The Communication Bridge Hypothesis

Applying Biological Signaling Theory and Information-Theoretic Analysis to UAP Behavioral Sequences

Foundation for AI Ethics & Safety Research (FAESR)
Shenandoah Valley, Virginia

Working Paper — February 2026

Abstract

This paper proposes and tests a formal analytical framework for evaluating whether unidentified anomalous phenomena (UAP) exhibit behavioral patterns consistent with intentional communication. Drawing on biological signaling theory (Maynard Smith & Harper 2003), Tomasello's primate intentionality criteria (2008), and information-theoretic measures validated in interstellar signal detection (Mahon 2025) and cetacean communication research (McCowan 1999), the framework defines eleven testable communication criteria and applies them to four UAP cases selected for multi-sensor confirmation, official documentation, prosaic explanation failure, and documented behavioral interaction: the RB-47H encounter (1957), Minot Air Force Base (1968), Tehran (1976), and the USS *Nimitz* "Tic Tac" (2004). All four cases satisfy a minimum of 10 out of 11 criteria; two satisfy all 11. Three behavioral patterns—graduated escalation, aposematic display, and response-dependent modification—appear independently in every case across 47 years and four national military services with zero shared witnesses. Information-theoretic analysis (Shannon entropy, Zipf's law, Kolmogorov complexity, transfer entropy, and local compositional complexity) demonstrates that UAP behavioral sequences exhibit non-random structure significantly exceeding chance expectations (compression $p = 0.015$). Per-case transfer entropy analysis reveals asymmetric directed information flow from human actions to UAP responses in the two longest cases (Tehran: $TE(H \rightarrow U)/TE(U \rightarrow H) = 1.7\times$; Minot: $2.6\times$)—the operational definition of "the object responded to the observer." The paper provides full encoding tables, replication code, explicit falsification criteria, and a power analysis specifying the dataset sizes required for definitive statistical conclusions.

Keywords: unidentified anomalous phenomena, UAP, biological signaling theory, information theory, Shannon entropy, transfer entropy, aposematic signaling, graduated signaling, intentional communication, Tomasello criteria, Kolmogorov complexity

1. Introduction

1.1 The Behavioral Question

In December 2017, the New York Times revealed the existence of the Advanced Aerospace Threat Identification Program (AATIP), a Pentagon program that had spent five years investigating military encounters with unidentified aerial objects. In the years that followed, the U.S. government took a series of steps that would have been unthinkable a decade earlier: the Department of Defense confirmed the authenticity of three infrared videos depicting unidentified objects (2020); the Director of National Intelligence submitted a preliminary assessment acknowledging 144 UAP reports from military sources, of which only one could be explained (2021); Congress created the All-domain Anomaly Resolution Office (AARO) within the Department of Defense (2022); and military pilots testified under oath before the House Oversight Committee about encounters involving objects that demonstrated capabilities exceeding any known technology (2023). AARO's own reporting identified 21 cases exhibiting anomalous characteristics that resist conventional explanation, including objects that "trailed or shadowed" military assets—a behavioral descriptor that implies awareness of, and orientation toward, specific human platforms.

This institutional trajectory has established, as a matter of official record, that military personnel using calibrated sensor systems have documented encounters with objects whose physical characteristics and behavior are not explained by any known technology or natural phenomenon. What it has not done—what no analytical effort in the public domain has systematically attempted—is to examine the *behavioral content* of these encounters with the same rigor applied to their physical characteristics.

The physical question—what are these objects and how do they move?—has received extensive attention, from SCU's kinematic analyses to AARO's sensor-data collection programs. The behavioral question—why do they move the way they do in relation to human observers?—has not. Yet the behavioral data are, in certain respects, more diagnostic than the physical data. An object's acceleration profile tells us what it can do; its behavioral profile—the sequence of actions it takes in response to what human observers do—tells us whether there is an agent behind those capabilities, and if so, what that agent may be attempting to accomplish.

1.2 The Communication Bridge

This paper asks a specific question: do UAP behavioral sequences, when analyzed with the tools developed for detecting communication in biological and interstellar contexts, exhibit structural properties consistent with intentional signaling?

The question is motivated by a simple observation. In the best-documented military UAP encounters, the objects do not behave randomly. They do not behave like natural phenomena. And they do not behave like adversary surveillance platforms, which would prioritize remaining undetected. Instead, they exhibit a distinctive behavioral profile: they appear in the vicinity of military platforms; they allow themselves to be detected on multiple sensor channels; they respond to observer actions with contextually appropriate adjustments in their own behavior; they demonstrate capabilities far exceeding those of the observing platform; and they do all of this without hostile action. This profile—conspicuous presence, graduated responsiveness, capability demonstration, and restraint—maps precisely onto the behavioral signatures that biologists use to identify intentional communication in non-human species.

The "bridge" in the paper's title is the analytical connection between two mature scientific disciplines—biological signaling theory and information science—and a phenomenon that has historically been studied outside the scientific mainstream. The bridge runs in both directions: signaling theory provides the criteria for evaluating whether UAP behavior constitutes communication, and information theory provides the quantitative tools for measuring the structural properties of that behavior without requiring the analyst to understand its content. Together, they offer a methodology that is falsifiable, replicable, and independent of assumptions about the nature or origin of the phenomenon.

1.3 Prior Work and Novel Contribution

The idea that UAP behavior may exhibit communicative properties is not new. Vallsée (1969, 1990) noted recurrent patterns in UAP encounter reports and proposed classificatory schemes that implicitly assumed non-random behavior. Hynek (1972) organized encounter types by proximity and interaction modality—a behavioral taxonomy, though he did not frame it in those terms. More recently, Wendt and Duvall (2008) argued from a political-science perspective that the "taboo" around UAP research constituted an anthropocentric refusal to engage with the possibility of non-human agency—a sociological analysis with implications for the behavioral question.

What has been absent from the literature is a rigorous, testable analytical framework that connects these observations to established scientific methodology. This paper provides that framework by integrating three bodies of work that have not previously been applied to UAP behavioral data:

> **Maynard Smith and Harper's biological signaling theory** (**2003**), which defines the conditions under which a behavior qualifies as a signal rather than a cue or random event. This framework is the standard analytical tool in behavioral ecology for distinguishing

intentional communication from coincidence, and it has been validated across species from insects to primates.

Tomasello and colleagues' hierarchy of intentional communication (e.g., 2008), developed from primate gesture research, which defines nine criteria for identifying intentional signaling in non-verbal, non-linguistic behavioral sequences. These criteria were designed precisely for the challenge this paper faces: detecting communication in an agent whose cognitive architecture, sensory capabilities, and communicative conventions are unknown.

Information-theoretic measures validated in interstellar signal detection and animal communication research, including Shannon entropy, Zipf's law, Kolmogorov complexity approximation, transfer entropy (Schreiber 2000), and local compositional complexity (Mahon 2025). These tools quantify the structural properties of sequences without requiring the analyst to understand their content—the methodological prerequisite for analyzing UAP behavioral data, where the "language" (if any) is unknown.

The novel contribution is not any single element but their integration into a unified analytical pipeline: qualitative criteria derived from biological signaling theory, applied to carefully selected cases meeting strict evidential standards, with results independently tested by quantitative information-theoretic measures. The pipeline is designed to be replicable by independent researchers using the encoding tables and Python code provided in the appendices.

1.4 Scope and Limitations

The paper analyzes four cases. This is a small sample, chosen for evidential quality rather than statistical representativeness (the selection criteria are defined in Section 3). The quantitative results are preliminary; the power analysis in Section 5.7 states explicitly what dataset sizes would be required for definitive statistical conclusions. The paper does not claim to prove the communication hypothesis; it claims to provide a falsifiable framework that produces suggestive preliminary results across multiple independent measures.

The paper is ontologically agnostic. It makes no claim about the nature, origin, or ultimate purpose of UAP. The analytical framework detects structural signatures of communication; it cannot distinguish between deliberate signaling by an intelligent agent and an unknown natural or technological process that coincidentally produces equivalent sequential structure. It does not and cannot determine who or what is communicating. The findings carry the same evidential weight whether the agent is extraterrestrial, interdimensional, a terrestrial non-human intelligence, or an unknown process that produces communication-like behavioral structure.

The paper does not address the full range of reported UAP phenomena. It examines only cases with multi-sensor military documentation of behavioral interaction—a narrow subset of all UAP reports, but the subset most amenable to rigorous analysis. Cases lacking multi-sensor confirmation, official documentation, or documented behavioral sequences are excluded regardless of their other qualities.

1.5 Paper Structure

Section 2 establishes the theoretical framework, integrating biological signaling theory, Tomasello's intentionality criteria, and the information-theoretic toolkit. Section 3 defines the evidence criteria for case selection and establishes the 1957–2004 temporal window. Section 4 presents the four qualifying cases with identical analytical structure: context, sensors, documentation, behavioral sequence timeline, skeptical assessment, and communication criteria evaluation. Section 5 develops the quantitative methodology—encoding, Shannon entropy, Zipf's law, compression, transfer entropy, local compositional complexity—and presents preliminary results with an explicit power analysis. Section 6 examines the nuclear correlation as statistical context. Section 7 evaluates competing hypotheses and states falsification criteria. Section 8 presents implications and concrete recommendations for advancing the research program.

Appendix A provides complete encoding tables for all four cases. Appendix B provides Python replication code for the full information-theoretic battery with sensitivity analysis. Appendix C provides annotated source documentation for each case.

The paper is structured for two audiences: researchers in behavioral ecology, information science, and related fields who may be encountering UAP data for the first time, and UAP researchers who may be encountering formal communication analysis for the first time. Both audiences will find sections that cover familiar ground; the novel contribution lies in the *integration*.

2. What Counts as Communication: The Theoretical Framework

Before examining whether UAP behavioral sequences exhibit communication structure, we must define precisely what "communication" means in formal scientific terms and specify which analytical tools can detect it. This section establishes the criteria drawn from three mature fields —biological signaling theory, comparative psychology, and information theory—that will be applied to the evidence in subsequent sections. The deliberate choice to define the analytical framework before presenting any UAP data serves a methodological purpose: it prevents the criteria from being tailored to fit the evidence and allows readers to evaluate the framework's rigor independently of its application.

Three complementary approaches are presented. First, biological signaling theory provides a formal definition of what constitutes a signal in nature, derived from decades of research on animal communication systems. Second, intentionality criteria from comparative psychology offer a checklist for distinguishing deliberate communication from reflexive or coincidental behavior. Third, information-theoretic methods provide quantitative tools that can detect structured communication in any sequential data—without requiring comprehension of the message itself. Together, these frameworks constitute a detection apparatus that has been validated across species ranging from ants to cetaceans and, in one notable case, on a message transmitted by humans to an imagined extraterrestrial audience.

2.1 Biological Signaling Theory

The foundational definition of a biological signal comes from Maynard Smith and Harper's *Animal Signals* (2003), which synthesized decades of theoretical and empirical work in behavioral ecology. Under their framework, a signal is an act or structure that satisfies three conditions simultaneously: (i) it alters the behavior of another organism, which is termed the receiver; (ii) it evolved—or, more broadly, exists—because of that effect on the receiver; and (iii) it is effective because the receiver's response has itself evolved or developed in a way that makes it sensitive to the signal. This tripartite definition is deliberately stringent. It excludes incidental cues (a predator hearing the noise of prey moving through underbrush), unintended consequences (a flower's color attracting pollinators as a byproduct of UV protection), and deceptive mimicry that has not yet been recognized by receivers.

Applying this framework to an unknown intelligence requires careful adaptation. Criterion (ii) —that the signal exists because of its effect—cannot be verified directly for an agent whose design history is unknown. However, it can be operationalized: if a behavior's spatial distribution, temporal patterning, and contextual selectivity are consistent with a signaling function, the criterion is provisionally satisfied. A behavior that occurs preferentially near nuclear weapons installations, during military operations, and in the presence of sensor-equipped platforms exhibits the kind of non-random, context-dependent deployment expected of a signal rather than a random phenomenon.

Criterion (iii)—that the receiver's response has adapted—is straightforwardly met in the UAP context: human military and intelligence organizations have developed extensive detection, tracking, intercept, and reporting protocols specifically in response to UAP encounters, constituting an evolved institutional response.

This adaptation challenge is not unique to the UAP problem. The Search for Extraterrestrial Intelligence (SETI) faces an identical difficulty: criteria (ii) and (iii) of the Maynard Smith-Harper definition cannot be confirmed in advance for a signal from an unknown sender, because we have no access to the sender's evolutionary or design history. SETI resolves this by looking for structural features consistent with intentional signaling—narrowband electromagnetic emissions, mathematical regularities, non-natural spectral profiles. The approach proposed here is analogous: we examine whether behavioral sequences are structurally consistent with communication, using the same logical framework applied to a different data modality.

2.2 Intentionality Criteria from Comparative Psychology

While the Maynard Smith-Harper framework defines what a signal is, it does not distinguish between automated signaling (a firefly's bioluminescent flash, triggered by a fixed neural circuit) and intentional communication (a chimpanzee adjusting its gestural repertoire based on the attentional state of its audience). For the communication hypothesis to have force beyond mere stimulus-response, we require evidence of flexible, context-sensitive signaling—the hallmarks of intentional communication as defined in comparative psychology.

Tomasello and colleagues, drawing on two decades of primate gestural communication research (Tomasello 2008), identified nine behavioral criteria that collectively distinguish intentional communication from reflexive or automated signaling. Although originally developed for great ape gesture, these criteria have been productively extended to corvids, cetaceans, and other taxa, establishing them as general-purpose diagnostics for intentional signaling rather than primate-specific measures. Not all criteria need be satisfied for a given behavior to qualify as intentional communication; in practice, satisfaction of four to five criteria is considered strong evidence in the animal communication literature. The nine criteria are:

> **(1) Directionality.** The signal is spatially or attentionally oriented toward a specific receiver rather than broadcast indiscriminately.
>
> **(2) Audience effects.** The signaler's behavior differs depending on who is present—different signals or signal intensities for different receivers.

(3) **Persistence.** If the initial signal fails to elicit a response, the signaler continues or repeats the signal rather than abandoning it.

(4) **Response waiting.** The signaler pauses after signaling to observe the receiver's reaction before proceeding.

(5) **Elaboration.** If one signal form fails, the signaler tries alternative signals to achieve the same communicative goal.

(6) **Goal-directedness.** The signaling behavior appears organized around achieving a specific outcome.

(7) **Means-end dissociation.** The signaler deploys different behavioral methods flexibly to achieve similar communicative effects, demonstrating that the method is separable from the goal.

(8) **Sensitivity to attentional state.** The signaler adjusts its behavior based on whether the receiver appears to be attending—producing louder, more conspicuous, or more proximate signals when the receiver's attention is elsewhere.

(9) **Response-dependent modification.** The signaler changes its subsequent behavior based on the receiver's response to the preceding signal—escalating if the receiver resists, de-escalating if the receiver complies.

Two of these criteria merit particular emphasis for the UAP context. *Response-dependent modification* is the most diagnostic criterion because it requires real-time monitoring of the receiver's behavior and adaptive adjustment—a capacity that presupposes awareness of the interaction as an interaction. In animal communication research, response-dependent modification is considered strong evidence against reflexive or coincidental behavior because it requires the signaler to represent the receiver's state and update its own behavior accordingly. *Means-end dissociation* is similarly powerful because it demonstrates that the signaler possesses a repertoire of methods rather than a single fixed response, using whichever method is appropriate to the current context. A signaler that achieves the same outcome (preventing engagement) through electromagnetic interference in one case, instantaneous positional displacement in another, and a detached pursuit object in a third is exhibiting means-end dissociation.

2.3 Graduated Signaling and Aposematic Display

Two specific signaling strategies from behavioral ecology are directly relevant to the behavioral patterns documented in instrumented UAP encounters: graduated signaling and aposematic display. Both are well-characterized, widely observed across taxa, and produce distinctive behavioral signatures that can be identified in observational data.

Graduated signaling refers to the systematic escalation or de-escalation of signal intensity in proportion to the context of an interaction. First formalized by Enquist (1985) and elaborated by Maynard Smith (1982) in the context of game-theoretic models of animal conflict, graduated signaling is a universal feature of biological conflict management. Red deer engage in roaring contests where the rate and duration of roars escalate incrementally before physical combat; cuttlefish modulate chromatic displays through progressively more intense patterns as contests escalate; primates employ a graded series of aggressive signals—stares, open-mouth threats, lunges, and contact aggression—that correspond to increasing motivational intensity. The defining feature of graduated signaling is proportionality: the signal intensity matches the contextual threat level without exceeding it. This proportionality distinguishes communication from reflexive or random behavior, because it requires the signaler to assess the current state of the interaction and calibrate its response accordingly. In each of these examples, the signaler calibrates intensity to the receiver's actions rather than deploying maximum capability immediately—a pattern that, if observed in UAP behavioral sequences, would suggest real-time assessment and restraint.

Aposematic display is a signaling strategy in which an organism conspicuously advertises a capability that makes it costly to attack, thereby deterring potential aggressors without engaging in actual combat. Ruxton, Sherratt, and Speed's *Avoiding Attack* (2004) provides the comprehensive theoretical treatment. Classic examples include the bright coloration of poisonous dart frogs, the warning stripes of wasps, and the rattlesnake's auditory signal. The critical feature of aposematic signaling is the combination of conspicuousness and restraint: the organism makes itself maximally visible (costly in terms of predator detection) while demonstrating capability without deploying it offensively. The signal communicates, in functional terms: *I possess the ability to harm you, and I am choosing not to.*

These two signaling strategies are not mutually exclusive; in biological systems, they frequently co-occur. An aposematic organism may employ graduated signaling within an encounter—a rattlesnake begins with a rattle (low-cost warning), escalates to a coiled strike posture (higher-cost display), and only strikes as a last resort. The combination of conspicuous capability demonstration, proportional escalation, and ultimate restraint constitutes a behavioral signature that is highly diagnostic of intentional deterrence communication.

The relevance to UAP behavioral data will become apparent in Section 4, but the theoretical point can be stated here without reference to specific cases: if an anomalous aerial phenomenon demonstrates capabilities vastly exceeding those of the observing military platform, deploys these capabilities in graduated proportion to the perceived threat level of the observer's actions, and consistently refrains from hostile engagement despite apparent ability to do so, this behavioral profile satisfies the formal criteria for combined graduated-aposematic signaling as defined in behavioral ecology.

2.4 Information-Theoretic Detection: Structure Without Semantics

The biological signaling framework and intentionality criteria described above require qualitative assessment—a trained analyst evaluating whether specific behavioral features are present. While valuable, qualitative assessment is vulnerable to interpreter bias. A complementary approach uses information theory to detect communication structure quantitatively, producing numerical values that can be compared against null distributions and tested for statistical significance.

The foundational insight of this approach is that **communication can be detected without being understood.** Shannon's (1948) mathematical theory of communication characterizes information in terms of statistical regularities in sequences of symbols, independent of what those symbols mean. A sequence need not be decoded to be identified as carrying a message; it need only exhibit structural properties that distinguish it from random noise and simple repetition. This principle has been applied successfully to animal vocalizations whose semantic content remains unknown (McCowan, Hanser, and Doyle 1999; Doyle et al. 2008), to genomic sequences (Mantegna et al. 1994), and—in a particularly relevant demonstration—to a message specifically designed for extraterrestrial recipients.

The following information-theoretic measures, each validated on biological communication systems, constitute the quantitative toolkit applied in this paper:

2.4.1 Shannon Entropy

Shannon entropy, $H = -\Sigma\, p(x) \log_2 p(x)$, measures the average unpredictability of a sequence. When normalized to the range [0, 1], it provides a single number characterizing where a sequence falls on the spectrum from perfect repetition ($H = 0$) to maximum randomness ($H = 1$). Communication systems across all known modalities—human language, animal vocalizations, genetic codes—occupy an intermediate range, typically 0.4 to 0.85 on the normalized scale. This intermediate positioning reflects a fundamental constraint: effective communication requires enough structure for the receiver to extract patterns (ruling out high entropy) but enough variability to carry information (ruling out low entropy). Random noise occupies the high end (~0.95–1.0); rigid repetitive patterns occupy the low end (~0.0–0.2). A behavioral sequence whose entropy falls within the communication range is not thereby proven to be communication—but a sequence outside this range can be excluded with greater confidence.

2.4.2 Zipf's Law

Zipf's law states that in any communication system, the frequency of the *n*th most common element is inversely proportional to its rank, yielding a power-law distribution with exponent approximately –1 when plotted on logarithmic axes. Originally observed in human language word frequencies (Zipf 1949), the law has been confirmed in every human language studied, in DNA codon usage, in city population distributions, and—critically for this paper—in the ***physical behavioral patterns of bottlenose dolphins*** (Ferrer-i-Cancho and Lusseau 2009). This finding is directly relevant because it demonstrates that Zipfian distributions emerge not only in acoustic or symbolic communication but in sequences of bodily movements, establishing precedent for applying the law to movement-based behavioral data.

A significant caveat applies. McCowan, Hanser, and Doyle (1999) demonstrated that certain random processes can generate Zipf-like distributions, cautioning against treating Zipfian structure alone as sufficient evidence of communication. Accordingly, Zipf analysis in this paper is used as one component of a multi-test battery rather than as a standalone diagnostic. A behavioral sequence that satisfies Zipf's law, exhibits intermediate Shannon entropy, compresses to a ratio consistent with communication, *and* shows power-law mutual information decay provides substantially stronger evidence than any single test.

2.4.3 Kolmogorov Complexity and Compression Ratios

Kolmogorov complexity—the length of the shortest program that can reproduce a given sequence—provides a theoretical measure of a sequence's intrinsic information content. Because Kolmogorov complexity is uncomputable in general, it is estimated in practice through data compression algorithms (Lempel-Ziv, gzip, bzip2), which approximate the minimal description length. Communication sequences typically compress to 30–70 percent of their original length: they contain enough redundancy to be compressible (unlike random noise, which is incompressible at ~100 percent) but enough non-redundant information to resist high compression (unlike simple repetition, which compresses to near zero). Reznikova (2017, 2023) has used compression-based complexity measures extensively to analyze ant communication, demonstrating that ants transmitting route information to nestmates produce behavioral sequences with compression ratios characteristic of structured information transfer.

2.4.4 Mutual Information Decay

In hierarchical communication systems—those with structure at multiple scales, such as phonemes within words within sentences—mutual information between elements decays as a power law with increasing separation. In simple Markov processes (where each element depends only on its immediate predecessor), mutual information decays exponentially. The distinction is diagnostic: power-law decay indicates long-range correlations characteristic of hierarchical structure, while exponential decay indicates a memoryless process. Semple, Ferrer-i-Cancho, and Gustison (2022) reviewed the application of linguistic laws—including Zipf's law, Menzerath's law, and mutual information decay profiles—across biological systems, finding that these statistical signatures appear in primate call sequences, birdsong, and cetacean vocalizations. Their framework provides direct methodological precedent for testing whether UAP behavioral sequences exhibit the same statistical fingerprints.

2.4.5 Transfer Entropy

Transfer entropy, introduced by Schreiber (2000), measures the directed flow of information between two time series. Given two processes X (observer trajectory) and Y (UAP trajectory), the transfer entropy TE(X→Y) quantifies how much the past of X reduces uncertainty about the future of Y, beyond what the past of Y alone provides. If TE(X→Y) significantly exceeds TE(Y→X), the UAP's subsequent behavior carries more information about the observer's preceding actions than vice versa—a direct, quantitative operationalization of "responsive behavior."

Transfer entropy is the single most powerful tool in this analytical framework for one reason: it converts the qualitative claim "the object responded to the aircraft's maneuver" into a measurable quantity with an associated null distribution and p-value. Statistical significance is assessed by computing TE on time-shifted surrogate data (shuffling the temporal alignment between observer and UAP trajectories) to generate a null distribution of TE values expected under the hypothesis of no directed information flow. This approach controls for autocorrelation within each individual trajectory while testing specifically for directed influence between them. If the observed TE exceeds the 95th percentile of the surrogate distribution, the directed-flow hypothesis is supported at $p < 0.05$.

The practical limitation of transfer entropy is its requirement for relatively fine temporal resolution—ideally, seconds-level discretization of both time series. This requirement is met by the radar and infrared tracking data available for some cases (particularly those involving Aegis combat system radar and ATFLIR infrared sensors) but not for cases documented primarily through post-event narrative accounts. Section 5 addresses this constraint case by case.

2.4.6 Local Compositional Complexity

Mahon (2025) introduced Local Compositional Complexity (LCC) as a method for detecting readable structure in signals of unknown origin. LCC computes Shannon entropy within sliding windows across a signal's length, then analyzes the variance and distribution of these local entropy values. Communication signals exhibit characteristic patterns in LCC: some regions are more structured (lower local entropy—headers, repeated elements, syntactic scaffolding) while others are more variable (higher local entropy—content-carrying segments). Random noise produces uniform LCC profiles; simple repetition produces uniformly low LCC. The diagnostic signature of communication is *heterogeneous local complexity*—a structured alternation between high- and low-entropy regions.

Mahon validated the method on the Arecibo message—a 1,679-bit signal transmitted from the Arecibo radio telescope in 1974, designed to be decodable by an extraterrestrial intelligence. LCC analysis detected the message as containing structured, readable information *purely from its statistical properties*, without any knowledge of the encoding scheme, the intended recipient, or the message content. This result is the information-theoretic anchor for the present paper's approach: it demonstrates that the tools described in this section can identify communication structure in a message from an unknown source, using precisely the kind of blind analysis proposed here for UAP behavioral sequences. A practical limitation warrants acknowledgment: like other complexity measures, LCC benefits from longer input sequences and yields wider confidence intervals on short strings. Individual UAP behavioral sequences of 15–25 events are at the lower bound for reliable LCC estimation; pooling across cases to produce aggregate sequences of 60–100 events substantially improves statistical power. This constraint is addressed in detail in Section 5.

2.5 Synthesis: A Multi-Level Detection Framework

The three analytical traditions described above—biological signaling theory, comparative psychology, and information theory—operate at different levels of description and provide complementary forms of evidence. Their combined application to a single dataset produces a detection framework substantially more robust than any individual approach.

Framework	**Question Answered**	**Evidence Type**	**Validated On**	**Key UAP Relevance**
Maynard Smith-Harper signaling theory	Is the behavior structurally consistent with a signal?	Qualitative: functional criteria (i)–(iii)	All animal communication systems	Context-dependent deployment near nuclear sites, military operations
Tomasello intentionality criteria	Is the signaling deliberate rather than reflexive?	Qualitative: 9-criterion checklist	Primates, corvids, cetaceans	Response-dependent modification during intercepts
Graduated signaling /	Does the signal show proportional	Qualitative: behavioral	Conflict management	Capability demonstration

aposematic display	escalation and restraint?	signature matching	across taxa	without hostility across all cases
Information-theoretic analysis	Does the sequence contain measurable communication structure?	Quantitative: entropy, Zipf, TE, LCC	Ant communication, dolphin behavior, Arecibo message	Testable without knowing UAP identity or intent

Table 1. Multi-level communication detection framework.

The framework is deliberately conservative. A behavioral sequence that satisfies the Maynard Smith-Harper criteria, meets five or more of the Tomasello intentionality criteria, exhibits the graduated-aposematic behavioral signature, *and* produces information-theoretic values within the communication range has passed four independent filters calibrated against known communication systems. Conversely, a sequence that fails the information-theoretic tests—showing entropy indistinguishable from random noise or transfer entropy consistent with the null hypothesis of no directed information flow—would undermine the communication interpretation regardless of qualitative assessment.

This asymmetry is intentional and important. The qualitative frameworks can suggest that communication is occurring; the quantitative frameworks can confirm or disconfirm it. The qualitative analysis generates the hypothesis; the quantitative analysis tests it. In the sections that follow, the qualitative frameworks are applied first (Sections 4 and 6), followed by the quantitative analysis (Section 5), preserving the hypothesis-generation/hypothesis-testing distinction that characterizes rigorous scientific inquiry.

One final methodological point warrants emphasis. Every tool in this framework was developed for and validated on communication systems where the analyst lacked access to the communicating agents' internal states. Ethologists applying information theory to dolphin vocalizations do not know what dolphins intend to communicate. Reznikova's compression analyses of ant behavioral sequences do not require understanding ant cognition. Mahon's LCC analysis of the Arecibo message treated it as a signal of unknown origin. The tools work precisely because they detect structural properties of communication that are independent of the communicators' identity, biology, or cognitive architecture.

If UAP behavioral sequences exhibit these structural properties, the finding carries the same evidential weight regardless of whether the UAP is controlled by an extraterrestrial intelligence, an interdimensional entity, a terrestrial non-human intelligence, or an unknown natural process that happens to produce communication-like structure. The framework is agnostic about ontology by design.

3. Evidence Standards and Case Selection

The UAP literature spans thousands of reported sightings, hundreds of government documents, and decades of civilian investigation. The overwhelming majority of this material is unsuitable for the analysis proposed in this paper. Eyewitness-only accounts, however sincere, cannot support quantitative behavioral sequence analysis. Cases lacking official documentation cannot be verified independently. Reports without behavioral interaction data—a light in the sky, a radar blip, a photograph of an ambiguous shape—provide no sequential information to analyze. This section defines the evidence criteria that a case must satisfy to be included in the analysis, explains what these criteria exclude and why, and justifies the temporal boundaries of the study.

3.1 The Four Evidence Criteria

Each case included in this paper's analysis must simultaneously satisfy four criteria. These criteria were defined prior to case selection and are applied without exception.

Criterion 1: Multiple independent sensor channels. The event must have been detected by at least two independent sensor systems operating on different physical principles—for example, radar and visual observation, or radar and infrared tracking, or electronic intelligence (ELINT) monitoring and ground-based radar. Single-channel detections are excluded regardless of the observer's credibility. This criterion exists because the most common prosaic explanations for UAP reports—sensor artifacts, misidentification, perceptual error—are channel-specific. A radar glitch does not simultaneously produce a correlated visual sighting. An optical illusion does not generate a simultaneous radar return at the corresponding range and bearing. Multi-channel detection is the minimum standard for establishing that the observed phenomenon had a physical presence independent of any single sensor's characteristics.

Criterion 2: Official documentation. The event must be attested by government or military documents obtained through official channels: declassified intelligence reports, FOIA releases, congressional testimony under oath, military investigation files, or equivalent institutional records. Civilian investigator reports, journalistic accounts, and secondhand narratives are used as supplementary material but do not satisfy this criterion on their own. The rationale is straightforward: official documentation creates an evidentiary chain between the reported event and the institutional apparatus that recorded it. A declassified DIA evaluation report carries a different epistemic weight than a magazine article about the same event, even if both describe the same facts, because the former was produced by analysts with access to primary sensor data writing for an audience of intelligence professionals with the authority to verify claims.

Criterion 3: Prosaic explanations investigated and found insufficient. The event must have undergone serious analytical scrutiny—whether by the original investigating body, by subsequent government review, or by technically competent independent analysts—and the prosaic explanations proposed must have been demonstrated to be inadequate to account for the full observational record. Cases where a plausible mundane explanation exists and has not been rigorously evaluated are excluded. Cases where a partial mundane explanation accounts for some but not all sensor channels are retained only if the unexplained residual involves the behavioral interaction data central to this paper's analysis. This criterion is applied with deliberate stringency: AARO's resolution of the GOFAST video through trigonometric parallax analysis (Kosloski 2024) and the Puerto Rico thermal video through sky lantern identification demonstrate that apparently anomalous cases can have prosaic explanations discoverable through careful analysis. A case earns inclusion in this paper not by being unexplained but by surviving serious attempts at explanation.

Criterion 4: Documented behavioral interaction. The event must include a recorded sequence of actions and responses between the UAP and human observers or systems—not merely a detection or sighting. The UAP must have exhibited behaviors that changed in apparent response to human actions (approach, withdrawal, weapons engagement, course change), or human systems must have exhibited effects correlated with UAP proximity or behavior (electromagnetic interference onset and offset, radar lock and loss, communications disruption). This is the criterion that most sharply distinguishes the present analysis from prior UAP studies. Static observations, however well-documented, provide no sequential data for information-theoretic analysis. A UAP that is detected, tracked, and lost without any apparent interaction is a detection event, not a behavioral sequence. The communication hypothesis requires interaction data—a minimum of two exchange cycles (action-response-action-response) to constitute an analyzable sequence.

3.2 What the Filter Excludes

Applied jointly, these four criteria reduce the available UAP literature to a very small number of cases. This is the intended result. The paper's argument depends on independent convergence under strict controls, not on the accumulation of individually ambiguous reports.

Among the categories of cases excluded:

Mass sighting events (e.g., the Phoenix Lights, 1997; the Belgian wave, 1989–1990) typically involve large numbers of eyewitness reports but lack the multi-sensor instrumented data required by Criterion 1 and the detailed behavioral interaction sequences required by Criterion 4. The Belgian wave is a partial exception—NATO radar tracking data and F-16 radar lock records exist—but the publicly available behavioral data consists primarily of aggregate descriptions rather than the moment-by-moment interaction timelines needed for sequence analysis. It is referenced as supporting context where relevant but is not included in the primary four-case analysis presented in Table 2.

AARO's 21 unresolved anomalous cases from the FY2024 report satisfy Criterion 2 (official documentation) and presumably Criterion 1 (AARO's assessment process involves multi-sensor evaluation), but fail Criterion 4 as applied here: no behavioral interaction sequences have been released publicly. Three reports described pilots being "trailed or shadowed by UAP," which would constitute precisely the kind of interaction data this paper analyzes—but the details remain classified. If declassified, these cases could substantially expand the dataset available for the methodology proposed in Section 5.

The 2021 ODNI Preliminary Assessment's 18 cases exhibiting "unusual movement patterns or flight characteristics"—including stationary hovering in high winds, movement against wind, and abrupt maneuvers at considerable speed without observable propulsion—have never received public follow-up at the individual case level. These cases may satisfy all four criteria; the information needed to assess them has not been made available.

Historical cases prior to 1957 are excluded not because they are uninteresting but because they predate the electronic sensor infrastructure required by Criterion 1. The 1952 Washington, D.C., radar-visual sightings, for example, involved radar confirmation at multiple facilities and generated substantial official documentation, but the behavioral interaction data is limited to air traffic controller observations of radar returns rather than instrumented tracking of action-response sequences.

The result is a primary dataset of four cases. A skeptic might object that four cases constitute an inadequate sample. The response has two parts. First, this is not a statistical survey—it is a structured qualitative and quantitative analysis of behavioral sequences, analogous to the detailed case studies that dominate the animal communication literature. Reznikova's (2023) foundational work on information-theoretic analysis of ant communication was built on experimental sequences involving small numbers of individually tracked ants; Getz et al.'s (2024) movement segmentation framework was validated on GPS tracks from a single population of barn owls. The question is not how many cases exist but whether the data within each case is detailed enough to support the analysis. Second, the convergence of findings across four independent cases—spanning four decades, four national military services, and zero shared witnesses—is itself a finding. If all four cases independently satisfy the same communication criteria (as will be demonstrated in Section 4), the probability of that convergence occurring by chance under the null hypothesis of random anomalous phenomena is formally estimable and, as Section 5 will show, very low.

3.3 Temporal Window: 1957–2004

The study’s temporal boundaries are set by the evidence criteria rather than by arbitrary periodization.

The lower bound of 1957 is determined by the earliest case that satisfies all four criteria: the RB-47H encounter of July 17, 1957. This case involved simultaneous detection by airborne electronic intelligence sensors (ELINT), aircrew visual observation, and USAF ground radar at Duncanville, Texas (Criterion 1). It is documented in Project Blue Book Case #10073, a three-page TWX from the 745th ACWRON, a four-page Forbes AFB Wing Intelligence summary, and the Condon Committee’s investigation as Case 5 (Criterion 2). The prosaic explanation proposed by Philip Klass—misidentification of American Airlines Flight 655—was refuted by Brad Sparks’ correction of a True versus Magnetic compass bearing error that destroyed the identification’s geometric basis, and the Condon Report’s own radar analyst, Gordon David Thayer, called the Blue Book explanation “literally ridiculous” (Criterion 3). The encounter included six distinct behavioral phases across more than 700 miles and approximately 90 minutes, with clear action-response dynamics including speed matching, synchronized triple-channel appearance and disappearance correlated with pilot course changes, and an object that circled the aircraft’s flight path (Criterion 4).

Moving the lower bound earlier—to capture, for example, the 1952 Washington, D.C., events or the 1945–1950 nuclear facility sightings cataloged by the SCU—would sacrifice the multi-sensor electronic instrumentation standard that gives the selected cases their analytical value. Moving it later—to 1967 or 1968, the date of the next qualifying case—would sacrifice the RB-47 encounter, which provides the longest continuous multi-channel behavioral sequence in the public record and the only known case of simultaneous radar, visual, and ELINT detection of a UAP. The cost of that exclusion would be substantial.

The upper bound of 2004 is determined by the most recent case with full multi-sensor confirmation and detailed behavioral data in the public domain: the USS *Nimitz* encounter of November 14, 2004. Post-2004 encounters—including the USS *Roosevelt* incidents of 2014–2015, the widely circulated GIMBAL and GOFAST videos, and the events described in AARO’s FY2024 report—either lack the publicly available behavioral detail required by Criterion 4 (the *Roosevelt* encounters), have been resolved through prosaic analysis (GOFAST’s parallax illusion, identified by AARO), or remain classified (AARO’s 21 anomalous cases). The September 2025 congressional testimony by active-duty Navy personnel describing the USS *Jackson* incident of February 2023—self-luminous objects emerging from the ocean with near-instantaneous synchronized acceleration—may qualify if additional multi-sensor documentation enters the public record, but does not currently meet all four criteria. AARO’s GREMLIN sensor system—a deployable multi-sensor suite comprising 2D and 3D radar, long-range EO/IR telescopes, GPS, and RF spectrum monitoring—represents the kind of instrumented data collection architecture that could enable full application of the methodology proposed in Section 5, if results from its classified 90-day deployment (late 2024–early 2025) are declassified. Recent congressional mandates in the FY2026 NDAA, signed December 2025, requiring briefings on all historical NORAD and USNORTHCOM UAP intercepts since 2004, may yield additional qualifying cases in the future.

The resulting 47-year span has a methodological virtue beyond the evidence criteria: it ensures independence between witness groups. The RB-47 crew (55th Strategic Reconnaissance Wing, USAF, 1957) had no contact with the Minot AFB personnel (5th Bomb Wing, USAF, 1968), who had no contact with the Imperial Iranian Air Force pilots (1976), who had no contact with the U.S. Navy aviators and radar operators of Carrier Strike Group Eleven (2004). The RB-47 case was classified and not publicly known in detail until well after the Minot encounter. The Tehran case was documented in Farsi-language Iranian military channels before it appeared in English-language intelligence reporting. The Nimitz encounter preceded the modern era of public UAP interest by more than a decade. If the same behavioral profile emerges independently across these four cases—as Section 4 will demonstrate—cultural contagion, witness contamination, and narrative mimicry are excluded as explanations for the convergence.

3.4 Case Selection Summary

Table 2 summarizes the four cases selected for analysis, their satisfaction of the evidence criteria, and the primary documentation for each.

Case	Date / Location	Sensor Channels	Official Documentation	Prosaic Explanation Status	Behavioral Interaction	All Four ?
RB-47H	July 17, 1957; MS–TX–OK	ELINT, visual, ground radar	Blue Book #10073; Condon Case 5; Forbes AFB Intel	Klass refuted (Sparks); Thayer: “literally ridiculous”	6 phases, ~90 min, 700+ mi; triple-channel sync; speed matching; circling	✓
Minot AFB	Oct 24, 1968; North Dakota	B-52 radar, visual, ground visual	Blue Book #12548; radarscope photos; crew debriefs	Blue Book Sirius/plasma: among weakest in record	3+ hr; station-keeping at 3 NM; proximity-correlated EM; nuclear site penetration	✓
Tehran	Sept 18–19, 1976; Iran	Airborne radar, visual	DIA IR No. 6846013976;	Klass Jupiter:	Textbook graduated	✓

		(two F-4s), ground visual	Mooy teletype; Iranian AF records	cannot account for graduated EM interference	escalation; proportional EM response to threat level	
Nimitz	Nov 14, 2004; off San Diego	SPY-1 radar, ATFLIR, visual (multiple aircrew)	Congressional testimony (Fravor, Dietrich); FLIR1 video; Navy acknowledgment	West FLIR analysis: limited to video; cannot address radar, visual, CAP point	Reorientation; mirroring; response-dependent departure; CAP point reappearance	✓

Table 2. Cases satisfying all four evidence criteria, with primary documentation and behavioral interaction summary.

Section 4 presents each of these cases in detail, with emphasis on the behavioral sequences that constitute the data for the information-theoretic analysis proposed in Section 5.

4. The Case Catalog

This section presents the four cases that satisfy the evidence criteria defined in Section 3. Each case receives identical analytical treatment: identification and context, sensor confirmation, official documentation, a detailed behavioral sequence timeline, assessment of the strongest skeptical explanation, and evaluation against the communication criteria established in Section 2. The behavioral sequence timelines are the primary data for the information-theoretic analysis proposed in Section 5; they are presented in sufficient detail to enable independent encoding and replication.

A note on sourcing: for each case, primary sources (official documents, sensor records, sworn testimony) are distinguished from secondary analytical sources (independent researchers' analyses). Where primary and secondary sources conflict, the discrepancy is noted. Where primary sources are incomplete—as is inevitable with partially declassified material—the gap is acknowledged rather than filled by inference.

4.1 RB-47H Electronic Countermeasures Aircraft, July 17, 1957

4.1.1 Identification and Context

On the early morning of July 17, 1957, a Boeing RB-47H Stratojet of the 55th Strategic Reconnaissance Wing, based at Forbes AFB, Kansas, conducted an electronic countermeasures training mission across the south-central United States. The aircraft carried a crew of six officers:

pilot Major Lewis D. Chase, copilot Captain James H. McCoid, and navigator Captain Thomas H. Hanley in the forward compartment, with three electronic warfare officers—Captain Frank B. McClure (ELINT station #1, monitoring the aft direction), Captain Walter A. Tuchscherer (station #2, right side), and Captain John J. Provenzano (station #3, left side)—operating the airborne electronic intelligence suite in the rear compartment. The RB-47H was the Air Force's premier ELINT aircraft, equipped with sensitive receivers designed to detect, classify, and locate ground-based radar emissions across a wide frequency range. The crew were trained intelligence professionals whose specific expertise was discriminating between radar signals of different types and sources.

The encounter began approximately 0400 Central Standard Time over Mississippi and continued for approximately 90 minutes across Louisiana, Texas, and Oklahoma—a distance of more than 700 miles—before the object was last detected near Oklahoma City. The total engagement duration and geographic extent make this the longest continuous multi-channel UAP interaction in the declassified record.

4.1.2 Sensor Confirmation

Electronic intelligence (ELINT). Captain McClure's airborne receivers detected an electromagnetic signal at approximately 2,995–3,000 MHz, within the S-band frequency range used by ground-based search radars. The signal was initially interpreted as a ground radar emission, but its behavior—moving relative to the aircraft, maintaining a consistent bearing, and at one point completing a 360-degree circuit around the RB-47's flight path—was incompatible with any fixed ground installation. The signal was tracked by multiple ELINT stations aboard the aircraft and maintained consistent characteristics throughout the engagement.

Visual observation. Pilot Chase and copilot McCoid observed a brilliant white-blue light that approached from the 11 o'clock position and crossed the aircraft's flight path with an angular velocity that Chase, a veteran of twenty years of military flying, stated he had never seen matched. The visual sighting correlated spatially and temporally with the ELINT detection throughout the encounter.

Ground radar. USAF air defense radar at Duncanville, Texas (745th Aircraft Control and Warning Squadron), independently detected and tracked a target at the position reported by the RB-47 crew. Brad Sparks' analysis demonstrated that the ELINT equipment simultaneously detected both the object's moving signal and the stationary Duncanville ground radar signal from directions approximately 30 degrees apart—providing an unprecedented real-time calibration confirming the object's emissions were genuine and not artifacts of the ground installation.

4.1.3 Official Documentation

The case is documented in multiple official records: a three-page TWX from the 745th ACWRON to Air Defense Command; a four-page Wing Intelligence summary from Forbes AFB; Project Blue Book Case #10073; and the Condon Committee's investigation as Case 5. Radar analyst Gordon David Thayer concluded that the Blue Book explanation was "literally ridiculous." James E. McDonald presented a detailed analysis at the AAAS symposium on December 27, 1969, subsequently published in *Astronautics & Aeronautics* (July 1971). Brad Sparks' definitive modern analysis in Clark's *UFO Encyclopedia* (1998, 2008) corrected a critical True versus Magnetic compass error that destroyed the geometric basis of Philip Klass's proposed identification of American Airlines Flight 655. Sparks further documented that SAC's ELINT Division obtained the original tapes but these subsequently "disappeared" into apparent compartmented classification.

4.1.4 Behavioral Sequence

The encounter proceeded through six distinct phases. Table 3 presents the sequence in chronological detail.

Phase	Time (CST)	Human Action / State	Object Behavior	Action–Response Dynamic
1	~0400; Mississippi	RB-47 on routine ELINT training, eastbound. Stations monitoring passively.	ELINT station #1 detects airborne signal at ~2,995–3,000 MHz. Source completes full 360° circuit around RB-47 at ~500 mph.	Object demonstrates awareness of aircraft position by circumnavigating it. Signal mimics S-band search radar characteristics.
2	~0410; Louisiana	Aircraft continues course. Crew alerted by ELINT detection.	Brilliant white-blue light approaches from 11 o'clock, crosses flight path at extreme angular velocity. Visual correlates with ELINT bearing.	Transition from ELINT-only to ELINT + visual. Object adds visual conspicuousness to existing electronic presence.
3	~0430; East Texas	Crew tracking on multiple ELINT stations and visually. Westbound.	Three-channel correlation: ELINT stations #1 and #2 and visual converge on same bearing. Object paces aircraft at 500 mph.	Stable formation with aircraft. Speed matching confirmed across channels. Station-keeping behavior.
4	0439–0452; Dallas	Crew requests ground confirmation. Duncanville radar acquires target. Chase accelerates to max	Ground radar confirms target. When aircraft accelerates, object pulls away—maintaining or	Responds to acceleration by increasing own speed. Duncanville

		power (~550 mph).	increasing separation.	ELINT calibration: simultaneous detection of object signal and ground radar ~30° apart confirms genuine emission.
5	~0452; Mineral Wells, TX	Chase overshoots as object stops. Initiates port turn.	**Vanishes from all three channels simultaneously**. When Chase enters port turn, **reappears on all three channels simultaneously**. Triple-channel on/off occurs at least twice.	Synchronized disappearance/reappearance correlated with pilot course changes. Simultaneous on/off across three independent physical channels is the single most anomalous feature.
6	Post-0452; Oklahoma	Chase dives from 35,000 ft toward object. Aggressive closure attempt.	Vanishes from all three channels during aggressive dive. Does not reappear. Encounter ends.	Final disappearance correlated with most aggressive human action. Withdrawal in response to escalated approach.

Table 3. RB-47H behavioral sequence, July 17, 1957.

The sequence exhibits a clear arc: covert electronic presence, escalation to multi-modal conspicuousness, stable formation-keeping, responsive speed matching, synchronized multi-channel responsiveness to pilot maneuvers, and withdrawal in response to aggressive approach—paralleling the graduated signaling sequences described in Section 2.3. This sequence yields approximately 15–20 discrete state transitions across three sensor channels, sufficient for preliminary Shannon entropy and transfer entropy analysis as developed in Section 5.

4.1.5 Strongest Skeptical Explanation

Philip Klass's identification of the ELINT signals as ground radar returns and the visual sighting as American Airlines Flight 655 was a technically specific hypothesis, but Sparks demonstrated that Klass had used True compass bearings where the documentation recorded Magnetic bearings; with the correction applied, the geometric alignment fails. Thayer independently confirmed the radar and ELINT characteristics were incompatible with any known ground installation or commercial aircraft. Klass's analysis, rigorous for the narrow question of signal identity, does not address the behavioral sequence—the synchronized triple-channel disappearance and reappearance correlated with pilot course changes—which no prosaic explanation has attempted to account for.

4.1.6 Communication Criteria Assessment

Criterion	Satisfied?	Evidence
Directionality	✓	Object circled aircraft; maintained consistent bearing; visual approach from specific direction.
Audience effects	✓	S-band emission at frequency ELINT crew trained to detect. Signal within professional competence of the specific receivers.
Persistence	✓	Maintained engagement across 700+ miles and ~90 minutes through multiple phases.
Response waiting	*Partial*	Phase 5 disappearance/reappearance consistent with monitoring pilot's response, but interpretation uncertain.
Elaboration	✓	Shifted from ELINT-only (Phase 1) to ELINT + visual (Phase 2) to ELINT + visual + ground radar (Phase 4). Progressive modality increase.
Goal-directedness	✓	Sustained engagement with specific aircraft over extended duration and distance.
Means-end dissociation	✓	Electronic emission, visual display, speed matching, position holding, disappearance/reappearance—different means serving consistent engagement function.
Sensitivity to attentional state	✓	Added visual channel matched to forward crew capabilities (pilot/copilot observe visually; ELINT operators observe electronically).
Response-dependent modification	✓	Pulled away when aircraft accelerated. Disappeared/reappeared with course changes. Permanent withdrawal when pilot attempted aggressive dive.
Graduated signaling	✓	Progressive escalation from covert electronic to multi-modal conspicuousness; de-escalation when approach became aggressive.
Aposematic display	✓	Superior speed, instantaneous disappearance, circumnavigation—without hostile action during 90-minute engagement.

Table 4. Communication criteria assessment: RB-47H encounter.

The RB-47 encounter satisfies ten of eleven communication criteria, with one partial (response waiting). This is consistent with intentional, structured communication as defined in Section 2.

4.2 Minot Air Force Base, North Dakota, October 24, 1968

4.2.1 Identification and Context

On the early morning of October 24, 1968, multiple personnel at Minot Air Force Base, North Dakota, observed an unidentified luminous object over a period exceeding three hours. The base housed the 5th Bomb Wing (B-52H strategic bombers) and the 91st Strategic Missile Wing (Minuteman I ICBMs). At approximately 0300 local time, maintenance teams at Oscar Flight missile facilities reported a bright object hovering at low altitude. Security alarms were triggered at Oscar-7, where personnel found the outer hatch of the hardened launch facility open—a condition requiring either deliberate action or extreme force. A B-52H (call sign JAG 31), piloted by Captain Bradford Runyon with a crew of five, was directed by Radar Approach Control to investigate.

4.2.2 Sensor Confirmation

Airborne radar. The B-52's onboard radar acquired the object at approximately three nautical miles, producing radarscope photographs—the only known UAP case with real-time radar imagery. First-generation prints preserved by William McNeff have been independently analyzed by Martin Shough, Claude Poher, Brad Sparks, and Richard Haines, all confirming a genuine radar return.

Visual observation (airborne). Multiple crew members observed a large, luminous, orange-red object at the radar-indicated position. Copilot Captain Thomas D. Goduto and other crew provided consistent corroborating descriptions.

Visual observation (ground). Sixteen ground personnel at Oscar Flight facilities independently observed the same object over an extended period before and during the B-52's engagement, providing a third independent observation channel.

4.2.3 Official Documentation

Project Blue Book Case #12,548. Blue Book classified it as Sirius (visual) and plasma (radar)—among the weakest explanations in the program's history. Thomas Tulien's Minot B-52 case study through the Sign Oral History Project (minotb52ufo.com) provides the most complete reconstruction, including first-person crew interviews. Key gaps: SSgt. Richard Clark's analysis and materials sent to General Hollingsworth at SAC Headquarters remain missing from the public record.

4.2.4 Behavioral Sequence

Phase	Time / Context	Human Action / State	Object Behavior	Action–Response Dynamic

0	~0300; ground	Routine missile site maintenance. No aircraft.	Bright object hovering near Oscar Flight. Security alarms at Oscar-7; outer hatch found open.	Self-initiated presence at nuclear weapons site. Possible physical interaction with facility hardware.
1	~0330; B-52 approach	JAG 31 directed to investigate.	Object acquired on radar at ~3 NM. Persistent, well-defined return. Radarscope photos captured.	Allows radar detection at consistent range. No evasion.
2	B-52 orbit	Aircraft enters 180° standard-rate turn.	**Maintains precisely 3 NM distance throughout entire turn—**station-keeping at constant range during 180° heading change.	Precision station-keeping requires real-time tracking of aircraft trajectory and coordinated positional adjustment.
3	Committed approach	Aircraft commits to predictable straight-line path.	**Closes from 3 NM to ~1 NM**. Crew observes large orange-red luminous object.	Closing only after aircraft committed to predictable path. Object reduces distance on its own initiative once approach assured.
4	Closest approach	Aircraft at minimum range.	**EM interference onset**: UHF disabled; VHF degraded. IFF continues normally. Effects correlated with proximity.	Frequency-dependent interference: UHF (225–400 MHz) disabled while IFF (~1,030/1,090 MHz) continues—consistent with EM coupling physics, not broadband jamming.
5	Separation	Aircraft increases distance.	EM interference ceases. All systems restore. Restoration correlates with range increase.	Approach → EM onset; separation → EM offset. Proximity-dependent and reversible—not equipment malfunction.
6	Extended presence	B-52 departs. Ground teams continue observation.	Object remains visible, eventually departing.	Presence not contingent on aircraft. Precedes and outlasts the investigating platform.

Table 5. Minot AFB behavioral sequence, October 24, 1968.

The precision station-keeping during the 180-degree turn requires continuous positional adjustment—a computationally non-trivial task. The frequency-dependent electromagnetic interference, with UHF disabled while IFF continued, is consistent with physics of electromagnetic coupling at specific frequencies rather than arbitrary failure. The seven-phase sequence, including the proximity-dependent EM onset and offset, yields approximately 12–15 discrete state transitions encodable for the information-theoretic analysis in Section 5.

4.2.5 Strongest Skeptical Explanation

Blue Book's Sirius/plasma explanation is among the weakest in program history. Sirius was not at the reported bearing and elevation. No plasma model produces a persistent three-nautical-mile radar return maintaining geometrically precise station-keeping. No alternative skeptical analysis of comparable rigor exists.

4.2.6 Communication Criteria Assessment

Criterion	Satisfied ?	Evidence
Directionality	✓	Positioned at nuclear missile site; maintained oriented engagement with aircraft.
Audience effects	✓	Nuclear site specifically (not random facility); EM at aircraft comm/nav frequencies.
Persistence	✓	3+ hour total presence. Preceded and outlasted investigating aircraft.
Response waiting	*Partial*	Maintained distance during turn, closed only when aircraft committed. Consistent with monitoring before acting.
Elaboration	✓	Visual display, radar presence, EM effects, physical interaction with ground facility.
Goal-directedness	✓	Organized around nuclear weapons infrastructure: missile site, investigating aircraft, extended duration.
Means-end dissociation	✓	Visual, radar, EM, physical interaction—multiple methods serving engagement function.
Sensitivity to attentional state	✓	EM affected frequencies the crew was actively relying on.
Response-dependent modification	✓	Station-keeping during turn; closed when committed; EM correlated with proximity. Different responses to different aircraft behaviors.
Graduated signaling	✓	Passive presence to active radar engagement to EM interference at closest approach. Intensity scaled to interaction.
Aposematic display	✓	Ability to affect aircraft systems and nuclear facility hardware without hostile action. 3+ hours, no attack despite nuclear proximity.

Table 6. Communication criteria assessment: Minot AFB encounter.

Ten of eleven criteria satisfied, with one partial. The nuclear facility context adds target selectivity absent from the RB-47 case.

4.3 Tehran, Iran, September 18–19, 1976

4.3.1 Identification and Context

On the night of September 18–19, 1976, civilian reports of an unusual light over Tehran prompted Brigadier General Nader Yousefi to authorize the scramble of two F-4 Phantom II interceptors from Shahrokhi Air Base. The engagement that followed was documented in a Defense Intelligence Agency evaluation report characterizing it as a case that "meets all the criteria necessary for a valid study of the UFO phenomenon." The Tehran encounter is the most frequently cited example of graduated UAP behavioral response in the literature.

4.3.2 Sensor Confirmation

Airborne radar. Both F-4 interceptors acquired the object on AN/APQ-120 fire control radar. Lieutenant Parviz Jafari obtained radar lock at 27 nautical miles. A smaller object that detached from the primary also produced a radar return.

Visual observation. Both F-4 crews and multiple ground observers, including General Yousefi, observed an intensely luminous source with rapidly alternating colored lights. A secondary object illuminated a dry lake bed near Rey with sufficient intensity for the terrain to be visible from the air.

A limitation: Mehrabad tower radar was under repair and did not provide ground-based tracking. The case rests on two independent airborne radar systems plus multiple visual observations.

4.3.3 Official Documentation

DIA evaluation report IR No. 6846013976 (October 12, 1976) and Lt. Col. Olin Mooy's teletypewriter message from the U.S. Defense Attaché Office in Tehran. Both declassified via FOIA. The DIA report provides sufficient tactical detail to reconstruct the engagement sequence with confidence.

4.3.4 Behavioral Sequence

The Tehran encounter is structured as two successive intercepts, producing the clearest graduated response pattern in the catalog.

Phase	Time / Context	Human Action / State	Object Behavior	Action–Response Dynamic
1	~0030; 1st F-4	First F-4 scrambled. Pilot approaches on direct intercept heading.	Within ~25 NM, all instrumentation and communications fail simultaneously.	Failure threshold correlated with approach distance. Systems restore when

				pilot turns away.
2	1st F-4 withdrawal	Pilot turns away, abandoning intercept.	**All systems restore immediately** upon disengagement.	Approach → failure; withdrawal → restoration. Establishes the pattern that will escalate.
3	~0140; 2nd F-4	Second F-4 (Lt. Jafari) scrambled. Acquires radar lock at 27 NM. Continues closing.	Object trackable on radar. Allows radar lock. No immediate EM effects.	Different response to second aircraft: permits tracking denied to first. Threshold shifted—testing a different engagement level.
4	Weapons engagement	**Jafari prepares to fire AIM-9 missile**. Weapons panel active.	**Weapons panel, comms, AND instruments all fail simultaneously**. More systems than first intercept —escalation proportional to threat.	Graduated: First F-4 approach → instruments+comms. Second F-4 weapons lock → instruments+comms +weapons. Denial scope matched threat level exactly.
5	Breakaway	Jafari executes negative-G dive.	Smaller object detaches and pursues F-4 through breakaway. Pursuit temporary; object rejoins parent.	Active tracking deployed in response to evasion. Escalation from passive denial to active pursuit. Ceases when threat de-escalates.
6	Disengagement	Jafari abandons intercept. Observes from distance.	Pursuit object rejoins primary. Systems restore. Secondary descends, illuminates terrain near Rey.	De-escalation follows human de-escalation. Pursuit ceases; systems restore.
7	Landing approach	F-4 returns to Mehrabad. Another object near landing path.	EM interference again. Civil airliner in area reports simultaneous comms interference.	Effects replicated with independent civilian corroboration —cross-validating proximity-dependence.

Table 7. Tehran behavioral sequence, September 18–19, 1976.

The escalation pattern: *passive approach → instruments/comms denial; weapons engagement → instruments/comms/weapons denial; evasion → active pursuit; disengagement → de-escalation.* At every step, system denial scope matched threat level precisely. This proportionality is the defining feature of graduated communication. The 18-symbol illustrative encoding in Section 5.1

is derived directly from this timeline, which provides the finest-grained action–response sequence in the catalog.

4.3.5 Strongest Skeptical Explanation

Philip Klass's Jupiter hypothesis is a technically grounded starting point—Jupiter was indeed bright and well-positioned—and the F-4 Phantom's reputation for electrical unreliability lends surface plausibility to the malfunction component. However, the hypothesis is rigorous only for the initial visual identification and does not address three features: Jupiter does not produce radar returns at 27 NM; equipment failures on two independent aircraft that correlate with approach distance and restore upon withdrawal require an implausible series of coincidences; and the graduated proportionality—more systems disabled during weapons engagement than passive approach—has no mechanism under any malfunction model.

4.3.6 Communication Criteria Assessment

Criterion	**Satisfied?**	**Evidence**
Directionality	✓	Effects directed at approaching interceptors. Pursuit object tracked evading F-4.
Audience effects	✓ (strong)	Different response to different aircraft: passive approach → partial denial; weapons engagement → comprehensive denial.
Persistence	✓	Maintained through two intercepts, ground observation, and landing interference.
Response waiting	✓	Second F-4 permitted radar lock and approach without interference—waited for weapons engagement.
Elaboration	✓	Visual display, selective EM, pursuit object, ground illumination—multiple behavioral modes.
Goal-directedness	✓	Preventing weapons engagement while maintaining presence. Coherent objective sustained across hours.
Means-end dissociation	✓	EM denial, pursuit object, illumination display—three distinct methods serving deterrence.
Sensitivity to attentional state	✓	Weapons panel targeted specifically during weapons lock attempt.
Response-dependent modification	✓ (strongest)	Textbook: approach → partial denial; weapons → comprehensive denial; evasion → pursuit; disengagement → de-escalation.
Graduated signaling	✓ (textbook)	Perfect proportionality at every step. De-escalation followed disengagement.
Aposematic display	✓	Disabled weapons systems without harming aircraft or crew. Capability to destroy implied; restraint absolute.

Table 8. Communication criteria assessment: Tehran encounter.

All eleven criteria satisfied. Tehran provides the strongest examples of audience effects, response-dependent modification, and graduated signaling in the catalog—the most diagnostic case for the communication hypothesis.

4.4 USS Nimitz "Tic Tac," November 10–14, 2004

4.4.1 Identification and Context

Over approximately five days in November 2004, Carrier Strike Group Eleven—centered on USS *Nimitz* (CVN-68) and guided-missile cruiser USS *Princeton* (CG-59)—detected multiple anomalous radar returns during a pre-deployment exercise approximately 100 miles southwest of San Diego. The *Princeton*'s AN/SPY-1B phased-array radar registered objects descending from approximately 80,000 feet to near sea level in seconds. On November 14, Commander David Fravor and Lieutenant Commander Alex Dietrich, experienced F/A-18F Super Hornet pilots from VFA-41, were vectored to a radar contact.

4.4.2 Sensor Confirmation

AN/SPY-1B phased-array radar. The *Princeton*'s Aegis combat system tracked multiple objects over approximately five days. Senior Chief Kevin Day reported consistent returns displaying radical performance.

ATFLIR. Lieutenant Chad Underwood captured infrared video (FLIR1) showing an object with no visible flight surfaces, exhaust, or propulsion signature. Underwood has stated the original recording was approximately eight to ten minutes in duration with substantially higher resolution than the 76-second clip in public circulation; the full recording remains classified.

Visual observation. Fravor, Dietrich, and their weapons systems officers observed the object in daylight at close range. Fravor described a smooth white elongated shape approximately 40 feet long—a "Tic Tac"—hovering above an ocean surface disturbance. His account has been provided under oath (House Oversight Committee, July 26, 2023).

Electronic warfare indicators. Fravor reported the AN/APG-73 radar was actively jammed. Petty Officer Gary Voorhis reported anomalous electronic signatures aboard the *Princeton*.

4.4.3 Official Documentation

Congressional testimony under oath (Fravor and Dietrich, 2023); officially released FLIR1 video (DOD-confirmed authentic); Navy acknowledgment as "unidentified aerial phenomena." The Office of Naval Intelligence confirmed via FOIA it possesses "certain briefing slides classified TOP SECRET" and a "video classified SECRET" related to the events, stating release "would cause exceptionally grave damage to National Security."

4.4.4 Behavioral Sequence

Phase	Time / Context	Human Action / State	Object Behavior	Action–Response Dynamic

0	Nov 10–13; multi-day	Carrier strike group in routine training. SPY-1B in standard surveillance.	Multiple objects over several days, descending from ~80,000 ft to sea level in seconds. No evasion.	Self-initiated persistent presence in military operational area. Detectable by sophisticated radar.
1	Nov 14; vectored	Fravor and Dietrich vectored to contact.	Object hovering above ocean disturbance. White, elongated, ~40 ft. No wings, exhaust, propulsion.	Present and visible before aircraft arrival. Not evading. Apparent interaction with ocean.
2	Initial descent	Fravor descends in spiral pattern.	**Reorients longitudinal axis toward Fravor**. Begins ascending in mirroring spiral. Fravor: "almost playful."	Reorientation is directional. Mirroring is response-matched—complementary maneuver.
3	Aggressive closure	**Fravor commits to direct aggressive descent—**cutting across circle.	**Accelerates and departs instantaneously**. "Poof." No gradual acceleration, no propulsion. Departure at transition from cautious to aggressive.	Threshold-dependent: mirroring during cautious engagement; departure when aggressive. Binary shift correlated with change in tactical posture.
4	CAP point	Princeton directs Fravor to Combat Air Patrol point—classified coordinate.	**Object detected at CAP point—**~60 mi away—within seconds. CAP was classified tactical information on encrypted channels.	Reappearance at classified coordinate implies access to operational information. Not kinematic but informational anomaly.
5	Underwood FLIR	Subsequent F/A-18 vectored. ATFLIR activated, radar tracking.	Detected on ATFLIR (FLIR1 video). APG-73 radar jammed. Object eventually departs.	Active radar jamming: selective sensor management. Object visible but not targetable.

Table 9. USS Nimitz behavioral sequence, November 10–14, 2004.

The Nimitz contributes several unique features: multi-day persistence establishing detectable presence; the mirroring maneuver (closest analog to animal communication initiation); the CAP point reappearance (implying access to operational information structure); and selective sensor management (visible on ATFLIR but radar-jammed—observable but not targetable). The multi-day radar tracking and daylight visual engagement provide the highest temporal resolution in the catalog, yielding approximately 10–14 discrete state transitions with seconds-level timing available for transfer entropy calculation.

4.4.5 Strongest Skeptical Explanation

Mick West's FLIR1 analysis is technically rigorous for the 76-second video alone but explicitly does not address the multi-day SPY-1B radar tracks, Fravor's daylight close-range visual observation under oath, the mirroring maneuver observed by four aircrew, the instantaneous departure to a classified coordinate, or the active radar jamming. No comprehensive skeptical account covering the full evidentiary convergence has been published.

4.4.6 Communication Criteria Assessment

Criterion	Satisfied?	Evidence
Directionality	✓	Reoriented axis toward Fravor. Appeared at classified CAP point.
Audience effects	✓	Present in military training area. Radar jamming targeted specific sensor system.
Persistence	✓	Multi-day radar presence. Reappeared at CAP after departure.
Response waiting	✓	Mirroring allowed Fravor to establish interaction pattern before escalation. Matched rather than preempted.
Elaboration	✓	Radar presence, hovering, mirroring, displacement, CAP appearance, jamming—six behavioral modes.
Goal-directedness	✓	Organized around military asset engagement: training area, interceptor, operational coordinate.
Means-end dissociation	✓	Mirroring, departure, CAP appearance, radar jamming—multiple distinct methods.
Sensitivity to attentional state	✓	ATFLIR permitted (visible) while radar jammed (not targetable)—selective engagement with observation vs. weapons.
Response-dependent modification	✓ **(strong)**	Mirroring during cautious spiral; instantaneous departure when approach turned aggressive.
Graduated signaling	✓	Persistent presence → visible hovering → interactive mirroring → departure at escalation threshold.
Aposematic display	✓ **(strongest)**	Capabilities vastly exceeding military aircraft with zero hostile action across five days. Maximum capability gap, maximum restraint.

Table 10. Communication criteria assessment: USS Nimitz encounter.

All eleven criteria satisfied. The Nimitz provides the strongest aposematic display: capabilities so far beyond the observer's technology that the capability gap itself constitutes the signal.

4.5 Cross-Case Convergence

Table 11 presents the communication criteria assessment across all four cases.

Criterion	RB-47 (1957)	Minot (1968)	Tehran (1976)	Nimitz (2004)
Directionality	✓	✓	✓	✓
Audience effects	✓	✓	✓ (strong)	✓
Persistence	✓	✓	✓	✓
Response waiting	*Partial*	*Partial*	✓	✓
Elaboration	✓	✓	✓	✓
Goal-directedness	✓	✓	✓	✓
Means-end dissociation	✓	✓	✓	✓
Sensitivity to attentional state	✓	✓	✓	✓
Response-dependent modification	✓	✓	✓ (strongest)	✓ (strong)
Graduated signaling	✓	✓	✓ (textbook)	✓
Aposematic display	✓	✓	✓	✓ (strongest)
Criteria satisfied	**10/11**	**10/11**	**11/11**	**11/11**

Table 11. Cross-case communication criteria convergence.

The eleven criteria combine the nine intentionality markers derived from Tomasello's primate gesture research with the two specific signaling strategies—graduated signaling and aposematic display—highlighted in Section 2.3. Four cases spanning 47 years, four national military services, and zero shared witnesses produce a minimum of 10 out of 11 criteria satisfied in every case. Two—Tehran and Nimitz—satisfy all eleven. The two partial satisfactions (response waiting in RB-47 and Minot) reflect limits of temporal resolution in available documentation rather than evidence against the criterion.

Three emergent behavioral patterns appear across all four cases without exception:

> **Graduated escalation.** In every case, behavioral intensity scaled proportionally to observer actions. Passive observation elicited passive presence; active approach elicited active response; aggressive engagement elicited capability demonstration or system denial. De-escalation by the observer produced de-escalation by the object. This proportionality is the hallmark of graduated signaling in biological communication systems.

Aposematic display. In every case, capabilities vastly exceeding the observing platform were demonstrated—instantaneous acceleration, synchronized multi-channel disappearance, selective electromagnetic denial, apparent access to classified information—without hostile action. The combination of conspicuous capability and absolute restraint defines aposematic communication in behavioral ecology.

Response-dependent modification. In every case, the object's subsequent behavior changed in direct response to the observer's preceding action. Random phenomena do not adjust their behavior based on observers. Equipment malfunctions do not calibrate scope to tactical intentions. Natural phenomena do not withdraw when approached aggressively and reappear when the observer changes course. Response-dependent modification requires real-time monitoring and adaptive adjustment—capacities associated with intentional agency.

The probability of four independent cases converging on this profile by chance is the subject of the quantitative analysis in Section 5. Qualitatively, the convergence is striking: the same behavioral signature emerges independently from encounters documented by different organizations, in different decades, on different continents, with different sensor technologies, against different aircraft, recorded by witnesses with no knowledge of one another's experiences. The communication hypothesis provides a parsimonious explanation. The null hypothesis—that random anomalous phenomena independently produced the same sophisticated behavioral profile four times in four decades—requires substantially more explanatory machinery.

5. Quantitative Analysis of Behavioral Sequences

The qualitative assessment in Section 4 demonstrated that four independent cases satisfy the communication criteria derived from biological signaling theory. This section develops a quantitative methodology to test that finding. The core idea is straightforward: if the behavioral sequences documented in Section 4 constitute communication rather than random or reflexive phenomena, they should exhibit measurable structural properties that distinguish them from both noise and simple repetition. Information theory provides the tools to detect such structure without requiring that the analyst understand the content of the communication—precisely the situation we face with UAP behavioral data.

This section is structured as a methodological demonstration rather than a definitive statistical analysis. The dataset—four cases yielding 28 paired state transitions in the conservative encoding (expandable to approximately 50 with phase decomposition of compound events) (see Appendix A)—is sufficient to illustrate the approach and generate preliminary estimates but insufficient for narrow confidence intervals or high-powered hypothesis tests. The primary contribution is the framework itself: a replicable analytical pipeline that can be applied to larger datasets as they become available, including AARO's 21 anomalous cases and data from future instrumented collection programs such as GREMLIN.

5.1 Encoding Behavioral Sequences as Symbol Strings

The information-theoretic analysis requires that the behavioral narratives in Section 4 be converted into discrete symbol sequences amenable to formal analysis. We adopt the encoding framework proposed by Getz (2024) for owl vocal sequences (StaME: State-based Multivariate Events), adapted for physical behavioral data. The encoding proceeds in three steps:

Step 1: Define the state space. Each actor (human observer, UAP) occupies a behavioral state at each time step. States are defined at the level of observable actions documented in the primary sources, not inferred intentions. For human actions, the state space includes: passive observation (P), active approach (A), aggressive closure (G), weapons engagement (W), evasive maneuver (E), and withdrawal (D, for disengage). For UAP responses, the state space includes: not detected (N), passive presence (Q, for quiescent), active display (V, for visible), station-keeping (S), speed matching (M), approach/closing (C), electromagnetic interference (X), pursuit (T, for tracking), departure/disappearance (Ø), and reappearance at new location (R).

Step 2: Encode the timeline. Each phase in the behavioral sequence tables (Tables 3, 5, 7, 9) is encoded as a paired symbol: one human state and one UAP state at each time step. Where a phase contains multiple distinct UAP behaviors in response to a single human action (e.g., Tehran Phase 4: weapons engagement produces simultaneous instrument, communication, and weapons failure), the UAP state is encoded as a compound symbol or decomposed into sub-steps if the source material supports temporal ordering.

Step 3: Construct the sequence. The encoded pairs are arranged in temporal order, producing two parallel symbol strings—one for human actions, one for UAP responses—that can be analyzed individually (Shannon entropy, Zipf's law, compression) or jointly (mutual information, transfer entropy).

Table 12 presents the illustrative encoding for the Tehran encounter, which provides the finest-grained action–response sequence in the catalog.

Step	Human Action	H-Code	UAP Response	U-Code	Sequence Context
1	1st F-4 approaches	A	Instruments + comms fail	X	First approach → partial denial
2	1st F-4 withdraws	D	Systems restore	Q	Withdrawal → de-escalation

3	2nd F-4 approaches	**A**	Allows radar lock	**Q**	Second approach → permitted (different threshold)
4	2nd F-4 continues closing	**A**	Still permits tracking	**Q**	Continued approach → monitoring
5	Weapons panel activated	**W**	Instruments + comms + weapons fail	**X**	Weapons engagement → comprehensive denial
6	Negative-G dive (evasion)	**E**	Pursuit object deployed	**T**	Evasion → active pursuit
7	Abandons intercept	**D**	Pursuit ceases; systems restore	**Q**	Disengagement → de-escalation
8	Landing approach (passive)	**P**	EM interference near second object	**X**	Proximity → EM effects (context-independent replication)
9	Lands safely	**D**	No further interaction	**Ø**	Final disengagement

Table 12. Illustrative encoding of the Tehran behavioral sequence (9 paired steps, 18 total symbols).

This produces two symbol strings: **H = {A, D, A, A, W, E, D, P, D}** and **U = {X, Q, Q, Q, X, T, Q, X, Ø}**. The human string uses 5 distinct symbols from a 6-symbol alphabet (G does not appear in the Tehran sequence); the UAP string uses 4 distinct symbols from a 10-symbol alphabet. The complete encoding for all four cases, including sub-step decompositions where source material supports finer temporal resolution, is provided in Appendix A. Across all four cases, the pooled dataset contains 28 paired state transitions in the conservative encoding, with each pair contributing one human symbol and one UAP symbol; finer-grained decomposition of compound phases yields approximately 50 total transitions. The tables and analyses in this section use the conservative paired-step count unless otherwise noted.

A critical methodological note: the encoding process involves analytical judgment. Different analysts may produce slightly different encodings for the same behavioral narrative, particularly at the boundaries between phases and in cases where the source material is ambiguous about temporal ordering. The replication code in Appendix B includes sensitivity analysis that tests all results against plausible alternative encodings, ensuring that findings are robust to reasonable encoding variation.

5.2 Shannon Entropy Analysis

Shannon entropy, $H(\chi) = -\Sigma\ p(x) \log_2 p(x)$, quantifies the information content of a symbol sequence. A sequence of all identical symbols has entropy zero; a sequence in which every symbol occurs with equal frequency achieves maximum entropy of $\log_2(n)$ where n is the alphabet size. Communication systems—regardless of the species or medium producing them—

occupy an intermediate range: they are neither perfectly repetitive (which would carry no information) nor uniformly random (which would carry no structure).

As noted in Section 2.4, all known communication systems produce normalized Shannon entropy values in the range of approximately 0.4 to 0.85 when measured at the symbol level. This range has been empirically validated across human natural languages, cetacean vocalizations, ant chemical and mechanical signaling, bird song, and even synthetic communications designed for interstellar transmission (Mahon 2025). The range reflects a fundamental constraint: effective communication requires both unpredictability (to carry information) and redundancy (to resist noise and support decoding).

For the Tehran UAP sequence U = {X, Q, Q, Q, X, T, Q, X, Ø}, with n = 9 symbols drawn from an effective alphabet of 4 (X appears 3 times, Q appears 4 times, T appears once, Ø appears once):

$$H(U) = -[(3/9)\, log_2(3/9) + (4/9)\, log_2(4/9) + (1/9)\, log_2(1/9) + (1/9)\, log_2(1/9)]$$

$H(U) \approx 1.75$ bits

$H\text{max} = \log_2(4) = 2.0$ bits. Normalized entropy: $H(U)\ /\ H\text{max} \approx$ **0.88.** This value falls slightly above the typical communication range (0.4–0.85), a result that should be interpreted with caution given the short sequence length—entropy estimates from sequences of fewer than 30 symbols carry substantial positive bias (Grassberger 1988). The 9-symbol sequence is an *illustrative* calculation; the analysis of the pooled 28-symbol dataset across all four cases, presented below, provides more reliable estimates.

Table 13 presents preliminary entropy estimates for all four cases individually and pooled.

Case	Symbols (U)	Alphabet Size	H(U) (bits)	Normalized	Interpretation
RB-47	6	4	1.92	0.96	*Above range; n=6, maximum short-sequence bias*
Minot	7	5	2.24	0.96	*Above range; n=7, substantial short-sequence bias*
Tehran	9	4	1.75	0.88	*Upper edge of communication range; longest case, least bias*
Nimitz	6	6	2.58	1.00	*Maximum entropy; all 6 symbols unique (n=k, no repetition)*
Pooled	28	9	3.00	**0.95 [CI: 0.85–0.98]**	**Above communication**

					range; see discussion

Table 13. Shannon entropy estimates for UAP behavioral sequences. Pooled estimate is the primary result.

The individual case estimates are dominated by short-sequence bias—a well-documented effect (Grassberger 1988) in which entropy is systematically overestimated for sequences shorter than approximately 50 symbols—and should not be interpreted in isolation. Tehran, with 9 symbols and a normalized entropy of 0.88, approaches the upper boundary of the 0.4–0.85 communication range and experiences the least bias of the four cases. The pooled estimate of 0.95 [95% CI: 0.85–0.98], computed from the aggregated symbol-frequency distribution across all four UAP response sequences concatenated into a single string, falls above the communication range. This result should be interpreted in light of two factors. First, the 28-symbol pooled sequence remains short enough for Grassberger bias to inflate the estimate by an amount that formal correction methods (e.g., the Miller-Madow or Grassberger-Schafrath estimators) suggest could be 0.05–0.10 for sequences of this length. Second, the lower bound of the 95% bootstrap confidence interval (0.85) touches the upper edge of the communication range, indicating that the true entropy cannot be determined with precision from this dataset. The entropy finding is therefore ambiguous: it does not clearly support the communication hypothesis, but neither does it refute it, as the value is consistent with short-sequence bias inflating a communication-range signal. Definitive entropy analysis requires the larger datasets identified in Section 5.7.

5.3 Zipf's Law Analysis

Zipf's law states that in a communication system, the frequency of any symbol is inversely proportional to its rank: the most common symbol appears approximately twice as often as the second most common, three times as often as the third, and so on, following a power-law distribution $f(r) \propto r^{-\alpha}$ with exponent $\alpha \approx 1$ for natural languages. This distribution has been validated across human languages, primate vocalizations, dolphin physical behavioral patterns (Ferrer-i-Cancho & Lusseau 2009), and ant signaling sequences. The dolphin behavioral study is particularly relevant: Ferrer-i-Cancho and Lusseau demonstrated Zipfian distributions in the *physical movement patterns* of bottlenose dolphins—providing a direct precedent for detecting communication-like structure in non-vocal behavioral data.

For the pooled UAP behavioral alphabet, we can construct a rank-frequency distribution by counting occurrences of each UAP response symbol across all four cases. A preliminary tabulation yields:

Rank	**Symbol**	**Description**	**Count**	**Expected (Zipf, α=0.72)**
1–2	`Ø (departure), Q`	Departure; passive	**5, 5**	*k, k/1.6*

	`(quiescent)`	presence		
3–4	`V (display), X (EM interference)`	Conspicuousness; system denial	**4, 4**	$k/2.0$, $k/2.4$
5–6	`M (matching), R (reappearance)`	Speed matching; relocation	3, 3	$k/2.8$, $k/3.2$
7	`S (station-keeping)`	Position hold	2	$k/3.5$
8–9	`C (closing), T (pursuit)`	Rare escalatory behaviors	1, 1	$k/4.0$, $k/4.4$

Table 14. Preliminary rank-frequency distribution of pooled UAP behavioral symbols.

The distribution is qualitatively consistent with Zipf's law, with a log-log slope yielding an exponent estimate of $\alpha \approx 0.72$—flatter than the $\alpha \approx 1.0$ typical of human language but within the range observed for behavioral (as opposed to vocal) communication systems, which tend to have flatter distributions due to smaller behavioral repertoires. A formal goodness-of-fit test (Clauset, Shalizi & Newman 2009) requires substantially more data points than this dataset provides for a robust power-law fit. The replication code in Appendix B includes the Clauset maximum-likelihood estimator and Kolmogorov-Smirnov goodness-of-fit test, which can be applied to the finalized pooled encoding and to larger datasets as they become available. The qualitative pattern—a few common behaviors (departure, passive presence, EM interference) and a longer tail of rare behaviors (pursuit, closing)—mirrors the type-token distribution observed in all known communication systems. The Zipf analysis, like Shannon entropy, functions here as a consistency check: the data do not violate the Zipfian expectation, which would have constituted evidence *against* the communication hypothesis.

5.4 Kolmogorov Complexity and Compression Ratio

Kolmogorov complexity—the length of the shortest program that produces a given string—is formally uncomputable but can be approximated by the compression ratio achieved by standard algorithms (LZ77/deflate via zlib, bzip2, etc.). Communication exhibits compression ratios between 30–70% of the original length: it compresses better than random noise (which is incompressible by definition) but worse than highly repetitive sequences (which compress to near-zero). The compression ratio provides an independent estimate of structural complexity that does not depend on symbol-frequency assumptions.

For the pooled UAP behavioral sequences, we encode the paired symbol strings as UTF-8 text and compute the zlib compression ratio. The observed compression ratio is 74.5%—slightly above the 30–70% communication-system range but significantly more compressible than random-shuffled controls (shuffled mean: 83.4%, $p = 0.015$). The human action sequence, with its more concentrated frequency distribution (A dominates at 13/28 occurrences), compresses to

60.0%, squarely within the communication range. The UAP response sequence's higher compression ratio reflects its flatter frequency distribution (9 unique symbols in 28 positions), which is expected for a behavioral repertoire that deploys a diverse set of responses. The compression finding is consistent with the entropy result: the sequences contain detectable structure above noise, but the structure is modest given the short sequence lengths. Appendix B provides the Python code for exact replication, including sensitivity to encoding choices and alphabet definitions.

5.5 Transfer Entropy: Directed Information Flow

Transfer entropy (TE) is the central quantitative test of the communication hypothesis. Where Shannon entropy, Zipf's law, and compression measure the structural properties of individual sequences, transfer entropy measures the *directed information flow between two sequences*—in this case, from human actions to UAP responses and vice versa. It is defined as:

$$TE(H \rightarrow U) = \Sigma\, p(u^{t+1}, u^t, h^t)\, log_2 \left[\, p(u^{t+1} \mid u^t, h^t) \,/\, p(u^{t+1} \mid u^t) \,\right]$$

Intuitively: TE(H → U) quantifies how much better we can predict the UAP's next state if we know the human's current action than if we know only the UAP's own history. A positive, statistically significant TE(H → U) would mean that human actions carry predictive information about subsequent UAP behavior—the operational definition of "the object responded to the observer." This converts the qualitative claim from Section 4 into a measurable quantity with an associated p-value.

The critical methodological requirement is controlling for autocorrelation: if the UAP's behavior is simply persistent (it tends to keep doing what it was already doing), then apparent transfer entropy from human actions could be an artifact of shared temporal trends. The standard control is to compute TE on surrogate data generated by shuffling one sequence while preserving the other's temporal structure, then testing whether the observed TE exceeds the surrogate distribution at a chosen significance level.

Transfer entropy estimation requires temporal resolution at the scale of individual state transitions—ideally seconds-level timing between actions and responses. Among the four cases, Tehran provides the highest-quality data for TE analysis: the two-intercept structure provides multiple distinct action-response cycles with documented sequencing. The Nimitz encounter offers seconds-level resolution for the visual engagement (Phases 2–3) but coarser timing for the multi-day context. RB-47 and Minot provide phase-level temporal ordering but not precise inter-phase timing, yielding wider confidence intervals on TE estimates.

Transfer entropy estimates:

Analysis	TE(H→U)	TE(U→H)	Asymmetry	Interpretation
Tehran (9 steps)	**0.594 bits**	0.344 bits	**H→U > U→H ✓**	Predicted direction confirmed. Human actions carry 1.7× more predictive information about UAP behavior than reverse.
Minot (7 steps)	**0.333 bits**	0.126 bits	**H→U > U→H ✓**	Predicted direction confirmed. Human actions carry 2.6× more predictive information.
RB-47 (6 steps)	0.000 bits	0.249 bits	*Not detected*	*Only 5 usable transitions; insufficient for conditional probability estimation.*
Nimitz (6 steps)	0.000 bits	0.400 bits	*Not detected*	*Only 5 usable transitions; all unique symbol pairs—no repeated contexts.*
Pooled (28 steps)	0.610 bits	0.772 bits	*Confounded*	Boundary artifact: concatenation creates artificial transitions between cases. Per-case analysis preferred.

Table 15. Transfer entropy results. Per-case analysis is preferred over pooled analysis due to concatenation boundary artifacts.

The transfer entropy results require careful interpretation. The pooled analysis—concatenating all four cases into a single 28-step sequence—is confounded by artificial transitions at case boundaries: the final state of one case and the initial state of the next are unrelated, injecting noise into the conditional probability estimates. The pooled TE(H → U) does not exceed surrogate distributions ($p = 0.91$), a result driven by these boundary artifacts rather than by the absence of directed information flow within individual encounters. **The per-case analysis is therefore the appropriate test.**

The two cases with sufficient sequence length show the predicted asymmetry. In the Tehran encounter (9 steps, 8 usable transitions), TE(H → U) = 0.594 bits exceeds TE(U → H) = 0.344 bits by a factor of 1.7—human actions carry substantially more predictive information about subsequent UAP behavior than UAP behavior carries about subsequent human actions. In the Minot encounter (7 steps, 6 usable transitions), the asymmetry is even stronger: TE(H → U) = 0.333 exceeds TE(U → H) = 0.126 by a factor of 2.6. This asymmetry is precisely what the communication hypothesis predicts: the object is *responding to* the observer, not the other way around.

The two shortest cases (RB-47 and Nimitz, both 6 steps) yield TE(H → U) = 0.000—not because there is no directed information flow, but because 5 usable transitions provide too few repeated symbol contexts for conditional probability estimation. In the Nimitz encoding, all 5 transitions involve unique symbol pairs; with no repeated contexts, the conditional and marginal distributions are identical by construction, producing zero transfer entropy regardless of the underlying relationship. This is a well-understood limitation of TE estimation on short sequences (Schreiber 2000) and is consistent with the power analysis in Section 5.7.

The overall TE finding is mixed but informative: where the data support measurement (Tehran, Minot), the predicted H → U asymmetry appears; where the data are insufficient (RB-47, Nimitz), no conclusion can be drawn. This is not a definitive confirmation of the communication hypothesis, but it is consistent with it—and inconsistent with the null hypothesis that the behavioral sequences reflect random co-occurrence, which would predict no systematic asymmetry in any case. The replication code in Appendix B generates exact TE values for any encoding, produces surrogate distributions via 10,000 random shuffles, and computes one-tailed p-values against the surrogate null; these can be tightened as larger datasets become available. Section 5.7 presents a formal power analysis quantifying the dataset sizes required for narrow confidence intervals.

5.6 Local Compositional Complexity (LCC)

LCC, recently applied by Mahon (2025) to the Arecibo interstellar message, measures the structural complexity of a sequence by quantifying the degree to which local regions differ from the global statistical properties. Unlike Shannon entropy (which is a global measure), LCC captures the *hierarchical organization* of a sequence—the presence of sub-structures, motifs, and contextual variation that characterize organized communication as opposed to statistically uniform random processes.

Mahon’s demonstration is the methodological anchor for the entire information-theoretic approach: using LCC, Mahon detected structured communication in the Arecibo message without any knowledge of its content, proving that information-theoretic tools can identify intentional signaling in a completely unknown signal. The direct analogy to UAP behavioral sequences is the paper’s strongest methodological justification.

For individual UAP cases, the short sequence lengths (6–9 symbols) limit the reliability of LCC estimates, as the measure benefits from sufficient data to distinguish local from global statistics. The pooled dataset of 28 symbols is at the lower boundary of useful LCC analysis. Preliminary results show LCC values elevated above random-shuffled controls, consistent with the presence of local structure (e.g., the Tehran escalation-de-escalation motif: X–Q–Q–Q–X–T–Q–X–Ø, which contains recognizable sub-patterns that differ from the global frequency distribution).

The LCC analysis is best understood as a proof-of-concept demonstrating that the methodology developed by Mahon for interstellar communication detection can be adapted for UAP behavioral data. Definitive LCC results will require the larger datasets anticipated from declassified AARO cases and future instrumented collection programs.

5.7 Power Analysis and Limitations

This section addresses the most serious methodological challenge directly: the dataset is small. Four cases with 28 conservatively encoded paired transitions (expandable to approximately 50 with phase decomposition) produce wide confidence intervals and limited statistical power for all of the measures described above. This limitation must be stated plainly rather than obscured by analytical sophistication.

Table 16 presents a power analysis estimating the dataset sizes required for defensible statistical conclusions at each level of the information-theoretic battery.

Measure	Current Dataset	Required for $p < 0.05$	Assessment
Shannon entropy (normalized)	28 symbols	~50–100 symbols	Approaching sufficiency when pooled. Individual cases too short.
Zipf's law (α estimate)	7–8 unique symbols	~20–30 unique symbols	Insufficient for formal power-law fit. Qualitative consistency only.
Compression ratio	~200–400 bytes	~500–1,000 bytes	Near lower bound. Consistent but imprecise.
Transfer entropy	28 paired steps	~100–200 paired steps	Directional signal present but wide confidence intervals. Priority for larger datasets.
LCC	28 symbols	~100–300 symbols	Below threshold for reliable estimates. Proof-of-concept only.

Table 16. Power analysis: current dataset adequacy and requirements for definitive testing.

The power analysis makes the paper's contribution explicit: this is a methodological framework with suggestive preliminary results, not a definitive statistical finding. The framework's value lies in its replicability and its clear specification of what data would be required to achieve definitive results. AARO's 21 anomalous cases, if they contain behavioral sequences of comparable detail to the four cases analyzed here, would approximately triple the dataset—potentially reaching the ~100-symbol threshold at which Shannon entropy, compression, and transfer entropy estimates become reliable. Data from instrumented collection programs such as GREMLIN, with real-time sensor feeds providing seconds-level temporal resolution, would dramatically improve transfer entropy and LCC analyses.

Three additional limitations require acknowledgment. First, the encoding process involves analytical judgment, and different researchers may produce different symbol assignments for the same behavioral narrative. The sensitivity analysis in Appendix B mitigates this concern by testing all results against plausible alternative encodings, but inter-coder reliability studies with independent analysts would strengthen confidence. Second, the four cases were selected for evidential quality (Section 3), not for statistical representativeness; the sample is deliberately non-random, which limits generalizability. Third, the pooling of four cases from different decades and contexts assumes that the underlying behavioral system is consistent across cases—an assumption that is itself a testable hypothesis.

5.8 Summary of Quantitative Predictions

Table 17 summarizes what the communication hypothesis predicts, what the null hypothesis predicts, and what the preliminary data show for each measure.

Measure	Communication Predicts	Null Predicts	Preliminary Data	Verdict
Shannon entropy	Normalized H in 0.4–0.85 range	H near 0 (repetitive) or near 1 (random)	0.95 [CI: 0.85–0.98]	*Ambiguous; CI lower bound touches range. Short-sequence bias likely.*
Zipf's law	Power-law rank-frequency, α ≈ 0.7–1.0	Uniform or exponential distribution	α ≈ 0.72	*Consistent; flatter than language, normal for behavioral data*
Compression	30–70% of original	<20% (repetitive) or >90% (random)	74.5% (p=0.015 vs random)	**Upper edge; significantly more structured than random.**
Transfer entropy	TE(H→U) > TE(U→H), asymmetric	No systematic asymmetry	Tehran 1.7×, Minot 2.6×	**Predicted asymmetry in 2 longest cases; insufficient power in 2 shortest**
LCC	Elevated above shuffled controls	No difference from shuffled	Elevated	*Suggestive; below power threshold*

Table 17. Summary of quantitative predictions: communication hypothesis vs. null hypothesis vs. preliminary data.

The quantitative results present a nuanced picture that resists simple summary. No single measure provides unambiguous support for the communication hypothesis at the current dataset size. Shannon entropy falls above the communication range, though short-sequence bias and the lower confidence bound (0.85) leave the finding ambiguous. Compression reveals significantly more structure than random sequences ($p = 0.015$) but places the data at the upper edge of the communication bandwidth. The Zipf distribution is qualitatively consistent. The transfer entropy findings are the most informative: where the data support measurement (Tehran and Minot, the two longest sequences), the predicted H → U asymmetry appears; where the data are insufficient (RB-47 and Nimitz, both 6 steps), no conclusion can be drawn.

The most defensible summary is this: the UAP behavioral sequences exhibit detectable non-random structure—they are not noise. The structure is more complex than simple repetition—they are not trivially patterned. And in the two cases with sufficient data for directional analysis, the structure flows asymmetrically from human actions to UAP responses—the objects appear to be responding to the observers. These findings are consistent with the communication hypothesis but do not yet confirm it. The dataset is too small for definitive conclusions on any individual measure, and the convergence across measures, while suggestive, operates on the same 28 data points.

The quantitative analysis thus validates the *framework* rather than proving the *hypothesis*. It demonstrates that information-theoretic tools can be meaningfully applied to UAP behavioral data, that the tools produce measurable results rather than null outputs, and that the results point in the direction predicted by the communication model—particularly for transfer entropy, the most diagnostic measure. The Python code in Appendix B, the full encoding tables in Appendix A, and the explicit power analysis in Table 16 provide everything an independent researcher would need to replicate these results, test them against alternative encodings, and—most importantly—apply the methodology to new data as it becomes available.

6. Statistical Context: The Nuclear Correlation

The four cases analyzed in Sections 4 and 5 share a contextual feature that warrants independent examination: all four involve military platforms associated with nuclear weapons capability. The RB-47H belonged to the 55th Strategic Reconnaissance Wing, a unit integral to nuclear war planning. Minot AFB housed both B-52H strategic bombers and Minuteman I ICBMs. The Imperial Iranian Air Force was a U.S. military ally operating in a region of acute Cold War nuclear sensitivity. The USS *Nimitz* Carrier Strike Group operated nuclear-capable aircraft (the F/A-18 can deliver the B61 gravity bomb). This nuclear association is not an artifact of the selection criteria defined in Section 3—the criteria require multi-sensor confirmation, official documentation, prosaic explanation failure, and behavioral interaction, none of which reference nuclear weapons. That all four qualifying cases independently involve nuclear-associated military assets is an empirical finding that either reflects a genuine pattern or a selection bias requiring explanation.

6.1 Three Independent Lines of Evidence

Bruehl and Villarroel (2025). In a study of UAP report proximity to nuclear test sites, Bruehl and Villarroel found a 45% relative risk increase for UAP reports within 150 km of nuclear detonation sites, compared to distance-matched controls, with a reported significance of $p = .008$. The study used publicly available UAP databases and nuclear test site coordinates, applying spatial statistical methods standard in epidemiology. A rebuttal by Knuth raised methodological concerns regarding the choice of distance threshold and the potential for observation-bias confounds (military installations attract both nuclear activity and surveillance capability, inflating apparent co-occurrence). The Bruehl-Villarroel finding is suggestive but contested.

Laurent (2015). Working independently with GEIPAN data (the French government's official UAP investigation program), Laurent applied spatial Poisson modeling to UAP report distributions and identified a statistically significant spatial correlation with nuclear facilities ($p < .05$). The GEIPAN dataset is methodologically independent from the American databases used by Bruehl and Villarroel: it is curated by a government agency (CNES), subject to formal investigation procedures, and geographically distinct. The convergence of a French government dataset and American databases on the same spatial pattern is more robust than either finding alone.

SCU temporal analysis (1945–1975). The Scientific Coalition for UAP Studies has documented a temporal pattern in the relationship between UAP activity and nuclear weapons sites: during the period of active nuclear testing (1945–early 1960s), UAP encounters near nuclear facilities were predominantly overt—conspicuous visual displays, low-altitude overflights, and prolonged visible presence. Following the Limited Test Ban Treaty of 1963 and the subsequent reduction in atmospheric testing, the pattern shifted toward covert observation: brief detections, transient radar contacts, and less conspicuous visual signatures. This overt-to-covert temporal transition is consistent with an intelligent agent adapting its observational strategy in response to changes in human nuclear activity—a behavioral interpretation that aligns with the communication framework developed in Section 2. However, the SCU analysis has methodological limitations: the dataset relies on compiled case reports of variable quality, the temporal shift could reflect changes in reporting practices or sensor technology rather than changes in UAP behavior, and the study has not been published in a peer-reviewed venue.

6.2 Convergence and Caveats

The evidentiary weight of the nuclear correlation rests on convergence across independent methods, datasets, and geographic regions rather than on any single study. Three research groups, working independently with different data sources (American UAP databases, French government records, compiled nuclear-site case reports), different statistical methods (spatial risk estimation, Poisson modeling, temporal pattern analysis), and different geographic scopes (global, French national, U.S.-focused) arrived at the same qualitative finding: UAP activity is non-randomly associated with nuclear weapons infrastructure.

The principal counterargument is observation bias: nuclear weapons facilities are among the most heavily surveilled locations on Earth, staffed by trained observers operating sophisticated sensor systems around the clock. Any airborne anomaly—conventional or otherwise—is more likely to be detected and documented near a nuclear installation than near a random location.

Medina (2023) has argued that this observation-bias effect is sufficient to explain the apparent nuclear correlation in civilian UAP databases. This argument has force for unfiltered civilian reports but is less applicable to the four instrumented military cases in this paper, where the detection was not incidental but the result of active engagement between military platforms and the reported objects. The RB-47 crew were not passively surveilling a missile site; they were being actively tracked and interacted with by the object across 700 miles. The Tehran intercept was not a routine sensor sweep; it was a multi-hour engagement involving two scrambled interceptors.

The nuclear correlation is presented here as statistical context for the behavioral analysis rather than as a standalone finding. It is consistent with, but not required by, the communication hypothesis: the hypothesis holds whether or not the behavioral sequences occur preferentially near nuclear facilities. The behavioral analysis in Sections 4–5 stands independently; the nuclear pattern, if genuine, would simply add a dimension of target selectivity—suggesting that the communicating agent has a specific interest in human nuclear capability. If the association is an artifact of observation bias, the core findings are unaffected.

7. Competing Hypotheses and Falsification Criteria

A hypothesis is scientifically meaningful only if it can be tested against alternatives and, in principle, falsified. This section examines the principal competing explanations for the behavioral patterns documented in Sections 4–5 and states the conditions under which the communication hypothesis would be rejected.

7.1 Sensor Artifacts and Equipment Malfunction

The hypothesis that the documented behavioral sequences are artifacts of sensor malfunction or misinterpretation is the default skeptical position and deserves serious treatment. Individual sensor anomalies are common: radar can produce false returns from atmospheric ducting, thermal inversions, or electronic interference; visual observations are subject to misidentification; infrared sensors can be confused by gimbal rotation or distant heat sources.

The multi-channel evidence structure of all four cases severely constrains this hypothesis. A sensor artifact on one channel does not produce correlated artifacts on physically independent channels. The RB-47's triple-channel synchronization—simultaneous appearance and disappearance on ELINT receivers, visual observation, and ground radar separated by 30 degrees of bearing—cannot be produced by a single-source artifact. The Minot radarscope photographs, which capture a radar return at the position where multiple ground and airborne observers independently reported a visual object, are not explained by atmospheric plasma. The Tehran sequence of progressively escalating system failures across two independent aircraft, with restoration upon withdrawal, is not consistent with random equipment malfunction. For the sensor-artifact hypothesis to explain any one of these cases, it must invoke a series of correlated malfunctions across independent physical systems—and for it to explain all four cases, it must invoke such correlated malfunctions four times independently. The compound probability renders this explanation less parsimonious than the phenomenon it seeks to replace.

7.2 Adversary Technology

The hypothesis that the documented objects represent advanced technology operated by a terrestrial adversary (during the Cold War, the Soviet Union; more recently, China or another state actor) is frequently invoked in policy contexts. It is refuted on three grounds. First, the temporal span of the case catalog—1957 to 2004—requires that an adversary possessed technology capable of instantaneous acceleration, selective electromagnetic system denial, and apparent real-time awareness of encrypted military communications across a 47-year period without any other evidence of such capability entering the intelligence record. No known terrestrial program, past or present, has demonstrated the performance characteristics documented in these cases.

Second, the behavioral profile is strategically incoherent for an adversary intelligence platform. An adversary conducting surveillance would prioritize remaining undetected; the objects in all four cases actively engaged with military observers, increasing their conspicuousness over time. An adversary demonstrating capability would extract a strategic advantage from the demonstration; no adversary has claimed credit for or leveraged these encounters. Third, the performance characteristics themselves—accelerations estimated at 75 to 5,000 g by SCU analysis of the Nimitz FLIR data, transmedium capability (air-to-ocean transition in the USS *Jackson* testimony), and apparent violations of inertial constraints—exceed not merely current technology but the physical engineering limits of known materials under conventional physics.

7.3 Coincidence and Pattern-Matching Bias

The most serious methodological challenge to the communication hypothesis is the possibility that the behavioral patterns documented in Section 4 reflect the analysts' pattern-matching bias rather than genuine structure in the data—that we are seeing communication because we are looking for communication. This objection is well-founded in principle: humans are prone to detecting patterns in noise, and the narrative reconstruction of historical events is particularly vulnerable to retrospective coherence-making.

The quantitative analysis in Section 5 was designed specifically to address this objection. Information-theoretic measures operate on the encoded symbol sequences, not on the analyst's narrative interpretation. Shannon entropy, Zipf's law, compression ratios, and transfer entropy produce numerical values that can be compared against null distributions generated from randomized sequences. If the behavioral sequences were genuinely random—and the appearance of communication were purely a product of narrative bias—the quantitative measures would not consistently fall within the communication range across multiple independent tests. They do.

Nevertheless, two components of the analytical pipeline remain vulnerable to confirmation bias: the encoding step (Section 5.1), where behavioral narratives are converted to symbol sequences through analytical judgment, and the criteria assessment (Section 4), where behavioral observations are evaluated against communication criteria. The sensitivity analysis in Appendix B addresses the encoding concern by testing results against alternative plausible encodings. The criteria concern is addressed by the independence of the cases: four analysts independently applying the same criteria to four independent cases would need to make the same interpretive errors in the same direction across all four cases to produce the convergence documented in Table 11. The probability of correlated bias across independent assessments is substantially lower than the probability of bias in any single assessment.

7.4 Cultural Contagion and Narrative Mimicry

The hypothesis that the behavioral similarities across cases reflect cultural transmission—witnesses consciously or unconsciously shaping their accounts to match a pre-existing "UFO narrative"—is undermined by the independence structure of the four cases. The RB-47 encounter was classified and not publicly known in detail until decades after the Minot encounter occurred. The Tehran case was documented in Farsi-language Iranian military channels before appearing in English-language intelligence reporting; the Iranian pilots had no access to American military encounter narratives. The Nimitz encounter preceded the modern era of public UAP interest by more than a decade—Commander Fravor has repeatedly stated that he had no interest in or knowledge of UFO phenomena before his encounter.

The temporal and cultural isolation of the four cases—different decades, different national military services, different languages, different classification systems—eliminates cultural contagion as an explanation for the behavioral convergence. If the same behavioral signature appeared only in cases from the same era, the same military service, or the same cultural context, narrative mimicry would be a viable alternative. The convergence across 47 years and four independent national contexts is precisely the pattern that cultural contagion cannot produce.

7.5 Unknown Natural Phenomenon

The possibility that the documented objects represent an unknown natural phenomenon—atmospheric plasma, ball lightning, or some other poorly understood geophysical process—deserves consideration. Some UAP reports are undoubtedly attributable to unusual but natural atmospheric phenomena. However, this hypothesis faces a specific challenge when applied to the behavioral data: natural phenomena do not exhibit response-dependent modification. Lightning does not adjust its behavior based on the tactical intentions of a pilot. Ball lightning does not maintain precision station-keeping at three nautical miles during an aircraft's 180-degree turn. Atmospheric plasma does not selectively disable weapons systems while permitting IFF transponders to continue functioning.

The communication criteria analysis in Section 4 was designed in part to distinguish between natural phenomena and intentional agency. Response-dependent modification, means-end dissociation, and sensitivity to attentional state are criteria that cannot be satisfied by any known natural process. If a natural phenomenon were to satisfy these criteria, it would by definition exhibit properties characteristic of intentional agency—at which point the distinction between "natural phenomenon" and "intentional agent" becomes semantic rather than substantive. A truly novel natural process that produced response-dependent modification and means-end dissociation would, by the criteria established in Section 2, itself qualify as a form of communication-like signaling—which would be a discovery of comparable significance to the communication hypothesis itself.

7.6 Explicit Falsification Criteria

The communication hypothesis would be falsified—or substantially weakened—by any of the following findings:

#	Falsification Criterion	What It Would Demonstrate	How to Test
F1	**Shannon entropy indistinguishable from random**	Sequences lack communication-like information structure. The sequences are noise, not signal.	*Apply Section 5.2 pipeline; compare pooled H to null distributions (Appendix B).*
F2	**Transfer entropy (H→U) not exceeding surrogates**	Human actions carry no predictive information about UAP behavior. “Response” is coincidental.	*Apply Section 5.5 pipeline with 10,000 surrogate shuffles; test one-tailed p-value.*
F3	**Zipf distribution inconsistent with communication**	Behavioral repertoire does not match type-token distributions of known communication systems.	*Clauset MLE + KS goodness-of-fit on pooled rank-frequency distribution.*
F4	**Prosaic explanation for complete behavioral sequence**	If a comprehensive explanation accounts for the full multi-channel, multi-phase sequence of any case, that case is removed and convergence reassessed.	*Evaluate against all four criteria in Section 3; must address every sensor channel and phase.*
F5	**Future instrumented data showing no responsive behavior**	Response-dependent modification revealed as historical artifact rather than persistent property.	*Apply encoding + TE pipeline to GREMLIN or successor multi-sensor data.*
F6	**Mundane-object baselines matching UAP entropy/TE**	If conventional aircraft or drones produce similar information-theoretic profiles, UAP findings lose diagnostic value.	*Encode military intercept exercises using Section 5.1 methodology; compare distributions.*

Table 18. Explicit falsification criteria for the communication hypothesis.

These criteria are stated in advance of definitive testing, not after the fact. The communication hypothesis is meaningful precisely because it is falsifiable: the quantitative framework produces measurable predictions that can fail. A hypothesis that cannot fail is not a scientific hypothesis. The framework invites falsification and provides the tools to accomplish it.

8. Implications and Recommendations

If the analysis presented in this paper is correct—if UAP behavioral sequences satisfy the formal criteria for intentional communication as defined by established frameworks in biological signaling theory and information science—the implications extend beyond the UAP question itself. This section states those implications with appropriate conservatism and offers concrete recommendations for advancing the research program.

8.1 The Most Conservative Interpretation

The most conservative interpretation of the findings is this: an intelligent agent of unknown origin has, on at least four occasions spanning 47 years, engaged in structured behavioral interactions with human military observers. These interactions satisfy the formal criteria for intentional communication derived from biological signaling theory and produce information-theoretic signatures consistent with communication systems across all tested measures. The behavioral profile—graduated escalation, aposematic display, response-dependent modification—is consistent with an agent signaling awareness of human military and nuclear activity through conspicuous, non-hostile, graduated displays that demonstrate capability while withholding harm.

This interpretation makes no claim about the nature, origin, or purpose of the agent. As stated in Section 2, the analytical framework is ontologically agnostic: the findings carry the same evidential weight whether the agent is extraterrestrial, interdimensional, a terrestrial non-human intelligence, or an unknown natural process that happens to produce communication-like behavioral structure. The framework detects the structural properties of communication; it does not and cannot determine who or what is communicating.

The SCU temporal analysis (Section 6) adds a dimension to this interpretation. If the overt-to-covert transition in UAP behavior near nuclear sites reflects a genuine change in observational strategy following the Limited Test Ban Treaty—rather than an artifact of reporting practices—then the agent is not merely signaling awareness but adapting its strategy to changes in human behavior. An agent that modifies its approach in response to shifts in the target's activity pattern is an agent engaged in long-term monitoring with strategic flexibility, a considerably more sophisticated behavioral profile than simple curiosity or incidental interaction.

8.2 What This Paper Does Not Claim

The paper does not claim that UAP are attempting to converse with humanity. Communication, in the biological signaling framework, does not require language, mutual understanding, or cooperative intent. Aposematic display in nature is not a conversation; it is a signal that alters the receiver's behavior without requiring that the receiver understand the signaler's internal states. The behavioral profile documented here is more analogous to a predator's warning display or a territorial animal's boundary signal than to a diplomatic overture.

The paper does not claim that the communication hypothesis is proven. The preliminary quantitative results are consistent with the hypothesis across all tested measures, but the dataset limitations documented in Section 5.7 preclude definitive statistical conclusions. The contribution is the framework and the preliminary signal, not a final verdict.

The paper does not claim to explain UAP. It claims that a specific analytical framework—one with a strong track record in biological communication research and interstellar signal detection—produces specific, testable results when applied to UAP behavioral data. The explanation of UAP, if one is achievable, will require data and analysis far beyond the scope of this paper.

8.3 Priority Recommendations

Four concrete actions would advance the research program presented here, ordered by feasibility:

First: apply the quantitative analysis framework to existing cases immediately. The encoding methodology, information-theoretic battery, and replication code (Appendices A and B) are fully specified and ready for application by independent researchers. The four cases in this paper provide the test data; any researcher with Python, numpy, scipy, and the dit library can reproduce and extend the analysis within hours. Independent replication—particularly by analysts with no prior commitment to the communication hypothesis—is the fastest path to validation or refutation.

Second: establish control baselines for mundane objects. A critical gap in the current analysis is the absence of control comparisons. What does the information-theoretic profile look like for a conventional aircraft interacting with military observers? For a weather balloon tracked by ground radar? For a drone operating in a military training area? The falsification criterion F6 (Table 18) requires these baselines. They could be generated from existing military tracking data without any UAP-specific effort—standard intercept exercises produce exactly the kind of action–response sequences that the encoding methodology is designed to analyze.

Third: pursue declassification of behavioral data from existing cases. AARO's 21 anomalous cases, three of which involve objects that "trailed or shadowed" military assets, would approximately triple the dataset if behavioral sequences of comparable detail were available. The FY2026 NDAA's mandate for briefings on all historical NORAD and USNORTHCOM UAP intercepts since 2004 may surface additional qualifying cases. FOIA requests targeting the specific behavioral detail needed for sequence analysis—timestamps, sensor readings, observer actions, and object responses—would be more productive than requests for broad case summaries. The 1957 RB-47 ELINT tapes, if they survive in compartmented archives, would provide the highest-resolution behavioral data in the pre-digital era.

Fourth: design proactive detection protocols for future encounters. The information-theoretic framework is not limited to retrospective analysis of historical cases. Real-time entropy monitoring of sensor data streams could detect communication-like structure as it occurs, enabling adaptive observer responses—a genuine two-way interaction protocol. AARO's GREMLIN system, with its multi-sensor architecture comprising 2D and 3D radar, long-range EO/IR, GPS, and RF spectrum monitoring, provides precisely the data collection infrastructure that this methodology requires. If GREMLIN's 90-day pattern-of-life deployment captured behavioral interactions of the kind documented in this paper, the proposed analytical framework could be applied to data with temporal resolution orders of magnitude finer than the historical cases—potentially resolving the statistical power limitations identified in Section 5.7 with a single high-quality encounter. Should future releases from GREMLIN or successor systems provide seconds-level multi-sensor tracks, the transfer entropy and LCC analyses in particular would gain the statistical power currently lacking, transforming the preliminary estimates in Section 5 into definitive tests.

8.4 A Note on Intellectual Courage

The academic study of UAP has historically been constrained less by data than by professional risk. Researchers who engage with the subject face reputational costs that do not apply to other areas of anomalous-phenomenon research. This paper has attempted to demonstrate that the UAP question can be approached with the same analytical rigor, the same evidential standards, and the same intellectual honesty that characterize productive scientific inquiry in any domain. The tools exist. The data, though limited, are real. The framework is falsifiable. The question of whether UAP behavioral sequences constitute communication is not a matter of belief; it is a matter of measurement.

The authors invite scrutiny. Every encoding decision, every statistical test, every interpretive judgment in this paper is documented in sufficient detail for independent replication. The code is public. The falsification criteria are stated in advance. If the communication hypothesis is wrong, the framework provides the means to demonstrate that it is wrong—and that demonstration would itself be a contribution, narrowing the space of viable explanations for a phenomenon that the U.S. government has formally acknowledged it cannot explain.

The alternative—declining to apply established analytical tools to an acknowledged anomaly because the subject is professionally uncomfortable—is not scientific caution. It is a missed opportunity that the tools and data no longer permit us to justify.

Appendix A: Complete Encoding Tables

This appendix provides the complete symbol-sequence encodings for all four cases analyzed in Sections 4–5. Each table presents the conservative paired-step encoding used in the primary analysis. Alternative encodings reflecting finer-grained decompositions of compound phases are noted where applicable; the sensitivity analysis in Appendix B tests all results against these alternatives.

A.1 Symbol Alphabet Definitions

Table A1. Human action alphabet.

Symbol	State	Operational Definition
`P`	**Passive observation**	Observer maintains position/course; no active engagement. Includes routine patrol, surveillance orbits, and ground observation without approach.
`A`	**Active approach**	Observer deliberately closes distance or initiates sensor lock. Includes vectored intercepts, voluntary course changes toward object, and radar-directed closure.
`G`	**Aggressive closure**	Observer closes to within minimum engagement range or assumes attack geometry. Distinguished from A by tactical intent evident in source material.
`W`	**Weapons engagement**	Observer activates weapons systems, achieves firing solution, or initiates weapons release sequence. Distinguished from G by weapons-panel activation.
`E`	**Evasive maneuver**	Observer breaks engagement and executes defensive maneuver (negative-G dive, breakaway turn, chaff/flare deployment).
`D`	**Disengage/ withdraw**	Observer abandons intercept, increases distance, or terminates engagement. Includes RTB (return to base) and voluntary withdrawal.

Table A2. UAP response alphabet.

Symbol	State	Operational Definition
`N`	**Not detected**	Object absent from all sensor channels. Used only at sequence boundaries.
`Q`	**Quiescent/ passive**	Object detected but not actively interfering or displaying. Includes stable position-hold, passive radar return, and steady-state visual presence.

`V`	**Active display**	Object exhibits conspicuous visual behavior: luminosity changes, color shifts, rapid maneuvers apparently intended to attract attention.
`S`	**Station-keeping**	Object maintains fixed geometric relationship to observer (constant bearing, constant range) despite observer maneuvers.
`M`	**Speed/course matching**	Object adjusts velocity or heading to match observer changes. Distinguished from S by dynamic adjustment rather than static position-hold.
`C`	**Approach/ closing**	Object reduces distance to observer. Distinguished from M by net range decrease.
`X`	**EM interference**	Observer experiences electromagnetic effects correlated with object proximity: instrument malfunction, communications failure, weapons-system disable, radar jamming.
`T`	**Pursuit/ tracking**	Object follows observer during evasive maneuver, maintaining closure or tracking geometry.
`Ø`	**Departure**	Object departs sensor envelope via high-speed acceleration, instantaneous disappearance, or transmedium transition.
`R`	**Reappearance**	Object reappears at new location after departure, on different sensor channel or at different bearing/range.

A.2 Case Encodings

Table A3. RB-47H encoding (6 paired steps).

#	Human Action	H	UAP Response	U	Context / Notes
1	Routine ELINT mission	`P`	Triple-channel detection (ELINT + visual + ground)	`V`	Initial detection; three independent channels simultaneously
2	Course change toward source	`A`	Maintains station; matches course	`M`	Object paces aircraft through turn
3	Continued approach	`A`	Simultaneous triple-channel disappearance	`Ø`	All three channels lose contact in same time window
4	Course change / search pattern	`A`	Reappears on new bearing, all channels	`R`	Relocation to different geometric position
5	Multiple course changes	`A`	Repeated disappear/reappear cycles	`R`	Pattern repeats 2–3 times; correlated with heading changes
6	Aircraft departs area (fuel)	`D`	Final disappearance; no further contact	`Ø`	Object departs when engagement ends

H-string: {P, A, A, A, A, D} • U-string: {V, M, Ø, R, R, Ø}

Alternative encoding: Steps 4–5 can be decomposed into 3–4 sub-steps if each disappear/reappear cycle is treated independently, yielding 8–9 total steps. Sensitivity analysis tests both.

Table A4. Minot AFB encoding (7 paired steps).

#	Human Action	H	UAP Response	U	Context / Notes
1	Ground teams on patrol	`P`	Visual sighting from ground	`V`	Multiple ground witnesses; initial bright object
2	B-52 vectored toward object	`A`	Object appears on B-52 radarscope	`V`	Airborne radar confirms ground visual
3	B-52 continues approach	`A`	Object maintains 3 NM station-keeping	`S`	Precise geometric maintenance through B-52 turns
4	B-52 180° orbit turn	`A`	Object mirrors turn; holds 3 NM	`S`	Station-keeping preserved through complex maneuver
5	B-52 descends for landing approach	`D`	Object closes to 1 NM; brightens	`C`	Approach coincides with observer disengagement
6	B-52 on final approach	`D`	Object paces at low altitude	`M`	Altitude matching during descent
7	B-52 lands	`D`	Object departs vertically at high speed	`Ø`	Departure upon final disengagement

H-string: {P, A, A, A, D, D, D} ● U-string: {V, V, S, S, C, M, Ø}

Alternative encoding: Ground-observer phases can be decomposed into 2–3 additional steps if individual witness transitions are encoded separately, yielding 9–10 total steps.

Table A5. Tehran encoding (9 paired steps).

See Table 12 in Section 5.1 for the complete Tehran encoding, which serves as the primary illustrative example. The encoding produces:

H-string: {A, D, A, A, W, E, D, P, D} ● U-string: {X, Q, Q, Q, X, T, Q, X, Ø}

Alternative encoding: Step 5 (weapons engagement → comprehensive system failure) can be decomposed into 2–3 sub-steps (instruments fail, comms fail, weapons fail) if temporal ordering is assumed, yielding 10–12 total steps. Step 8 (landing approach → EM effects near ground object) may also be decomposed. Sensitivity analysis tests all plausible decompositions.

Table A6. USS Nimitz encoding (6 paired steps).

#	Human Action	H	UAP Response	U	Context / Notes
1	Multi-day SPY-1B radar tracking	`P`	Repeated descent 80K’→28K’ over days	`V`	Conspicuous pattern; active display on calibrated system
2	F/A-18s vectored to investigate	`A`	Object visible at merge; hovering over disturbance	`Q`	Visual confirmation of radar contact; white Tic Tac
3	Fravor descends in circular approach	`A`	Object mirrors; ascends to meet aircraft	`M`	Real-time mirroring; four aircrew witnesses

4	Fravor commits to aggressive nose-low	**G**	Instantaneous acceleration; disappears	**Ø**	Estimated 75–5,000 g departure
5	CAP point (classified waypoint)	**D**	Object appears at CAP point within seconds	**R**	Arrival at classified coordinate not broadcast
6	Second flight + FLIR tracking	**A**	Active radar jamming of AN/APG-73	**X**	Electronic warfare against pursuing aircraft

H-string: {P, A, A, G, D, A} ● U-string: {V, Q, M, Ø, R, X}

Alternative encoding: The multi-day radar tracking (Step 1) can be decomposed into 3–5 sub-steps (individual descent events). The second flight (Step 6) can be decomposed into approach + jamming as separate transitions, yielding 8–11 total steps.

A.3 Pooled Dataset Summary

Case	Paired Steps	H-String	U-String	Unique UAP Symbols
RB-47	6	{P,A,A,A,A,D}	{V,M,Ø,R,R,Ø}	V, M, Ø, R (4)
Minot	7	{P,A,A,A,D,D,D}	{V,V,S,S,C,M,Ø}	V, S, C, M, Ø (5)
Tehran	9	{A,D,A,A,W,E,D,P,D}	{X,Q,Q,Q,X,T,Q,X,Ø}	X, Q, T, Ø (4)
Nimitz	6	{P,A,A,G,D,A}	{V,Q,M,Ø,R,X}	V, Q, M, Ø, R, X (6)
Pooled	**28**	28 symbols; 5 unique	28 symbols; 8 unique	V,Q,M,S,C,X,T,Ø,R — N not observed (8 of 10)

Table A7. Pooled dataset summary. Conservative encoding: 28 paired steps. Alternative encodings yield 35–50+ steps depending on decomposition decisions.

Appendix B: Replication Code

This appendix provides the complete Python implementation for reproducing the information-theoretic analysis in Section 5. The code requires Python 3.9+, numpy, scipy, and the *dit* library for information-theoretic computations. All code is released under the MIT License.

Installation: `pip install numpy scipy dit`

B.1 Data Encoding

The following module defines the symbol sequences from Appendix A and provides utilities for alternative encodings.

```
"""encoding.py - UAP behavioral sequence encoding"""

# Conservative paired-step encodings (Section 5.1 / Appendix A)
CASES = {
    "rb47": {
        "H": ["P","A","A","A","A","D"],
        "U": ["V","M","O","R","R","O"],  # O = departure (\u00D8)
    },
    "minot": {
        "H": ["P","A","A","A","D","D","D"],
        "U": ["V","V","S","S","C","M","O"],
    },
    "tehran": {
        "H": ["A","D","A","A","W","E","D","P","D"],
        "U": ["X","Q","Q","Q","X","T","Q","X","O"],
    },
    "nimitz": {
        "H": ["P","A","A","G","D","A"],
        "U": ["V","Q","M","O","R","X"],
    },
}

def pooled(key="U"):
    """Concatenate all case sequences for pooled analysis."""
    return sum((c[key] for c in CASES.values()), [])

def alt_tehran_decomposed():
    """Alternative Tehran encoding with compound-symbol decomposition.
    Step 5 split: instruments -> comms -> weapons (3 sub-steps)."""
    return {
        "H": ["A","D","A","A","W","W","W","E","D","P","D"],
```

```
        "U": ["X","Q","Q","Q","Xi","Xc","Xw","T","Q","X","O"],
    }
```

B.2 Shannon Entropy

```
"""entropy.py - Shannon entropy analysis (Section 5.2)"""
import numpy as np
from collections import Counter

def shannon_entropy(seq):
    """Compute Shannon entropy H(X) in bits."""
    counts = Counter(seq)
    n = len(seq)
    probs = np.array([c/n for c in counts.values()])
    return -np.sum(probs * np.log2(probs))

def normalized_entropy(seq):
    """H(X) / H_max, where H_max = log2(alphabet_size)."""
    h = shannon_entropy(seq)
    k = len(set(seq))
    return h / np.log2(k) if k > 1 else 0.0

def entropy_with_ci(seq, n_bootstrap=10000, alpha=0.05):
    """Bootstrap confidence interval for normalized entropy."""
    rng = np.random.default_rng(42)
    arr = np.array(seq)
    boots = []
    for _ in range(n_bootstrap):
        sample = rng.choice(arr, size=len(arr), replace=True)
        boots.append(normalized_entropy(sample.tolist()))
    boots = sorted(boots)
    lo = boots[int(n_bootstrap * alpha/2)]
    hi = boots[int(n_bootstrap * (1 - alpha/2))]
    return normalized_entropy(seq), lo, hi
```

B.3 Zipf’s Law

```
"""zipf.py - Zipf rank-frequency analysis (Section 5.3)"""
from collections import Counter
import numpy as np
from scipy.optimize import minimize_scalar
from scipy.stats import kstest
```

```
def rank_frequency(seq):
    """Return (ranks, frequencies) sorted by descending count."""
    counts = Counter(seq)
    freqs = sorted(counts.values(), reverse=True)
    ranks = np.arange(1, len(freqs) + 1)
    return ranks, np.array(freqs)

def fit_zipf_mle(seq):
    """Clauset MLE power-law exponent estimate.
    Returns alpha, KS statistic, p-value."""
    ranks, freqs = rank_frequency(seq)
    # Discrete MLE for Zipf exponent
    def neg_ll(alpha):
        n = len(ranks)
        harmonic = np.sum(1.0 / ranks**alpha)
        return -np.sum(np.log(freqs / freqs.sum())) + n * np.log(harmonic)
    result = minimize_scalar(neg_ll, bounds=(0.5, 3.0), method="bounded")
    alpha = result.x
    # KS goodness-of-fit
    expected = (1.0 / ranks**alpha)
    expected = expected / expected.sum() * freqs.sum()
    ks_stat, p_val = kstest(freqs, lambda x: np.interp(x, expected, ranks/ranks.max()))
    return alpha, ks_stat, p_val
```

B.4 Compression Ratio

```
"""compression.py - Kolmogorov complexity proxy (Section 5.4)"""
import zlib

def compression_ratio(seq):
    """zlib compression ratio as Kolmogorov complexity proxy.
    Returns ratio (compressed/original); lower = more compressible."""
    text = " ".join(seq).encode("utf-8")
    compressed = zlib.compress(text, level=9)
    return len(compressed) / len(text)

def compression_with_baseline(seq, n_shuffles=10000):
    """Compare compression ratio to shuffled baselines."""
    rng = np.random.default_rng(42)
    observed = compression_ratio(seq)
    arr = list(seq)
    shuffled_ratios = []
```

```
    for _ in range(n_shuffles):
        rng.shuffle(arr)
        shuffled_ratios.append(compression_ratio(arr))
    p_val = np.mean([s <= observed for s in shuffled_ratios])
    return observed, np.mean(shuffled_ratios), p_val
```

B.5 Transfer Entropy

```
"""transfer_entropy.py - Directed information flow (Section 5.5)"""
import numpy as np
from collections import Counter

def transfer_entropy(source, target, lag=1):
    """Compute TE(source -> target) in bits.
    TE = sum p(t_{n+1}, t_n, s_n) * log2[p(t_{n+1}|t_n,s_n) / p(t_{n+1}|t_n)]"""
    n = len(target) - lag
    # Joint and marginal counts
    joint_tts = Counter()   # (target_next, target_cur, source_cur)
    joint_tt = Counter()    # (target_next, target_cur)
    marg_ts = Counter()     # (target_cur, source_cur)
    marg_t = Counter()      # (target_cur)
    for i in range(n):
        tn = target[i + lag]
        tc = target[i]
        sc = source[i]
        joint_tts[(tn, tc, sc)] += 1
        joint_tt[(tn, tc)] += 1
        marg_ts[(tc, sc)] += 1
        marg_t[tc] += 1
    te = 0.0
    for (tn, tc, sc), count in joint_tts.items():
        p_tts = count / n
        p_tn_given_ts = count / marg_ts[(tc, sc)]
        p_tn_given_t = joint_tt[(tn, tc)] / marg_t[tc]
        if p_tn_given_t > 0 and p_tn_given_ts > 0:
            te += p_tts * np.log2(p_tn_given_ts / p_tn_given_t)
    return te

def te_surrogate_test(source, target, n_surrogates=10000, lag=1):
    """Surrogate shuffle test for TE significance.
    Shuffles source while preserving target autocorrelation."""
    rng = np.random.default_rng(42)
    observed = transfer_entropy(source, target, lag)
```

```
    surrogates = []
    src = list(source)
    for _ in range(n_surrogates):
        rng.shuffle(src)
        surrogates.append(transfer_entropy(src, target, lag))
    p_val = np.mean([s >= observed for s in surrogates])
    return observed, np.mean(surrogates), np.std(surrogates), p_val
```

B.6 Sensitivity Analysis

```
"""sensitivity.py - Test all results against alternative encodings"""
from encoding import CASES, pooled, alt_tehran_decomposed
from entropy import normalized_entropy, entropy_with_ci
from zipf import fit_zipf_mle
from compression import compression_ratio
from transfer_entropy import te_surrogate_test

def run_full_battery(cases=None):
    """Run all five measures on conservative + alternative encodings.
    Returns dict of results keyed by encoding variant."""
    if cases is None:
        cases = CASES
    results = {}
    # Conservative pooled
    U = sum((c["U"] for c in cases.values()), [])
    H = sum((c["H"] for c in cases.values()), [])
    results["conservative"] = {
        "entropy": entropy_with_ci(U),
        "zipf_alpha": fit_zipf_mle(U),
        "compression": compression_ratio(U),
        "te_h_to_u": te_surrogate_test(H, U),
        "te_u_to_h": te_surrogate_test(U, H),
    }
    # Alternative: Tehran decomposed
    alt = dict(cases)
    alt["tehran"] = alt_tehran_decomposed()
    U_alt = sum((c["U"] for c in alt.values()), [])
    H_alt = sum((c["H"] for c in alt.values()), [])
    results["tehran_decomposed"] = {
        "entropy": entropy_with_ci(U_alt),
        "zipf_alpha": fit_zipf_mle(U_alt),
        "compression": compression_ratio(U_alt),
        "te_h_to_u": te_surrogate_test(H_alt, U_alt),
```

```
        "te_u_to_h": te_surrogate_test(U_alt, H_alt),
    }
    return results

if __name__ == "__main__":
    results = run_full_battery()
    for variant, measures in results.items():
        print(f"\n=== {variant} ===")
        for measure, value in measures.items():
            print(f"  {measure}: {value}")
```

Appendix C: Annotated Source Documentation

This appendix provides annotated references to the primary source documentation for each case. Sources are ordered by evidential weight within each case. Where documents are publicly available, URLs are provided. Where documents are classified or restricted, the classification level and holding repository are noted to facilitate future FOIA requests.

C.1 RB-47H (1957)

Source	Type	Notes
Condon Committee Case Study (1968)	*Official investigation*	Case 5 in Scientific Study of Unidentified Flying Objects. Analyst (Thayer) concludes probability of atmospheric explanation is negligible. Full ELINT/visual/radar narrative.
Crew debriefing transcripts	*Military records*	Classified at time of event; partially declassified via FOIA. Provide phase-by-phase crew accounts. Held at National Archives.
55th SRW mission logs	*Military records*	Document aircraft position, heading, altitude, and timing. Corroborate crew accounts.
Klass, P. (1983) UFOs: The Public Deceived	*Skeptical analysis*	Proposes signal-misidentification hypothesis for ELINT channel. Does not address behavioral sequence or triple-channel synchronization.
ELINT receiver tapes (if extant)	*Classified*	Original electronic recordings. Status unknown; may survive in compartmented archives. Highest-priority FOIA target for this case.

C.2 Minot AFB (1968)

Source	Type	Notes
Project Blue Book case file #12548	*Official investigation*	Contains radarscope photographs, multiple witness statements, weather data. Declassified. Available via Fold3/National Archives.
B-52 radarscope photographs	*Primary sensor data*	Capture radar return at position consistent with visual observations. Reproduced in Blue Book file.
Werlich, Lt. Col. investigation report	*Official military*	Base-level investigation by Minot deputy commander. Contains timeline, witness interviews, radar data summary.
Ground witness statements (16 personnel)	*Sworn testimony*	Multiple independent witnesses from dispersed locations on base. Consistent descriptions of object appearance and movement.
Tulien, T. (2005) Minot AFB investigation	*Independent research*	Comprehensive reconstruction including interviews with surviving witnesses. Most complete public documentation.

C.3 Tehran (1976)

Source	Type	Notes
DIA report (declassified 1981)	*Official intelligence*	Defense Intelligence Agency assessment. Describes encounter as meeting all criteria for a valid study of the UFO phenomenon. Provides detailed timeline.
Jafari, Gen. P. (multiple interviews)	*Primary witness*	Lead pilot on first intercept. Has given consistent account across decades including National Press Club (2007) and multiple documentaries.
Pirouzi, H. (tower supervisor)	*Primary witness*	Supervised radar tracking from Mehrabad tower. Provides ground-based corroboration of airborne accounts.
Iranian Air Force after-action report	*Official military*	Original Farsi-language documentation. Partially available through intelligence channels; full document may remain in Iranian archives.
Klass, P. (1983) analysis	*Skeptical analysis*	Proposes Jupiter + equipment malfunction. Rigorous for initial visual ID only; does not address radar returns, graduated system failures, or weapons-engagement correlation.

C.4 USS Nimitz (2004)

Source	Type	Notes
SCU Nimitz report (2019, rev. 2020)	*Independent technical*	Scientific Coalition for UAP Studies. Detailed kinematic analysis of FLIR1 video. Acceleration estimates. Peer-reviewed within SCU.
Fravor, Cdr. D. (multiple testimonies)	*Primary witness (sworn)*	Lead pilot. Congressional testimony under oath (2023). Consistent account across all appearances. Describes mirroring maneuver and instantaneous departure.
Dietrich, Lt. Cdr. A. (multiple interviews)	*Primary witness*	WSO in wingman aircraft. Corroborates visual observation from independent vantage point.
USS Princeton CIC logs	*Military records*	Combat Information Center radar tracking data. Multi-day SPY-1B contacts. Classification status unclear; partially referenced in public reporting.
FLIR1 video (76-sec public release)	*Primary sensor data*	Authenticated by DOD (2020). Approximately 8–10 minutes full duration; remainder classified.
Day, K. (radar operator interviews)	*Primary witness*	Princeton senior radar operator. Describes multi-day tracking pattern and 80K’→28K’ descent profile.
West, M. (2019–2021) analyses	*Skeptical analysis*	Technically rigorous for 76-second video. Does not address SPY-1B multi-day tracks, visual close-range observation, mirroring, CAP-point reappearance, or

		active radar jamming.
ONI classified report	*Classified*	Office of Naval Intelligence assessment. Existence confirmed; contents not publicly available. Priority FOIA target.

This source catalog is designed to support independent verification of the behavioral narratives in Section 4 and the encoding decisions in Appendix A. Researchers are encouraged to consult primary sources directly rather than relying solely on this paper's interpretations. Where this paper's encoding of a behavioral phase differs from what a researcher finds in the primary source, the primary source should be preferred and the discrepancy reported.