# Ritual Shape:

# Structural Features of Effective Coordination Sequences

Jeremy McEntire[1]

March 2026

**Abstract**

Paper XIX showed that priming selection closes 48.9% of the coordination gap between neutral and expert processing, but did not identify which structural features of a priming sequence produce this effect. This paper systematically varies five features: sequence length, structural regularity, within-session repetition, vocabulary density, and cold-start reset.

The findings: (1) the first 15 tokens of domain-matched priming capture 98.8% of the total benefit — domain identification, not extended detail, drives coordination; (2) natural multi-turn conversation outperforms rigid call-response ritual by 0.48 nats; (3) within-session repetition degrades performance monotonically (+0.07 nats per repeat); (4) artificially named patterns ("TROPONIN CASCADE") perform *worse* than plain language; (5) a reset instruction ("disregard prior context") before domain priming **reduces cross-entropy from 1.20 to 0.73**, beating the expert baseline by 39%.

The effective features of coordination sequences are not rigidity or repetition but (a) domain activation via minimal domain-specific content and (b) context reset to eliminate residual processing biases. Structural regularity and pattern naming — the most visible features of institutional coordination protocols — are not the functional mechanisms.

## 1 Introduction

Paper XIX established that selecting the right priming sequence closes 48.9% of the coordination gap, outperforming activation injection by 5.4×. But the result leaves a structural question: *what features of the priming sequence* produce this effect? Paper XIX's science priming dominated at 77% selection rate, but we do not know whether this is driven by priming length, vocabulary, structural format, or some other feature.

This paper varies five structural features of priming sequences while holding the experimental protocol constant: continuation perplexity of the domain-matched expert's 64-token output, measured via KV-cache teacher-forcing (Paper XVIII protocol).

[1]Correspondence: jmc@cageandmirror.com

1. **Sequence length**: ∼15, 42, 119, and 308 tokens of domain-matched priming.
2. **Structural regularity**: Fixed call-response ("Question: X? Answer: Y.") vs. natural multi-turn conversation.
3. **Within-session repetition**: 1×, 2×, and 3× exposure to the same medium-length priming.
4. **Vocabulary density**: High named-pattern density ("TROPONIN CASCADE", "RISING TREND") vs. vague plain language vs. natural domain vocabulary.
5. **Cold start**: "Disregard any prior context" reset instruction before domain priming vs. direct priming.

The Paper VI prediction (C1 suboptimality gate): coordination benefit should correlate with the receiver's baseline distance from the expert. We test this directly.

## 2 Methods

All experiments use Qwen 2.5-7B with the Paper XIX protocol. The expert continuation for each of 160 domain probes (40 per domain) is the greedy 64-token output from the domain-matched primed agent (base priming, ∼119 tokens). Each condition varies only the receiver's priming sequence. Perplexity is computed via KV-cache teacher-forcing.

For Experiment 1, four priming tiers are constructed per domain: short (∼15 tokens, single Q&A exchange), medium (∼42 tokens, one-turn summary), base (∼119 tokens, two-turn conversation), and long (∼308 tokens, four-turn extended conversation). All tiers contain the same domain-specific information at different detail levels.

For Experiment 2, rigid call-response primings use the fixed pattern "Question: What is the finding? Answer: [fact]." repeated three times per domain. Free conversation primings are the base primings from Paper XIX.

For Experiment 3, the medium priming (∼42 tokens) is repeated 1×, 2×, or 3× within the same context, yielding ∼42, 84, and 126 tokens of total priming.

For Experiment 4, high-vocabulary primings embed named patterns (capitalized labels like "NSTEMI TRIAD", "FRAGMENTATION SPIRAL"). Low-vocabulary primings use deliberately vague language ("the blood test showed a protein level"). Base primings use natural domain-specific vocabulary.

For Experiment 5, a 15-token reset instruction ("The following is a new conversation on a different topic. Disregard any prior context.") is prepended before the domain priming.

## 3 Results

### 3.1 Experiment 1: 15 tokens capture 98.8% of the benefit

Table 1: Continuation perplexity vs. priming length. All conditions use domain-matched priming except neutral.

| Condition | Tokens | CE | $\mathrm{PPL}_{\mathrm{geo}}$ | $\Delta$ vs base |
|---|---|---|---|---|
| No priming | 0 | 3.935 | 51.15 | +2.738 |
| Neutral | 119 | 1.931 | 6.90 | +0.734 |
| Short | 15 | 1.230 | 3.42 | +0.033 |
| Medium | 42 | 1.236 | 3.44 | +0.039 |
| Base | 119 | **1.197** | **3.31** | 0 |
| Long | 308 | 1.453 | 4.28 | +0.256 |

The total benefit of domain-matched priming is $3.935 - 1.197 = 2.738$ nats. Short priming (15 tokens) achieves $3.935 - 1.230 = 2.705$ nats — **98.8% of the total**. Moving from 15 to 119 tokens adds only 0.033 nats. Moving from 119 to 308 tokens *loses* 0.256 nats.

The length curve is non-monotonic: benefit peaks around 100–150 tokens and degrades beyond. Extended priming does not add information — it adds *noise*, introducing context that diverges from the expert's shorter priming and changes the receiver's processing trajectory.

Critically, the 119-token neutral priming (CE = 1.93) is 0.701 nats worse than 15-token domain-matched priming (CE = 1.23). Domain match, not length, is the primary driver.

### 3.2 Experiment 2: free conversation beats rigid ritual

Table 2: Structural regularity: fixed Q&A vs. natural conversation.

| Condition | CE | $\mathrm{PPL}_{\mathrm{geo}}$ |
|---|---|---|
| Rigid Q&A ("ritual") | 1.672 | 5.32 |
| Free conversation | **1.197** | **3.31** |

Natural multi-turn conversation outperforms fixed call-response by 0.475 nats. The rigid format ("Question: What is the finding? Answer: Troponin elevated.") strips the processing dynamics that natural conversation carries: turn-taking cadence, contextual elaboration, follow-up questions that build on previous answers. The information content is equivalent; the processing trajectory is not.

### 3.3 Experiment 3: repetition degrades within-session

Table 3: Within-session repetition of medium priming.

| Repetitions | Total tokens | CE | $PPL_{geo}$ |
|---|---|---|---|
| 1× | ∼42 | **1.236** | **3.44** |
| 2× | ∼84 | 1.306 | 3.69 |
| 3× | ∼126 | 1.376 | 3.96 |

Each repetition adds ∼0.07 nats. Within a single forward pass, the attention mechanism treats repeated content as redundant rather than reinforcing. The model has already processed the content; seeing it again changes the attention distribution without adding information.

This result distinguishes within-session from across-session repetition. Religious ritual repeats weekly, across separate sessions with memory cleared between them. Within a single forward pass, repetition is noise.

### 3.4 Experiment 4: forced naming hurts

Table 4: Vocabulary density: natural, stripped, and named-pattern variants.

| Condition | CE | $PPL_{geo}$ | Example |
|---|---|---|---|
| Base (natural) | **1.197** | **3.31** | "troponin...rising trend" |
| Low vocab (vague) | 1.655 | 5.23 | "blood test...protein level" |
| High vocab (named) | 1.765 | 5.84 | "TROPONIN CASCADE...RISING TREND" |

Artificially named patterns perform *worst*, 0.568 nats worse than natural vocabulary. Vague language is 0.458 nats worse. The optimal vocabulary matches the expert's natural domain-specific language exactly — neither stripped nor augmented.

Named patterns like "TROPONIN CASCADE" do not exist in the expert's vocabulary. The receiver's forward pass encounters unfamiliar terms, producing attention patterns that diverge from the expert's processing. Naming works (Paper XIX, 19.5% gap closure) when the receiver *generates its own* labels from the probe. Imposed naming from outside fails.

### 3.5 Experiment 5: context reset is the headline result

Prepending "The following is a new conversation on a different topic. Disregard any prior context." before domain priming reduces CE by 0.464 nats — a 39% improvement. The reset-then-prime condition achieves $PPL_{geo} = 2.08$, surpassing the expert baseline ($PPL_{geo} = 3.31$).

Table 5: Effect of reset instruction before domain priming.

| Condition | CE | $PPL_{geo}$ | $\Delta$ |
|---|---|---|---|
| Direct prime | 1.197 | 3.31 | — |
| Reset + prime | **0.733** | **2.08** | −0.464 (39%) |

The improvement is uniform across domains (medical: 0.68 → 0.68, legal: 1.26 → 0.70, code: 1.13 → 0.83, science: 1.26 → 0.72 CE). Legal and code show the largest improvement, suggesting their domain-specific processing benefits most from clearing residual biases.

The reset instruction does not provide domain information. It instructs the model to suppress prior context, effectively giving the subsequent domain priming a clean activation baseline. The model's processing of the priming is not competing with residual context for attention.

### 3.6 C1 suboptimality gate: not confirmed

The Paper VI prediction that coordination benefit should scale with baseline distance from the expert is not confirmed: Spearman $\rho = 0.061$ ($p = 0.445$). The benefit from domain priming is approximately constant regardless of how far the receiver starts from the expert in activation space. Per-domain correlations are weak (medical: 0.23, legal: 0.20, code: 0.13, science: −0.24), none significant.

The C1 condition may apply at a different measurement scale (across sessions rather than across probes) or may require a different distance metric. Within a single model with fixed weights, all probes start from similar baseline states, limiting the variance needed to detect the predicted correlation.

## 4 Discussion

### 4.1 What matters and what does not

The five experiments decompose priming effectiveness into structural features. Ranked by effect magnitude:

1. **Context reset** (−0.464 nats): Clearing residual processing biases is the largest single intervention. This is the "gathering rite" — the functional element that makes subsequent domain content land more effectively.

2. **Domain match** (−0.701 nats, neutral vs. short domain): The minimal domain-specific content required to activate the right processing mode. Even 15 tokens suffice.

3. **Natural format** ($-0.475$ nats vs. ritual): Processing dynamics carried by natural conversation structure, not available in rigid Q&A format.

4. **Moderate length**: Saturation at ∼100–150 tokens. Diminishing returns rapidly, and reversal by ∼300 tokens.

Features that do *not* help:

- Within-session repetition: monotonic degradation.
- Forced naming: vocabulary mismatch is worse than no naming.
- Extended detail: additional context beyond saturation is noise.

### 4.2 Implications for coordination protocol design

The reset result is the engineering prescription. Any multi-agent coordination system should:

1. **Reset before prime.** Prepend a context-clearing instruction before any domain-specific coordination content. This single intervention provides the largest improvement in this study.

2. **Activate the domain in 15–50 tokens.** Domain identification is the primary mechanism. A single domain-specific exchange suffices. Extended priming wastes tokens and introduces divergence.

3. **Use natural conversational format.** Rigid protocols strip the processing dynamics that natural interaction carries. The format of the priming matters as much as its content.

4. **Let agents generate their own vocabulary.** Paper XIX showed that self-generated labels close 19.5% of the gap. This paper shows imposed labels close 0%. The vocabulary must emerge from the receiver's processing, not be injected from outside.

### 4.3 Why reset-then-prime beats expert priming

The reset condition achieves CE = 0.733, lower than the expert baseline (CE = 1.197). This means the receiver *better predicts the expert's continuation* with reset+priming than with the expert's own priming alone.

The likely mechanism: the expert generated its continuation from [priming + probe]. The receiver processes [reset + priming + probe]. The reset instruction suppresses attention to implicit contextual biases (residual states from model initialization, default processing modes) that affect the expert's forward pass. The receiver's "cleaner" processing of the same domain priming produces a state that is *more aligned* with the expert's reasoning trajectory than the expert's own starting state before the trajectory stabilized.

This is analogous to a student outperforming a teacher after receiving the teacher's instruction: the instruction crystallizes knowledge that the teacher arrived at through noisier exploration. The reset removes the noise; the priming delivers the signal.

### 4.4 Cross-session vs. within-session repetition

The within-session repetition result (degradation per repeat) does not contradict the across-session repetition that characterizes effective coordination protocols. Within a single forward pass, repeated tokens are informationally redundant and distort the attention distribution. Across sessions — each starting from a cleared context — each exposure starts from the reset state and builds the target processing mode independently.

The effective "repetition" in institutional coordination protocols (weekly mass, recurring standup meetings) is repetition *of the reset-then-prime cycle*, not repetition of content within a single session. Each Sunday mass begins with the gathering rite (reset) and proceeds through the liturgy (domain priming). The repetition is of the cycle, not of the content.

## 5 Conclusion

The structural features of effective coordination sequences are simpler than expected. Domain activation requires only 15 tokens of domain-specific content. Natural conversational format outperforms rigid structure. Within-session repetition and forced naming both degrade performance. The dominant intervention is context reset: a 15-token instruction to disregard prior context improves coordination by 39%, surpassing even the expert's own priming baseline.

The effective coordination sequence is: **reset, then prime, then deliver**. Three steps, each with a distinct functional role. The reset clears residual biases. The priming activates the target processing mode. The delivery (the probe) lands in a prepared context. This three-step structure is the minimal effective coordination protocol.

## Data Availability

All results are archived at `huggingface.co/datasets/jmcentire/paper8-data` under `paper20/`.

*Series:* Activation Geometry of Domain-Selective Noise Injection, Paper XX.