

Cargo and Revelation: John Frum and the Anatomy of a Technology Religion

Justin H. Kuiper, CISSP

Technology Worship · RP-2 · v0.1 preprint · Curiosity Killing the Yeti · 2026

Cargo and Revelation: John Frum and the Anatomy of a Technology Religion

1. The Airstrip in the Jungle

On the island of Tanna in what is now Vanuatu, men drill with wooden rifles under the tropical sun. They march in formation wearing the letters “USA” painted across their bare chests. They have built bamboo control towers beside cleared airstrips, complete with coconut-shell “headphones” through which they call to the sky. At the center of their ritual geography stands a bamboo flagpole flying a red cross — the symbol of John Frum, the prophet who promised that the cargo would come again.

This is not theater. This is not a tourist attraction, though it has become one. This is the material form of belief: the conviction that faithful reproduction of the *forms* associated with overwhelming technological power will summon that power back. The American military brought an avalanche of goods to the Pacific during World War II — prefabricated buildings, vehicles, radios, medicine, food in quantities beyond local comprehension. Then they left. The cargo cults that emerged in their wake represent humanity’s most visible attempt to reverse-engineer transcendent technological power through ritual imitation.

The image is so striking, so seemingly primitive, that it has become a cliché in technological discourse. We invoke “cargo cult” as shorthand for misguided imitation, for form without substance, for people who don’t understand what they’re doing. Richard Feynman imported the metaphor into science in 1974; we now speak casually of “cargo cult programming” and “cargo cult management.” But the ease of the metaphor should make us suspicious. What if the cargo cultists of Melanesia were not *primitive* but *paradigmatic* — not an outlier in human technological engagement but a *model* of what we all do when confronted with power we cannot explain?

This paper argues that the cargo cults — particularly the John Frum movement — constitute the clearest anthropological documentation of *technology religion*: the transformation of technological

awe into eschatological ritual. By examining the historical record of what actually happened on Tanna, extracting the structural anatomy of cargo cult belief, and tracing how that structure recurs in modern technological culture, we can name the pattern that governs how societies relate to technologies they deploy but do not understand. The cargo cultists were not wrong to recognize something transcendent in what they witnessed. They were trying to solve the hardest problem any culture faces: how to maintain agency in the presence of overwhelming, incomprehensible power.

2. What Actually Happened on Tanna

The John Frum movement did not spring fully-formed from wartime contact. Its roots predate the American military presence, emerging in the late 1930s as a millenarian response to decades of colonial and missionary pressure on traditional Tanna society. Jean Guiart's 1952 documentation in *Oceania* establishes the movement's pre-war origins: "John Frum" appeared first as a spirit voice promising the restoration of custom (*kastom*) and the expulsion of European authority (Guiart 1952, 166). The name itself may derive from "John from America" or may be a folk etymology for a phrase in a local language; certainty is impossible, but the ambiguity is itself revealing — the movement's origin was already syncretic, already a negotiation between indigenous apocalyptic expectation and the overwhelming presence of white technological power.

World War II transformed this indigenous millenarianism into something unprecedented. Between 1942 and 1945, the American military turned Melanesia into a vast logistics operation. Peter Worsley documents the sheer scale of the transformation: "The Americans landed in the New Hebrides in 1942 with wealth beyond the dreams of the natives... Refrigerators, radios, radar equipment, huge transport planes and 'liberty' ships by the score... It was a revelation" (Worsley 1957, 153). For peoples whose primary contact with industrial civilization had been through colonial plantation labor and missionary Christianity — both systems of deprivation and control — the Americans represented something different: white men who were generous, who paid well, who treated natives with a degree of respect unknown under European colonialism, and who commanded technology that seemed to violate every rule of production the islanders knew.

The cargo arrived by air and sea without apparent *creation*. The Americans did not plant gardens or raise pigs; they performed rituals (filling out forms, speaking into radios, guiding planes with lights) and goods *appeared*. Worsley quotes a New Guinea informant: "The Europeans said they 'worked' for their goods, but we knew this was a lie. We saw that they got them from some other source, which they kept secret" (Worsley 1957, 117). This was not stupidity but rational inference from available evidence: the visible labor of the Americans bore no recognizable relationship to the goods they received.

Then, in 1945, the Americans left. The bases closed, the cargo stopped, and the colonial order reasserted itself. On Tanna, the John Frum movement crystallized into its enduring form: a prophet-mediated cult promising that fidelity to certain rituals and the maintenance of *kastom*

would bring John Frum (now explicitly identified with American power) back with the cargo. Guiart documents the movement's structure in the late 1940s: designated villages as ritual centers, Friday as the sacred day (rejecting the missionary Sunday), the red cross as the central symbol, and periodic prophetic updates about the cargo's imminent arrival (Guiart 1952, 170-173).

Peter Lawrence's documentation of similar movements in the Madang District of New Guinea reveals the same pattern: "The people built dummy airstrips, warehouses, and radio masts in the hope that these would attract cargo planes" (Lawrence 1964, 203). They lit fires to guide planes at night. They drilled with wooden rifles, mimicking the military formations they had witnessed. They were not being irrational; they were performing *empirical* imitation of the forms associated with cargo arrival. If ritual fidelity to observable form had worked for generations to secure yam harvests and maintain cosmic order, why wouldn't it work for summoning cargo?

The historical record is crucial because it establishes what cargo cults actually *are*, not what they have become in metaphorical usage. They are: (1) responses to contact with technological power so overwhelming it cannot be integrated into existing explanatory frames; (2) attempts to reproduce the *observable correlates* of that power through ritual; (3) eschatological movements promising that faithful reproduction will restore or surpass the original encounter; (4) syntheses of indigenous religious structures with elements of the foreign power system (Christianity, military hierarchy, American nationalism). They are not stupidity. They are what happens when a society tries to maintain symbolic agency in the face of causal opacity.

3. The Anatomy of a Cargo Cult

The anthropological literature, particularly Worsley's structural analysis, allows us to extract a consistent anatomy shared across cargo cult movements. Five features recur with such regularity that they constitute the *invariant structure* of technology religion:

First: Arrival of incomprehensible power from outside the explanatory frame. The defining precondition is technological encounter that exceeds the culture's causal models. Lindstrom notes: "Cargo — Western goods and technology — was conceived as having a supernatural or mysterious origin" (Lindstrom 1993, 12). The key word is *conceived*; the conception was forced by epistemic gap, not by choice. When the Americans flew thousands of tons of materiel into jungle bases using aircraft larger than anything the islanders had imagined, powered by fuel systems they could not see, manufactured through processes they had never witnessed, the *reasonable* inference was supernatural origin. Arthur C. Clarke's third law — "Any sufficiently advanced technology is indistinguishable from magic" — is not a quip but an epistemological fact. The cargo cultists encountered technology sufficiently advanced that the boundary between natural and supernatural causation became meaningless.

Second: A prophet or mediator who interprets the power and promises its return.

No cargo cult forms without a prophet figure who can narrativize the encounter and prescribe the response. John Frum on Tanna, Yali in the Madang region, numerous other named and unnamed figures across Melanesia — all serve the same function: they are the hermeneutic bridge between the incomprehensible encounter and actionable community response. Worsley emphasizes their role: “The prophets are... intermediaries between the people and the powers that control the cargo” (Worsley 1957, 241). They do not merely report revelations; they provide *theodicy* — explanations for why the cargo came, why it left, what conditions govern its return. John Frum promised that the Americans would return if the people rejected European ways and maintained *kastom*. The prophet translates technological awe into moral instruction.

Third: Ritual imitation of form rather than causal mechanism. This is the feature that makes cargo cults visually striking and metaphorically portable. Lawrence documents the material culture of cargo ritual in obsessive detail: “The people built repl

icas of airstrips, complete with bamboo control towers... Wooden headphones were carved so that ‘operators’ could communicate with the spirits who controlled the cargo planes” (Lawrence 1964, 205). The logic is sympathetic magic at industrial scale: if control towers and headphones were *present* when cargo arrived, then reproducing control towers and headphones should *summon* cargo.

But sympathetic magic is not the right frame. What the cargo cultists were doing was *empirical imitation* — reproducing the observable correlates of a desired outcome when the causal mechanism is opaque. Worsley points out that this is structurally identical to how traditional ritual works: “In the old days, if you wanted yams, you performed the yam rituals. If you wanted pigs to multiply, you performed pig magic. Now, if you wanted cargo, you performed cargo ritual” (Worsley 1957, 119). The form had always been: identify the ritual associated with abundance, perform it with fidelity, receive abundance. The content changed (airstrips instead of yam gardens), but the *structure* was continuous with centuries of religious practice.

The tragedy — and this is what makes cargo cults poignant rather than merely absurd — is that the imitation was *almost* right. Control towers and radio communication *are* causally connected to cargo arrival. Drilling and discipline *are* part of how military logistics work. The cargo cultists identified real correlations. What they could not access was the mechanism *inside* the radio, the fuel chemistry *inside* the plane, the manufacturing supply chain *behind* the goods. They reproduced the surface with extraordinary fidelity. They could not reproduce the depth.

Fourth: Deferred eschatology — the cargo will come if fidelity holds. Every cargo cult movement promises *imminent* return of the cargo, and every movement survives the failure of that promise through eschatological deferral. Guiart documents multiple prophecies on Tanna of John Frum’s return “next Friday” or “next month” or “when the people are pure enough” (Guiart 1952, 174). When the prophecy fails, the movement does not collapse; instead, the prophet or his successors reinterpret the condition. The people were not pure enough. The ritual was not performed with

sufficient fidelity. The Europeans' continued presence pollutes the spiritual channel. The structure is identical to every deferred eschatology in religious history — Christianity's perpetually-delayed Second Coming, Marxism's perpetually-deferred communist utopia. The prophet's authority survives prophetic failure because the failure can always be moralized: *we* failed the prophecy, not the prophecy us.

Lindstrom notes that this is not cynicism but *necessary structure*: “Deferral allows the movement to persist through disconfirmation... The promise remains forever on the horizon, pulling the community forward” (Lindstrom 1993, 48). Without deferral, a cargo cult would be falsified the first time the cargo failed to arrive and the movement would dissolve. With deferral, each failure becomes an opportunity to *purify* practice, to intensify fidelity, to wait with greater discipline. The movement becomes unfalsifiable — which is to say, it becomes a religion.

Fifth: In-group identity organized around the awaited power. Cargo cults reorganize social belonging around the technological encounter and its expected return. Worsley emphasizes that the movements were not merely religious but *political*: “They represented rejection of European authority and assertion of native autonomy” (Worsley 1957, 255). To participate in the John Frum movement was to reject the mission church, to refuse plantation labor, to assert that *kastom* (indigenous custom) had equal or greater power than European ways. The movement created sharp in-group/out-group boundaries: those who believed versus those who collaborated with the colonial order, those who waited for John Frum versus those who accepted European explanations.

This is the feature that makes cargo cults socially durable. They are not just epistemological responses to technological encounter but *identity formations*. Lawrence documents how cargo cult villages became parallel societies with their own governance, their own ritual calendars, their own moral economies (Lawrence 1964, 218-220). To abandon the cult was not merely to change one's beliefs but to betray one's community. The technological eschatology became the ground of belonging.

These five features — incomprehensible external power, prophetic mediation, ritual imitation of form, deferred eschatology, identity reorganization — constitute the *invariant anatomy* of a technology religion. They are not unique to Melanesia. They are the pattern that emerges when technological power arrives faster than a culture's ability to integrate it causally, and when that gap is filled by ritual rather than explanation.

4. Cargo Cult Science

In 1974, Richard Feynman delivered a commencement address at Caltech that permanently transplanted the cargo cult metaphor into the epistemology of science. Feynman had never been to Melanesia; his knowledge of cargo cults came from popular accounts and anthropological summaries. But he saw in their ritual logic a precise analogy to a failure mode he had observed throughout his

career: research that reproduces the *form* of scientific method without its *substance*.

The key passage deserves full quotation:

“In the South Seas there is a cargo cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they’ve arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas — he’s the controller — and they wait for the airplanes to land. They’re doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn’t work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they’re missing something essential, because the planes don’t land” (Feynman 1974/1985, 340-341).

Feynman’s target was research that mimics scientific *procedure* — controlled experiments, statistical analysis, peer review, publication — without the underlying *discipline* that makes those procedures productive. His examples included rat-running experiments that failed to control for prior learning, educational psychology studies that reported results without checking whether students actually *learned* anything, and medical research that cherry-picked positive outcomes while ignoring negative trials. In each case, the researchers were “doing everything right” in the procedural sense: running experiments, analyzing data, publishing results. But they were missing “something essential” — what Feynman called “a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty” (Feynman 1974/1985, 341).

That “something essential” is *mechanism-seeking*. Real science does not merely correlate inputs and outputs; it proposes causal mechanisms and tests them through adversarial design. Cargo cult science skips mechanism — it looks for “what works” without asking *why* it works, and therefore cannot distinguish true causal relationship from spurious correlation or experimental artifact. The cargo cultists reproduced the form of the airstrip without understanding aerodynamics, fuel, navigation, or supply chains. Cargo cult scientists reproduce the form of experiment without understanding how experimental design isolates causal variables.

But Feynman’s metaphor does something more subtle than just criticize bad methodology. It implicitly argues that *imitation of form is a natural response to epistemic opacity*. The cargo cultists were not stupid; they were doing the best they could with available information. They observed correlations (airstrips correlate with cargo arrival), and they reproduced the observable elements of those correlations with impressive fidelity. What they lacked was access to the *invisible mechanism* — the engineering, logistics, and industrial capacity that made airstrips functional rather than decorative.

Cargo cult science is the same structure in a different domain. Researchers who lack deep under-

standing of causal mechanism fall back on *procedural fidelity*: if I follow the experimental protocol, if I use the right statistical tests, if I publish in the right journals, the results will be valid. The form is reproduced faithfully. But without mechanism-seeking discipline — without asking “why would this intervention *cause* this outcome?” and designing tests to falsify competing causal stories — the form is empty. The planes don’t land.

Feynman’s migration of the metaphor was powerful because it named a failure mode that was (and remains) endemic in research that wants the *status* of science without the *discipline* of science. But it also risked trivializing the cargo cults themselves. The Melanesian movements were not merely methodological errors; they were existential responses to cultural crisis. Feynman’s usage abstracts the pattern from its context — which makes it portable, but also makes it easier to miss what the pattern actually *is*: the structure of technological

religion when understanding lags behind encounter.

5. Our Cargo Cults

We are building bamboo control towers.

The contemporary technological landscape is thick with cargo cult structure — not because we are primitive or irrational, but because we are deploying systems whose causal mechanisms we do not and often *cannot* understand, and we are filling the epistemic gap with ritual. The five features of cargo cult anatomy map onto modern technological practice with uncomfortable precision.

First: Incomprehensible power arriving from outside the explanatory frame. Large language models — the technology this paper is written *with* — are the paradigmatic case. A model trained on hundreds of billions of tokens, using transformer architectures with billions of parameters, optimized through gradient descent at scales that require warehouse-sized compute clusters, produces outputs that are fluent, contextually appropriate, and sometimes genuinely useful. But *why* a particular prompt produces a particular completion is not mechanistically legible to the people deploying the model, and often not to the people who trained it. We cannot point to specific parameters and say “this is where the model learned subject-verb agreement” or “this is the circuit that handles causal reasoning.” The model is a black box whose interior is not merely complex but *epistemically opaque* — we can observe inputs and outputs, but the transformation between them is not tractable to human understanding.

This is not a defect of current models that better interpretability research will fix. It is a structural property of systems whose behavior emerges from optimization over parameter spaces too large for exhaustive human audit. Anthropic’s recent mechanistic interpretability work has identified some interpretable features in smaller models (Anthropic 2023), but the gap between “we found a circuit that detects indirect object syntax” and “we understand why this model produces this essay” is the gap between isolating one gear in a watch and understanding why the watch keeps time. The watch

still works; we deploy it; we trust it. But we do not understand it in the way we understand, say, a combustion engine — where every part has a known function and the causal chain from fuel to motion is mechanistically traceable.

When organizations deploy these systems at scale — for customer service, for content generation, for code assistance, for decision support — they are deploying *incomprehensible technological power*. It works. We do not know why it works in the mechanistic sense. We proceed anyway, because the alternative is competitive disadvantage.

Second: Prophets who interpret the power and promise its return. The prophet role in contemporary AI discourse is filled by a diffuse class: research lab leaders, public intellectuals, venture capitalists, and popular science communicators who mediate between the technical reality of AI systems and the public (and institutional) understanding of those systems. They perform the same function John Frum performed on Tanna: they translate incomprehensible power into narrative, promise, and moral imperative.

Sam Altman’s repeated framing of AGI as “probably closer than most people think” functions as prophetic promise — it is eschatological, it is deferred, and it reorganizes institutional behavior in the present around an expected future transformation (Altman 2023). Effective Altruism’s “AI safety” discourse similarly prophecies: not cargo arrival but *cargo catastrophe* if ritual is not performed correctly (alignment, governance, capability control). The prophecy structure is identical whether the promised event is utopian or apocalyptic — both demand present reorganization around future technological transformation whose mechanism we do not control and whose timeline we cannot specify.

This is not to say these figures are charlatans (some may be; most are earnest). It is to say they occupy the *structural role* of prophet in a cargo cult: they stand between the incomprehensible system and the community that must decide how to organize itself in relation to it. They translate. They promise. They warn. They provide narrative where mechanism is unavailable.

Third: Ritual imitation of form without causal mechanism. This is where contemporary practice becomes most legible as cargo cult structure. Organizations that do not understand how AI systems work nevertheless adopt the *forms* associated with responsible AI deployment: ethics boards, bias audits, red-teaming exercises, alignment roadmaps, responsible AI principles. These are not inherently empty — some instances involve genuine mechanistic engagement. But in many cases they reproduce the *appearance* of rigor without the *substance* of causal control.

Consider the standard “bias audit” process: an organization tests its model on a demographic-fairness benchmark, identifies disparate outcomes across protected groups, and reports the disparity as a metric. What the organization rarely does — because it is *hard*, because it requires mechanistic understanding the organization often lacks — is identify *why* the disparity exists at the parameter level and intervene on the causal mechanism. Instead, the ritual is: measure, report, acknowledge.

The form is “we are addressing bias.” The substance is “we documented a correlation we cannot mechanistically explain or fix.” The dashboard updates. The planes do not land.

Or consider “alignment theater”: the ritual of declaring a model “aligned with human values” based on RLHF (reinforcement learning from human feedback) training that optimizes the model to produce outputs human raters prefer. The form is “we aligned the model to human preferences.” The substance is “we shifted the model’s output distribution toward patterns human raters scored highly *in the training context*, and we have no mechanistic guarantee this generalizes, no causal story for why preference-matching in training implies safety in deployment, and no control mechanism if the model encounters a context where human preference and actual safety diverge.” We built the airstrip. We are confident in our airstrip-building methodology. We do not know why planes land at airstrips.

The ritual is not *useless* — RLHF demonstrably changes model behavior in ways users often prefer. But the gap between “this intervention produces outputs we like better” and “this intervention *causes* the model to be safe/aligned/trustworthy” is the gap between correlation and mechanism. We are optimizing for form (preferred outputs) without causal access to substance (what internal model structure produces those outputs and whether that structure is robust to distributional shift).

Fourth: Deferred eschatology. The perpetually-deferred arrival is AGI itself — the moment when AI systems reach human-level general intelligence, or superintelligence, or “transformative AI,” depending on which prophet you follow. The prophecy timeline is perpetually six months to ten years away, always close enough to reorganize present behavior, always far enough to survive near-term non-arrival.

Yudkowsky’s timeline has shortened and lengthened cyclically over two decades (Yudkowsky 2001–2023). OpenAI’s charter promises AGI “within our lifetime” without specifying a decade (OpenAI 2018). DeepMind’s leadership oscillates between “sooner than you think” and “still a long way off” depending on audience and funding cycle (Hassabis, various). The eschatology remains structurally deferred — close enough to matter, far enough to excuse present non-arrival.

And like classic cargo cult eschatology, each deferral is moralized: we are not ready, we have not aligned properly, we have not solved interpretability, we have not established sufficient governance. The delay is not evidence against the prophecy but evidence of *our failure to prepare adequately*. The cargo will come when we are worthy — when our rituals are pure enough, our alignment fidelity sufficient, our runway well-maintained.

Fifth: In-group identity organized around the awaited power. The AI discourse has produced sharp identity boundaries: “accelerationists” versus “safetyists,” “doomers” versus “optimists,” “alignment researchers” versus “capabilities researchers.” These are not merely technical disagreements; they are *moral communities* organized around different stances toward the expected transformation. To identify as an AI safety researcher is not merely to work on a technical problem

but to belong to a community that shares an eschatology (AI catastrophe is the default outcome without intervention), a moral imperative (working on safety is altruistic; working on capabilities is reckless), and a social structure (EA movement, rationalist community, x-risk discourse networks).

Similarly, corporate adoption of AI has become an identity marker: organizations signal their modernity, their innovation-readiness, their competitive fitness by deploying AI systems — often before they understand what those systems do or whether deployment serves any legible business need. The form matters more than the function. To *not* deploy AI is to signal backwardness, to risk out-group status in an innovation-driven economy. The technology becomes the ground of belonging.

None of this is mockery. The claim is not that AI researchers are stupid or that organizations are irrational. The claim is *structural*: when technological power arrives faster than our ability to integrate it causally, we respond with the tools we have — ritual, narrative, identity, eschatology. We reproduce the forms associated with control (ethics boards, alignment protocols, governance frameworks) because we cannot access the mechanisms. We trust the prophets who mediate between the system and ourselves because we cannot read the system directly. We defer the promised transformation because we cannot specify its conditions mechanistically. We organize our communities and institutions around the

awaited power because the power itself has become the axis of meaning in our technological moment.

This is not a failure of intelligence. It is a predictable response to an epistemic condition: **we have built systems whose behavior we cannot causally explain, and we have deployed them anyway, and we are now reasoning about them using the cognitive tools available when causal access is unavailable.** Those tools are the same ones the Tanna islanders used: pattern-matching, ritual fidelity, prophetic interpretation, eschatological hope, and in-group identity. The structure is identical. The airstrips are more sophisticated. The cargo still does not come on schedule.

6. The Discipline of Substance

How do we tell genuine understanding from cargo-cult form-mimicry? Feynman gave us the first-pass answer in 1974: **real science has a mechanism for self-correction that cargo cult science lacks.** Real science specifies conditions under which it could be proven wrong; cargo cult science reproduces the appearance of method (equations, citations, controlled experiments) without the disciplined humility that allows evidence to overturn cherished conclusions. Real science treats prediction-failure as signal; cargo cult science treats it as noise, as temporary setback, as evidence we need *more fidelity to the ritual* rather than evidence the ritual itself might be wrong (Feynman 1974).

Applied to AI: genuine understanding of a system means we can specify *what would falsify our*

confidence in it. If we claim a model is “aligned,” what observation would prove us wrong? If we claim a safety protocol is sufficient, what outcome would demonstrate insufficiency? If we claim we understand why a model produces a particular output, can we intervene on the mechanism and predict the downstream effect? If we cannot answer these questions, we are in cargo-cult territory — we have reproduced the form of confidence without the substance of causal access.

This demands institutional discipline. It demands **rejecting theater**. A bias audit that documents disparate outcomes without identifying causal mechanisms is theater — it performs the ritual of accountability without delivering the substance. A red-teaming exercise that tests a model against a static adversarial dataset without modeling distributional shift is theater — it performs the ritual of robustness-checking without delivering robustness. An alignment protocol that optimizes for human preference-matching in training without specifying how that generalizes out-of-distribution is theater — it performs the ritual of safety without delivering safety.

The alternative is **mechanistic humility**: we specify what we do not know, we articulate the limits of our causal models, we treat deployment as an experiment whose failure modes we are prepared to learn from rather than a ritual whose fidelity guarantees success. This does not mean we stop deploying AI systems — the competitive and operational pressures are real, and in many contexts deployment precedes full understanding by necessity. But it means we **stop pretending deployment is the same as understanding**, and we build institutions that can learn from the gap between the two.

It also demands **rejecting prophetic authority** — not the expertise of researchers and builders, but the claim that any individual or institution has privileged access to the eschatological timeline, the conditions of AGI arrival, the moral structure of the transformation to come. The prophets on Tanna promised the cargo would come if ritual fidelity held. The cargo did not come. The modern AI prophet promises AGI will arrive (or catastrophe will arrive, or utopia will arrive) if we perform the right rituals — alignment, governance, capability control, accelerationism, deceleration. We do not know the conditions. We cannot specify the timeline. We should stop organizing as if we do.

Finally, it demands **reclaiming “I don’t know” as a legitimate institutional posture**. Organizations deploying AI systems they do not mechanistically understand should say so — not as disclaimer buried in footnotes but as operational fact that shapes decision-making. Researchers working on problems (interpretability, alignment, robustness) where no causal solution yet exists should say so, rather than offering ritual assurances that form-matching (RLHF, adversarial training, transparency dashboards) delivers substance. Policymakers regulating systems they do not understand should say so, and should build regulatory structures that assume ignorance rather than pretend to omniscience.

The airstrip problem on Tanna was not that the islanders were stupid. The problem was that they **mistook form for substance** — they believed fidelity to the observable pattern (clearing the

strip, building the tower, performing the drills) would summon the underlying power (American logistics, industrial production, geopolitical will). It did not. The planes did not come. The cargo did not arrive.

Our airstrips are more sophisticated. We build them in Python and PyTorch. We staff them with PhDs. We measure their performance on benchmarks. We audit their bias. We red-team their vulnerabilities. We declare them aligned. We trust them with decisions that matter — hiring, lending, content moderation, code generation, medical diagnosis. And when they fail, when they produce outcomes we do not understand and cannot predict, we do not ask whether the airstrip model was wrong. We ask whether we need *more* airstrips, *better* drills, *higher fidelity* to the ritual.

The discipline of substance is the discipline of **asking the harder question**: Do we actually understand the causal mechanism? Can we specify the conditions under which our confidence would be proven wrong? Are we building systems or building rituals? Are we flying planes or painting “USA” on our chests and hoping?

The technology is real. The power is real. The planes exist. But they do not land because we built a runway out of bamboo. They land because we understand metallurgy, aerodynamics, fuel chemistry, navigation, logistics, and geopolitical economy. **Substance precedes form.** Understanding precedes deployment. And when we deploy before we understand — as we often must, as the Tanna islanders often had no choice but to do in the face of overwhelming external power — we owe ourselves the honesty to admit it.

We worship what we build when we do not understand it. The way out is not to stop building. The way out is to **stop worshipping** — to trade eschatology for engineering, prophecy for mechanism, ritual for experiment, and the identity of the believer for the discipline of the skeptic. The cargo cult ends when we stop waiting for the planes and start learning to build them.

6. Beyond Tanna: The Wider Melanesian Pattern

The John Frum movement on Tanna was not an isolated aberration. It was one instance of a recurrent structure — a pattern that appeared across Melanesia wherever incomprehensible technological power arrived without its causal substrate. The airstrip ritual, the prophet-mediator, the deferred arrival, the community organized around waiting: these features emerged independently in societies separated by hundreds of miles of ocean, speaking unrelated languages, with no direct communication. The structure itself is the signal. When opacity meets power, humans everywhere respond with the same cognitive tools.

The earliest documented case is the **Vailala Madness** of 1919, in the Papuan Gulf region of what is now Papua New Guinea. F. E. Williams, the Territory of Papua’s government anthropologist,

recorded what he initially termed collective hysteria: villagers reported visions of the dead returning aboard steamships loaded with European cargo — tinned food, tools, cloth, rifles. They abandoned traditional ceremonies, built wharves and warehouses for the arriving ships, and appointed “prophets” who claimed direct communication with the dead ancestors who would pilot the vessels. The movement featured glossolalia (speaking in tongues), trance states, and the systematic destruction of sacred objects from the pre-contact ritual system (Williams 1923).

Williams interpreted it as pathology — mass delusion triggered by cultural stress. He missed the structure. The Vailala participants had observed European steamships arriving with cargo. They had observed Europeans performing incomprehensible rituals (church services, administrative paperwork, flag ceremonies) that appeared to precede cargo arrival. They inferred a causal link and reverse-engineered it. The dead ancestors, in their cosmology, were the source of all wealth and power; Europeans clearly had access to that ancestral power (how else to explain their material abundance?); therefore Europeans must have discovered the ritual key that the living had lost. The “madness” was rational inference under causal opacity. The ships never came. The movement persisted for years, intensifying ritual fidelity when prophecy failed (Williams 1923, 1934).

A generation later, in the Madang region of northeast New Guinea, the **Yali movement** emerged with near-identical features. Yali Singina, a decorated World War II veteran who had served with Australian forces, returned home convinced he had glimpsed the source of European power. He had visited Australian cities, seen factories, observed cargo being produced rather than merely arriving. But he could not access the causal mechanism — industrial production, commodity exchange, wage labor — because those mechanisms were embedded in institutional and economic structures he had no framework to parse. What he *could* observe was the ritual: Europeans read books, wrote letters, attended meetings, performed ceremonies. He concluded the cargo came from ancestors who favored those who performed the correct rituals, and Europeans had simply discovered the right ones (Lawrence 1964).

Yali established a movement that replicated European administrative forms: offices, filing systems, written records, hierarchical titles. Followers believed that meticulous fidelity to paperwork rituals would unlock ancestral cargo. They built “cargo sheds” and waited. The anthropologist Peter Lawrence, who documented the movement exhaustively, noted the same pattern Williams had seen: **when prophecy failed, ritual intensified rather than dissolved**. Yali’s followers did not conclude the model was wrong; they concluded they had not yet performed the ritual with sufficient precision. More meetings. More paperwork. More fidelity to form. The cargo did not come (Lawrence 1964).

The pattern recurs even within Tanna itself, in a movement that runs parallel to John Frum: the **Prince Philip Movement**. Starting in the 1960s, villagers in Yaohnanen identified Prince Philip, Duke of Edinburgh, as a divine figure — the pale-skinned son of the mountain spirit who had traveled across the sea, married a powerful queen, and would one day return to Tanna bearing cargo.

The movement organized around maintaining ritual connection to Philip: they kept a photograph of him in a place of honor, they performed ceremonies marking his birthday, they sent him gifts (a traditional pig-killing club, which he reciprocated with a photograph of himself holding it). When Philip died in 2021, the movement transferred allegiance to his son, Charles, now King Charles III (Bonnemaïson 1994; Tabani 2008).

The Philip movement shares the John Frum structure exactly: a prophet-figure mediating between incomprehensible external power and the community; rituals designed to maintain relationship with that power; deferred eschatology (Philip will return, or his son will, or his grandson; the timeline is indefinite but the promise is certain); and community identity organized around the waiting. The fact that Prince Philip was a real person, unlike John Frum, does not alter the structure. The movement is not *about* Philip as an individual; it is about Philip as a symbol of external technological power that might be persuaded to flow toward Tanna if the ritual relationship is maintained correctly. The planes exist. The ships exist. The wealth exists. The question is how to make them land *here* (Tabani 2008).

What these cases demonstrate is that **cargo cultism is not a cultural pathology specific to Melanesia**. It is a rational response to a particular epistemic condition: the arrival of overwhelmingly powerful technology without its causal explanation. The same condition obtained in the Papuan Gulf in 1919, in Madang in 1945, on Tanna in 1940 and again in 1960. The societies were different. The languages were different. The prior cosmologies were different. The response was identical because the underlying problem was identical: **how do you make sense of power you cannot explain?**

You reverse-engineer it. You observe the visible forms associated with the power (ships, planes, administrative rituals, literacy, flag ceremonies). You infer that performing those forms will summon the power. You appoint prophets who claim privileged interpretive access. You defer the arrival because causal opacity means you cannot specify the conditions mechanistically — you can only trust that sufficient fidelity will eventually work. And when it does not work, you do not abandon the model; you intensify the ritual, because **the alternative is admitting you have no causal access at all**, and humans rarely choose that alternative when the stakes are high (Festinger, Riecken & Schachter 1956).

This is the pattern we should be watching for in ourselves. Not the specific forms (we are not building airstrips out of bamboo), but the *structure*: incomprehensible technological power, reverse-engineered into ritual, organized around prophets, sustained by deferred eschatology, binding communities through shared identity as the faithful. Wherever you see that structure, you are looking at cargo cultism — regardless of how sophisticated the airstrip appears, regardless of how credentialed the prophets are, regardless of how many peer-reviewed papers support the ritual.

The Melanesian cases are valuable not because they are exotic, but because **the structure is easier**

to see when the airstrips are made of bamboo. When the airstrips are made of GPUs and the rituals are written in PyTorch, we have a harder time recognizing the pattern. But the pattern is the same. The discipline is the same: distinguish form from substance, ritual from mechanism, prophecy from engineering. The Vailala Madness ended when younger generations gained causal access to industrial production through education and wage labor. The Yali movement faded when participants could explain European wealth without invoking ancestral favor. The John Frum movement persists because causal opacity persists — Tanna remains outside the industrial economy, and the planes still do not land for *them*, even though the planes clearly exist.

Our cargo cults will persist as long as we deploy systems we do not understand. The way out is the same way it has always been: **build causal access**. Stop reproducing forms. Start building mechanisms. And when you cannot build mechanisms yet, admit it — do not perform the ritual and call it science.

7. The Epistemics of Opacity

Why does incomprehensible power reliably produce ritual rather than inquiry? Why do humans confronted with causal opacity so often default to pattern-matching, prophetic interpretation, and ritual fidelity instead of systematic investigation of underlying mechanisms? The answer is not irrationality. The answer is that **under conditions of extreme causal opacity, ritual is often the correct inference**.

Consider the epistemic position of a Tanna islander in 1940. American forces arrive. They possess material wealth orders of magnitude beyond anything in local experience. They arrive in machines (planes, ships, trucks) whose operation is completely opaque — no visible power source, no legible mechanism, behaviors that violate every prior model of how objects move through the world. They perform elaborate rituals: marching in formation, raising and lowering flags at fixed times, saluting, sitting at desks shuffling paper, speaking into radios. Cargo follows.

From the islander’s perspective, what is the rational inference? **The rituals cause the cargo.** There is a perfect correlation: Americans perform rituals, cargo appears. Americans stop performing rituals (they leave), cargo

stops appearing. The correlation is consistent, repeated, predictive. The only missing piece is the mechanism linking ritual to cargo — but that is *exactly what causal opacity means*. You cannot observe the mechanism. You can only observe the correlation.

Now consider the alternative: attempting to investigate the mechanism. Where would you start? You cannot open the plane’s engine and understand it — the components are unfamiliar, the materials unknown, the principles of combustion and aerodynamics inaccessible without centuries of prior physics and engineering. You cannot trace the supply chain — it extends across oceans, through factories you have never seen, organized by economic and logistical systems you have no

framework to parse. You cannot interview the Americans about causation in a shared language — the language barrier is total, and even if it weren't, the Americans themselves might not be able to explain the full causal chain from ore extraction to aircraft assembly to cargo delivery. Most soldiers cannot.

Arthur C. Clarke's third law — “Any sufficiently advanced technology is indistinguishable from magic” — is not a metaphor. It is an epistemological fact. When the gap between your causal model and the actual mechanism is sufficiently large, **you cannot distinguish advanced technology from supernatural intervention** on the basis of observation alone. Both produce effects that violate your prior models. Both operate through mechanisms you cannot access. The only difference is that advanced technology has a *naturalistic* explanation you do not yet possess, while supernatural intervention has no naturalistic explanation at all. But that difference is not observable from your epistemic position. You are making an inference about unobservable mechanisms, and both hypotheses fit the evidence equally well.

Faced with this situation, **ritual replication is the rational strategy**. If you believe the correlation between ritual and cargo is causal, then reproducing the ritual is the correct way to test the hypothesis. You build the airstrip. You raise the flag. You march in formation. You wait. If cargo arrives, the hypothesis is confirmed. If it does not, you have not falsified the model — you have only discovered that you have not yet identified the *complete* set of necessary conditions. Maybe you need to perform the ritual with more precision. Maybe there are additional rituals you have not observed. Maybe the timing matters. Maybe the gods (or ancestors, or Americans) are testing your faith.

This is the structure Festinger, Riecken, and Schachter documented in *When Prophecy Fails* (1956). They studied a millenarian group that predicted a specific date for the end of the world. When the date passed and the world did not end, the group did not dissolve. Instead, the members **intensified their commitment**. They reinterpreted the prophecy: their faith had delayed the apocalypse; the date had been a test; the prophecy was metaphorical rather than literal. Disconfirmation did not weaken belief; it triggered elaborate rationalization to preserve the core model.

Why? Because abandoning the model has a cost. If you abandon the cargo cult model, you are left with no explanation for the correlation you observed. The Americans *did* perform rituals. Cargo *did* appear. If the rituals did not cause the cargo, what did? You do not know. You have no alternative causal story. Admitting ignorance is epistemically honest, but it is also **strategically costly** when the stakes are high. If there is even a small probability that the ritual works, and the cost of performing the ritual is low (time, effort, social coordination), and the benefit of success is enormous (cargo, wealth, power), then **continuing to perform the ritual is the rational choice** under uncertainty.

This is why cargo cults persist in the face of repeated failure. It is not because the participants

are irrational. It is because **they are making rational inferences under radical uncertainty about high-stakes outcomes**. The same logic governs much of human behavior around uncertain, high-consequence events: religious ritual (perform the ceremony and maybe God will answer; fail to perform it and definitely face divine disfavor), medical desperation (try the unproven treatment because doing nothing guarantees death), financial speculation (buy the asset because missing the upside is psychologically costlier than losing the investment). These are not cognitive errors. They are utility-maximizing strategies under deep uncertainty when the cost of a false negative (failing to act when action would have succeeded) exceeds the cost of a false positive (acting when action is futile).

The problem arises when **this logic is applied in contexts where causal access is actually possible but costly**. The Tanna islanders could not realistically gain causal access to industrial production in 1940 — the knowledge infrastructure required (literacy, mathematics, physics, engineering, economics) was generations away. For them, ritual was not just a rational inference; it was the *only* available strategy. But in contexts where causal access *is* achievable — where you could, with effort, learn the mechanism instead of replicating the form — choosing ritual over investigation is a failure of discipline.

This is the epistemological knife edge modern technological societies walk. We are surrounded by systems whose mechanisms we do not understand: machine learning models, financial instruments, pharmaceutical interventions, supply chains, software stacks. For most of these systems, causal access is *possible* — the knowledge exists, the explanations are published, the mechanisms are legible to those who invest the effort to learn them. But the effort is high. The expertise is specialized. The time investment is steep. And the correlation between ritual fidelity (follow the API documentation, run the training script, take the recommended dosage, deploy the standard configuration) and successful outcome is strong.

Under these conditions, **we face the same choice the cargo cultists faced**: invest heavily in causal understanding (learn the mechanism, trace the dependencies, build the mental model that would let you predict behavior from first principles), or reproduce the ritual (follow the recipe, trust the expert, copy the pattern). The difference is that **we have no excuse**. The Tanna islanders could not learn aeronautical engineering in 1940. We can learn how transformers work in 2024. The causal explanations exist. The textbooks are available. The trade-off is effort, not impossibility.

When we choose ritual over mechanism *despite* having access to causal explanations, we are not responding to insurmountable epistemic opacity. We are responding to the *cost* of resolving that opacity. We are choosing the low-effort strategy (follow the form) over the high-effort strategy (understand the substance) because the low-effort strategy *usually works*. And it does usually work — most of the time, copying the pattern is sufficient. Most users of a deep learning library do not need to understand backpropagation. Most drivers do not need to understand internal combustion. Most patients do not need to understand pharmacokinetics.

But “usually works” is not the same as “always works,” and the failure modes matter. When a system fails in a novel way — when the pattern breaks, when the recipe produces an unexpected outcome, when the ritual does not summon the cargo — **causal understanding is the only route to diagnosis and repair**. If you have only ever reproduced the form, you cannot fix the mechanism. You can only intensify the ritual and hope. You can retrain the model with more data (more marching, more flag-raising). You can add more layers (more ceremony, more fidelity to form). You can try different hyperparameters (different ritual timings, different offerings). But you cannot *reason* about what went wrong, because you do not know how the system works when it goes *right*.

This is why **cargo cultism is a tax on technological societies at scale**. Individual cargo cultism (following forms without understanding mechanisms) is often locally rational. But when cargo cultism becomes the *dominant* mode of engagement with technology — when most users, most developers, most decision-makers are operating at the ritual level — the society loses the ability to respond coherently to systemic failures. The people who *do* understand the mechanisms are a thin layer. The rest are performing the ritual, and when the ritual stops working, they have no fallback.

The Melanesian cargo cults reveal the structure because the failure is visible: the planes do not land, the cargo does not come, the gap between ritual and outcome is undeniable. In modern technological societies, the gap is often hidden. The ritual *does* usually work. The model *does* usually train. The drug *does* usually have the listed effects. The correlation holds. And so the illusion of understanding persists until it breaks — and when it breaks, we discover how many people were operating on ritual alone.

The discipline, then, is this: **assume you are in a cargo cult unless you can demonstrate otherwise**. Can you state the mechanism? Can you predict novel behavior from first principles? Can you specify falsification conditions — observations that would prove your model wrong? If the answer is no, you are performing a ritual. That does not mean stop (sometimes ritual is the best available option). But it does mean: **know that you are performing a ritual, and do not mistake it for understanding**.

The Tanna islanders knew they did not understand. They knew they were inferring causation from correlation. They knew the mechanism was opaque. That is why they called on prophets, why they waited for revelation, why they intensified ritual in the face of disconfirmation. They were honest about the limits of their causal access. The modern failure mode is different: we perform rituals while *believing* we understand mechanisms. We mistake successful pattern-matching for causal knowledge. We confuse “I can make it work by following these steps” with “I know why these steps work.” And that confusion is **epistemologically more dangerous** than honest ignorance, because it forecloses the question of whether deeper understanding is necessary.

8. The Prophet and the Roadmap

Every cargo cult has a prophet: a mediator between the human community and the supernatural source of cargo. John Frum is never seen directly; he communicates through intermediaries who receive his messages and transmit them to the faithful. Yali was the prophet of the Madang movement, legitimated by his direct exposure to the cargo source (Australia) and his special knowledge of its operations. The Vailala prophets channeled ancestors and received their instructions for ritual performance. The prophet's function is to **translate opacity into practice** — to tell the community what rituals will summon cargo, when the cargo will arrive, and how to interpret events that seem to disconfirm the prophecy.

The prophet serves four structural functions:

1. **Legitimation:** The prophet has access the community lacks. He has seen the cargo source, received revelation, been chosen by supernatural powers. His authority derives from this asymmetric access.
2. **Interpretation:** When events are ambiguous — which they always are — the prophet interprets them. A plane flying overhead is a sign. An illness is a test. A failed ritual is a lesson. The prophet converts ambiguity into meaning.
3. **Deferral:** When the cargo does not arrive, the prophet defers. Not yet. The time is not right. We must purify ourselves further. The prophecy is not falsified; it is **postponed**. Deferral protects the belief system from disconfirmation.
4. **Prescription:** The prophet tells the community what to do. Build the airstrip. Destroy the idols. March in formation. Wait. The prophet's prescriptions feel actionable even when the underlying model is false. They provide a sense of agency under conditions of helplessness.

Now consider the figure of the **technology founder** in contemporary culture — specifically, founders of companies promising transformative technological capability. The structural parallels are precise.

Legitimation: The founder has access the community lacks. She has seen the technology work. She has private demonstrations not available to the public. She has been chosen by investors who have performed due diligence. Her authority derives from this asymmetric access. “Trust me, I’ve seen what it can do.”

Interpretation: When events are ambiguous, the founder interprets them. A demo failure is a minor setback. A missed milestone is a pivot. A competitor's success is validation of the category. A lawsuit is persecution by incumbents. The founder converts ambiguity into narrative.

Deferral: When the promised capability does not arrive, the founder defers. “We’re six months away.” “Version 2 will fix this.” “We need one more round of funding.” “The technology works; it’s the regulatory environment.” The roadmap — that formal document of prophetic deferral —

is the **cargo cult calendar**, always pointing to a date beyond the present, always receding as it approaches.

Prescription: The founder tells the community what to do. Invest. Evangelize. Adopt early. Wait for the product to mature. Pre-order. Sign up for the beta. The founder’s prescriptions feel actionable even when the underlying technology is not ready — or does not exist at the claimed capability level.

The roadmap is particularly revealing. In cargo cult prophecy, the arrival of cargo is always imminent but never present. John Frum will return *soon*. The ancestors are sending cargo *now* — look for the signs. The specific date is either unprovable (in the mythic past or eschatological future) or perpetually deferred when it passes without fulfillment.

The technology roadmap has identical structure. “Full self-driving by 2020.” Then 2021. Then 2022. Then “we’re confident in our timeline.” Then “safety is our priority; we won’t rush.” The roadmap is always detailed enough to seem concrete and always malleable enough to survive disconfirmation. When capability does not arrive, the roadmap is updated. When critics point to the update, the response is: “that’s normal in technology development; we’re iterating.” But the function of the roadmap is not to predict; it is to **maintain commitment under uncertainty**. It is the prophetic calendar.

I want to be careful here. The claim is not that technology founders are charlatans, or that roadmaps are lies, or that technological development is fake. Genuine technological progress exists. Real capabilities emerge. Founders sometimes deliver. The claim is that **the social structure around technology development replicates cargo-cult epistemics** because the conditions for cargo cultism are satisfied: opacity (most people cannot evaluate the technology’s actual state), high investment (careers, capital, status), cohesive community (the “tech ecosystem”), and available alternative explanations for disconfirmation (pivots, iterations, regulatory obstacles, market timing).

Under these conditions, distinguishing the genuine prophet (the founder with real technology approaching real capability) from the cargo-cult prophet (the founder with compelling ritual but no substantive mechanism) is **extremely difficult from outside the system**. The external observer sees the same signals: confidence, roadmaps, demos, investor support, deferrals. The difference lies in the underlying causal mechanism — and the mechanism is opaque.

This is Feynman’s point, restated: the cargo cult scientist and the real scientist look similar from outside. Both have labs. Both publish papers. Both give talks. Both use technical language. The difference is that one has captured a causal mechanism and the other has only captured forms. But the audience cannot tell the difference until the plane lands or fails to land, and even then, the cargo cultist will explain the failure as a refinement opportunity.

The test for the observer — and for the founder, the investor, the employee, the customer — is: **can**

I state falsification conditions? Can I name a concrete outcome that would make me conclude the technology does not work as claimed? If no outcome would change my mind, I am in a cargo cult. If every disconfirmation is reinterpreted as iteration, pivot, learning, or external obstruction, I am in a cargo cult. If the roadmap never arrives but I continue investing anyway, I am in a cargo cult.

The discipline of substance, when applied to technology prophecy, is the demand for falsifiability. What would it take for you to stop believing? Name it. Write it down. Hold yourself to it. That is the test.

9. Our Cargo Cults, Extended

Section 5 introduced the claim that contemporary technological culture exhibits cargo-cult structures. Let us now be more precise about the mechanisms, because the claim is structural, not dismissive. The point is not to mock technology practitioners but to identify **where the gap between form and substance emerges and how to close it**.

The dashboard as ritual instrument. Modern organizations are saturated with dashboards: real-time visualizations of metrics, KPIs, OKRs, engagement rates, conversion funnels, uptime percentages, model accuracy scores. Dashboards feel scientific. They present numbers. Numbers feel objective. The numbers move, and movement feels like progress or regress, and response to movement feels like management.

But consider: what causal model connects the numbers on the dashboard to the outcomes the organization actually cares about? In many cases, the answer is: **the causal model is vague, contested, or absent**. The dashboard shows that metric X increased. Did the thing we did cause metric X to increase? Did the increase in metric X cause any outcome we value? Often, these questions are not asked because the dashboard ritual is complete in itself. We checked the dashboard. The dashboard showed improvement. We are doing well.

This is cargo cult management. The form (monitoring metrics, displaying numbers, meeting to review the dashboard) is replicated without the substance (a validated causal model connecting metrics to outcomes). The dashboard is the bamboo radio: it looks like the thing that works, it provides a sense of connection to unseen powers (the market, the algorithm, the users), but it does not actually transmit causally useful information unless the causal chain has been established.

The test: Can you articulate the mechanism by which the metric you are monitoring connects to the outcome you value? Can you state falsification conditions — what dashboard reading would convince you your strategy is wrong? If not, the dashboard is ritual.

“Alignment” and safety theater. In AI development, “alignment” refers to the project of

ensuring that AI systems do what humans want them to do. It is a real research program addressing a real problem. But the *term* “alignment” also functions as a **ritual marker of institutional legitimacy**. A company that says “we prioritize alignment” sounds responsible. A researcher who works on “alignment” has a morally elevated status. Conferences have alignment tracks. Job postings specify alignment expertise.

The risk is that the word becomes disconnected from the research. An organization can claim to prioritize alignment without having any concrete mechanism for achieving it, without falsifiable claims about what alignment means, without research programs that would produce alignment if successful. The word becomes a form — a ritual utterance that establishes membership in the community of responsible actors — without necessarily corresponding to substance.

This is not a claim that alignment research is fake. Some alignment research is substantive, technically grounded, and connected to real mechanisms. The claim is that the discourse *around* alignment exhibits cargo-cult structure: the form (using the word, funding the team, publishing the blog post) can be replicated without the substance (technical progress on the actual problem). The test is: **can the organization state falsification conditions for its alignment claims?** If its system is shown to behave in misaligned ways, does it update, or does it reinterpret the failure as a training opportunity? If every failure is absorbed as iteration, the safety practice is ritual.

Benchmark worship. Machine learning develops through benchmarks: standardized datasets and tasks on which models are evaluated. Benchmarks are useful. They provide reproducible comparisons. They track progress. They create competition that drives improvement.

But benchmarks can also become cargo. When the goal shifts from “build a system that solves the problem the benchmark was designed to represent” to “achieve a high score on the benchmark,” the causal chain breaks. Models are optimized to perform well on the benchmark rather than on the underlying task. The benchmark score becomes the goal rather than a proxy for the goal. Improvements on the benchmark may not transfer to real-world performance.

This is Goodhart’s Law operationalized: when a measure becomes a target, it ceases to be a good measure. But from a cargo-cult perspective, the structure is identical to Melanesian ritual: we observe that high-benchmark models come from prestigious institutions; we optimize for high benchmarks; we present benchmark scores as evidence of capability. The form (benchmark performance) is replicated. The substance (real-world problem-solving) is not guaranteed to follow.

The test: Does your benchmark score predict performance on the task you actually care about? Can you state what real-world outcome corresponds to a one-point improvement in benchmark accuracy? If the benchmark is not validated against deployment conditions, the benchmark is ritual.

The ceremony of the demo. Technology companies present demos: live or recorded demonstrations of technology performing impressive tasks. The demo is powerful. Seeing is believing. The

demo compels investment, partnership, media coverage, talent recruitment.

But the demo can be cargo. A demo shows a system performing under controlled conditions: carefully selected inputs, favorable environments, human oversight concealed behind the curtain. The demo does not show the failure cases, the edge conditions, the hours of cherry-picking required to find a demonstration-worthy output. The demo shows the form (impressive performance) without revealing the substance (what the system actually does at scale, under deployment conditions, without human curation).

Some demos are honest: they demonstrate real capability that will hold at deployment. Some demos are theater: they demonstrate a best-case scenario that does not generalize. From outside, the observer cannot tell the difference — because the demo is designed to prevent that distinction. The demo is the bamboo tower: it looks like the capability that produces cargo, and it is impressive, and it summons investment, but whether it is connected to real cargo-delivery depends on mechanisms the audience cannot inspect.

The test: What would the system do on a randomly selected input rather than a curated one? What is the failure rate? What are the boundary conditions? If these questions cannot be answered, or if answers are resisted as “missing the point,” the demo is ritual.

The common structure across these examples is: **form substitutes for substance when the causal chain is opaque and the form is sufficient to satisfy social or institutional requirements**. Organizations need to look like they are managing; dashboards satisfy this need regardless of causal validity. Organizations need to look responsible; safety language satisfies this need regardless of technical substance. Researchers need to show progress; benchmark scores satisfy this need regardless of real-world transfer. Companies need to attract investment; demos satisfy this need regardless of deployment reliability.

This is not conspiracy or fraud (usually). It is **cargo-cult epistemics emerging from structural incentives under opacity**. The participants are not lying; they are performing rituals that satisfy legitimate social needs. The tragedy is that the rituals can be performed without the substance, and under opacity, no one — including the performers — may be able to tell the difference.

10. The Discipline of Substance

The cargo cults of Melanesia could not have escaped their epistemological trap alone. The causal chain between industrial production and cargo delivery was too long, too distributed, too opaque for any observer on Tanna or in the Papuan Gulf to reconstruct from first principles. Escape required **access to the mechanism**: education, exposure, integration into the broader economic system, explanatory frameworks that connected ritual to production to transportation to delivery. Escape

was not a matter of “being smarter” or “thinking harder.” It was a matter of gaining access to information that was structurally unavailable.

We are not in the same position. We have access. We can inspect mechanisms. We can demand evidence. We can state falsification conditions. We can distinguish form from substance — if we have the discipline to do so.

The discipline of substance is a practice, not a disposition. It is a set of concrete tests that anyone can apply to distinguish genuine causal understanding from cargo-cult form-mimicry. Here are the tests:

Test 1: Can you state falsification conditions?

A belief is cargo-cult if no possible observation would change it. For any claim — technological, institutional, strategic, personal — ask: what would convince me this claim is false? Name a specific outcome. If you cannot name one, or if every outcome you consider is reinterpretable as consistent with the claim, you are in a cargo cult.

Falsification conditions must be specific enough to be observable. “If the system fails” is not a falsification condition; “failure” can always be reinterpreted. “If the system produces output X under conditions Y” is a falsification condition. Specificity is the discipline.

Test 2: Do you understand the mechanism, or only the form?

Cargo cults replicate form because the mechanism is opaque. The test is: can you explain *why* the system produces the outputs it produces? Can you trace the causal chain? Can you predict what would happen if you intervened at a specific point in the chain?

You do not need to understand every mechanism. You are allowed to trust experts. But you should be able to distinguish: “I understand why this works” from “I trust that someone understands why this works” from “I believe this works because it has the form of things that work.” The third posture is cargo cult. The first two are legitimate, but they have different epistemic statuses, and conflating them is the failure mode.

Test 3: Does disconfirmation update you or reinforce you?

When predictions fail, do you conclude that the underlying model is flawed, or do you conclude that the prediction was off for exogenous reasons while the model remains sound? Both responses are sometimes appropriate. But if you notice that *every* disconfirmation reinforces the model — if every failure is interpreted as “we need more X,” “the timing was off,” “external factors intervened,” “this is actually good because it shows we’re learning” — you are in a cargo cult. The model has become unfalsifiable in practice.

The discipline is to notice when you are explaining away rather than updating. Write down predictions. Note outcomes. Track your update history. If you never update, you have stopped

learning.

Test 4: Is the ritual complete in itself, or does it require external validation?

Some practices are means to ends: they are validated by whether they produce the claimed outcome. Other practices are rituals: they are validated by being performed correctly, regardless of outcome. Cargo-cult practices are rituals that present as means. They claim to produce outcomes, but they are actually validated by performance.

The test is: if the outcome did not occur, would you conclude the practice failed, or would you conclude the practice was performed inadequately? If the practice can never fail — if poor outcomes are always attributed to inadequate performance rather than flawed practice — the practice is ritual. Rituals are fine; rituals serve social and psychological functions. But calling rituals “technology” or “strategy” or “science” is cargo cult.

Test 5: Can you distinguish your belief from its social function?

Beliefs serve social functions: they signal membership, establish status, coordinate action, provide meaning. These functions are legitimate. But they can become untethered from truth. You can believe something because believing it serves social purposes, not because it is true.

The test is: would you believe this if believing it had no social benefit? If holding a different belief would cost you status, membership, or identity, you should be suspicious of your confidence. This does not mean the belief is false. It means you cannot trust your own judgment because your judgment is compromised by incentive. Acknowledge the compromise. Seek external validation. Be especially careful when the belief is both socially rewarded and conveniently unfalsifiable.

Institutional implications.

These tests are individual, but the discipline of substance must also be institutional. Organizations — companies, universities, governments, research groups — develop cargo-cult practices when incentive structures reward form over substance. The CEO who demands dashboards creates dashboard rituals. The funder who demands publications creates publication rituals. The regulator who demands compliance creates compliance rituals.

The institutional discipline of substance is: **reward outcomes, not forms; demand falsification, not narrative; track predictions, not explanations; audit practices against results.** This is difficult. Forms are visible; substance is often invisible until deployment. Narratives are comforting; falsification is threatening. Explanations are flexible; predictions are rigid. But the alternative is institutional cargo cultism — organizations that replicate the forms of success without achieving the substance.

The through-line.

This paper is part of a series on technology and belief. The series examines the structures of

thought that govern how humans relate to technological power — structures that are often invisible to the thinkers inside them. Cargo cultism is one such structure: the inference that form causes outcome, the replication of form in hope of reproducing outcome, the persistence of belief despite disconfirmation.

The cargo cults of Melanesia are not exotic curiosities. They are mirrors. They show us the structure of our own thinking when we encounter power we do not understand. The airstrips on Tanna and the dashboards in conference rooms are different in content but identical in structure: they are rituals performed in hope of summoning cargo, forms replicated in absence of causal understanding.

The discipline of substance is the antidote. Not intelligence; disciplined habits of thought. Not skepticism; constructive demand for mechanism. Not cynicism; the practice of writing down predictions and checking them against reality. The cargo cult fails because it mistakes form for cause. The discipline succeeds because it insists on inspecting the mechanism, stating falsification conditions, and updating when evidence demands.

John Frum has not returned to Tanna. The bamboo control towers are maintained anyway. The discipline of substance is the decision not to build bamboo towers — or, if we find ourselves in the tower, to ask whether the radios work.

References

- Clarke, A. C. (1973). Hazards of prophecy: The failure of imagination. In *Profiles of the Future: An Inquiry into the Limits of the Possible* (Rev. ed.). Harper & Row.
- Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When Prophecy Fails: A Social and Psychological Study of a Modern Group That Predicted the Destruction of the World*. University of Minnesota Press.
- Feynman, R. P. (1974). Cargo cult science. *Engineering and Science*, 37(7), 10–13. [Caltech commencement address]
- Guiart, J. (1952). John Frum movement in Tanna. *Oceania*, 22(3), 165–177.
- Lawrence, P. (1964). *Road Belong Cargo: A Study of the Cargo Movement in the Southern Madang District, New Guinea*. Manchester University Press.
- Lindstrom, L. (1993). *Cargo Cult: Strange Stories of Desire from Melanesia and Beyond*. University of Hawaii Press.
- Williams, F. E. (1923). *The Vailala Madness and the Destruction of Native Ceremonies in the Gulf Division*. Territory of Papua Anthropology Report No. 4. Government Printer.

Worsley, P. (1957). *The Trumpet Shall Sound: A Study of "Cargo" Cults in Melanesia*. MacGibbon & Kee.
