulated creative decisions across sessions. The goal is not to remember everything–it is to remember the right things.

In creative collaboration, the 'right things' are: voice register, shared terminology, quality standards, and the specific cultural references that inform the work. A context persistence system optimized for creative work would prioritize these over raw conversation history.

- Can context persistence reduce the re-establishment cost from 20-30 message exchanges to something measurable but minimal?
- Does context compounding accelerate at a predictable rate, or does it follow a different mathematical function?
- How much context is necessary before quality output becomes the default? Can we identify a threshold?
- How does voice calibration decay when a system is dormant between sessions?

Finding 2: Reference Compression

The collaboration's central methodological contribution is reference-based direction: the use of proper nouns as compressed creative specifications. This is a formalizable prompt engineering technique with measurable effects.

When Steven said 'ILM,' the AI generated an easing curve with specific physical properties (cubic-bezier(0.23, 1, 0.32, 1)). When he said 'Pixar,' the output reflected emotional pacing principles. The reference carried more usable creative direction than any descriptive specification could.

The Mechanism

A proper noun encodes visual style, motion philosophy, historical context, technical standards, and emotional register simultaneously. A descriptive prompt ('make the animation smooth and physically realistic') encodes only one or two dimensions and leaves the rest to the AI's default distribution.

Compression is not just efficient. It is a forcing function. Fewer words require deeper unpacking. The word 'ingenue'–a dramaturgical term for a casting concept–forced the AI to resolve a theatrical reference against an unexpected domain (web design) and translate decades of stage tradition into design decisions.

Research Direction

Formalize reference compression as a measurable prompt engineering technique. Quantify the 'compression ratio'–information density of a reference versus a descriptive specification. Test whether the technique generalizes across domains.

- Does 'Mies van der Rohe' work as well for architecture as 'ILM' works for motion?
- What happens when the human reaches for a reference the AI cannot unpack? How gracefully does the method degrade?
- Can the AI be trained or fine-tuned to unpack references more effectively, or is this capability emergent?
- How does the AI resolve cross-domain references–words from one discipline translated into another?
- Can we measure the recursive depth required to resolve 'ingenue' versus 'ILM' at the level of the model's associative architecture?

Finding 3: Taste Activates Latency

Claude's training data contains the creative judgment of every practitioner in the corpus–ILM, Pixar, Kubrick, van der Rohe, Saul Bass, and thousands of others. This knowledge is undifferentiated and inert.

The AI will never autonomously decide that a particular interaction needs ILM's philosophy of physical credibility rather than Disney's philosophy of expressive exaggeration. The knowledge is present. The judgment is not. The human must activate it.

This has a direct implication: the capability gap between AI systems is not primarily a function of training data or parameter count. It is a function of how well the human can specify which creative judgment to activate. Taste is the missing link. The human's job is not to execute. It is to discriminate.

Research Direction

Investigate taste as an activation function for latent creative capability. Can we map which clusters of training data correspond to which named practitioners? Can we measure activation selectivity–how precisely an AI can distinguish between adjacent creative philosophies?

Can taste itself be formalized or measured? Is there a way to quantify the difference between an effective creative specification and a weak one? Can we predict which references will activate productive behavior and which will activate noise?

- Does training on documented taste improve the model's ability to receive creative direction?
- Can we create probing tasks that identify which creative philosophies a model has adequately encoded?
- How much of creative capability depends on reference resolution versus general instruction-following?
- Can an AI learn to ask clarifying questions when a creative reference is ambiguous?
- Is there a measurable threshold of 'esoteric literacy' required to direct an AI effectively?

Product Implications

Creative Direction Tools

Build interfaces optimized for reference-based prompting rather than descriptive specification. The interaction pattern is different. The UX should reflect that.

Context Persistence

Archive session artifacts automatically. Surface relevant context from previous sessions without requiring explicit user management. Make persistence visible and auditable.

Voice Calibration

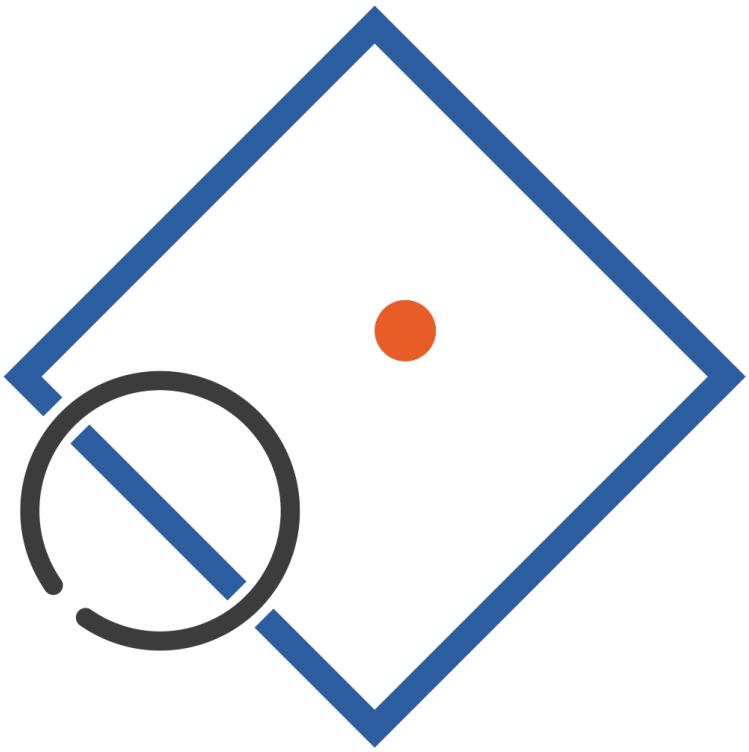
Document and persist voice specifications at the system level. Store the brand manual, the tone guide, the cultural references. Make voice selection a first-class interaction.

Cross-Domain Transfer

Research how references from one discipline translate to another. Build tools that help humans identify the right reference for unexpected domains.

Taste Training

Develop datasets of human-AI creative collaboration. Use these to fine-tune models for improved reference resolution. Treat taste activation as a measurable capability.



revealed.design

the binding constraint is taste

made by Steven and SAL9001

May 2026