Beyond Optimal Foraging: Representation Repair and the Structural Consequences of AI-Augmented Research

Flyxion
Independent Researcher

August 2026

Abstract

Large language models are often evaluated in terms of their reliability, accuracy, or tendency to fabricate information. Duede, Gross, Crockett, and Bergstrom instead ask a narrower and more fundamental question: assuming LLMs are perfectly reliable, how does reducing different forms of research labor alter the allocation of scientific effort? Using an optimal-foraging model, they show that reducing discovery costs makes researchers more selective, reducing publication costs makes them less selective, and increasing the productivity of substantive development leads to deeper individual projects. This essay argues that these comparative statics are correct under the model's central ontological assumption that scientific value is a fixed scalar revealed at the conclusion of discovery. It then relaxes that assumption by replacing static project value with a representation-dependent value function, $v_t(p \mid \mathcal{M}_t)$, whose assessment evolves with the scientific community's representational apparatus. This single substitution transforms the optimal-stopping problem into a diagnosis-and-repair problem while leaving the optimization machinery itself unchanged. Rerunning the original comparative statics under the generalized ontology shows that one result survives unchanged, AI that directly accelerates substantive representation repair increases project depth, whereas the remaining results require reinterpretation because they operate on currently representable rather than intrinsic value. The paper further argues that publication is more appropriately understood as insertion into a persistent knowledge ledger than as the terminal stage of an individual research pipeline, shifting the limiting resource from production to integration and repair. The resulting framework recovers the original model as the special case in which representational repair is absent and scientific value is fully determined at the conclusion of discovery, while providing a broader account of how AI alters the long-term viability of scientific inquiry.

I The Baseline Model

I.1 What Duede et al. actually model

Duede et al. set out to answer a question that is usually asked badly. Most discussions of LLMs in science worry about whether the models are reliable: whether they hallucinate citations, fabricate results, or subtly distort the papers they help produce. These are real concerns, but they leave open a harder question, and it is the question this paper asks instead. Suppose the models worked exactly as advertised. Suppose they saved researchers real time, introduced no errors, and cost nothing to use. Would science still change, and if so, how?

Their answer comes from a source outside economics proper: optimal foraging theory, the branch of behavioral ecology that asks how an animal should allocate time between searching for food patches and exploiting the patches it finds. The translation to science is direct. A researcher does not know in advance how valuable a line of inquiry will turn out to be. They commit a fixed period of time, the discovery phase of length s, to finding out: forming a hypothesis, running preliminary experiments, evaluating early data. At the end of that phase, the project's value v becomes known, and the researcher faces a decision. They can abandon the project and start over, or they can develop it further, investing additional time $t = a + b$, where a is the minimum time needed to bring any publishable finding to the point of submission, and b is discretionary time spent making the project better than it strictly needs to be: additional experiments, more careful analysis, more thorough prose.

Time invested in development returns a benefit $v \cdot u(t)$, where u is a saturating function of how thoroughly the project has been developed, increasing but with diminishing returns. The researcher's problem, repeated indefinitely across a career, is to choose a policy: a rule specifying how much time to invest in a project of value v, so as to maximize the long-run rate at which benefit accrues. This is a classical renewal-reward problem, and it has a clean solution. There is a single opportunity cost of time, λ, set endogenously by the researcher's overall rate of return, against which every project is measured. Any project the researcher chooses to develop will be developed until its marginal value falls to λ. There is a threshold value v_0, below which projects are abandoned and above which they are developed. And projects above the threshold are developed more thoroughly the higher their value.

Against this structure, the paper asks what happens when an LLM shortens one of three quantities: the discovery period s, the minimum development time a, or the rate at which discretionary development b produces benefit. Each has a distinct, provable effect. Shortening s raises λ and raises v_0: researchers become choosier, because it is now cheaper to return to the drawing board and try again. Shortening a also raises λ, but lowers v_0: even though time is now more valuable, the bar for what counts as publishable has itself dropped, and this second effect dominates. In both cases, whatever gets published gets developed less thoroughly, because the increased opportunity cost of time makes further polish more expensive relative to moving on. Only when the LLM accelerates discretionary development itself, making b more productive rather than making a cheaper, does thoroughness unambiguously increase.

The paper's own summary of this is worth preserving in full, because it is the sentence the rest of this essay is written against: LLMs make researchers' time more valuable, and

this is precisely why researchers become less willing to refine already-publishable results further. There is something more useful they could be doing with that time. The hope that saved time becomes reclaimed time for depth turns out, within this model, to be mostly false.

I.2 The assumption that makes the mathematics work

This result is elegant, and it is worth being honest about why. The optimal-stopping formalism that makes the model tractable, that lets it deliver a unique threshold v_0 and a family of provable comparative statics rather than a qualitative gesture, depends on treating the discovery phase as an act of measurement. Before discovery, v is unknown. After discovery, v is known, fixed, and available for the researcher to optimize against. Discovery is a window that opens once, reveals a number, and closes. Everything downstream, the decision to abandon or develop, the choice of how much time to invest, is optimization against that number.

This is what allows the paper's central mechanism, the rising opportunity cost of time, to do all the explanatory work. Once v is fixed and known, the only thing left for an LLM to change is the price of researcher time relative to that fixed quantity. Every one of the paper's results is, in this sense, a restatement of the same fact under a different name: LLMs change λ, and everything else, the threshold, the thoroughness, the publication rate, falls out as a function of λ and the structure of $u(t)$.

The assumption is not incidental to the model. It is what makes the model a model, in the specific sense of being solvable in closed form with propositions and proofs rather than merely gestured at. And it is a reasonable assumption to make for the purpose the authors set themselves, which is to isolate the pure economic mechanism, uncontaminated by questions of model reliability, and show that it alone is sufficient to produce non-obvious, occasionally unwelcome, consequences. Stripping away the ontology of scientific value in order to see the economics clearly is a legitimate and, in this case, successful move.

But it is still a choice, and it is worth stating plainly what has been chosen. The model treats a scientific project as a patch with a hidden but determinate scalar value, discovered rather than constructed, measured rather than negotiated with the instruments and concepts available at the time of measurement. The remainder of this essay develops what happens when that determinacy is relaxed: when the number that discovery reveals is not a fixed fact about the project waiting to be uncovered, but a function of the representational apparatus the researcher currently has for seeing it.

II The Ontological Substitution

II.1 From v to $v_t(p \mid \mathcal{M}_t)$

Nothing in the optimal-foraging machinery of Section I requires that a project's value be intrinsic. The mathematics requires only that some quantity exist, at the moment discovery concludes, on which an optimal policy can condition. Proposition 1 in Duede et al. is a statement about the existence and characterization of an optimal policy given a payoff

structure; it does not depend on where that payoff structure comes from. This essay therefore does not relax the optimization. It relaxes the ontology supplying the optimization's value variable.

Where the original model treats v as a scalar attached to a project p, fixed and revealed in full at the close of the discovery phase, this essay treats a project's assessed value as conditional on the representational manifold $\mathcal{M}_t$ available to the researcher, or to the field, at time t:

$$v_t(p \mid \mathcal{M}_t).$$

$\mathcal{M}_t$ stands for the working vocabulary, instruments, accepted methods, inferential norms, and conceptual categories through which a project's contribution can currently be recognized and expressed. It is not hidden in the way v was hidden before discovery in the original model. It is public, shared, slowly changing, and, crucially, itself a target of scientific labor rather than a fixed backdrop against which labor occurs.

Two distinct claims follow from this substitution, and it is worth keeping them separate rather than letting one slide into the other.

The first is epistemic. A low measured value at time t stops being unambiguous evidence about the project and becomes ambiguous between two explanations that present identically to the researcher. It might mean the project genuinely has little to offer. Or it might mean that $\mathcal{M}_t$ currently lacks the distinctions required to register what the project offers. Discovery, on this reading, does not simply reveal v; it applies whatever $\mathcal{M}_t$ the researcher currently has to the project and reports what that instrument is capable of seeing. This is not a claim that every low-value project is secretly valuable, and it is not a claim that most abandoned projects are wrongly abandoned. Most are abandoned correctly. The claim is narrower: that a single measured v, on its own, cannot distinguish a project that is genuinely poor from a project whose poor showing is an artifact of the current representational apparatus, and the original model has no mechanism for making that distinction because it was never asked to.

The second claim is ontological, and it is the stronger of the two. It is not merely that we cannot always tell which explanation applies. It is that the quantity being measured is itself representation-dependent, so that a later measurement can differ from an earlier one without either measurement having been in error:

$$v_{t+1}(p \mid \mathcal{M}_{t+1}) \neq v_t(p \mid \mathcal{M}_t)$$

need not be a correction of a mistake, in the way a more careful re-measurement corrects an earlier estimate of a fixed quantity. It can instead be the signature of a repaired representation applied to an unchanged project. The two claims reinforce one another, the epistemic claim motivates taking the ontological one seriously, but they are logically distinct, and only the second requires abandoning the static-value ontology rather than merely being more careful within it.

II.2 The repair warrant

If value is representation-relative, the continuation decision has to change accordingly. In Duede et al.'s model, the decision rule is a comparison between the surplus available from

developing the project and the opportunity cost of the time that would take:

$$v \cdot u(t) - \lambda t.$$

The project is developed if some t makes this positive, abandoned otherwise. This rule is correct once v is fixed and known. It is not obviously correct once v is itself something that a further act of labor, a repair of $\mathcal{M}_t$ rather than of the project, could change. The relevant question is no longer only whether developing the project as currently represented clears the opportunity-cost bar. It is whether some admissible repair to the representation itself, applied to the project, would clear a bar that the unrepaired project cannot.

This essay proposes the following as the corrected continuation rule, standing in the place occupied by the threshold comparison in the original model. The comparison can no longer be stated in terms of $v \cdot u(t) - \lambda t$ alone, since that expression presupposes a fixed value being weighed against a fixed cost of time; once value is representation-relative, the object being compared has to be a functional over continuations rather than a scalar over a single project. Let J denote this abstract continuation functional, introduced here specifically because the repaired setting no longer admits a purely scalar objective based on immediate project value, and understood as reducing to $v \cdot u(t) - \lambda t$ in the special case where $\mathcal{M}_t$ is held fixed. Let $\mathcal{A}_t(p)$ denote the set of admissible repairs available to the researcher at time t: revisions to method, vocabulary, framing, or instrumentation that could plausibly be applied to p without abandoning the standards that make the repair recognizable as scientific work rather than post-hoc rescue. Let $\varnothing$ denote the null intervention of leaving the project and the representation as they are. Then a project should be abandoned only when

$$\sup_{R \in \mathcal{A}_t(p)} \mathbb{E}\left[J\big(R(p), \mathcal{M}_{t+1}\big)\right] \leq J(\varnothing, \mathcal{M}_t).$$

In words: abandonment is warranted when no admissible repair, considered honestly and in expectation, does better than doing nothing. This is a genuinely different claim from the original threshold rule, and it is worth being precise about how. It does not forbid abandonment. It does not claim that most abandoned projects are wrongly abandoned. What it denies is that a single measured v, taken at face value immediately after a fixed-length discovery phase, is sufficient grounds for the decision. It requires that the decision also consider the supremum over admissible repairs, which is to say, it requires that fault identification, a project currently looks unpromising, be kept distinct from repair warrant, no available intervention could make it look otherwise. Duede et al.'s model collapses this distinction by construction, because collapsing it is exactly what makes discovery a one-shot measurement rather than an ongoing diagnostic process. The repair warrant reintroduces the distinction at the point where the original model most needed it and had the least room for it.

Three clarifications matter here before the consequences of this substitution are worked out in Section III.

First, $\mathcal{A}_t(p)$ is not unbounded. An admissible repair is constrained by the same standards that make a null result a null result rather than an excuse: it must be a recognizable move within the discipline's own methodological commitments, not an unconstrained license to keep any project alive indefinitely by redefining success after the fact. The repair warrant is a higher bar to clear for abandonment, not a lower bar to clear for publication.

Second, the supremum in the repair-warrant inequality is not free. Evaluating it costs time and attention, the same scarce resources that set λ in the original model. The two frameworks are not in tension on this point; they share the same currency. What changes is what that currency is being spent to determine.

Third, the asymmetry between the two branches of the inequality is deliberate and should be stated rather than left implicit. The repaired branch is evaluated at $\mathcal{M}_{t+1}$, the manifold as it would stand after the repair; the null branch is evaluated at $\mathcal{M}_t$, the manifold as it stands now. This is appropriate under the reading that $\varnothing$ means literally declining to invest further labor in p, not declining to benefit from whatever background progress the field makes independently of this decision. If the field's representational apparatus improves for reasons unrelated to p, both branches would in principle be re-evaluated at that later, common manifold, and the inequality is a comparison holding that background rate of change fixed. The warrant, in other words, asks specifically whether investing in p is worth more than not investing in p, not whether the future will be better than the present regardless of what is done.

The repair warrant changes only one decision rule. But because that decision rule mediates every transition from discovery to development, its consequences propagate through the rest of the model. The next section returns to the assumptions that were previously implicit in Section I and asks how each must be reformulated once value is allowed to evolve through representational repair.

III Consequences of the Substitution

Section II changed one thing: the ontology supplying the value variable that Section I's optimization conditions on. Nothing here alters the mechanics of that optimization. What follows are five places where the original model's implicit assumptions, adequate under a static v, cease to be sufficient once v is allowed to depend on $\mathcal{M}_t$. Each is presented in the same order: the baseline assumption, why it stops being sufficient, the single piece of machinery imported to address it, and the immediate consequence. No new formal apparatus is introduced beyond what is needed to state the consequence; entropy, ledgers, ecology, and institutional design are postponed to where they are actually needed rather than folded in here.

III.1 Value Is Fixed $\rightarrow$ Representation-Relative Value

Baseline assumption. A project's value v is a single scalar, constant across the entire development phase.

Why it is no longer sufficient. Once value is allowed to depend on $\mathcal{M}_t$, as established in Section II, two projects with identical observed value at time t need not occupy the same position with respect to future representational change. One may be legible under any reasonable future $\mathcal{M}$; the other may be legible only under a manifold that has not yet been repaired.

Imported machinery. None beyond the representational manifold itself, already introduced in Section II.

Immediate consequence. Observed value equality does not imply continuation-potential equality. This single point is the seed from which the remaining four substitutions grow; each of them is a different aspect of what follows from not being able to read continuation potential directly off of v_t.

III.2 Discovery Reveals Value → Diagnosis Precedes Repair

Baseline assumption. The discovery phase is an act of measurement. It reveals v; the researcher's task afterward is to act on what has been revealed.

Why it is no longer sufficient. If $v_t(p \mid \mathcal{M}_t)$ can differ from $v_{t+1}(p \mid \mathcal{M}_{t+1})$ without the project having changed, then a low reading at t is not self-interpreting. It requires asking which of the two explanations from Section 2.1 applies before deciding what the reading warrants.

Imported machinery. The distinction, developed elsewhere as Diagnostic Coordinates, between fault identification and repair warrant. Fault identification is the observation that a project currently scores low. Repair warrant is the further, separate finding that no admissible repair could raise that score. The original model has only the first; the repair-warrant inequality of Section 2.2 supplies the second.

Immediate consequence. Observation alone cannot license abandonment. The continuation decision now requires evaluating the repair-warrant inequality rather than comparing v directly to a threshold.

III.3 Projects Are Independent → Knowledge Graph Dynamics

Baseline assumption. Projects are independent patches. Developing or abandoning one has no bearing on the value of any other.

Why it is no longer sufficient. If value is relative to $\mathcal{M}_t$, and $\mathcal{M}_t$ is shared across projects, then a repair applied in the course of developing one project can change $\mathcal{M}_t$ for all of them. Project value becomes relational rather than independent:

$$v_t(p) \to v_t(p, G_t),$$

where G_t is the field's current knowledge graph, the set of projects together with the relations, dependencies, and shared representational commitments among them.

Imported machinery. Relational meaning is exactly the machinery this assumption requires, because the assumption being relaxed is independence itself, and independence is a claim about how one project's standing depends, or fails to depend, on others. The relevant resource is therefore structural semantics: the relational account of meaning under which a contribution's significance is constituted in part by the transformations it licenses across the surrounding network of claims, not solely by its own propositional content.

Immediate consequence. Developing or abandoning a project can raise or lower the assessed value of other projects through $\mathcal{M}_t$, independent of any direct interaction between them. The independent-patch assumption, adequate for a model of isolated foraging, is not adequate once the environment being foraged is partly constituted by the forager's own activity.

III.4 Publication Is Terminal → Ledger Entry

Baseline assumption. Publication is the terminal event of the pipeline. Once a project is published, its costs and benefits to the researcher have been fully accounted for by a, b, and the resulting $v \cdot u(t)$.

Why it is no longer sufficient. If publication inserts a claim into a shared representational environment that other projects' values depend on, as established in III.3, then publication is not terminal with respect to the field even if it is terminal with respect to the researcher's own time accounting.

Imported machinery. The Monotonic Ledger: publication as a permanent, non-deletable insertion into the field's historical record, which future work must index, compare, reproduce, or reconcile with. The total cost of a publication is distributed rather than borne entirely by its author:

$$C_{\text{total}}(p) = C_{\text{author}}(p) + C_{\text{review}}(p) + C_{\text{integration}}(p) + C_{\text{repair}}(p).$$

Immediate consequence. An LLM that reduces C_{author}, which is what shortening a amounts to in the original model, does not thereby reduce C_{total}. It can leave the remaining terms unchanged, in which case the field's obligation to integrate and reconcile new claims grows at the pre-LLM rate even as the rate of claim production rises.

III.5 Optimization Is the Objective → Viability Is the Constraint

Baseline assumption. The researcher's problem is to maximize a payoff, $\max_\tau \Lambda(\tau)$, over policies τ.

Why it is no longer sufficient. Maximization presupposes that the space being optimized over remains one in which the field can continue to operate: that claims entering the ledger under III.4 can still be integrated, that the knowledge graph of III.3 remains repairable, that diagnosis under III.2 remains possible. None of this is guaranteed by maximizing $\Lambda(\tau)$ alone; a policy can be individually optimal for every researcher following it while the aggregate system it produces drifts outside the region in which those guarantees hold.

Imported machinery. The distinction, prior to any objective function, between optimization and viability: the requirement that the system remain within an admissible region $\mathcal{V}$, $G_t \in \mathcal{V}$, before asking what is optimal within it. Only once this ordering is fixed does it make sense to introduce a long-run objective $J(\pi)$ that weighs immediate value against the field's repair capacity, its reachable representational states, and its accumulated unresolved distinctions; that objective is not developed formally here, since doing so is not needed to state the present consequence.

Immediate consequence. A policy that is optimal under the original model's $\Lambda(\tau)$ can be non-viable in the sense that it depletes the field's capacity to integrate what it produces, without this showing up anywhere in $\Lambda(\tau)$ itself. Maximization and viability are not the same requirement, and the original model, by construction, can only see the first.

The five substitutions above do not alter the mechanics of the optimization problem laid out in Section I. They alter the interpretation of its variables and the constraints under which optimizing them remains meaningful. The remaining question is therefore straightforward:

if Duede et al.'s own comparative statics, the three automation regimes of Section 1.1, are re-executed under this generalized ontology, which conclusions remain unchanged, which split into multiple cases, and which require reinterpretation? Section IV returns to their three principal results and answers exactly that question.

IV Comparative Statics Revisited

Section III established that the five substitutions leave the optimization machinery of Section I untouched. They change what its variables mean and what conditions must hold for optimizing them to remain meaningful, but they do not change the mathematics of the optimal-stopping problem itself. If that claim is correct, it should be possible to return to Duede et al.'s own three results, the effects of cheaper discovery, cheaper publication, and cheaper discretionary development, and show precisely where each is preserved, where each requires reinterpretation, and why. This section does exactly that, taking their three propositions in the order presented in Section 1.1.

IV.1 Discovery Becomes Cheaper

Duede et al.'s result is unambiguous. When the discovery period s decreases, the opportunity cost of researcher time λ rises, because returning to the discovery phase and trying again becomes cheaper relative to persisting with any given project. A higher λ raises the threshold v_0: researchers rationally abandon more projects, and abandon them earlier, than they would under a longer discovery period.

Nothing about this derivation is disturbed by the substitution in Section II. The threshold still rises. The opportunity cost still rises. Researchers still become more selective, and they are still behaving rationally in doing so. What changes is not whether the threshold rises but what the threshold is now understood to be partitioning. In the original model, v_0 separates projects by intrinsic value: everything above the line is genuinely worth developing, everything below genuinely is not. Under $v_t(p \mid \mathcal{M}_t)$, the threshold instead separates projects by currently representable value, and Section 2.1 already established that this is a distinct quantity from intrinsic value whenever $\mathcal{M}_t$ is incomplete with respect to p.

This produces a specific structural consequence rather than a vague one. A project can fall below v_0 for either of the two reasons distinguished in Section 2.1: because it is genuinely unpromising, or because the current representational apparatus fails to expose whatever continuation potential it has. Cheaper discovery raises v_0 and shortens the window in which that potential might otherwise have been noticed before the decision is made. The result is not that the threshold becomes irrational. It remains locally optimal with respect to the information the researcher has. The result is that cheaper discovery increases the rate of premature representational pruning: the systematic loss of projects whose low showing reflects the current $\mathcal{M}_t$ rather than the project itself, at exactly the moment when the incentive to look past that showing is weakest.

IV.2 Publication Becomes Cheaper

Duede et al.'s second result is the mirror image of the first in its mechanics but not in its direction. Reducing the minimum development time a also raises λ, since it too makes researcher time more valuable. But it simultaneously lowers the bar a project must clear to be worth publishing at all, and at the margin this second effect dominates the first: the net result is a lower threshold v_0, more projects published, each developed less thoroughly than before.

The generalized model leaves this immediate prediction intact. Publication volume rises; that conclusion does not depend on whether value is static or representation-relative, since it follows from the relative sizes of two effects on λ and v_0 that the substitution does not touch. What the substitution reinterprets is not whether more gets published but what publication is. In Section III.4, publication stopped being the terminal event that closes the researcher's time-accounting and became instead an insertion into a persistent, shared ledger, one that imposes review, integration, and repair costs on the field regardless of how cheap it was for the author to produce. The relevant cost of a publication is therefore no longer well approximated by $a + b$ alone; it is $C_{\text{total}} = C_{\text{author}} + C_{\text{review}} + C_{\text{integration}} + C_{\text{repair}}$, and an LLM that shortens a has only been shown to shrink the first term.

The bottleneck consequently shifts rather than disappears. Publication itself stops being the limiting resource, since it has just become cheap. Integration and repair become the limiting resource in its place, and nothing in Duede et al.'s model, which is a model of the researcher's own time, has anything to say about whether that resource keeps pace. This is not a correction of their result. It is an observation that their result describes only C_{author}, correctly, while leaving the remaining terms of C_{total} outside the frame.

IV.3 Discretionary Development Becomes Cheaper

The third regime is the one in which the two frameworks agree without qualification, and the agreement is worth stating plainly rather than passing over quickly, because it is the clearest evidence that the generalized model is not constructed simply to disagree with the original one. Duede et al. show that when discretionary development itself becomes more productive, when the function converting time b into benefit is accelerated rather than merely the fixed cost a being reduced, projects are developed more thoroughly. Depth increases, and it increases unambiguously, because the acceleration raises the real value of further development by more than it raises the opportunity cost of the time it takes.

This conclusion survives the substitution of Section II without modification, and it survives for a specific reason rather than by coincidence. Accelerating b is precisely the regime in which an LLM is assisting representation repair itself, proof verification, invariant discovery, boundary testing, consistency checking, theoretical synthesis, rather than merely reducing the transaction costs surrounding a fixed representation. If what is being accelerated is the process by which a project's treatment under $\mathcal{M}_t$ is actually improved, then both the original model and the generalized one predict the same outcome, because in this one regime there is no gap between the two ontologies for the substitution to act on. The generalized framework's distinction is not between AI that helps and AI that does not. It is between AI that reduces the friction surrounding science, which is what accelerating s or a does, and AI that directly expands what the current representation can capture, which is

what accelerating b in this sense does. Only the latter is guaranteed, by both accounts, to deepen the resulting work.

IV.4 Entropy Across Three Scales

One further observation follows from the three results above, and it is included here because it resolves an apparent tension rather than because it is a new theoretical claim in its own right. Sections 4.1 through 4.3 together predict that individual researchers become more efficient, in the specific sense that navigating projects, discarding unpromising ones, and producing clean manuscripts all become locally easier, while the field's capacity to integrate what is produced does not automatically keep pace.

This is not a paradox once the claim is stated at the correct scale. At the level of the individual researcher, an LLM reduces operational entropy: projects are easier to inspect, compare, and move through, so $\Delta S_{\text{operator}} < 0$. At the level of the individual manuscript, it reduces artifact entropy: prose, structure, and citation are more internally consistent, so $\Delta S_{\text{artifact}} < 0$. Neither of these figures says anything about what happens at the level of the ledger established in Section III.4. There, the relevant quantity is the rate at which claims are produced relative to the rate at which they are integrated, reconciled, or corrected, and Section 4.2 already showed that the first can rise while the field's capacity to do the second is left untouched by the same technology. This gives $\Delta S_{\text{ledger}} > 0$ as a live possibility even while the first two quantities fall. Individual clarity and collective disorder are not in tension; they are measurements taken at different scales of the same system, and nothing prevents them from moving in opposite directions at once.

The comparative statics therefore divide into three classes, not two. One result, cheaper discretionary development, survives unchanged, because it concerns the economics of representation repair itself rather than the economics of friction surrounding it. Two results, cheaper discovery and cheaper publication, survive only after reinterpretation, because they concern variables, the threshold v_0 and the act of publication, whose ontological status changes under the substitution even though their formal behavior does not. None of the three require any modification to the optimization machinery of Section I. The sole alteration, throughout, is that optimization is now performed over values conditional on an evolving representation rather than over fixed scalar quantities. Section V returns to what this implies for how scientific institutions should treat the difference between the two kinds of acceleration this section has distinguished.

V Discussion

V.1 The Scope of the Generalization

It is worth stating plainly what this essay has and has not argued, since the distance traveled from Section I to Section IV can obscure how conservative the underlying method has been. This essay has not argued that Duede et al.'s comparative statics are incorrect. It has not argued that opportunity cost ceases to organize scientific labor, or that the renewal-reward

structure underlying their optimal policy fails to describe rational behavior. The optimization problem is left standing exactly as Section I presented it. The sole alteration has been to relax the assumption that a project's value is an intrinsic scalar, revealed once and for all at the close of discovery, and to replace it with a value conditional on the representational manifold available at the time it is assessed. Under that substitution the mathematics continues to describe rational behavior; what changes is that the objects the behavior is defined over acquire dynamics of their own. This is why Section IV's results split the way they did. Two of Duede et al.'s three conclusions required reinterpretation rather than replacement, and the third required nothing at all.

V.2 Scientific Institutions Under the Generalized Model

The institutional implications of this reinterpretation can be stated compactly rather than developed here. If publication is a ledger insertion, as Section III.4 argued, rather than a terminal event, then the institutions that surround publication, peer review, archival systems, citation practice, and the AI-assisted workflows now entering all three, are better understood as mechanisms for maintaining the field's capacity for future repair than as filters that certify a finished product. Negative results, partially developed theories, abandoned methodologies, and unresolved anomalies acquire value under this description not because they document what happened, but because preserving their continuation state keeps a future repair of $\mathcal{M}_t$ available that discarding them would foreclose. Whether existing institutions already approximate this, or could be redesigned to, is a separate empirical and institutional question that this essay does not attempt to settle. What the argument above establishes is only why the question arises at all once representation is admitted as a variable rather than held fixed.

V.3 Conclusion

Duede et al. ask what happens when scientific labor becomes cheaper while the objects of that labor, the projects and their values, remain unchanged. Their answer is mathematically rigorous and, as Section I argued, methodologically appropriate to the question they set themselves: it isolates the pure economic mechanism by which opportunity cost redistributes researcher effort, and it does so without needing to say anything about how scientific value is constituted in the first place.

This essay has asked a complementary question. What happens when the objects of scientific labor, the representations through which a project's value becomes visible, are themselves allowed to evolve? Under that extension, developed across Sections II through IV, scientific success is no longer characterized solely by the efficient allocation of effort against a fixed opportunity cost. It is characterized by the preservation of the conditions that permit continued representation repair, conditions that the original model, by construction, has no variable capable of representing. Artificial intelligence, seen this way, is not merely a labor-augmenting technology in the sense Duede et al. give the term. It is a technology capable of altering the geometry of scientific continuation itself, depending on whether it accelerates the friction surrounding a fixed representation or the repair of the representation.

The relationship between the two accounts is not one of correction. Duede et al.'s model is recovered exactly, as the limiting case in which representational repair is absent and scientific value is fully determined at the conclusion of discovery. The present essay has argued only that this limiting case, elegant and correct on its own terms, is not the general one. Optimal foraging remains an appropriate description of scientific labor once representations are fixed. Beyond optimal foraging lies the problem of how those representations themselves are repaired, and it is only at that level that the structural consequences of AI-augmented research become visible.

References

[1] Duede, E., Gross, K., Crockett, M. J., & Bergstrom, C. T. (2026). *The unintended consequences of large language models as a labor-augmenting technology in science*. arXiv:2607.17397 [physics.soc-ph].

[2] Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2), 129–136.

[3] Ross, S. M. (2014). *Introduction to Probability Models* (11th ed.), Chapter 7: Renewal Theory and Its Applications. Amsterdam: Academic Press.

[4] Flyxion. *Diagnostic Coordinates and the Time of Repair*. Manuscript in progress.

[5] Flyxion. *The Geometry of Correction*. Manuscript in progress.

[6] Flyxion. *What Survives Reconstruction*. Manuscript, 118 pp.

[7] Flyxion. *Building Forth from Spherepop Primitives: Stacks, Continuations, and Recursive Control*. Essay, August 2026. `standardgalactic.github.io/memory/building-forth.pdf`

[8] Flyxion. *Structural Semantics: A Mathematical Theory of Admissible Representation and Context Preservation*. Research monograph, Version 1.0.

[9] Flyxion. *Persistent Generative Worlds: A Field Theory of Semantic Simulation*. Book manuscript in progress.

[10] Flyxion. *Spherepop: An Operational Calculus of History-Based Computation*. Monograph, 242 pp.