The Collapse of Proxy Integrity

Signal, Simulacrum, and the Structural Phase Change of Attention Platforms

Flyxion

2025

When a measure becomes a target, it ceases to be a good measure.

— Charles Goodhart, 1975

All that was directly lived has receded into a representation.

— Guy Debord, *The Society of the Spectacle*, 1967

Abstract

This essay argues that large-scale attention platforms have undergone a structural phase change in which the function of signals has shifted from *representing* value to *manufacturing* the appearance of it. What is conventionally diagnosed as a collection of abuses—deepfake content, coordinated metric inflation, synthetic performance—is better understood as a single systemic transformation: the progressive detachment of proxy signals from the processes they were constructed to track. Drawing on Goodhart's law generalized as a structural attractor and evolutionary selection pressure rather than a contingent failure, the essay distinguishes between manipulation of the proxy layer (metric gaming) and manipulation of the referent layer (synthetic performance), and shows that contemporary platforms have achieved the simultaneous degradation of both. The result is a signaling system that is no longer representational but generative: it produces the very signals it then measures, and in doing so loses bidirectional constraint between signal and world. The extraction architecture of platforms is shown to be not merely tolerant of this epistemic degradation but positively sustained by it under competitive conditions, which explains why the collapse is stable rather than self-correcting. The essay closes with a defense of temporal integrity—the condition under which effort, duration, and constraint are constitutive of meaning—and considers what structural features, specifically incomplete observability of the proxy layer, an honest platform would require.

Length: approximately 14,000 words.

2

Contents

1. The Mechanism Before the Examples

Begin with a minimal formal picture, because the argument requires it.

A *proxy* is a measurable quantity that stands in for something harder or more expensive to measure directly. A "like" is a proxy for interest or approval. A follower count is a proxy for audience reach or social credibility. A high view count is a proxy for content value. A viral spread is a proxy for cultural resonance. In each case, the proxy is adopted because the underlying thing—genuine interest, real credibility, actual value—resists direct measurement at scale. The proxy substitutes for it.

For a proxy to function, it must be *grounded*: there must be a reliable causal relationship between the proxy signal and the process it tracks. When someone clicks "like" on a post that moved them, the signal is grounded because the click is caused by the affective response. When a follower subscribes to an account because they anticipate future value, the signal is grounded because the subscription is caused by the anticipated quality. The signal *tracks* the process; the map *represents* the territory.

Grounding admits of degrees. A proxy can be weakly grounded—noisy, imperfect, subject to confounds—while still carrying genuine information about the underlying phenomenon. Most real-world proxies operate in this weakly grounded regime. The problem is not weak grounding but *detachment*: the condition in which the causal relationship between signal and underlying process has been severed or inverted, such that the signal moves independently of, or contrary to, the phenomenon it nominally measures.

The distinction between *representation* and *manufacture* maps directly onto this. A system is representational when its signals are caused by the phenomena they track, even imperfectly. A system is

generative—in the sense used here, which is not the sense of generative AI—when it produces signals that are no longer caused by the underlying phenomena but are instead produced independently and circulate as if they were. In a representational system, you observe the territory and produce a map. In a generative system, you produce the map directly, and the territory becomes irrelevant to the map's circulation.

This essay's central claim is that the major attention platforms have undergone a transition from the first regime to the second, and that this transition is not accidental, correctable, or the result of bad actors exploiting an otherwise sound system. It is a structural consequence of how the system was built, and it was, in the sense that will be made precise below, largely inevitable.

2. Grounding as Constraint: An Operational Criterion

The concept of grounding introduced above can be sharpened into an operational criterion that governs the remainder of the analysis.

A proxy is not merely grounded when it is *correlated* with an underlying process. Correlation is too weak a condition, because correlations can be preserved even when the causal structure that produced them has been severed—by coincidence, by common cause, or by the residue of an earlier relationship that no longer holds. The criterion must be stronger.

A proxy is grounded when it is *constrained* by the process it represents, in the precise sense that the proxy cannot be moved at scale without commensurate movement in the underlying process. Equivalently: a grounded proxy is one that cannot be cheaply moved without moving the underlying process. If it is possible, at low cost, to systematically alter the proxy signal while leaving the underlying pro-

cess unchanged, the proxy is not grounded. Conversely, if moving the proxy without engaging the underlying process incurs costs comparable to those of engaging the process itself, the proxy retains its grounding.

This framing yields a clean diagnostic that recurs throughout the analysis. Wherever signals can be moved independently of their supposed referents—wherever the cost of manufacturing the signal is substantially lower than the cost of generating the underlying process—grounding has failed. The diagnostic is structural, not moral: it concerns whether the signal is causally entangled with the phenomenon it purports to measure, regardless of the intentions of anyone involved.

This criterion also clarifies the distinction between weak grounding and detachment. In a weakly grounded system, the proxy is noisy but constrained: manipulation is possible but limited, costly, or detectable, and the signal retains informational content despite its imperfection. In a detached system, the constraint is broken: manipulation becomes cheap, scalable, and difficult to distinguish from genuine signal generation. The proxy ceases to track the process not because noise has increased, but because the causal link has been rendered optional.

The operational criterion allows the central thesis to be stated with precision. The collapse of proxy integrity is the transition from a regime in which proxies are causally constrained by the processes they represent to a regime in which proxies can be freely manipulated independently of those processes. What follows is not merely a degradation of signal quality but a reconfiguration of the system's causal architecture. Signals become autonomous variables within the system rather than reflections of external phenomena. It is this reconfiguration—not any particular instance of abuse—that constitutes the structural phase change this essay analyzes.

3. Goodhart's Law as Structural Attractor and Evolutionary Selection

The standard version of Goodhart's law says: when a measure becomes a target, it ceases to be a good measure (Goodhart, 1975). This is usually presented as a warning, an observation about the risks of naive optimization. But that framing understates the force of what is being claimed. The law does not say that proxy measures *tend to degrade* under optimization pressure. It says that they *cease to function*—that the optimization pressure applied to the proxy necessarily introduces behaviors that corrupt the proxy's relationship to the underlying variable. If that is right, then proxy degradation is not a risk to be managed but an *attractor*: the stable end-state toward which any system applying sustained optimization pressure to a proxy will converge.

Why is it an attractor rather than a risk? Because the logic is asymmetric. In a world where proxy signals are used to distribute rewards—attention, income, reach, status—any agent capable of influencing those signals faces an incentive to do so. If the signal is cheaper to manipulate directly than to earn through the underlying process, then manipulation will occur. The cost asymmetry is almost always present in digital systems, where producing a click, a follow, or a view is computationally trivial compared to the effort required to produce genuinely compelling content or authentic social connection. Once that asymmetry exists, optimization pressure selects for signal manipulation over genuine achievement. The proxy detaches from its referent not because agents are malicious but because the incentive structure makes detachment rational.

The further point, which is crucial, is that this process is self-reinforcing. As more agents optimize for the signal rather than the underlying process, the signal becomes a less reliable guide for other agents

who are still trying to use it as a measure of quality. The noise in the proxy increases. This makes genuine quality harder to identify through the proxy, which weakens the incentive to invest in genuine quality, which causes more agents to shift toward direct signal manipulation. The system spirals. The proxy degrades not linearly but through a feedback loop that accelerates the detachment. Marilyn Strathern's generalization of this principle—that the act of measurement transforms the thing measured once actors know they are being measured—identifies the same dynamic operating at the cultural level (Strathern, 1997).

This is not merely an internal drift. It is also a competitive selection pressure. Consider two platforms with similar initial conditions. One enforces strict coupling between proxy signals and underlying processes: it invests in detection, imposes friction on manipulation, and constrains the degree to which engagement metrics directly determine distribution. The other relaxes these constraints, allowing engagement signals to circulate with minimal verification and using them aggressively to drive ranking and amplification. The second platform will, all else equal, exhibit higher short-term engagement metrics. It surfaces content optimized for immediate interaction regardless of its grounding, and it attracts creators willing to optimize for those signals. Under an advertising-based revenue model, the second platform is more successful. It aggregates more attention, generates more interactions, and produces a larger surface for monetization.

In a competitive environment, the first platform is at a systematic disadvantage. Users migrate toward the platform that appears more active and more rewarding, even if the underlying signal environment is degraded. Platforms that maintain stronger forms of grounding are outcompeted by those that relax them. Proxy detachment is not merely an internal attractor; it is reinforced by external selection. The stable equilibrium is one in which detachment is pervasive, because systems that resist it are displaced. This closes a potential

objection. One might argue that platforms could, in principle, preserve proxy integrity through better design and enforcement. The selection argument shows why this is structurally unlikely: any such platform incurs costs its competitors do not. Without a change in the economic or regulatory environment, the equilibrium outcome is not reform but convergence on the same degraded state across the entire competitive landscape.

The political economy of platforms adds a final layer to this structure. Platforms do not merely fail to correct proxy degradation; they are indifferent or actively hostile to its correction. The advertising model monetizes attention rather than value: what matters to the revenue stream is that users are present and engaged, not that the content they consume is genuine or meaningful. A platform that successfully curates only high-quality, authentically produced content might generate a smaller, more stable audience. A platform that surfaces maximally engaging content regardless of its authenticity generates a larger, more volatile audience, which is more valuable for advertising purposes. The optimization target at the platform level is thus not aligned with proxy integrity, and the competitive dynamics described above ensure that any platform that tries to maintain that alignment will be penalized for doing so.

4. Inflation and Indistinguishability: The Proxy Layer's Epistemic Failure

With the theoretical structure established, the first of the two empirical phenomena can be identified precisely.

Metric gaming—the ecosystem of follow-for-follow exchanges, reciprocal like networks, coordinated engagement pods, and purchased interactions—is a direct manipulation of the proxy layer. Its practitioners are not trying to produce better content or to reach audiences

more effectively. They are trying to move the needle on the measurable signals that the platform uses to distribute further reach. The underlying process—the creation of genuinely interesting or useful content that attracts a real audience—is bypassed entirely. The signal is addressed in isolation from its referent. By the operational criterion established in Section 2, this is a clear case of proxy detachment: the cost of moving the signal has been decoupled from the cost of engaging the underlying process.

This is not a marginal phenomenon. Entire communities are organized around it. Their operating logic is explicit: members exchange follows, likes, and comments not because they are interested in one another's content but because they want their own numbers to increase. The social relationship is purely instrumental. The result is a dense cluster of interactions that, viewed from outside—or from the platform's algorithmic vantage—is indistinguishable from organic enthusiasm. The ranking system sees engagement; it does not see the social organization that produced it.

It is tempting to describe this as a problem of *inflation*: too many likes, too many followers, too many views relative to the underlying level of genuine interest. While inflation captures part of the phenomenon, it does not reach its epistemic core. Inflation alone does not destroy a metric's informational content. If inflated and non-inflated signals were reliably distinguishable, users could discount the former and continue to extract information from the latter. The problem is not simply that there are more signals than before, but that signals produced through fundamentally different processes—organic interest, coordinated exchange, automated amplification—are *observationally equivalent at scale*.

This indistinguishability is the critical failure mode. A follower count of one hundred thousand may represent a genuinely engaged audience, a network of reciprocal exchanges, a mixture of purchased and organic growth, or some combination of all three. The visible number

does not encode the provenance of the interactions that produced it. From the standpoint of an external observer, the signals are informationally collapsed: they no longer differentiate between distinct underlying processes.

Once indistinguishability is established, the metric ceases to function as a guide to action. A user cannot reliably use follower counts to identify creators worth attending to, because the mapping from count to quality is no longer stable. A creator cannot rely on engagement signals to infer audience response, because those signals may reflect coordinated amplification rather than genuine reception. The proxy has not merely been inflated; it has lost its capacity to discriminate between the phenomena it was designed to differentiate. This is a qualitative transformation, not a quantitative one: the metric has not become less accurate but structurally uninformative.

This loss of discrimination feeds back into the production environment. When metrics cannot be trusted to differentiate quality, actors rationally shift their efforts from improving quality toward manipulating the metrics themselves, because the latter is more directly connected to distribution outcomes. Indistinguishability thus accelerates the behaviors that produced it, reinforcing the collapse in a second feedback loop that operates independently of the first.

The platform's response to this phenomenon illustrates exactly the dynamic described in the previous section. Large platforms do attempt to detect what they call "coordinated inauthentic behavior" (Zuboff, 2019), and they do remove some of it. But aggressive enforcement is bounded by several constraints. First, these are technically real people performing technically real actions. Unlike bot networks generating fake accounts or scripted clicks, engagement pod members are human and executing genuine platform interactions. The signal they produce passes most behavioral filters designed to detect automation. Second, enforcement risks false positives that would penalize legitimate social activity—friends support-

ing each other, small communities amplifying shared work. The result is a stable equilibrium in which a persistent layer of manufactured engagement coexists with organic activity, with no reliable mechanism for distinguishing them at scale.

Bourdieu's analysis of symbolic capital is clarifying here as a coordinate, not a foundation (Bourdieu, 1979). He showed that what looks like organic social prestige often encodes accumulated advantages that are themselves structurally produced. The follow-for-follow ecosystem can be read as a crude mechanization of this dynamic: it manufactures symbolic capital through social coordination, bypassing the cultural processes that symbolic capital was supposed to reflect. The difference is that Bourdieu was describing structures that operated over lifetimes and across fields. The digital proxy layer permits the same operations to be executed in minutes, at scale, by anyone with a smartphone. What Bourdieu analyzed as a slow structural feature of social fields has become an instantly available tactical resource.

5. Temporal Compression as Referential Substitution: The Referent Layer

The second phenomenon operates at a different level. Where metric gaming manipulates the proxy—the signal layer—synthetic content manipulates the *referent*: the underlying process that the signal was supposed to track.

Consider what it means, in the human experience of watching a skilled dancer, to recognize skill. The recognition is not merely perceptual—it is not only that the movements are precise or beautiful. It is the implicit knowledge that those movements represent the accumulated residue of years of practice, failure, correction, and repetition. The craft is visible in the performance, but it is legible *because* you know about the temporal depth behind it. The skill is meaningful partly

because it is scarce, and it is scarce because it is costly to produce in time and effort. Stiegler's account of technical memory captures a part of this: he argued that skills constitute a form of externalized trace that carries the temporal structure of their production; when you encounter a master performer, you are encountering the trace of a long history of training (Stiegler, 1998). Sennett, examining craft from a different angle, arrives at a similar point: that the tacit knowledge embedded in skilled practice is inseparable from the duration of its accumulation (Sennett, 2008).

Now apply a filter. A parent points a phone at a child and taps a button. The resulting video shows the child executing a highly skilled professional routine, complete with precise footwork, fluid transitions, and stylistic details that would require years of training to produce. The child did not move. The movement was synthesized and overlaid. The video circulates, receives tens of thousands of views, and generates the same engagement signals as a genuine performance video.

What has happened can be stated precisely as *temporal compression*. A process that, under ordinary conditions, unfolds over an extended duration—training, practice, maturation—has been replaced by an instantaneous transformation that produces an output with similar surface properties. Temporal compression is not merely an acceleration of production. It is a substitution of process with artifact. In an accelerated but still grounded system, the underlying process remains intact, albeit executed more quickly. In a temporally compressed system, the process is removed entirely, and its output is synthesized directly. The relationship between input and output is no longer mediated by time, effort, or constraint.

This substitution has two consequences for the referent layer. First, it breaks the causal chain that links visible performance to underlying capability. The output no longer carries information about the process that produced it, because no such process occurred. The

performance now refers to nothing: not to a process of learning, not to years of effort, not to a developing human capability. It refers only to itself, a self-contained visual object that mimics the form of skill while severing the connection between form and history.

Second, temporal compression alters the effective distribution of apparent capability within the system. When the cost of producing the appearance of high-level skill approaches zero, the system becomes saturated with such appearances, regardless of the actual distribution of skill in the population. This is not merely an economic observation. It is an epistemic one: observers can no longer infer the existence or absence of underlying processes from the appearance of performance. The referent layer has been replaced by a simulation that preserves the observable features of the phenomenon while eliminating the conditions that made those features meaningful.

The content exists in a studied ambiguity that serves engagement. These videos are not always presented as fraudulent claims of real ability. But neither are they clearly labeled as synthetic. The ambiguity is functional: it invites interpretation, generates debate ("is this real?", "how did they do this?"), and sustains attention through unresolved uncertainty. What looks like a failure of transparency is actually, from an engagement-optimization standpoint, a feature. Unresolved questions about authenticity keep the viewer in a state of interpretive engagement longer than either a clearly fake or clearly real video would.

There is a specific category of synthetic content that carries additional weight in this analysis: videos of children performing complex religious recitations, of infants appearing to articulate sophisticated speech, of animals apparently behaving with improbable competence. These represent a further compression: not only is the temporal process of skill development eliminated, but the developmental processes of language acquisition, cognition, and maturation are bypassed entirely. The synthetic layer projects adult competencies

onto pre-competent subjects. This is not incidentally disturbing. The disturbance is structural. The meaning of a child singing beautifully depends on the knowledge that the child is a developing human being whose voice and musicianship are emerging over time. Remove that temporal knowledge, and what you have is a visual spectacle that refers to no process at all.

By the operational criterion established in Section 2, this is proxy detachment at the referent layer. The cost of producing the *appearance* of a long temporal process has been reduced to nearly zero, while the cost of the actual process remains unchanged. Signals can be moved—the video circulates as if it represented genuine skill—without moving the underlying process. The causal constraint between appearance and process has been broken.

6. The Total Inversion: Loss of Bidirectional Constraint

The argument now reaches its conceptual keystone.

Sections 4 and 5 have analyzed two distinct forms of proxy detachment. Section 4 addressed manipulation of the proxy layer: the signals of popularity and social credibility are manufactured independently of the underlying phenomena they nominally measure. Section 5 addressed manipulation of the referent layer: the apparent processes of skill, effort, and development are manufactured independently of actual training, time, and effort. One manufactures the map; the other manufactures the territory. These might seem like two separate problems, each requiring its own corrective. They are not. They are two expressions of the same structural move, and when both occur simultaneously—as they do, routinely, on contemporary platforms—the result is a total inversion of the representational relationship between signal and world.

To see why, consider what a functioning signaling system requires. At minimum, it requires that the proxy signals track some underlying process (grounding at the proxy layer), and that the underlying process refers to genuine human activity in the world (grounding at the referent layer). Both groundings have now been severed. The proxy layer has been colonized by manufactured engagement. The referent layer has been colonized by temporal compression and synthetic performance. A video of a synthetic child performance circulated through an engagement pod has its signals manufactured at both levels simultaneously. The view count does not track genuine audience interest, and the performance does not track genuine human capability. There is no representational relationship anywhere in the chain.

More precisely: the system has lost *bidirectional constraint.* In a grounded system, the proxy layer constrains interpretation of the referent by filtering attention—high engagement directs observers toward phenomena that are more likely to be valuable or interesting. The referent layer constrains interpretation of the proxy by providing a causal basis for the signals—the observed quality of a performance explains why it has attracted attention. This mutual constraint creates a feedback loop in which signals and processes reinforce one another. A viewer encountering a highly viewed performance can interpret the view count in light of the visible skill, and can interpret the skill in light of the collective response it has elicited. The system is not perfectly accurate, but it is self-correcting to some degree, because inconsistencies between proxy and referent can be detected and resolved.

When both layers become manipulable, bidirectional constraint is lost. The proxy no longer constrains interpretation of the referent, because high engagement may reflect coordinated amplification rather than genuine audience response. The referent no longer constrains interpretation of the proxy, because apparent skill or authenticity may be synthetically produced. Each layer becomes independent of

the other. The consequence is not merely increased noise but structural indeterminacy. There is no longer a stable basis for interpreting either signals or phenomena, because neither provides reliable information about the other. The user cannot resolve this ambiguity by gathering more information within the system. The uncertainty is not epistemic in the ordinary sense of insufficient data; it is structural, arising from the absence of a stable relationship between data and world.

The system is no longer *representational*. It is *generative* in a precise sense: it produces the signals it then measures, in a closed loop that has no external reference. This is not a rhetorical move. It is a structural description of what happens when both the proxy layer and the referent layer are simultaneously available for manufacture. The platform's engagement metrics are largely a measurement of their own internal dynamics, not a reflection of any external social or cultural reality.

The image that best captures this is neither a distorted map nor a false territory, but a system in which the map and the territory have been replaced by a single synthetic artifact that mimics both without being either. Debord's formulation—that all that was directly lived has receded into a representation—comes close (Debord, 1967). But even it preserves the residue of the lived. What is being described here is a further step: not the retreat of the lived into representation, but the elimination of the lived as a necessary condition of the representation's production. The spectacle no longer needs a referent to generate itself.

When this happens, the users operating within the system face a condition of radical interpretive uncertainty. They cannot distinguish synthetic from genuine performance by observation alone, because the synthetic mimics the surface properties of the genuine with high fidelity. They cannot use engagement signals to guide attention toward quality, because those signals are detached from both quality

and genuine audience response. Every visible indicator of value is simultaneously a potential artifact of the generative machinery.

7. The Extraction Layer: Why the System Does Not Collapse

If the preceding analysis is correct, a question becomes pressing: why does the system not fail? If the signaling environment has been gutted of semantic content, if users cannot rely on any visible indicator of quality or genuine social interest, why does engagement remain high and why do platforms continue to grow?

The answer is that value capture at the platform level is structurally decoupled from signal integrity. More than merely tolerant of epistemic degradation, the extraction architecture positively sustains it. Systems that increase engagement by degrading grounding are, under the advertising revenue model, positively selected. This is the feature of the system that makes the collapse stable rather than self-correcting, and it is the feature that distinguishes the present situation from a market failure that competition might eventually correct.

The advertising model does not pay for quality. It pays for presence. An advertiser buys access to an audience that is attending to the platform at a given moment. What that audience is attending to, and whether it is genuine or synthetic, is largely irrelevant to the transaction. A user watching a synthetic child-dance video and experiencing the particular affective state that video induces—amusement, wonder, the uncanny ambiguity about its reality—is just as monetizable as a user watching a genuine performance. In some respects more so: the video's structural ambiguity generates longer watch times and more comment activity, which increases the advertising surface. Synthetic content is not merely tolerated by the extraction model; it is, from the revenue standpoint, often preferable.

This creates what might be called *epistemic externalization*. The costs of proxy degradation are borne by users (who can no longer make meaningful judgments about content quality), by creators (whose genuine effort is devalued by synthetic competition), and by the broader culture (whose capacity to distinguish skill from simulation is gradually impaired). The benefits of proxy degradation accrue to the platform, in the form of higher engagement and larger advertising revenue. The platform is economically insulated from the damage its architecture causes. This is not an accident; it is what it means for an extraction model to be structurally indifferent to the quality of what it extracts from. Zuboff's analysis of surveillance capitalism identifies the same basic dynamic: the behavioral surplus that platforms harvest has value to advertisers independent of whether the behaviors themselves are meaningful, authentic, or produced under conditions of genuine human choice (Zuboff, 2019).

The creator monetization promise compounds this dynamic. Platforms offer creators the possibility of income through advertising revenue sharing, brand partnerships, and direct-support features. The promise is framed in the language of democratization: anyone can reach an audience, anyone can earn. But the distribution of outcomes follows a power law: a small fraction of creators captures the overwhelming majority of views, reach, and revenue, while the vast majority receive negligible returns (Wu, 2016). The system does not need most creators to succeed. It needs most creators to *believe* that success is within reach, and to continue producing content on that basis.

The analytics dashboards serve this function precisely. They present creators with an array of measurable variables—impressions, retention rates, click-through percentages, audience demographics—that create a strong impression of controllability. Pasquale has described a related dynamic: the opacity of algorithmic systems produces an experience of being governed by black boxes while simultaneously being told that one can optimize one's position within them (Pasquale,

2015). The dashboard gives creators a language for describing their work that sounds like skill refinement but behaves like probability management. If you understand your data, you can optimize your performance. That this is partially true is what makes it effective: small improvements in some of these variables do sometimes produce better outcomes. But the relationship between effort, optimization, and outcome is heavily stochastic. A large proportion of a post's success or failure is determined by factors outside the creator's control: timing, algorithmic sampling decisions, the behavior of engagement networks that may or may not amplify the post.

8. Deterministic Systems and Subjective Stochasticity: The Lottery Terminal

The characterization of the platform as a "video lottery terminal" can be refined by distinguishing between objective determinism and subjective stochasticity. At the level of the platform's internal operations, the system is not random. Content distribution is governed by algorithms that process large numbers of variables and produce deterministic outputs given those inputs. The system is, in principle, fully causally determined.

From the standpoint of the individual user or creator, however, the system is effectively stochastic. The relevant variables are not observable in full, their interactions are highly nonlinear, and they change continuously as the system evolves. Small differences in initial conditions—timing of a post, the behavior of a small subset of early viewers, the presence of competing content—can produce large differences in outcomes. The mapping from action (posting content) to outcome (distribution, engagement, monetization) is therefore opaque and unstable.

This opacity produces a phenomenological experience of randomness. Creators observe that similar pieces of content produce radically dif-

ferent outcomes, that careful optimization sometimes yields no improvement, and that occasional large successes occur without clear causal explanation. The system behaves, from their perspective, like a probabilistic machine. This is not the same as claiming the system is literally random. The claim is more precise: the system is irreducibly stochastic *from the user's epistemic position*, because the variables that determine outcomes are inaccessible, numerous, and rapidly changing. The stochasticity is a feature of the user's epistemic situation, not of the platform's operations. This distinction protects the analysis from the obvious objection that sufficiently skilled or informed creators might achieve predictable outcomes. They cannot, not because the system lacks causal structure, but because that structure is not accessible from within the user's position.

This is the structural basis of what Skinner would recognize as a variable-ratio reinforcement schedule (Skinner, 1938). You do not need frequent rewards; you need intermittent, unpredictable ones that are large enough to suggest that a pattern might exist. A single post that "takes off"—receives dramatically more distribution than usual, generates significant new followers, or produces meaningful income—can sustain months of continued effort and low-return content production. The unpredictability is not incidental. It is the mechanism by which attention and effort are captured. The slot machine does not need to pay out often; it needs to pay out occasionally, and in a way that sustains the belief that the next pull might be the one.

The phrase "video lottery terminal" is therefore not merely a metaphor. It is a phenomenological description of the user's epistemic situation once both the proxy layer and the referent layer have been destabilized. In a lottery, you cannot improve your expected value through skill, knowledge, or practice. You can only participate and wait. From the creator's perspective, once you cannot reliably predict which content will be amplified, and once the amplification signals no longer track genuine audience response, the system becomes functionally indistinguishable from a probabilistic machine. The interface presents

dashboards and analytics that suggest optimization is possible. The underlying dynamics behave like a draw.

The gap between the interface's promise of rationality and the actual stochasticity of outcomes is not an accident. It is what keeps participants engaged: it is easier to quit a slot machine that says it is a slot machine than one that tells you your luck will improve if you pull the lever more skillfully.

9. The Long-Term Consequences: Reversible and Irreversible Damage

The consequences of this structural phase change are not uniform in their severity or in their susceptibility to correction. Some of the damage is reversible, in the sense that a change in platform architecture or incentive structure could, in principle, restore what was lost. Some of it is not. The distinction is not merely practical but structural: reversible damage concerns the surface configuration of the system; irreversible damage concerns the epistemic and cultural infrastructure through which the system's outputs are interpreted.

9.1. Reversible: Trust in Metrics

The most immediately visible damage is the collapse of trust in platform metrics as meaningful signals. Follower counts, like totals, and view counts have already lost much of their informational credibility among sophisticated users. It is now widely understood that these numbers can be manufactured, that they are manipulated by coordinated networks, and that they correlate poorly with content quality or genuine audience interest. Some platforms have responded by hiding like counts or shifting toward more opaque composite signals—watch time, completion rates, network diversity—in an attempt to recover a closer relationship between signal and substance.

This is, in principle, recoverable. It requires architectural changes: decoupling rewards more firmly from easily manipulable counts, introducing friction into the proxy manipulation pipeline, and making engagement pod behavior visible rather than invisible in the ranking system. A platform that chose to prioritize signal integrity over raw engagement could, with sustained effort, begin to restore the informational content of its metrics.

9.2. Reversible: The Production Environment

The current production environment—in which synthetic content competes directly with authentic content for distribution, and often prevails because it is cheaper to produce—is also, in principle, recoverable. Labeling requirements for synthetic content, penalties for undisclosed AI or filter augmentation, or algorithmic demotions of unlabeled synthetic media could partially restore the production cost asymmetry. Regulations in some jurisdictions are beginning to move in this direction. The implementation is difficult and enforcement is imperfect, but the mechanism is identifiable, and the intervention is bounded in scope.

9.3. Irreversible: Epistemic Loss and the Breakdown of Inferential Pathways

The irreversible consequences are best unified under the concept of epistemic loss: not the loss of specific pieces of information, but a reduction in the system's overall capacity to support reliable inference about the relationship between signals and processes.

In a grounded system, observers can use available signals to form justified beliefs about underlying phenomena. They can infer that a widely viewed performance is likely to be engaging, that a highly followed account is likely to produce valued content, that a display of skill reflects sustained effort. These inferences are not infallible, but they are generally reliable enough to support navigation of the environment. In a system characterized by proxy detachment and refer-

ent substitution, these inferential pathways break down. Observers can no longer reliably infer process from appearance, or quality from engagement. The system ceases to function as an epistemic environment in which users can form well-grounded beliefs about what they are encountering.

This loss is cumulative and path-dependent. As users adapt to the degraded environment, they update their expectations and interpretive strategies. They become more skeptical of signals, more reliant on external sources of verification, or more resigned to operating under structural uncertainty. The inferential habits that a grounded signaling environment makes possible—and that it requires, in order to function—begin to atrophy.

9.4. Irreversible: The Recalibration of Perceptual Expectation

The epistemic loss is compounded by a perceptual one. Human aesthetic judgment is calibrated by experience. What seems impressive, beautiful, or skillful is partly a function of what you have been exposed to and what you have come to understand as the normal range of human achievement. Audiences who have grown up watching synthetically enhanced performances—children executing professional dance routines at the tap of a filter, babies articulating adult speech, athletes performing physically impossible feats—develop a perceptual baseline calibrated not to human possibility but to the full capability range of synthesis. Against that baseline, genuine human performance is structurally disadvantaged. A child who dances beautifully, at a level that would have been extraordinary before synthetic enhancement became ubiquitous, may now seem unremarkable. Not because the skill has changed, but because the perceptual frame has been stretched to accommodate the infinite range of synthetic production.

This shift is path-dependent in a specific sense. It is not like a miscalibrated scale that you can reset; it is more like acquiring a new perceptual vocabulary that changes what counts as salient. Once an audi-

ence has internalized a perceptual frame calibrated to synthetic possibility, restoring that frame to one calibrated to human constraint is not a technical problem. It requires a generational change in what is widely consumed, which cannot be achieved by platform policy alone, and which may not occur at all if synthetic content continues to predominate.

9.5. Irreversible: The Erosion of Temporal Integrity as Legible Value

The deepest irreversible consequence is the erosion of *temporal integrity* as a legible cultural value.

Temporal integrity is the condition under which effort, duration, and constraint are constitutive of meaning. It is not merely the observation that good things take time; it is the structural claim that the meaning of a performance, a craft, or an achievement depends on the temporal depth behind it. A master craftsperson's chair is not merely well-made; it is a trace of accumulated practice, of decisions made and revised across decades, of a relationship between maker and material that developed over time. A skilled dancer's performance is not merely precise; it is the visible residue of years of repetition, failure, and incremental refinement. The meaning is not separable from the process. The temporal structure of production is part of what the thing *means*.

When synthetic production eliminates the temporal requirement, it does not merely devalue the genuine skill. It begins to make the connection between process and meaning *illegible*. In a cultural environment saturated with synthetically compressed performances, the claim "this took years of effort" loses its informational force, because the appearance of years of effort can be manufactured trivially and the audience has no reliable way to verify the claim. The informational content of the process claim collapses into the same structural indeterminacy that characterizes the rest of the signaling environment.

This is a deeper loss than the devaluation of craft. It is the erosion of the conceptual apparatus through which temporal integrity is understood and communicated. Once an audience has lost the intuitive sense that appearance and process are connected—that what something looks like reflects what went into producing it—the concept of earned achievement becomes harder to communicate and harder to evaluate. This is not recoverable through platform redesign. It is a shift in the cultural epistemology of value. Once sufficiently widespread, it alters the interpretive framework through which all future content is experienced, and that framework is not accessible to correction through individual decisions or institutional policies aimed at the platform level.

10. What a Structurally Honest Platform Would Require

The conclusion drawn from this analysis need not be merely normative. The question of what platforms *should* value is less interesting than the question of what architectural features a platform would need in order to function as a genuinely representational system— one in which signals track processes, proxies remain grounded in their referents, and the signaling environment preserves its informational content.

The answer is not obvious, and it is not a design proposal. But the structural requirements follow from the diagnosis.

First: incomplete observability of the proxy layer. A system in which all proxy signals are fully visible, continuously updated, and directly tied to distribution incentives creates ideal conditions for their manipulation. Actors can observe the effects of their actions on the signals in real time and adjust their behavior accordingly, converging on strategies that maximize those signals irrespective of the underlying

processes. Goodhart's law, stated as a structural attractor, makes this outcome inevitable when observability is complete. Incomplete observability disrupts this feedback loop. When signals are partially hidden, delayed, aggregated, or decoupled from immediate rewards, the capacity to optimize them directly is reduced. Actors can no longer easily identify which actions produce which signal changes, making systematic manipulation more difficult.

This is not a matter of obscuring information for its own sake. It is a structural necessity. If every proxy is made fully transparent and directly actionable, it becomes a target for optimization and therefore subject to the attractor dynamics described in Section 3. Maintaining some degree of opacity is, paradoxically, a condition for preserving informational content. A structurally honest platform cannot maximize both transparency of proxy signals and their use as optimization targets. It must choose a balance in which signals remain informative to users while being sufficiently insulated from direct manipulation. This requirement follows from the structure of the problem, not from any preference about platform design.

Second: friction in production. The production cost asymmetry between synthetic and authentic content is a direct driver of synthetic dominance of the referent layer. Reducing this asymmetry does not mean making synthetic production harder for its own sake, but it does mean refusing to treat synthetic and authentic content as equivalent for distribution purposes. This requires labeling and verification infrastructure: technically, socially, and legally enforced transparency about what is and is not synthetically produced. Friction is not punitive; it is the restoration of a cost structure that was always implicit in the platform's representational promise.

Third: coupling between signal and process. The deepest requirement is the most difficult. Some mechanism by which the platform can verify that what it is distributing reflects genuine underlying processes is necessary for the restoration of grounding. This is partially

achievable for the proxy layer—provenance tracking, behavioral analysis, slow-burn signal systems that weight long-term engagement over short-term spikes, network topology analysis that detects coordinated exchange structures. For the referent layer, it is harder, but not impossible: verification of genuine human performance is a technically tractable problem in many domains, and as detection methods improve, the feasibility of coupling distribution to verified authenticity increases.

None of these features is costless. They would reduce engagement metrics in the short term. They would require investments in infrastructure that advertising revenue models do not naturally incentivize. They would require the platform to accept a smaller, more semantically coherent signaling environment in exchange for a larger but informationally degraded one. The fact that no major platform has moved decisively in this direction is not evidence that the features are technically unachievable. It is evidence that the extraction model is indifferent or hostile to signal integrity, for exactly the reasons analyzed in Sections 3 and 7.

This means that the choice is not primarily technical or even regulatory. It is structural. Platforms organized around advertising-based extraction will systematically resist the architectural requirements of honest signaling, because honest signaling is not what the extraction model monetizes, and because, under competitive conditions, platforms that maintain honest signaling are outcompeted by those that do not. A structurally honest platform would require a fundamentally different value-capture architecture—one in which the platform's revenue depends on the quality and authenticity of what it surfaces, rather than on the quantity of attention it can aggregate and sell.

Whether such platforms can compete, at scale, with the currently dominant extraction model is an open question. What is not open is the diagnosis: the collapse of proxy integrity is not a malfunction of

the current system. It is the system functioning exactly as its incentive structure predicts.

11. Conclusion: The Defense of Constraint

This essay has argued that what looks like a collection of platform abuses—deepfake filters, follow-for-follow gaming, synthetic child performances, lottery-like monetization promises—is better understood as a set of symptoms of a single structural transformation. The transformation is the migration from a representational signaling system, in which proxies track genuine social processes and appearances refer to genuine underlying activities, to a generative one, in which signals are produced independently of any external referent and the entire visible surface of the platform is increasingly a self-sustaining artifact of its own internal dynamics.

This transformation is driven by three interlocking forces. Optimization pressure on proxy metrics, which by Goodhart's logic is a structural attractor toward detachment and not a contingent failure. The production cost asymmetry between synthetic and authentic content, which floods the referent layer with appearance without process through temporal compression. And the extraction architecture of advertising-based platforms, which makes the system economically robust to epistemic degradation and selects, under competitive conditions, for platforms that exploit rather than resist it.

The defense of temporal integrity is, in this light, not a sentimental attachment to artisanal production or a conservative resistance to technological change. It is a structural claim about the conditions under which meaning is possible. Meaning—in performances, in crafts, in social relationships, in the signals through which we navigate a shared world—depends on the connection between surface and depth, between appearance and process, between the visible

trace and the history that produced it. Constraint is not opposed to meaning; it is constitutive of it. When that connection is severed at scale, the signals through which we try to identify value, guide attention, and recognize genuine human achievement become unreliable tools. We are left navigating by maps that no longer represent any territory.

The constraint that technological production removes—the constraint of time, effort, and accumulated practice—was never merely an obstacle to be overcome. It was the condition under which what was produced carried the meaning that made it worth producing. A platform that treats the elimination of that constraint as a feature is not offering a more efficient route to the same destination. It is offering a route to a destination where the destination no longer exists.

What is required is not nostalgia, but architecture. Honesty in a platform is not a moral disposition. It is a structural property: the property of a system whose signals are caused by the phenomena they claim to track, whose proxies are constrained by the processes they represent, and whose observability is calibrated to preserve rather than destroy the grounding relationship. Building such a system is harder and less profitable than building the kind we have. That is an exact description of why we have the kind we have.

References

[Bourdieu(1979)] Bourdieu, P. (1979). *La Distinction: Critique sociale du jugement*. Paris: Les Éditions de Minuit. [English edition: *Distinction: A Social Critique of the Judgement of Taste*, trans. R. Nice. Cambridge, MA: Harvard University Press, 1984.]

[Campbell(1979)] Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67–90.

[Debord(1967)] Debord, G. (1967). *La Société du spectacle*. Paris: Buchet-Chastel. [English edition: *The Society of the Spectacle*, trans. D. Nicholson-Smith. New York: Zone Books, 1994.]

[Goodhart(1975)] Goodhart, C. A. E. (1975). Problems of monetary management: The UK experience. *Papers in Monetary Economics*, 1. Sydney: Reserve Bank of Australia.

[Pasquale(2015)] Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.

[Sennett(2008)] Sennett, R. (2008). *The Craftsman*. New Haven: Yale University Press.

[Skinner(1938)] Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century-Crofts.

[Stiegler(1998)] Stiegler, B. (1998). *Technics and Time, 1: The Fault of Epimetheus*, trans. R. Beardsworth and G. Collins. Stanford:

Stanford University Press. [Original: *La Technique et le temps, vol. 1: La Faute d'Épiméthée*. Paris: Galée, 1994.]

[Strathern(1997)] Strathern, M. (1997). 'Improving ratings': Audit in the British university system. *European Journal of Education*, 32(3), 305–321.

[Wu(2016)] Wu, T. (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. New York: Knopf.

[Zuboff(2019)] Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.