

The Conveyance Hypothesis

A Mathematical Framework for Measuring Information Transfer Effectiveness

Todd Bucy

Version 4.0 — December 2025

Status: Hypothesis under active investigation

Abstract

In 1948, Claude Shannon deliberately excluded meaning from his mathematical theory of communication, noting that “the semantic aspects of communication are irrelevant to the engineering problem.” This exclusion was pragmatic—the technology to measure semantic content did not exist. Seven decades later, with the emergence of Large Language Models and high-dimensional embedding spaces, we may now possess the tools to address what Shannon set aside.

This paper proposes the Conveyance Hypothesis: that information transfer effectiveness between computational agents can be mathematically measured through the interaction of semantic investment, relational structure, processing capability, and shared context. We present a core equation, define measurable variables, report preliminary empirical observations, and specify conditions that would falsify the hypothesis.

We argue that the distinctions between “data,” “information,” and “knowledge” are not merely philosophical but operationally measurable. We define data as static and boundary-preserving and existing in 3D spacetime, information as dynamic and transformation-inducing, and knowledge as the meaning that exists in high-dimensional space. This framing draws on Actor-Network Theory’s concept of translation and boundary objects. We extend these qualitative analytical tools into quantifiable territory through geometric analysis of neural embedding spaces.

Key claim: We may be entering an “Age of Measurable Meaning”—not because we can define meaning philosophically, but because we can detect its effects mathematically.

Primary Equation:

$$C_{\text{pair}}(i \leftrightarrow j) = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \times P_{ij}$$

Key Prediction: Low-dimensional representations (128–256D) outperform high-dimensional ones for information transfer, and dimensional collapse (measured by β) negatively correlates with task performance.

Contents

1	Introduction: Shannon’s Deliberate Exclusion	4
1.1	Why This Matters	4
1.2	Why This Is Feasible Now	5
2	Theoretical Foundation	5
2.1	Shannon’s Deliberate Exclusion (Expanded)	5
2.2	Actor-Network Theory: Translation, Not Transmission	6
2.3	Boundary Objects	6
2.4	The Data vs. Information Distinction	6

3	The Mathematical Framework	6
3.1	Primary Conveyance Equation (v4.0)	6
3.2	Why Harmonic Mean?	7
3.3	Component Equations	7
3.4	Why Multiplicative?	7
3.5	Dimensional Richness Function	8
4	Key Metrics	8
4.1	D_{eff} (Effective Dimensionality) — PRIMARY METRIC	8
4.2	β (Beta) — DIAGNOSTIC METRIC (NOT Optimization Target)	9
4.3	Secondary Geometric Metrics	9
4.4	Task Performance — VALIDATION METRICS	10
5	Core Hypotheses (Under Investigation)	10
5.1	Low-Dimensional Hypothesis	10
5.2	β -Overfitting Hypothesis	10
5.3	Attention-Only Hypothesis	10
5.4	Bilateral Asymmetry Hypothesis	11
6	Context Amplification	11
6.1	Original Formulation (Superseded)	11
6.2	Current Formulation (v4.0)	11
6.3	Geometric Extension (For Advanced Analysis)	11
7	Temporal Dynamics	12
7.1	Multi-Turn Context Evolution	12
7.2	Self-Reinforcing Cycles	12
7.3	Threshold-Based Management	12
8	Zero-Propagation Principle	12
8.1	Definition	12
8.2	Implications	12
8.3	Dimensional Collapse as Zero-Propagation	13
9	Falsification Criteria	13
9.1	Strong Falsification Evidence	13
9.2	Weak Falsification Evidence	14
10	Relationship to Existing Theories	14
10.1	Shannon's Information Theory	14
10.2	Rogers' Innovation Diffusion Theory	14
10.3	Kolmogorov Complexity	14
11	Practical Applications (If Validated)	14
11.1	AI Development	14
11.2	AI Safety	15
11.3	Memory Systems	15
11.4	Human Communication (Speculative)	15

12 Current Evidence Status	15
12.1 Validated (Tier 1 Evidence)	15
12.2 Preliminary Observations (Require Validation)	15
12.3 Unvalidated (Theoretical)	16
12.4 Falsified (Revised in v3.9+)	16
13 Conclusion	16
A Quick Reference	16
B Version History	17

1 Introduction: Shannon’s Deliberate Exclusion

Claude Shannon’s foundational paper “A Mathematical Theory of Communication” (1948) begins with a remarkable constraint:

“The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning. . . These semantic aspects of communication are irrelevant to the engineering problem.” [1]

This exclusion was pragmatic—Shannon understood that meaning mattered, but the technology of 1948 made the problem of measurable semantic change intractable. His theory optimized for signal fidelity: how accurately can bits travel from sender to receiver? The interpretation of those bits remained outside the mathematical framework.

Warren Weaver made this exclusion explicit:

“The word information, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning. In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information.” [2]

A crucial clarification: What Weaver called “information” is what we now distinguish as **data**—static, transmittable representations existing in three-dimensional spacetime. **Information**, in the Conveyance Hypothesis framework, is the process by which data expands into high-dimensional knowledge space or compresses from knowledge into transmittable form. Information is not a thing but a transformation—the dynamic bridge between low-dimensional data and high-dimensional meaning. Failures in semantic transfer are due to this compression and decompression of high-dimensional meaning into 3-dimensional data that can be transmitted between bounded neural networks.

Shannon’s theory optimized for moving data through channels. The Conveyance Hypothesis addresses what happens when that data is the result of semantic transformation in high-dimensional spaces or induces transformation in high-dimensional spaces—the meaning transfer that Shannon deliberately set aside.

1.1 Why This Matters

The “Internet of Things” is rapidly transforming into the “Internet of Agents.” Autonomous systems—from robotic platforms to financial trading algorithms to multi-agent orchestration frameworks—are proliferating across every domain. These agents must communicate not only with humans but increasingly with each other.

When Agent *A* transmits a message to Agent *B*, how do we verify that semantic content was preserved? How do we diagnose failures when autonomous systems miscommunicate?

These questions require examining both sides of the transfer:

- **Sender side:** Does the output faithfully represent the agent’s internal state?
- **Receiver side:** Does the incoming message successfully integrate into the receiver’s knowledge structure?

We term these outputs **boundary objects**—the low-dimensional representations that agents create to externalize their high-dimensional internal states. Boundary objects are the only artifacts that can traverse the gap between agents; the high-dimensional knowledge within a bounded neural network cannot be directly transmitted through low-dimensional 3D spacetime.

This raises a fundamental question at the heart of the Conveyance Hypothesis: *How do we verify that the boundary objects created by an agent are faithful representations of its internal semantic state?*

Without a theoretical framework for measuring this fidelity—and for measuring how successfully boundary objects induce appropriate semantic transformation in receiving agents—we lack the tools to interpret what happens as information moves between bounded networks.

The stakes are significant. Multi-agent systems coordinating physical actions, managing critical infrastructure, or making consequential decisions require reliable semantic transfer. As agent-to-agent communication becomes foundational to technological infrastructure, the absence of measurement frameworks becomes an increasingly urgent gap.

Understanding how information transforms as it moves between disparate bounded networks—each with its own internal geometry and representational structure—requires theoretical grounding. The Conveyance Hypothesis aims to provide that foundation.

1.2 Why This Is Feasible Now

Three technological developments suggest the question may now be tractable:

1. **High-dimensional embedding spaces:** Modern language models encode semantic relationships in geometric structures. Words, sentences, and documents occupy positions in high-dimensional space where distance correlates with semantic similarity. This gives us a measurable substrate for semantic content.
2. **Observable internal representations:** Unlike human minds, artificial neural networks offer visibility into their internal states—hidden layer activations, attention patterns, and geometric properties can be examined as information processes through the system. This visibility is imperfect; billions of parameters, polysemantic neurons, and distributed representations present significant interpretive challenges. The black box is not transparent, but it is instrumentable in ways impossible with biological cognition.
3. **Controlled experimental conditions:** AI-to-AI communication provides a laboratory setting where both sender and receiver internal states are observable, enabling bilateral measurement impossible with human subjects. We can track information from origin to destination.

These capabilities enable a research program that Shannon’s era could not pursue: the mathematical characterization of semantic transfer effectiveness.

2 Theoretical Foundation

2.1 Shannon’s Deliberate Exclusion (Expanded)

In 1948, Claude Shannon explicitly excluded meaning from his theory:

“The semantic aspects of communication are irrelevant to the engineering problem.”

This wasn’t naïve—Shannon understood meaning mattered, but 1948 technology couldn’t measure it. His theory optimized for **signal fidelity**: how accurately can bits travel from sender to receiver?

The Conveyance Hypothesis addresses what Shannon excluded: the effectiveness of semantic transformation between agents.

Aspect	Shannon Theory	Conveyance Hypothesis
Concern	Signal fidelity	Semantic effectiveness
Measures	Bit error rate, channel capacity	D_{eff} , bilateral C_{pair}
Question	Did the signal arrive intact?	Did meaning transform the receiver?
Agents	Sender \rightarrow Channel \rightarrow Receiver	Internal geometries interacting

Table 1: Comparison of Shannon Theory and Conveyance Hypothesis

2.2 Actor-Network Theory: Translation, Not Transmission

Drawing from Bruno Latour’s Actor-Network Theory, we recognize that information transfer is **translation**, not transmission:

“There is no society, no social realm, and no social ties, but there exists translations between mediators that may generate traceable associations.”

Key insight: Information doesn’t flow like water through pipes. Every transfer is a creative transformation where **both sender and receiver are modified** through their interaction.

2.3 Boundary Objects

Following Susan Leigh Star’s work, effective transfer between agents with different internal structures requires **boundary objects**—intermediate representations that maintain coherence across contexts.

In our framework, C_{ext} (external shared context) represents boundary objects that:

- Preserve sufficient structure for reconstruction
- Compress high-dimensional internal states into transferable form
- Expand appropriately in the receiver’s distinct geometric space

2.4 The Data vs. Information Distinction

Data Static patterns, boundary-preserving, exists without observers

Information Dynamic transformation, boundary-crossing, requires agent interaction

Data becomes information only when it transforms an agent’s internal state. A book on a shelf contains data; reading it creates information through the transformation of the reader’s understanding.

3 The Mathematical Framework

3.1 Primary Conveyance Equation (v4.0)

Bilateral conveyance effectiveness between agents i and j :

$$C_{\text{pair}}(i \leftrightarrow j) = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \times P_{ij} \quad (1)$$

Components:

Symbol	Name	Meaning
C_{pair}	Bilateral Conveyance	Overall transfer effectiveness between two agents
Hmean	Harmonic Mean	Captures bilateral constraint (weakest link dominates)
C_{out}	Output Capacity	Sender’s ability to encode meaning
C_{in}	Input Capacity	Receiver’s ability to integrate meaning
$f_{\text{dim}}(D_{\text{eff}})$	Dimensional Function	Richness of semantic representation
P_{ij}	Protocol Compatibility	How well agents’ interfaces match $[0, 1]$

Table 2: Components of the Primary Conveyance Equation

3.2 Why Harmonic Mean?

The harmonic mean is chosen deliberately:

$$\text{Hmean}(a, b) = \frac{2ab}{a + b} \quad (2)$$

Property: The harmonic mean is dominated by the smaller value.

- If $C_{\text{out}} = 0.9$ and $C_{\text{in}} = 0.1$, then $\text{Hmean} = 0.18$
- Excellent sender + poor receiver = poor conveyance

This matches intuition: a brilliant lecturer teaching in a language students don’t understand achieves low conveyance regardless of lecture quality.

3.3 Component Equations

Individual agent conveyance capacity decomposes as:

$$C_{\text{out}}(i \rightarrow j) = \frac{W \times R \times H}{T} \quad (3)$$

$$C_{\text{in}}(j \leftarrow i) = \frac{W \times R \times H}{T} \quad (4)$$

Variable Definitions:

Variable	Name	Concept	How to Measure
W	Semantic Investment	Computational allocation	Hidden state activation patterns
R	Relational Discovery	Geometric positioning quality	Graph-theoretic embedding properties
H	Computational Frame	Processing throughput	Attention efficiency, layer utilization
T	Temporal Investment	Total computational budget	Token count, processing time

Table 3: Variable Definitions for Component Equations

3.4 Why Multiplicative?

The framework uses **multiplication** rather than addition because:

- $W = 0$ (zero semantic content) \rightarrow Nothing to transfer \rightarrow Zero conveyance
- $R = 0$ (zero relational structure) \rightarrow No geometric organization \rightarrow Zero conveyance
- $H = 0$ (zero processing) \rightarrow Cannot utilize signals \rightarrow Zero conveyance

- $P_{ij} = 0$ (zero protocol match) \rightarrow Cannot communicate \rightarrow Zero conveyance

A single zero produces zero output. This “zero-propagation” principle explains why communication failures are often catastrophic rather than gradual—one essential missing component collapses the entire transfer.

3.5 Dimensional Richness Function

The dimensional function scales conveyance by how much semantic structure is preserved:

$$f_{\text{dim}}(D_{\text{eff}}) = \left(\frac{D_{\text{eff}}}{D_{\text{target}}} \right)^{\alpha_{\text{dim}}} \quad (5)$$

Where: D_{eff} = Effective dimensionality (measured via PCA)

D_{target} = Target dimensionality for the layer/system

$\alpha_{\text{dim}} \in [0.5, 1.0]$ (empirically determined scaling factor)

Interpretation:

- $D_{\text{eff}}/D_{\text{target}} = 1.0 \rightarrow$ Full dimensional preservation $\rightarrow f_{\text{dim}} = 1.0$
- $D_{\text{eff}}/D_{\text{target}} = 0.5 \rightarrow$ Half dimensions preserved $\rightarrow f_{\text{dim}} \approx 0.5\text{--}0.7$
- $D_{\text{eff}}/D_{\text{target}} \rightarrow 0 \rightarrow$ Dimensional collapse $\rightarrow f_{\text{dim}} \rightarrow 0$

4 Key Metrics

4.1 D_{eff} (Effective Dimensionality) — PRIMARY METRIC

Definition: The number of independent semantic dimensions preserved during processing, computed via PCA with 90% variance threshold.

```
def compute_d_eff(embeddings, variance_threshold=0.90):
    """
    Count dimensions capturing 90% of variance.

    CRITICAL: L2 normalize embeddings first to prevent
    magnitude artifacts from dominating variance.
    """
    # Center and compute covariance
    centered = embeddings - embeddings.mean(axis=0)
    cov = centered.T @ centered / (len(embeddings) - 1)

    # Eigendecomposition
    eigenvalues = np.linalg.eigvalsh(cov)[::-1] # Descending

    # Cumulative variance ratio
    cumvar = np.cumsum(eigenvalues) / eigenvalues.sum()

    # Count dimensions below threshold
    d_eff = np.searchsorted(cumvar, variance_threshold) + 1

    return d_eff
```

Why 90% threshold?

- Established compromise between signal preservation and noise reduction
- Components beyond 90% typically capture noise/artifacts, not semantics
- Robust across diverse domains (neural activity, manifold learning, NLP)

Target Values:

Nominal Dimension	Healthy D_{eff}	Collapse Warning
512D	≥ 34	< 20
256D	≥ 20	< 12
128D	≥ 12	< 8
64D	≥ 8	< 5
24D	≥ 20	< 12

Table 4: Target Values for Effective Dimensionality

4.2 β (Beta) — DIAGNOSTIC METRIC (NOT Optimization Target)

Definition: Collapse indicator measuring dimensional compression during processing.

$$\beta = \frac{D_{\text{eff}}^{\text{input}}}{D_{\text{eff}}^{\text{output}}} \quad (6)$$

CRITICAL: β is a **diagnostic warning signal**, not something to optimize. High β indicates information loss through dimensional collapse.

Interpretation:

β Value	Status	Meaning
< 2.0	Healthy	Low collapse, good generalization expected
2.0–2.5	Warning	Moderate collapse, acceptable but monitor
2.5–3.0	Concerning	High collapse, likely overfitting
> 3.0	Critical	Severe collapse, investigate immediately

Table 5: Interpretation of Beta Values

Empirical Finding: β shows strong negative correlation with task performance ($r \approx -0.92$ in preliminary experiments). Lower β = better generalization.

4.3 Secondary Geometric Metrics

Metric	Symbol	Target	Meaning
Mean k-NN Distance	d_{nn}	0.10–0.15	Moderate clustering
Label Consistency	LC	≥ 0.87	Neighbors share semantic categories
Boundary Sharpness	σ_{boundary}	0.40–0.50	Balanced separation

Table 6: Secondary Geometric Metrics

Metric	Target	Use Case
F1 Score	≥ 0.90 (strong), ≥ 0.85 (acceptable)	Classification
Recall@10	≥ 0.85	Retrieval
Perplexity	Lower is better	Language modeling

Table 7: Task Performance Validation Metrics

4.4 Task Performance — VALIDATION METRICS

Primary validation is always downstream task performance, not geometric metrics.

Geometric metrics are diagnostic. They help explain *why* task performance is good or bad, but task performance is ground truth.

5 Core Hypotheses (Under Investigation)

5.1 Low-Dimensional Hypothesis

Prediction: External shared context (C_{ext}) performs better at 128–256 dimensions than higher dimensions.

Rationale:

- Forcing low dimensions makes geometric relationships carry semantic meaning
- High dimensions allow magnitude artifacts to dominate
- BDH architecture independently arrived at $d = 256$ as optimal bottleneck

Falsification: High-dimensional representations (512D+) consistently outperform low-dimensional across tasks.

5.2 β -Overfitting Hypothesis

Prediction: $\beta \in [1.5, 2.0]$ correlates with better generalization; higher β indicates overfitting.

Rationale:

- Dimensional collapse destroys information needed for generalization
- Over-compressed representations memorize rather than generalize

Preliminary Evidence: $r(\beta, F1) \approx -0.92$ from limited experiments

Falsification: β shows positive correlation with task performance across domains.

5.3 Attention-Only Hypothesis

Prediction: Attention mechanisms outperform rigid boundary scaffolding for dimensional preservation.

Preliminary Evidence:

- Attention-only: $D_{\text{eff}} = 34$ (preserved)
- Boundary scaffolding: $D_{\text{eff}} = 6$ (collapsed, -83% loss)

Falsification: Boundary scaffolding systematically wins in controlled A/B tests.

5.4 Bilateral Asymmetry Hypothesis

Prediction: In adversarial or misaligned contexts, $C_{A \rightarrow B} \neq C_{B \rightarrow A}$ (asymmetric conveyance).

Implication: Misaligned AI systems might show high $C_{\text{AI} \rightarrow \text{Human}}$ (they understand us) but low $C_{\text{Human} \rightarrow \text{AI}}$ (we don't understand them).

Falsification: Bilateral measurements show no predictive value for alignment detection.

6 Context Amplification

6.1 Original Formulation (Superseded)

Earlier versions used exponential context amplification:

$$C_{\text{pair}} = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times C_{\text{ext}}^{\alpha} \times P_{ij} \quad (7)$$

Where $\alpha \in [1.5, 2.0]$ (super-linear amplification)

This predicted that context quality has super-linear effects—doubling context quality more than doubles conveyance effectiveness.

6.2 Current Formulation (v4.0)

Framework v4.0 replaces exponential C_{ext} with dimensional preservation function:

$$C_{\text{pair}} = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \times P_{ij} \quad (8)$$

Key insight: Context amplification occurs through **dimensional preservation**, not exponential scaling. Good context maintains D_{eff} ; bad context causes dimensional collapse.

6.3 Geometric Extension (For Advanced Analysis)

When considering manifold structure, the complete formulation includes curvature and geodesic effects:

$$\begin{aligned} C_{\text{pair}}^{\text{geometric}} &= \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \\ &\quad \times \exp\left(-\frac{\lambda}{\tau^2}\right) && \text{[curvature penalty]} \\ &\quad \times \exp\left(-\frac{\text{dist}_M^2}{2\sigma^2}\right) && \text{[geodesic distance decay]} \\ &\quad \times P_{ij} \end{aligned} \quad (9)$$

Where:

- τ = local reach (inverse maximum curvature)
- λ = curvature sensitivity parameter
- dist_M = geodesic distance on semantic manifold

Interpretation: Information flows efficiently through low-curvature regions (within semantic categories) and less efficiently across high-curvature boundaries (between categories).

7 Temporal Dynamics

7.1 Multi-Turn Context Evolution

In multi-turn interactions, context evolves over time:

$$C_{\text{ext}}(t) = f(C_{\text{ext}}(t-1), B_t) \quad (10)$$

Where B_t = boundary object at turn t .

Quality Trajectory:

$$\text{quality_trajectory}(t) = \prod_{k=1}^t q(B_k) \quad (11)$$

Where $q(B_k) \in [0, 1]$ = quality of boundary object k .

Critical insight: Quality is multiplicative across turns. A single low-quality exchange ($q \approx 0.3$) can severely degrade the entire trajectory.

7.2 Self-Reinforcing Cycles

Positive cycle: Good context \rightarrow better responses \rightarrow improved context $\rightarrow \dots$

Negative cycle: Poor context \rightarrow confused responses \rightarrow degraded context $\rightarrow \dots$

This explains why early interactions disproportionately determine outcomes—they set the trajectory’s initial slope.

7.3 Threshold-Based Management

To prevent catastrophic trajectory degradation:

```
if gap(B_t, expected) < theta_refine:
    # Minor gap: refine and continue
    B_t_prime = refine(B_t, feedback)

elif gap(B_t, expected) > theta_reset:
    # Major gap: reset to last good checkpoint
    context = restore_checkpoint(t_checkpoint)
```

8 Zero-Propagation Principle

8.1 Definition

Zero-propagation occurs when any essential component of conveyance equals zero:

$$\text{If } W = 0 \text{ OR } R = 0 \text{ OR } H = 0 \text{ OR } P_{ij} = 0 \text{ OR } D_{\text{eff}} \rightarrow 0 : \quad C_{\text{pair}} = 0 \text{ (categorical failure)} \quad (12)$$

8.2 Implications

Zero-propagation is **categorical failure**, distinct from “very poor” conveyance:

Example: A perfect lecture ($W = 1.0$, $R = 1.0$, $H = 1.0$) in a language no student speaks ($P_{ij} = 0$) achieves zero conveyance—not poor conveyance, but zero.

Condition	Result	Nature
All components > 0 but low	Low C_{pair}	Degraded but possible
Any component $= 0$	$C_{\text{pair}} = 0$	Impossible transfer

Table 8: Zero-Propagation Implications

8.3 Dimensional Collapse as Zero-Propagation

When D_{eff} collapses to near-zero, effective conveyance becomes impossible even with good W , R , H , P_{ij} values:

$$D_{\text{eff}} < \text{threshold} \rightarrow f_{\text{dim}}(D_{\text{eff}}) \rightarrow 0 \rightarrow C_{\text{pair}} \rightarrow 0 \quad (13)$$

This explains why memory poisoning attacks with only $\sim 10\%$ contamination can cause $\sim 95\%$ task failure—small contamination triggers dimensional collapse, which cascades to zero-propagation.

9 Falsification Criteria

A hypothesis must be falsifiable to be scientific. The Conveyance Hypothesis would be **falsified** by:

9.1 Strong Falsification Evidence

1. **β shows consistent positive correlation with performance across domains**
 - Current observation: $r \approx -0.92$ (negative)
 - Falsifying observation: $r > +0.5$ replicated across tasks
2. **High-dimensional C_{ext} systematically outperforms low-dimensional**
 - Current hypothesis: 128–256D optimal
 - Falsifying observation: 2048D+ consistently superior
3. **Boundary scaffolding beats attention-only in rigorous A/B tests**
 - Current observation: Attention-only preserves $D_{\text{eff}} = 34$; scaffolding collapses to $D_{\text{eff}} = 6$
 - Falsifying observation: Scaffolding wins majority of comparisons
4. **Bilateral conveyance measurements show zero predictive validity**
 - Current hypothesis: Asymmetric C_{pair} predicts misalignment
 - Falsifying observation: No correlation between C_{pair} asymmetry and outcomes
5. **P_{ij} compatibility shows no relationship to transfer success**
 - Current hypothesis: Protocol match enables transfer
 - Falsifying observation: Incompatible agents transfer equally well

9.2 Weak Falsification Evidence

- Single counterexamples (might be domain-specific)
- Mixed results without clear patterns
- Inability to measure proposed constructs reliably

10 Relationship to Existing Theories

10.1 Shannon’s Information Theory

Shannon	Conveyance
Channel capacity	Agent capacity ($W \times R \times H$)
Noise	Protocol mismatch ($1 - P_{ij}$)
Encoding	Boundary object creation
Decoding	Integration into receiver geometry
Bit error rate	Dimensional collapse (β)

Table 9: Mapping Shannon Concepts to Conveyance

Conveyance extends Shannon by adding semantic effectiveness to signal fidelity.

10.2 Rogers’ Innovation Diffusion Theory

Rogers	Conveyance
Adoption curves	Conveyance effectiveness over time
Opinion leaders	High-conveyance nodes in networks
Compatibility	P_{ij} protocol coefficient
Complexity	Inverse of D_{eff} preservation

Table 10: Mapping Rogers’ Concepts to Conveyance

Conveyance mathematizes Rogers’ qualitative descriptions of how innovations spread.

10.3 Kolmogorov Complexity

Kolmogorov	Conveyance
Minimum description length	Optimal boundary object compression
Incompressibility	Essential semantic structure
Algorithmic probability	Transfer success probability

Table 11: Mapping Kolmogorov Concepts to Conveyance

Conveyance operationalizes complexity concepts for agent-to-agent transfer.

11 Practical Applications (If Validated)

11.1 AI Development

- **Optimization targets:** Maximize D_{eff} rather than arbitrary metrics

- **Diagnostic tools:** Detect dimensional collapse before deployment
- **Architecture guidance:** Prefer attention-only over scaffolding
- **Training monitoring:** Watch geometric health during learning

11.2 AI Safety

- **Alignment detection:** Misaligned agents may show asymmetric C_{pair}
- **Early warning:** Geometric anomalies before behavioral symptoms
- **Interpretability:** Ground-truth about what information actually transferred

11.3 Memory Systems

- **Poisoning detection:** Dimensional collapse indicates contamination
- **Quality maintenance:** Monitor D_{eff} trajectory across interactions
- **Defense mechanisms:** Reset when D_{eff} drops below threshold

11.4 Human Communication (Speculative)

If the framework validates in AI systems, it may inform:

- Educational theory (why some teaching works)
- Organizational communication (why information gets lost in hierarchies)
- Cross-cultural understanding (how meaning transforms across contexts)

12 Current Evidence Status

12.1 Validated (Tier 1 Evidence)

- ✓ Dimensional richness correlates positively with utility
- ✓ β anti-correlates with utility ($r = -0.92$)
- ✓ Attention-only architecture preserves $D_{\text{eff}} = 34$
- ✓ Boundary scaffolding collapses to $D_{\text{eff}} = 6$ (-83% loss)
- ✓ L2 normalization prevents magnitude artifacts

12.2 Preliminary Observations (Require Validation)

- △ Low-dimensional (128–256D) outperforms high-dimensional
- △ $\beta \in [1.5, 2.0]$ optimal range
- △ BDH’s $d = 256$ bottleneck validates independently

12.3 Unvalidated (Theoretical)

- ? Bilateral asymmetry predicts misalignment
- ? Curvature-modulated conveyance
- ? Temporal amplification (T^β term)
- ? Human communication applications

12.4 Falsified (Revised in v3.9+)

- × Boundary-first approach (v3.7–3.8) — produced anti-utility
- × β as optimization target — now diagnostic only
- × ϕ (conductance) and κ (curvature) as primary metrics — deprecated

13 Conclusion

The Conveyance Hypothesis proposes that **semantic transfer effectiveness is mathematically measurable**. We offer:

1. **A core equation:** $C_{\text{pair}} = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \times P_{ij}$
2. **Measurable variables:** $W, R, H, T, D_{\text{eff}}, \beta, P_{ij}$
3. **Primary metric:** D_{eff} (effective dimensionality via PCA)
4. **Diagnostic metric:** β (dimensional collapse indicator)
5. **Falsification criteria:** Clear conditions that would disprove the hypothesis
6. **Preliminary validation:** β -utility anti-correlation, attention-only superiority

Status: This is a **hypothesis under investigation**, not a validated theory. The equations are mathematically coherent but require systematic empirical validation.

Core claim: Clean math \neq empirical reality. We have elegant equations that need rigorous testing. The Low-Dimensional Hypothesis, β -overfitting correlation, and bilateral asymmetry predictions all await experimental validation.

Shannon’s exclusion of meaning was wise for 1948. In 2025, with transformer architectures providing measurable embedding spaces, we may finally have the tools to include what he deliberately left out.

A Quick Reference

Primary Equation

$$C_{\text{pair}}(i \leftrightarrow j) = \text{Hmean}(C_{\text{out}}, C_{\text{in}}) \times f_{\text{dim}}(D_{\text{eff}}) \times P_{ij} \quad (14)$$

Component Equations

$$C_{\text{out}} = \frac{W \times R \times H}{T} \quad (15)$$

$$C_{\text{in}} = \frac{W \times R \times H}{T} \quad (16)$$

$$f_{\text{dim}}(D_{\text{eff}}) = \left(\frac{D_{\text{eff}}}{D_{\text{target}}} \right)^{\alpha_{\text{dim}}} \quad (17)$$

Variables

Symbol	Name	Range
W	Semantic Investment	$[0, 1]$
R	Relational Discovery	$[0, 1]$
H	Computational Frame	$[0, 1]$
T	Temporal Investment	$[0, \infty)$
D_{eff}	Effective Dimensionality	$[1, D_{\text{nominal}}]$
β	Collapse Indicator	$[1, \infty)$
P_{ij}	Protocol Compatibility	$[0, 1]$
α_{dim}	Dimensional Scaling	$[0.5, 1.0]$

Table 12: Variable Summary

Target Values

Metric	Target	Warning
D_{eff} (512D)	≥ 34	< 20
D_{eff} (256D)	≥ 20	< 12
β	< 2.0	> 2.5
d_{nn}	0.10–0.15	< 0.05 or > 0.25
LC	≥ 0.87	< 0.70
F1	≥ 0.90	< 0.85

Table 13: Target Values Summary

Critical Rules

1. **ALWAYS** L2 normalize embeddings before geometric analysis
2. **NEVER** optimize for β —it’s diagnostic only
3. **PRIMARY** validation is task performance, not geometric metrics
4. **WATCH** for dimensional collapse (D_{eff} dropping rapidly)

B Version History

Version	Date	Key Changes
v1.0	2024	Initial formulation with C_{ext}^α
v3.7–3.8	Oct 2025	Boundary-first approach (later falsified)
v3.9	Oct 2025	D_{eff} as primary metric, β inversion discovered
v4.0	Nov 2025	$f_{\text{dim}}(D_{\text{eff}})$ replaces C_{ext}^α , attention-only validated

Table 14: Version History

“The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.” — Claude Shannon, 1948

“The fundamental problem of conveyance is that of transforming at one agent a semantic structure that appropriately reorganizes another agent’s internal geometry.”
— The Conveyance Hypothesis, 2025

References

- [1] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.
- [2] Weaver, W. (1949). Recent contributions to the mathematical theory of communication. In Shannon, C. E. and Weaver, W., editors, *The Mathematical Theory of Communication*, pages 1–28. University of Illinois Press.