Compressed Truths: Finite-Slot AI Summaries and Costly Verification

Dheer Avashia

July 15, 2026

Preliminary draft. Comments are very welcome.

Abstract

This paper asks when a decision maker who observes a free, truthful compression of a finite evidence corpus will still pay to inspect the underlying evidence. AI review summaries provide the leading application: a consumer sees a finite-slot summary and may then open reviews from the same pool. Because the summary and later evidence share a source, Bayesian learning must be compression-consistent rather than treating them as independent experiments.

A scalar majority benchmark isolates the shared-pool mechanism. Raw-review updates are attenuated relative to a matched independent-summary benchmark, and verification is nonmonotone in signal precision: greater precision can first create and then eliminate review reading. The multidimensional model supplies the main economics. Reviews are sparse across experience attributes and the summary displays only the top-K dimensions by prevalence. Nondisplay then shifts the omitted dimension's coverage distribution downward in likelihood-ratio order, reducing its perceived chance of appearing in another review. Yet fixed display capacity leaves pivotal omitted dimensions unresolved even in arbitrarily large review pools. Finally, prevalence is not consumer relevance: it approximates the verification-minimizing slot allocation only when relevance per unit of prevalence is bounded, and otherwise has no positive uniform guarantee. Majority and top-K selection deliver these transparent closed forms; the posterior and Bellman construction remains exact under any known finite truthful compression kernel.

Contents

1 Introduction

Consumers increasingly encounter an AI-generated answer before deciding whether to inspect the sources from which it was produced. On product pages, the answer is a review summary and the sources are individual reviews; in AI-assisted web search, the answer is an overview and the sources are linked pages. This new information layer is behaviorally consequential. In a randomized field experiment, Agarwal and Sen (2026) find that, when a Google AI Overview appears, it reduces outbound organic clicks by 38% and raises the probability of a zero-click search by 33%. In product-information acquisition models such as Branco, Sun, and Villas-Boas (2012), a consumer pays to learn attributes and stops when the value of further information falls below its cost. AI changes that problem by making an initial synthesis nearly free while leaving source-level verification costly.

The central primitive is a truthful finite-output compression of a realized evidence corpus. The summary is neither an independent signal nor an unlimited report: it is generated from the same evidence that the consumer can later inspect at a cost. Hence the posterior must be compression-consistent. After observing the summary, the consumer updates both about the payoff-relevant state and about the raw evidence that remains hidden but must still rationalize the displayed output. The AI system is therefore modeled as an information policy—a known mapping from evidence into a compressed message—rather than as a separate strategic or behavioral agent.

This formulation defines a reusable post-compression verification problem: a state generates a finite source corpus, a compression rule maps that corpus into a negligible-cost message, and a Bayesian decision maker chooses whether to unpack sources sequentially before acting. Existing work studies its ingredients separately. Consumers search after seeing partially revealed product information (Gu and Wang, 2022); receivers acquire costly information after a sender's message (Wei, 2021; Matysková and Montes, 2023); platforms aggregate ratings (Dai et al., 2018); and individual reviews affect demand even after an aggregate rating is visible (Vana and Lambrecht, 2021). To my knowledge, the distinctive conjunction has not been modeled in economics: a free coarsening of a realized finite corpus followed by costly inspection of that same corpus, with every subsequent observation constrained by the coarsening already seen. The contribution is this economic object and the restrictions it places on rational search and learning, not a claim that costly information acquisition or coarse communication is itself new.

The main contribution is multidimensional. Product listings reveal search attributes such as price, size, and number of ports. Reviews speak to experience attributes: whether charging slows when a USB-C hub is fully loaded, whether a laptop keyboard is comfortable, or whether a restaurant is reliable for dietary restrictions. A finite-slot summary can display only a subset of these dimensions and may fold several narrow concerns into one broad cell. Even when the platform is fully truthful, omitted or folded dimensions can create residual verification value for consumers whose tastes load on them.

The distinction is especially sharp when a summary reports exact sentiment counts. In the binary benchmark, the total favorable count is sufficient for the latent scalar quality state, so opening a passively sampled constituent sign cannot add information once that count is known. Exact counts therefore make displayed singleton dimensions a finite-N null: residual review reading must be supported by omitted dimensions, folds, richer within-sign content, or a departure from the maintained sampling model.

The analysis delivers three findings. First, using the same finite pool for compression and verification changes Bayesian updating. Theorem 1 shows that an opened review receives less log-likelihood weight than in a matched independent-summary experiment, because the residual pool must continue to rationalize the summary. The summary's informational content is progressively depleted as reviews are opened. Proposition 1 and Corollary 1 then show that verification need not fall monotonically with precision: a more informative summary can move the decision toward the stopping margin before eventually becoming decisive.

Second, finite slots create a distinct multidimensional channel. Theorem 2 shows that top-K prevalence selection makes omission self-concealing. Nondisplay is not neutral: it shifts the omitted topic's coverage distribution downward and therefore reduces the perceived probability that another review will discuss it. Theorem 5 establishes that this channel does not vanish with evidence abundance. Large pools resolve displayed majority cells, but a pivotal omitted dimension with positive topic-arrival probability can continue to support costly verification.

Third, the platform's prevalence ranking and the consumer's relevance ranking need not agree. Theorem 4 gives a sharp approximation guarantee when residual verification value per unit of prevalence is bounded across dimensions, and proves that no positive uniform guarantee survives without that alignment restriction. The design implication is deliberately narrower than a welfare theorem: if the goal is to reduce costly follow-up, slots should be ranked by taste-weighted residual verification value rather than raw frequency alone.

Majority and top-K selection are transparent analytical benchmarks, not claims about a particular deployed interface. For any known finite truthful kernel Q_ϕ, the model integrates over hidden corpus realizations and yields an exact posterior, predictive transition, and within-product Bellman problem. Section 6 states this portability result and shows how a measured rule replaces the benchmark likelihood. Identification of that rule and of search costs, the across-product dynamic problem, and experimental validation are left to companion projects.

1.1 Related Literature

The paper sits at the intersection of consumer search, costly product-information acquisition, and information design. From canonical search models it inherits the stop-or-continue logic and the comparison to an outside opportunity (Stigler, 1961; McCall, 1970; Weitzman, 1979); from equilibrium and ordered-search models it inherits the idea that search behavior is shaped by the value of the next feasible action (Diamond, 1971; Wolinsky, 1986; Stahl, 1989; Olszewski and Weber, 2015; Armstrong, 2017; Choi, Dai, and Kim, 2018; Greminger, 2022). The paper is closest, however, to product-information acquisition models, in which consumers pay to learn about product attributes before choosing (Nelson, 1970; Branco, Sun, and Villas-Boas, 2012; Ke, Shen, and Villas-Boas, 2016; Ke and Villas-Boas, 2019; Gibbard, 2022). I use the same basic demand-side ingredients: observed product characteristics and price are known before costly learning; latent experience attributes enter utility through posterior expected quality; and the consumer compares the current product to an outside continuation value. Gu and Wang (2022) are a particularly close interface-level antecedent: selected attributes appear in an outer layer and a click reveals the product's remaining attributes. Here, by contrast, the free outer layer is an aggregate generated from a realized noisy corpus, and costly inspection reveals elements of that same corpus rather than the product's attributes directly. The departure is thus the information environment. Before paying to inspect reviews, the consumer sees a free AI summary that is a coarsening of the same finite review pool she can later unpack.

The modeling of the summary borrows from information-design and rating-design work, where a signal structure maps underlying information into messages (Nelson, 1974; Anderson and Renault, 2006; Kamenica and Gentzkow, 2011; Hopenhayn and Saeedi, 2026; Saeedi and Shourideh, 2026). The decision-margin logic also rests on standard value-of-information economics: Blackwell comparison ranks information structures, and information has zero gross value when it cannot change the decision (Blackwell, 1953; Chade and Schlee, 2002). Thus the decision-sufficiency result below should be read as an application of a classical principle to a new information environment, not as a new theorem about value of information per se. What is new is the compression-consistent posterior. The free summary and the later costly reviews are not independent experiments; they are two views of one finite pool, so every opened review updates beliefs about both quality and the hidden evidence that must still rationalize the summary.

That shared-pool restriction generates a belief-updating implication that is absent from a matched independent-summary benchmark. A review that contradicts the summary is still bad news, but its effect is rationally dampened because the unopened residual pool must remain summary-consistent. This is different from behavioral under-reaction or noisy probabilistic reasoning (Benjamin, 2019). The attenuation is fully Bayesian and disappears in an independent-summary model with the same summary precision; it is therefore a structural fingerprint of using the same finite pool for compression and later verification.

Existing theory already shows that free or public information can alter subsequent costly learning. Public communication may crowd out private acquisition (Chahrour, 2014); sender-provided information can deter, encourage, or redirect follow-up learning (Matysková and Montes, 2023; Cheng, 2023); and costly acquisition can be concentrated at intermediate levels of public-information quality (Yoon, 2015). In a particularly useful independent-signal benchmark, Liao and Szkup (2026) study a free initial signal followed by costly sequential signals and find that the conditional acquisition region expands with the precision of the additional signal. The mechanism here is different. A single primitive ρ simultaneously improves a free compression and the raw evidence available for costly verification, and the two observations are linked because they use the same finite pool. The resulting comparative static combines the usual value of a more precise raw signal with the growing decisiveness of the compression itself.

The multidimensional model connects the paper to disclosure and coarse communication, but the mechanism is different. Voluntary- and verifiable-disclosure models study strategic evidence transmission and the limits of disclosure when evidence can be selectively presented (Milgrom, 1981; Grossman, 1981; Dye, 1985; Seidmann and Winter, 1997; Giovannoni and Seidmann, 2007; Shin, 2003; Dziuda, 2011; Bizzotto, Rüdiger, and Vigier, 2020; Titova and Zhang, 2025). Here nondisplay is informative even without strategic concealment: under a known top-K prevalence rule, an omitted dimension is inferred to be less prevalent and therefore less likely to arrive in future reviews. Likewise, coarse-communication and persuasion models study limited message capacity and sender-provided information (Le Treust and Tomala, 2019; Aybas and Turkel, 2026; Wei, 2021; Matysková and Montes, 2023). In Wei (2021) and Matysková and Montes (2023), costly learning follows sender-provided information, but follow-up acquisition is an information-processing choice or an independent experiment. Here the follow-up observations are elements of the corpus that generated the initial message. The finite-slot constraint is used for a receiver-side question: whether a truthful summary leaves residual verification value, and which slots reduce costly follow-up for consumers with heterogeneous tastes.

Finally, the paper is related to rational-inattention and online-review evidence. Rational-inattention models endogenize information processing (Matějka and McKay, 2015; Caplin and Dean, 2015); here the platform supplies the compression and the consumer decides whether to unpack it. Empirical work on online reviews shows that reviews affect demand and that consumers use review text in addition to summary statistics (Chevalier and Mayzlin, 2006; Cheng, Guo, and Liu, 2021). Dai et al. (2018) study how platforms should aggregate heterogeneous ratings, while Vana and Lambrecht (2021) show that individual reviews retain purchase effects after controlling for the aggregate rating, especially when they resolve uncertainty or contrast with the aggregate. Recent field evidence further shows that platform information can change exploration (Fang et al., 2024), and early evidence on AI review summaries reports reduced engagement with individual reviews (Alavi and Nozari, 2025). These findings motivate, but do not provide, a compression-consistent model of the consumer's source-opening decision. The contribution here is a theory of truthful finite-slot summaries for multidimensional experience-good information acquisition. The companion numerical paper endogenizes $\bar{M}$ through the full across-option Bellman problem.

1.2 Roadmap

Section 2 states the common within-product environment. Sections 3–4 use scalar majority aggregation to isolate compression-consistent updating and the decision margin. Section 5 develops the core finite-slot multidimensional model and its omission, persistence, and design results. Section 6 explains which objects survive under a general known kernel and uses a three-cell rule as a short contrast. Section 7 disciplines magnitudes with an Amazon-informed illustration. Section 8 locates the across-product, identification, and experimental companions.

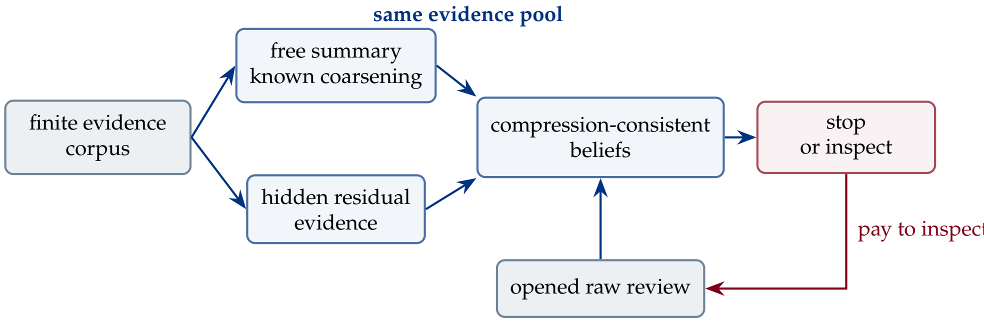

Figure 1: Post-compression verification. The AI summary and any later opened reviews are generated from the same finite evidence pool. Observing the summary therefore changes both beliefs about product quality and beliefs about what raw evidence remains hidden.

2 The Post-Compression Verification Problem

2.1 Economic environment

Fix a reached product j and suppress its index. At this point the consumer's problem is not whether to discover another product, but whether she knows enough about the current one. The product has $D \geq 1$ latent experience dimensions,

$$\theta = (\theta_1, \ldots, \theta_D) \in \{0,1\}^D.$$

The consumer observes product characteristics $\mathbf{x}$ and price P before verification. These observables resolve the search-good component of value and help determine the prior over latent experience quality. Her realized utility from choosing the product is

$$u(\theta;\gamma) = \underbrace{\mathbf{x}'\beta - \alpha P}_{\text{observed product value}} + \underbrace{\sum_{d=1}^{D} \gamma_d \theta_d}_{\text{latent experience value}}, \qquad \alpha > 0, \quad \gamma_d \geq 0, \quad \sum_{d=1}^{D} \gamma_d > 0. \tag{2.1}$$

Here β contains the utility coefficients on observed characteristics, α is price sensitivity, and $\gamma = (\gamma_1, \ldots, \gamma_D)$ records the consumer's tastes for the latent dimensions. If her posterior mean is $\mu = (\mu_1, \ldots, \mu_D)$, where $\mu_d = \mathbb{P}(\theta_d = 1 \mid \text{current information})$, expected product utility is

$$\delta(\mu;\gamma) = \mathbf{x}'\beta - \alpha P + \sum_{d=1}^{D} \gamma_d \mu_d. \tag{2.2}$$

The scalar benchmark sets $D = 1$; the core model allows $D > 1$.

The product generates a finite corpus $\mathcal{E}$ containing N inspectable evidence units. In the scalar benchmark, $\mathcal{E} = \mathcal{R} = \{r_1, \ldots, r_N\}$ is a pool of binary review signals about one dimension. In the multidimensional model, $\mathcal{E} = (R, M)$: R records dimension-level signs and M records which dimensions each sparse review discusses. These specializations are defined formally in Sections 3 and 5. In the binary-sign specializations, a sign about dimension d agrees with θ_d with probability $\rho_d \in (1/2, 1)$; when $D = 1$, I write this oriented precision as ρ.

Before any costly inspection, the platform maps the realized corpus into a free finite message. Let ω denote that message, let z collect public inputs to the summarization rule, such as product category or interface features, and let ϕ index the rule's known parameters. The compression kernel

$$\underbrace{Q_\phi(\omega \mid \mathcal{E}, z)}_{\text{known compression kernel}} = \mathbb{P}(\text{summary message is } \omega \mid \mathcal{E}, z; \phi) \tag{2.3}$$

is therefore the probability that the rule produces message ω from corpus $\mathcal{E}$. A deterministic summary is the special case in which this probability is one for a single message; the kernel notation also accommodates known random tie-breaking. The consumer knows Q_ϕ, observes z, and sees ω, but does not initially see the full corpus. Truthfulness means that the message is generated from the realized corpus through this known rule: the platform may select, threshold, or fold evidence, but it does not fabricate an observation outside the rule. Exogeneity means that the platform does not choose ϕ inside the consumer's verification problem.

In both the scalar and multidimensional cases the economic question is the same: after seeing ω, when is another costly evidence unit worth opening? A USB-C hub illustrates the distinction. The listing reveals the price and number of ports; reviews reveal whether charging slows when several devices are connected. The summary compresses those experiences into a few displayed claims, while the underlying reviews remain available at a reading cost.

Table 1: Search and Experience Components in the Within-Product Problem

Layer	Object in the model	Interpretation
Search-good attributes	$\mathbf{x}'\beta - \alpha P$	Observed from the listing before verification; not learned from reviews.
Latent experience quality	$\sum_d \gamma_d \theta_d$	Payoff-relevant experience quality; $D = 1$ gives the scalar benchmark.
Finite evidence corpus	$\mathcal{E}$	The N reviews or signals, compressed for free and individually inspectable at a cost.
Free compression	$Q_\phi(\omega \mid \mathcal{E}, z)$	The known rule that maps the realized corpus into message ω.
Review-signal technology	ρ or ρ_d	Oriented signal precision, scalar or dimension-specific; its measurement is deferred to the companion empirical framework.

Normalize immediate exit to zero. The outside of the within-product problem is summarized by $\bar{M} \in \mathbb{R}$, the expected continuation value of the best alternative product before any additional inspection of the current product. Thus the effective outside stopping payoff is $\max\{0, \bar{M}\}$. This paper conditions on $\bar{M}$ and solves the within-product verification problem. The companion numerical paper endogenizes this alternative-product value through across-product search.

Maintained assumption: authentic quality signals. Amazon's public description of its reviews system separates authenticity screening from AI summarization: the platform screens for fake reviews and then summarizes common themes and sentiment among eligible reviews.[1] The maintained assumption here is that this first stage leaves an authentic pool of quality signals. Thus every scalar or dimension-level sign is interpreted as a genuine experience report. The scalar precision ρ and dimension-specific precision ρ_d measure honest experiential agreement with true quality, not platform trust. The restriction $\rho, \rho_d \in (1/2, 1)$ keeps signals informative; the residual $1-\rho$ is idiosyncratic experience, taste heterogeneity, or lemon-unit variation among honest reviewers. A mixed or disagreeing summary cell is therefore interpreted as truthful reporting of genuine disagreement, not as evidence that the pool is fake or unreliable. Empirically, these precision objects are proxied by dispersion in verified reviews, ratings, returns, and category-level agreement. The benchmark additionally treats the eligible review pool as representative of the modeled experience distribution: conditional on latent quality and, in the multidimensional model, realized coverage, selection into the eligible pool does not further tilt the positive-negative sign distribution. Authenticity and representativeness are distinct maintained assumptions.

Interpretation: honest pool, honest compression. The environment is an honest pool feeding an honest compression. The summary may select dimensions and apply sentiment thresholds, but it does not fabricate evidence. The welfare question is therefore not whether the platform lies, but whether truthful compression preserves decision-relevant content. Any loss from omitted dimensions or folded cells is a structural limit of finite summarization.

Deferred margins: endogenous authenticity and reviewing selection. Three margins are deferred. First, one could model the authenticity filter explicitly as a map from the submitted corpus to an eligible corpus. Second, one could endogenize who chooses to review: verified purchase establishes an authentic transaction, not a representative posting decision. Third, one could let the platform choose the truthful coarsening rule strategically, which would turn the model into a persuasion problem with a no-fabrication constraint. This paper fixes those margins to isolate demand-side verification.

[1] See Amazon's discussion of trustworthy reviews at https://trustworthyshopping.aboutamazon.com/focus/trustworthy-reviews.

Decision maker and known primitives. The consumer is a Bayesian expected-utility maximizer. She knows her own taste weights, the relevant quality prior, the review-pool size, signal precisions, inspection-cost schedule, and summary rule; in the multidimensional model she also knows the coverage prior G, the distribution over which dimensions the reviews discuss. There is no separate discount factor within the finite verification episode: delay and cognitive effort are summarized by the inspection-cost schedule.

2.2 What the consumer sees and chooses

Let t denote the number of evidence units already inspected. The superscript W marks a within-product inspection cost. The marginal cost of inspection $t+1$ is

$$c^W(t+1) = \underbrace{\kappa^W}_{\text{first-inspection cost}} + \underbrace{\lambda^W t}_{\text{cost growth after } t \text{ inspections}}, \qquad \kappa^W > 0, \quad \lambda^W \geq 0, \tag{2.4}$$

where κ^W is the cost of opening the first evidence unit and λ^W is the incremental cost growth across successive inspections.

The within-product timing is:

1. The consumer arrives with prior vector $\mu_0 = (\mu_{1,0}, \ldots, \mu_{D,0})$, where $\mu_{d,0} = \mathbb{P}(\theta_d = 1 \mid \mathbf{x}, P, z)$, and outside value $\bar{M}$. In the scalar benchmark $D = 1$, so the prior is the scalar μ_0.

2. In the AI environment, she observes a free summary output of the finite review pool. In the scalar benchmark this output is denoted a; in the multidimensional core it is ω. In the no-summary benchmark, no such summary is shown.

3. She either stops, taking the best of exit, the current product, and the outside opportunity, or pays $c^W(t+1)$ to inspect evidence unit $t+1$.

4. After each inspection, she updates beliefs about quality and about the unopened corpus, then repeats the same stop-or-verify decision until the finite pool is exhausted.

Inspection is passive and without replacement. Conditional on the realized finite pool, each inspection selects uniformly from the unopened signals or reviews, and inspection order is independent of their realized contents. The consumer therefore cannot target a favorable review or a particular topic before paying the inspection cost. Thus the state variable is information about one product, and the outside opportunity enters only through $\bar{M}$. The Online Appendix collects a compact table of maintained assumptions and deferred margins.

What is general and what is specialized. Throughout, the summary rule is known, exogenous, and truthful. The integrated posterior and Bellman problem require a correctly specified finite joint law; binary signals, majority aggregation, and top-K selection are additional restrictions used only for closed forms and mechanism-specific comparative statics. Section 6 separates these general objects from the benchmark restrictions.

One verification problem for any known compression rule. After t inspections, let h_t denote the ordered history of evidence units the consumer has actually opened, with $h_0 = \emptyset$. Let $\mathscr{C}(h_t)$ be the set of full corpora whose opened components agree with that observed history. The consumer has not observed which member of $\mathscr{C}(h_t)$ is the realized corpus. She therefore integrates over all of them. For latent state θ, the probability of observing both summary ω and history h_t is the rule-specific integrated likelihood

$$\mathcal{L}_\theta^\phi(\omega, h_t) = \sum_{\mathcal{E}\in\mathscr{C}(h_t)} \underbrace{\mathbb{P}(\mathcal{E} \mid \theta, z)}_{\text{corpus probability}} \underbrace{Q_\phi(\omega \mid \mathcal{E}, z)}_{\text{summary compatibility}}, \tag{2.5}$$

where dependence of the left-hand side on the observed public variables z is suppressed to lighten notation. The sum is what makes updating compression-consistent: every candidate hidden corpus must fit the reviews already opened and must also be capable of generating the summary already seen.

Because θ has finite support, Bayes' rule gives the common posterior. Using ϑ to index candidate latent-state vectors in $\{0,1\}^D$,

$$\mathbb{P}(\theta \mid \omega, h_t, \mathbf{x}, P, z) = \frac{\mathbb{P}(\theta \mid \mathbf{x}, P, z)\, \mathcal{L}_\theta^\phi(\omega, h_t)}{\sum_{\vartheta\in\{0,1\}^D} \mathbb{P}(\vartheta \mid \mathbf{x}, P, z)\, \mathcal{L}_\vartheta^\phi(\omega, h_t)}. \tag{2.6}$$

In the specializations below, conditioning on z is suppressed whenever those public variables affect the compression rule but not the quality prior. Let $s_t = (\omega, h_t)$ denote the consumer's information state and let $\mu_t = (\mu_{1,t}, \ldots, \mu_{D,t})$ collect the posterior means implied by (2.6). Using (2.2), the value of stopping is

$$S(s_t; \gamma) = \max\left\{ \underbrace{0}_{\text{exit}}, \underbrace{\bar{M}}_{\text{best alternative}}, \underbrace{\delta(\mu_t; \gamma)}_{\text{current product}} \right\}. \tag{2.7}$$

Finally, let X_{t+1} denote the full content of the next passively sampled unopened evidence unit and let $s_{t+1}(X_{t+1})$ be the updated information state. Let $V(s_t)$ be the consumer's maximal expected payoff from state s_t. The common finite-state Bellman problem is

$$V(s_t) = \max\left\{ \underbrace{S(s_t; \gamma)}_{\text{stop now}}, \underbrace{-c^W(t+1) + \mathbb{E}[V(s_{t+1}(X_{t+1})) \mid s_t]}_{\text{inspect once, then continue optimally}} \right\}. \tag{2.8}$$

The specializations below change what the summary reveals and therefore what the consumer expects to learn from another review. They do not change her choice: at every history she compares the value of acting now with the value of paying once more to inspect evidence from the same pool.

3 Scalar Benchmark: Learning from a Shared Pool

The scalar benchmark is deliberately narrow. All review signals concern one latent experience dimension, such as whether a battery is reliable or whether a USB-C hub maintains charging speed under load. Its role is to deliver closed forms and decision-margin mechanics. It shows exactly how a truthful summary moves beliefs, when one more review can change the stopping action, and why scalar majority summaries become nearly decisive in large evidence pools. The core multidimensional model then studies the economically richer case in which finite summary slots omit or fold dimensions that some consumers still care about.

The scalar benchmark sets $D = 1$ in (2.1)–(2.2) and writes the single state, taste weight, and posterior as θ, γ, and μ, with $\theta \in \{0, 1\}$ and prior $\mu_0 = \mathbb{P}(\theta = 1 \mid \mathbf{x}, P) \in (0, 1)$. Thus the common expected utility specializes to

$$\delta(\mu) = \mathbf{x}'\beta - \alpha P + \gamma\mu,$$

where $\gamma > 0$. I use m for the number of scalar signals inspected, replacing the common history index t; all cost and outside-value objects remain those defined in Section 2. The only new primitive needed for this specialization is the scalar signal technology. The finite review pool is

$$\mathcal{R} = \{r_1, \ldots, r_N\}, \qquad r_k \in \{0, 1\},$$

with oriented signal precision

$$\mathbb{P}(r_k = 1 \mid \theta = 1) = \mathbb{P}(r_k = 0 \mid \theta = 0) = \rho, \qquad \rho \in (1/2, 1). \tag{3.1}$$

Ex ante, conditional on θ, the N signals are independent and identically distributed according to (3.1). Once the finite pool is realized, inspection samples from it uniformly without replacement as specified in the common timing. A positive signal is high-quality-indicating and a negative signal is low-quality-indicating.

3.1 A free majority summary

For concreteness, the benchmark overview is a randomized majority-rule summary with neutral tie-breaking. Majority is applied to the oriented binary signals defined above. If the underlying signals were systematically misoriented and the platform aggregated them without correcting orientation, majority rule would amplify that mistake; that non-benchmark case is excluded by the maintained assumption of known signal orientation. Let $Y = \sum_{k=1}^{N} r_k$, let $\iota_N = \mathbf{1}\{N \text{ is even}\}$, and let $\tau \in [0, 1]$ denote the probability that an exact tie is reported as favorable. The benchmark fixes $\tau = 1/2$. The summary rule is the known kernel

$$\mathbb{P}(a = 1 \mid Y) = \begin{cases} 1, & Y > N/2, \\ \tau, & \iota_N = 1 \text{ and } Y = N/2, \\ 0, & Y < N/2. \end{cases} \tag{3.2}$$

Under the assumptions used here–conditionally independent binary signals of equal precision, known orientation, and symmetric binary classification objectives–neutral majority is the canonical aggregation rule; richer environments with heterogeneous signal quality would instead imply weighted or threshold-shifted rules (Nitzan and Paroush, 1982). The summary should therefore be interpreted as a neutral, exogenous, costless aggregation of human signals that the consumer could otherwise process only by reading dispersed content.

However, what is not exogenous is the overview's informativeness, which is implied by the precision of the underlying human signals and the aggregation rule. Under majority aggregation, the induced overview precision is

$$\zeta(\rho) = \mathbb{P}(a = \theta \mid \theta) = \mathbb{P}_\rho(Y > N/2) \; + \; \begin{cases} \tau\, \mathbb{P}_\rho(Y = N/2), & N \text{ even}, \\ 0, & N \text{ odd}, \end{cases} \tag{3.3}$$

where $\mathbb{P}_\rho$ means that each raw signal is correct with probability ρ. Under neutral tie-breaking, the majority rule is symmetric, so $\mathbb{P}(a = 1 \mid \theta = 1) = \mathbb{P}(a = 0 \mid \theta = 0) = \zeta(\rho)$. The odd-$N$ case is the special case in which the tie mass is absent. For $N \geq 3$, the induced overview precision exceeds ρ whenever $\rho > 1/2$; for $N = 2$ neutral majority has the same precision as one raw signal. This is the standard Condorcet jury theorem logic (Austen-Smith and Banks, 1996; Boland, 1989). The consumer is assumed to know the aggregation rule and therefore knows the true overview precision. The model also assumes that the consumer knows the size of the summarized signal set, N, and that later inspection unpacks that set.[2][3] *A step-by-step derivation appears in Online Appendix C.2.1.* Because $\zeta(\rho)$ is increasing in ρ and nondecreasing in N under neutral majority, changes in the precision or abundance of underlying signals change the informativeness of the free overview and therefore the incentive for costly verification. The resulting comparative statics are central to the paper's analysis of when AI substitutes for within-option search. Online Appendix C.1 records the arbitrary-τ version; all main-text scalar objects below use the neutral benchmark $\tau = 1/2$. The rest of the scalar section retains this rule because it turns the integrated likelihood into binomial tails; Online Appendix C.4 gives the ternary threshold-band and general finite-kernel versions.

[2] In e-commerce settings, this corresponds to the review count displayed on product pages, which is visible on most major platforms. The consumer can therefore observe the size of the evidence pool being summarized before deciding whether to unpack it.

[3] Relaxing known precision by allowing consumers to hold priors over ζ would introduce an additional learning problem, because the overview would then update beliefs not only about quality but also about the information environment itself.

3.2 How the summary changes subsequent learning

The summary changes more than the consumer's initial belief. Because it was computed from the same finite pool she may later inspect, it also changes what she expects the next review to say. The no-summary environment provides the comparison ruler: there the consumer learns only from costly raw signals. With AI, she must condition jointly on the free summary and on every review opened afterward. In both environments, observables enter through the prior $\mu_0 = \mathbb{P}(\theta = 1 \mid \mathbf{x}, P)$, while the signal likelihood depends only on θ.

No-summary comparison. The matched counterfactual removes the free message and lets the consumer learn only by opening raw signals. It isolates the difference between ordinary Bayesian learning and learning from evidence already compressed into the summary. Online Appendices C.3.1, C.3.3, and C.3.4 give its posterior, predictive distribution, and dynamic recursion step by step.

3.2.1 Compression-consistent beliefs

Learning in the AI environment tracks both the overview outcome and what fraction of the underlying evidence has already been unpacked. The key object is the probability that the unseen part of the pool can still produce the summary already observed. Let

$$\tau = \frac{1}{2}$$

be the neutral tie-breaking probability, and let m denote the number of inspected signals so far, with y of them positive (high-quality-indicating). For $a \in \{0, 1\}$ and $p \in (0, 1)$,[4] let $B \sim \text{Bin}(N - m, p)$ and define

$$H_a^\tau(m, y; p) = \begin{cases} \underbrace{\mathbb{P}\big(B > N/2 - y\big)}_{\substack{\text{strict favorable-majority} \\ \text{completions}}} + \underbrace{\iota_N \tau \, \mathbb{P}\big(B = N/2 - y\big)}_{\text{favorable tie mass}}, & a = 1, \\ \underbrace{\mathbb{P}\big(B < N/2 - y\big)}_{\substack{\text{strict unfavorable-majority} \\ \text{completions}}} + \underbrace{\iota_N (1 - \tau) \, \mathbb{P}\big(B = N/2 - y\big)}_{\text{unfavorable tie mass}}, & a = 0. \end{cases} \tag{3.4}$$

The equality event has probability zero when N is odd, so the odd-N model is nested without any separate case. The two cells partition the signal pool: $H_1^\tau(m, y; p) + H_0^\tau(m, y; p) = 1$. The binomial recursion

$$H_a^\tau(m, y; p) = pH_a^\tau(m + 1, y + 1; p) + (1 - p)H_a^\tau(m + 1, y; p)$$

[4] Here p is only a generic Bernoulli success-probability argument used to write the binomial-consistency term compactly. In the actual model, the primitive raw-signal precision is ρ, so the relevant evaluations are $p = \rho$ under $\theta = 1$ and $p = 1 - \rho$ under $\theta = 0$.

also holds, because H_a^τ is the conditional expectation of the known summary kernel after the partial inspection history. Histories with zero total likelihood under both latent states are impossible and excluded from the reachable state space. Economically, $H_a^\tau(m, y; p)$ is the discipline that the displayed summary imposes on future evidence. After observing overview a and partial inspection history (m, y), it measures how likely the remaining uninspected signals are to complete the finite signal pool in a way that still agrees with the overview. At entry, the same object nests the overview precision:

$$H_1^\tau(0,0;\rho) = \zeta(\rho), \qquad H_0^\tau(0,0;\rho) = 1 - \zeta(\rho).$$

I keep the superscript H^τ in the scalar body even though all main-text results use the neutral benchmark $\tau = 1/2$, so that the parity convention remains visible in every posterior object.

Definition 1 (Reachable AI histories)**.** A within-product AI history (a, m, y) is reachable if $a \in \{0, 1\}$, $0 \le y \le m \le N$, and

$$H_a^\tau(m, y; \rho) > 0 \qquad \text{or} \qquad H_a^\tau(m, y; 1-\rho) > 0.$$

Equivalently, reachable histories are exactly those for which the overview realization and the inspected signal counts can arise with positive probability under at least one latent-quality state.

The likelihood of overview outcome a and inspection state (m, y) under $\theta = 1$ is

$$L_1(a, m, y) = \underbrace{\binom{m}{y}\rho^y(1-\rho)^{m-y}}_{\text{opened-signal likelihood}} \underbrace{H_a^\tau(m, y; \rho)}_{\text{summary consistency}}, \tag{3.5}$$

while under $\theta = 0$ it is

$$L_0(a, m, y) = \underbrace{\binom{m}{y}(1-\rho)^y\rho^{m-y}}_{\text{opened-signal likelihood}} \underbrace{H_a^\tau(m, y; 1-\rho)}_{\text{summary consistency}}. \tag{3.6}$$

Hence the exact posterior is, for every history (a, m, y) with positive probability,

$$\widetilde{\mu}^{AI}(a, m, y) = \frac{\overbrace{\underbrace{\mu_0 L_1(a, m, y)}}^{}_{\text{prior weight times likelihood under high quality}}}{\underbrace{\mu_0 L_1(a, m, y) + (1-\mu_0) L_0(a, m, y)}_{\text{total likelihood of the overview and inspection history}}}. \tag{3.7}$$

The posterior after seeing only the overview is the special case $\widetilde{\mu}^{AI}(a, 0, 0)$. In computation, the Bellman problem need only be evaluated on reachable histories. *A step-by-step derivation appears in Online Appendix C.3.2.*

To characterize continuation, define the probability that the next inspected signal is favorable conditional on state (a, m, y) and latent precision p:

$$\psi_a(m, y; p) = \underbrace{p}_{\substack{\text{raw positive} \\ \text{chance}}} \underbrace{\frac{H_a^\tau(m+1, y+1; p)}{H_a^\tau(m, y; p)}}_{\substack{\text{pool-consistency} \\ \text{correction}}}. \tag{3.8}$$

This ratio is defined whenever $H_a^\tau(m, y; p) > 0$; the derivation follows from conditional exchangeability of the underlying signals given θ.[5] Intuitively, $\psi_a(m, y; p)$ is the next-signal analogue of H: it is the probability that the next inspected signal is positive after conditioning on the fact that the remaining signal pool must still be compatible with the already observed overview. The unconditional probability that the next inspected signal is favorable is therefore

$$\pi^+(a, m, y) = \underbrace{\widetilde{\mu}^{AI}(a, m, y)\psi_a(m, y; \rho)}_{\text{next signal favorable if quality is high}} + \underbrace{\Big(1 - \widetilde{\mu}^{AI}(a, m, y)\Big)\psi_a(m, y; 1-\rho)}_{\text{next signal favorable if quality is low}}. \tag{3.9}$$

Equivalently, and without any division by a zero consistency probability, the same transition probability can be written directly as

$$\pi^+(a, m, y) = \frac{\begin{array}{c}\mu_0 \rho^y (1-\rho)^{m-y} \rho H_a^\tau(m+1, y+1; \rho) \\ + (1-\mu_0)(1-\rho)^y \rho^{m-y}(1-\rho) H_a^\tau(m+1, y+1; 1-\rho)\end{array}}{\mu_0 \rho^y (1-\rho)^{m-y} H_a^\tau(m, y; \rho) + (1-\mu_0)(1-\rho)^y \rho^{m-y} H_a^\tau(m, y; 1-\rho)}. \tag{3.10}$$

A step-by-step derivation of (3.8) *and* (3.9) *appears in Online Appendix C.3.3.*

3.3 The consumer's scalar verification problem

Conditional on a posterior μ, the stopping payoff is

$$S(\mu; \bar{M}) = \max\{0, \delta(\mu), \bar{M}\}. \tag{3.11}$$

The consumer now has all the information needed to choose. She can take the current product, leave for the outside option, or pay to reveal one more signal from the constrained

[5] Conditioning on the overview outcome and on y positives among m inspected signals, every remaining signal position is symmetric, so the probability that the $(m+1)$-th inspected signal is favorable equals the probability that one designated remaining signal is positive times the probability that the residual uninspected set of $N - m - 1$ signals remains consistent with the overview realizing at a. If $H_a^\tau(m, y; p) = 0$, the latent state associated with that value of p has zero posterior probability at the history, so the corresponding conditional ratio is payoff irrelevant. Equation (3.10) below gives the unconditional transition probability in a form that is defined on every reachable history.

residual pool. The main dynamic object in this paper is the post-summary Bellman problem. The no-summary recursion is the analogous benchmark obtained by replacing $\widetilde{\mu}^{AI}$ and π^+ with $\widetilde{\mu}$ and $\widetilde{\pi}^+$; it is stated in Online Appendix C.3.4 because it serves only as a comparison object here and becomes central in the companion numerical across-option paper.

In the AI environment, let $W(a, m, y; \bar{M})$ be the value after summary realization a, m inspected signals, and y positives. The terminal condition is

$$W(a, N, y; \bar{M}) = S(\widetilde{\mu}^{AI}(a, N, y); \bar{M}). \tag{3.12}$$

For $m < N$,

$$W(a, m, y; \bar{M}) = \max\left\{S(\widetilde{\mu}^{AI}(a, m, y); \bar{M}),\ \mathcal{K}(a, m, y; \bar{M})\right\}, \tag{3.13}$$

where

$$\begin{aligned} \mathcal{K}(a, m, y; \bar{M}) = -c^W&(m + 1) \\ &+ \pi^+(a, m, y)W(a, m + 1, y + 1; \bar{M}) \\ &+ (1 - \pi^+(a, m, y))W(a, m + 1, y; \bar{M}). \end{aligned} \tag{3.14}$$

Define the gross value of one additional signal at state (a, m, y) as

$$\Delta^A(a, m, y; \bar{M}) = \mathcal{K}(a, m, y; \bar{M}) + c^W(m + 1) - S(\widetilde{\mu}^{AI}(a, m, y); \bar{M}). \tag{3.15}$$

The corresponding ex ante post-entry AI value is

$$V^A(\bar{M}) = \sum_{a\in\{0,1\}} \mathbb{P}(a)W(a, 0, 0; \bar{M}),$$

while the no-summary value is $V^N(\bar{M}) = \widetilde{W}(0, 0; \bar{M})$, with $\widetilde{W}$ defined in Online Appendix C.3.4. The consumer continues within-product verification if and only if

$$c^W(m + 1) \leq \Delta^A(a, m, y; \bar{M}). \tag{3.16}$$

4 What Shared-Pool Compression Does

4.1 Shared-pool learning and attenuation

The first scalar result is not a stopping theorem. It isolates what is distinctive about the information structure itself. If the summary were merely an independent signal, later raw reviews would receive their usual Bayesian weight. Compression consistency changes that updating because the summary and the opened reviews are draws from the same finite pool.

The comparison benchmark is an *independent-summary* environment: the consumer observes a favorable or unfavorable summary signal that is statistically separate from the raw reviews she may later inspect. The next theorem places that benchmark and the shared-pool model in one statement.

Theorem 1 (Shared-pool attenuation and independent-summary separation). Fix $N > 1$, $\rho \in (1/2, 1)$, and neutral majority aggregation. Let $\zeta(\rho)$ be the resulting summary accuracy and define $\lambda^\rho = \log[\rho/(1-\rho)] > 0$.

(i) (Matched independent-summary benchmark.) Suppose

$$\mathbb{P}(a = 1 \mid \theta = 1) = \mathbb{P}(a = 0 \mid \theta = 0) = \zeta(\rho)$$

and, conditional on θ, a is independent of subsequently opened raw signals. After m inspections and y positives, log posterior odds are

$$\ell^I(a, m, y) = \log \frac{\mu_0}{1-\mu_0} + (2a-1) \log \frac{\zeta(\rho)}{1-\zeta(\rho)} + (2y-m)\lambda^\rho. \tag{4.1}$$

Thus every next raw realization $x \in \{0, 1\}$ changes log odds by exactly $(2x-1)\lambda^\rho$. Moreover, the summary is never spent: its log-odds contribution remains $(2a - 1) \log[\zeta(\rho)/(1-\zeta(\rho))]$ after every raw history.

(ii) (Shared-pool attenuation.) In the compression-consistent model, consider a favorable summary $a = 1$ and reachable history $(1, m, y)$. Write

$$n = N - m, \qquad s = \lfloor N/2 \rfloor + 1 - y, \qquad \ell^C(1, m, y) = \log \frac{\widetilde{\mu}^{AI}(1, m, y)}{1 - \widetilde{\mu}^{AI}(1, m, y)}.$$

At a strict-interior history, $1 \le s \le n-1$, every next raw update is strictly attenuated relative to part (i): for each $x \in \{0, 1\}$,

$$0 < (2x-1)\left[\ell^C(1, m+1, y+x) - \ell^C(1, m, y)\right] < \lambda^\rho. \tag{4.2}$$

In particular, a summary-disconfirming negative review satisfies

$$-\lambda^\rho < \ell^C(1, m+1, y) - \ell^C(1, m, y) < 0. \tag{4.3}$$

The same inequality holds after an unfavorable summary when $s^- = \lfloor N/2 \rfloor + 1 - (m-y)$ satisfies $1 \le s^- \le n-1$. In particular, a summary-disconfirming review has the correct sign but strictly less absolute log-likelihood weight than in the matched independent-summary benchmark. Disconfirmation is emphasized because it is behaviorally salient, not because confirmation follows a different mechanism.

(iii) (Depletion.) Once opened signals alone strictly imply the summary cell, the residual consistency term is spent. If $a = 1$ and $y > N/2$, then

$$\ell^C(1, m, y) = \log \frac{\mu_0}{1-\mu_0} + (2y-m)\lambda^\rho,$$

and the analogous identity holds when $a = 0$ and $m - y > N/2$. For even N, the tie-adjacent history is not yet fully spent:

$$H_1^\tau(m, N/2; p) = 1 - (1-\tau)(1-p)^{N-m} < 1$$

whenever $m < N$. The strict attenuation claim is therefore confined to the interior histories in part (ii); the boundary can be mechanically constrained by summary consistency or by final tie resolution.

Proof in Online Appendix C.6.

Why attenuation occurs. Independence makes the benchmark summary likelihood ratio a fixed additive term, so every raw signal receives its full weight forever. In the shared-pool model, the posterior also contains the likelihood that the hidden residual pool can rationalize the observed summary. A weighted-binomial-tail likelihood-ratio inequality shows that this conditioning attenuates both raw realizations at strict-interior histories. Once the opened history itself implies the summary, the residual likelihood equals one under both quality states and disappears.

The theorem is the observable fingerprint of using the same finite pool for compression and later verification. A contradictory review changes beliefs about quality and about what must remain hidden, while an independent public signal has no such path dependence.

Remark 1 (Remaining summary content). Let $q = 1 - \rho$ and define the remaining summary content at any nondegenerate AI history by

$$\mathcal{G}_a(m, y) = \log \frac{H_a^\tau(m, y; \rho)}{H_a^\tau(m, y; q)}. \tag{4.4}$$

The compression-consistent log odds can be written as

$$\ell^C(a, m, y) = \log \frac{\mu_0}{1-\mu_0} + (2y - m)\lambda^\rho + \mathcal{G}_a(m, y).$$

Thus, after opening a raw signal $x \in \{0, 1\}$, the log-odds revision is

$$\ell^C(a, m+1, y+x) - \ell^C(a, m, y) = (2x-1)\lambda^\rho + \mathcal{G}_a(m+1, y+x) - \mathcal{G}_a(m, y). \tag{4.5}$$

The first term is the ordinary raw-signal update; the second is the depletion or reinforcement of the summary's remaining informational content. When opened signals alone imply the displayed summary cell, $\mathcal{G}_a = 0$, so the summary is spent and later signals receive their ordinary Bayesian weight.

4.2 When does verification begin?

The Bellman problem answers whether the consumer continues, but its recursive solution does not by itself reveal the economics of starting to read. The tractable first question is

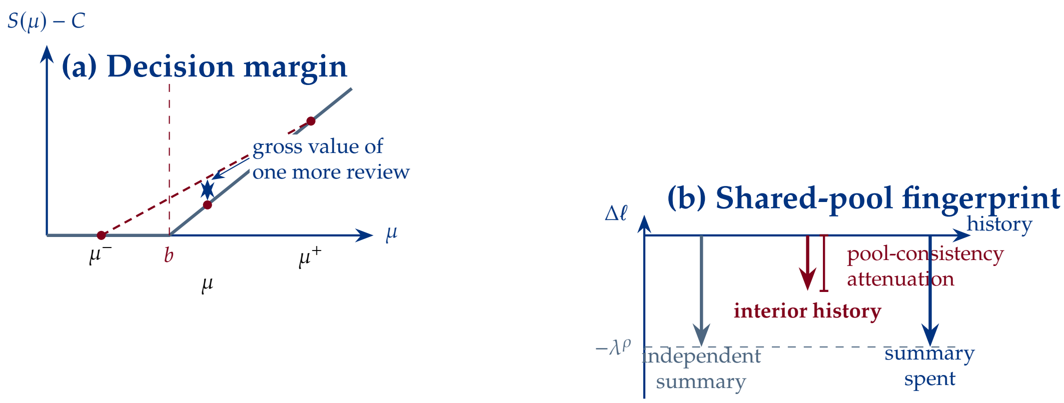

Figure 2: Scalar economics. Panel (a) shows why review opening is valuable only when successor beliefs cross the stopping margin: the gross value is the Jensen gap of the kinked stopping payoff. Panel (b) distinguishes compression of the same finite pool from an independent public signal. Pool consistency attenuates raw-review updates at interior histories; ordinary Bayesian weight returns once the summary constraint is spent.

therefore whether opening one review and then acting beats acting immediately. This feasible deviation preserves the central tradeoff between free compression and costly verification, delivers an exact initiation cutoff for the one-step problem, and provides a sufficient trigger for the unrestricted finite-horizon problem.

Throughout the analytical results, the outside value $\bar{M}$ is fixed. This is the key framing choice of the theory paper. The consumer is not solving the across-product search tree here; she is deciding whether the product-level summary is enough, or whether it is worth paying to inspect the underlying reviews of this product. The three stopping actions are exit with payoff 0, choose the current product with payoff $\delta(\mu)$, or take the outside opportunity with payoff $\bar{M}$.

This special case has a reservation-value flavor similar to Weitzman-style sequential search (Weitzman, 1979), but it is not the canonical Weitzman problem. Here, the consumer already sits inside one product's information problem: the free summary is a compression of a finite review pool, and the remaining choice is whether to unpack more of that same pool.

Belief updating remains available in closed form at every reachable history: for any (a, m, y), the posterior $\widetilde{\mu}^{AI}(a, m, y)$ is given by (3.7), where the H-function is a finite binomial tail plus the tie point mass when N is even. What does not remain closed form once multiple inspections are allowed is the recursively optimal continuation value. The one-step restriction removes that continuation recursion while preserving the post-summary comparison between compression and verification.

The primitive object is therefore the post-summary verification cutoff. A higher ρ makes the overview more informative, pushing the posterior toward an extreme and reducing the value of further search. But it also makes each additional underlying signal more informative if inspected. Those forces work in opposite directions, so the tractable object is the within-product initiation threshold.

Proposition 1 (One-step verification under finite-pool compression). Consider a single-product within problem with constant outside option $\bar{M}$, finite $N > 1$, neutral majority-rule aggregation, and at most one post-summary inspection at marginal cost κ^W. Fix overview realization $a \in \{0, 1\}$ and define

$$\mu^a = \widetilde{\mu}^{AI}(a, 0, 0), \qquad \mu^{a+} = \widetilde{\mu}^{AI}(a, 1, 1), \qquad \mu^{a-} = \widetilde{\mu}^{AI}(a, 1, 0),$$

and

$$\pi^a = \pi^+(a, 0, 0).$$

Let the current stopping payoff at posterior μ be

$$S(\mu; \bar{M}) = \max\{0, \delta(\mu), \bar{M}\}. \tag{4.6}$$

(i) Primitive cutoff. The primitive gross value of one post-summary inspection is

$$\mathcal{I}^a(\bar{M}) = \underbrace{\pi^a S(\mu^{a+}; \bar{M}) + (1 - \pi^a) S(\mu^{a-}; \bar{M})}_{\text{expected payoff after one review}} - \underbrace{S(\mu^a; \bar{M})}_{\text{payoff from stopping now}}. \tag{4.7}$$

Verification is optimal after overview realization a if and only if

$$\kappa^W \leq \mathcal{I}^a(\bar{M}). \tag{4.8}$$

(ii) Decision-margin representation. Define

$$C(\bar{M}) = \max\{0, \bar{M}\}, \qquad b(\bar{M}) = \frac{C(\bar{M}) - (\mathbf{x}'\beta - \alpha P)}{\gamma}. \tag{4.9}$$

Then, for every posterior $\mu \in [0, 1]$,

$$S(\mu; \bar{M}) = C(\bar{M}) + \gamma \big(\mu - b(\bar{M})\big)_+. \tag{4.10}$$

Moreover, posterior beliefs satisfy the one-step martingale identity

$$\mu^a = \pi^a \mu^{a+} + (1 - \pi^a)\mu^{a-}. \tag{4.11}$$

Hence the primitive one-step verification value can be written as the Jensen gap of the kinked stopping payoff:

$$\mathcal{I}^a(\bar{M}) = \gamma \left[\pi^a \big(\mu^{a+} - b(\bar{M})\big)_+ + (1 - \pi^a)\big(\mu^{a-} - b(\bar{M})\big)_+ - \big(\mu^a - b(\bar{M})\big)_+\right] \geq 0. \tag{4.12}$$

If $0 < \pi^a < 1$ and $\mu^{a-} < \mu^{a+}$, then $\mathcal{I}^a(\bar{M}) > 0$ if and only if the stopping margin lies strictly between the two possible next posteriors:

$$\mu^{a-} < b(\bar{M}) < \mu^{a+}. \tag{4.13}$$

If the stopping margin is weakly outside this interval, one-step verification has zero gross value.

(iii) Finite-N precision representation. For the favorable overview $a = 1$, let $q = 1 - \rho$, write $H_1(m, y; p) \equiv H_1^{1/2}(m, y; p)$, and define

$$D_N(\rho) = \mu_0 H_1(0, 0; \rho) + (1 - \mu_0) H_1(0, 0; q).$$

For the two possible next raw signals, define the joint likelihood components

$$L_+^H = \mu_0 \rho H_1(1, 1; \rho), \qquad L_+^L = (1 - \mu_0) q H_1(1, 1; q),$$
$$L_-^H = \mu_0 q H_1(1, 0; \rho), \qquad L_-^L = (1 - \mu_0) \rho H_1(1, 0; q).$$

Then the belief-unit value of one post-summary inspection is the primitive function

$$\Phi_N(\rho; b, \mu_0) = \frac{\left[L_+^H - b(L_+^H + L_+^L)\right]_+ + \left[L_-^H - b(L_-^H + L_-^L)\right]_+ - \left[L_+^H + L_-^H - b D_N(\rho)\right]_+}{D_N(\rho)}. \tag{4.14}$$

Consequently,

$$\text{one-step verification after } a = 1 \iff \frac{\kappa^W}{\gamma} \leq \Phi_N(\rho; b(\bar{M}), \mu_0). \tag{4.15}$$

For fixed N, b, and κ^W/γ, the precision values at which the weak reading region can begin or end are roots of

$$\Phi_N(\rho; b, \mu_0) = \kappa^W/\gamma,$$

holding the prior μ_0 fixed. Since H_1 is a finite binomial tail with the neutral tie correction, this is an exact finite root problem in ρ, piecewise over the kink regions in (4.14). The unfavorable-overview formula is obtained analogously by replacing H_1 with H_0.

Proof in Online Appendix C.6.

Proposition 1 is the scalar workhorse. Part (i) is the abstract value-of-information cutoff; part (ii) shows that the cutoff is only positive when the next raw signal can move beliefs across the stopping margin; part (iii) substitutes the majority-rule likelihoods and turns the same cutoff into a primitive precision-domain equation. This last representation is the scalar bridge to empirical work: if N, ρ, and b are known or estimated, the cutoff is computable. It also makes the nonmonotone force explicit: increasing ρ makes the free summary more decisive, but also makes an opened review more informative.

Two forces of precision. Under majority aggregation, higher ρ raises both the raw-signal log-likelihood ratio and the induced precision of the free overview. At low precision, neither information source can move the decision. At intermediate precision, an opened review can become decision-changing while the overview remains incomplete. At high precision, the overview itself becomes decisive. The resulting reading region need not therefore be monotone in the quality of the underlying evidence.

Corollary 1 (Nonmonotone verification in summary informativeness). There exist primitives in the scalar majority benchmark and a summary realization a with positive probability such that, as signal precision ρ rises from $1/2$ toward 1, one-step verification after the summary is first not optimal, then strictly optimal for an intermediate range of precisions, and finally not optimal again for all sufficiently high precisions.

Proof in Online Appendix C.6.

The claim is deliberately one of possibility rather than universal single-peakedness. Equation (4.15) identifies the entry and exit points as roots of the same finite-pool object whenever its interior value exceeds normalized cost; it does not impose that the reading set must contain only one connected interval. Economically, the result isolates a shared-source precision tradeoff that is absent when the free and costly signals are independent: better underlying evidence strengthens the source that can be opened, but also strengthens the compression that may make opening unnecessary.

Relation to unrestricted dynamics. The connection to unrestricted dynamics is immediate but asymmetric. Because opening once and stopping is feasible, $\Phi_N(\rho; b(\bar{M}), \mu_0) \geq \kappa^W/\gamma$ is a sufficient trigger for dynamic verification to start. The terminal-support condition in Online Appendix Theorem 2 supplies the complementary region in which no dynamic inspection plan can pay.

Online Appendix Corollary 8 records the corresponding one-step inequalities and scale normalization. Online Appendix Corollary 10 identifies the states on which the one-step cutoff is exact for unrestricted dynamics, while Online Appendix Theorem 2 and Corollary 2 derive the exact terminal posterior interval and the dynamic sandwich $\mathcal{C}^1 \subseteq \mathcal{C}^\infty \subseteq \mathcal{B}^R$.

The same geometry overturns the claim that free summaries must reduce review reading. A summary can move a consumer whose prior is safely on one side of the stopping margin into a state whose next-review posteriors straddle that margin. It then *creates* verification that would not occur in the corresponding one-step no-summary problem. The exact conditions are stated and proved in Online Appendix Proposition 3. In the unrestricted problem, the post-summary straddle remains a sufficient trigger; ruling out no-summary verification requires the stronger condition that the margin lie outside the full no-summary terminal-belief interval. Thus the prediction is not "AI always substitutes for reading," but that AI reallocates reading toward realizations that place the decision at risk.

The equivalent tent formula and decision-margin figure are in Online Appendix Corollary 3: inspection value peaks at the stopping margin and is zero when both next beliefs lie on the same side. The appendix also gives a strict-inclusion example, post-summary action regions, the $N = 3$ formulas, and the complementary fixed-N, high-precision limit.

Corollary 2 (Large evidence pools eliminate scalar verification). Fix $\rho \in (1/2, 1)$, $\mu_0 \in (0, 1)$, and binary majority aggregation, with any fixed tie-breaking rule when N is even. Let the post-summary inspection horizon $\bar{m}$ be finite and fixed independently of N, and suppose the marginal inspection cost is bounded below by $\underline{c} > 0$ for the first $\bar{m}$ possible inspections. Then, as $N \to \infty$, the gross improvement from any continuation plan using at most $\bar{m}$ inspections converges to zero after both binary summary realizations, uniformly over feasible outside-option beliefs. Hence there is N^* such that for every $N \geq N^*$, immediate stopping is optimal at the post-summary entry state for all feasible outside-option beliefs.

Proof in Online Appendix C.6.

Corollary 2 gives the empirically relevant large-pool version of the substitution logic. With many scalar signals summarized by a binary majority overview, the summary is nearly decisive, and a small number of later raw reviews moves the posterior by a vanishing amount. Thus a scalar Amazon-scale product with N near one hundred should exhibit essentially zero post-summary scalar verification. Persistent review reading after a large-pool summary is therefore not explained by residual scalar uncertainty in this limit; within the model, it points to the finite-slot multidimensional mechanisms below: omitted attributes, folded aspects, heterogeneous tastes, or residual topic arrival.[6]

Online Appendix Corollary 1 records a useful interface null: if a displayed singleton aspect reveals its exact favorable count from the same exchangeable pool, passively opening a constituent sign has zero residual scalar value. The conclusion does not apply to rounded counts or to the allocation of signs within a folded aspect. The null sharpens the multidimensional interpretation: post-summary reading must be supported by information that exact scalar counts do not already reveal.

The benchmark shows that scalar uncertainty cannot sustain post-summary reading at scale; the core model carries the same verification logic into the finite-slot environment where omitted and folded dimensions can.

5 Core Multidimensional Model

The scalar benchmark treats latent quality as a single object and establishes the decision-margin logic. The core model studies the case in which product quality is multi-attribute, consumers care unevenly across attributes, reviews are sparse across those attributes, and the AI summary is itself lower-dimensional. The key change is simple: the summary no longer compresses *all* of the evidence. It compresses only the dimensions that receive finite display slots, leaving the rest to costly inspection. The parameter K should be

[6] For three-cell threshold summaries, the same conclusion applies to decisive cells whose endpoint regions separate the high- and low-quality latent signal shares. With fixed interior share thresholds, a broad mixed band can instead contain one of the latent shares ρ or $1 - \rho$, so the mixed cell may not vanish asymptotically. The large-N corollary is therefore stated for the binary majority benchmark and decisive endpoint cells.

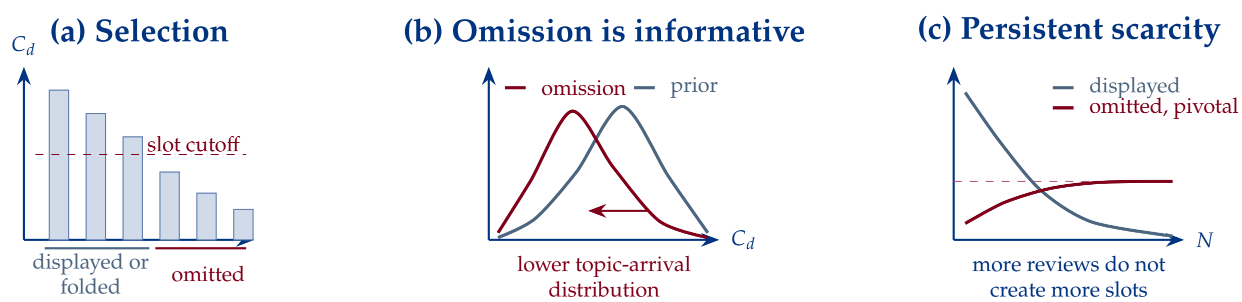

Figure 3: The finite-slot economic mechanism. Prevalence selection displays or folds only a subset of dimensions (panel a). Because selection uses the same corpus, nondisplay shifts the posterior coverage distribution toward lower topic arrival (panel b). As review volume grows, displayed majority cells become decisive, but a taste-relevant omitted dimension can retain positive verification value when display capacity remains fixed (panel c).

read as a display capacity: a platform may show at most six chips, for example, even when review text contains many more economically meaningful dimensions. The model therefore concerns which dimensions enter the displayed set, not the order in which those dimensions are arranged on the page.

That formulation is closer to how many summaries are actually consumed. A restaurant summary highlights food, service, and ambiance but may omit vegan options; a laptop summary highlights speed, battery, and build quality but may omit left-handed ergonomics; a USB-C hub summary highlights portability and compatibility but may fold charging slowdowns under a broad functionality chip. The economically relevant question therefore becomes: after the consumer has seen the summary, is one more unopened review likely to teach her something about a dimension she still cares about? To answer that question precisely, the consumer must update not only beliefs about latent quality, but also beliefs about what the remaining unopened reviews still contain.

This is not merely the scalar model with more notation. In the scalar model, compression changes the consumer's belief about one latent quality state and the feasible signs left in the finite pool. In the multidimensional model, compression also reallocates attention across attributes: covered dimensions become cheap to evaluate, while omitted dimensions remain costly and may or may not be recoverable from future reviews. The new economic force is therefore residual topic arrival, not just residual signal precision.

5.1 Sparse experience information

Return now to the common D-dimensional utility in (2.1), its taste vector γ, and the inspection cost in (2.4). The core model retains $D > 1$, sets $\mu_{d,0} = \mathbb{P}(\theta_d = 1 \mid \mathbf{x}, P) \in (0, 1)$, and assumes that latent dimensions are independent under the prior. What changes is the information technology, not the consumer's payoff: a review may discuss only some dimensions, and a finite-slot summary may display only some of what the corpus contains.

The product has a finite pool of underlying reviews indexed by $n = 1, \ldots, N$, where N

is known to the consumer. Review n is a sparse D-dimensional object

$$r_n = (r_{n1}, \dots, r_{nD}), \qquad r_{nd} \in \{0, 1, \varnothing\}, \tag{5.1}$$

where $r_{nd} = \varnothing$ means that review n contains no information about dimension d. Define the coverage indicator

$$m_{nd} = \mathbf{1}\{r_{nd} \neq \varnothing\}. \tag{5.2}$$

The finite review pool can therefore be summarized by a sparse signal matrix $R = \{r_{nd}\}_{n,d}$ together with its coverage matrix $M = \{m_{nd}\}_{n,d}$.

Conditional on latent quality and coverage, signs are dimension-specific:

$$\mathbb{P}(r_{nd} = \theta_d \mid \theta_d, m_{nd} = 1) = \rho_d, \qquad \mathbb{P}(r_{nd} = 1 - \theta_d \mid \theta_d, m_{nd} = 1) = 1 - \rho_d, \tag{5.3}$$

with $\rho_d \in (1/2, 1)$. If $m_{nd} = 0$, then $r_{nd} = \varnothing$ deterministically. Across reviews and dimensions, signs are conditionally independent given (θ, M).

The consumer need not know the exact sparse matrix M ex ante, but she knows the finite review count N and a prior distribution $G(M)$ over coverage matrices. The benchmark subjective prior treats latent quality and coverage as independent:

$$\mathbb{P}(\theta, M) = \mathbb{P}(\theta) G(M). \tag{5.4}$$

This factorization is the maintained prior restriction used in (5.10); a correlated quality-coverage prior can be accommodated by replacing the product prior with the appropriate joint distribution in the same integrated likelihood. The large-pool results below separately state when the data-generating coverage law must also be independent of latent quality. After observing the AI summary and opened reviews, the consumer updates jointly over latent quality and over what the remaining unopened reviews are likely to discuss.

5.2 Finite slots: prevalence selection and majority sentiment

Let $K < D$. The AI summary does not attempt to summarize the full D-dimensional review environment. Instead, it reports only the most prevalent dimensions in the finite review pool, up to a maximum of K. For any coverage matrix M, define the dimension-specific prevalence counts

$$C_d(M) = \sum_{n=1}^{N} m_{nd}, \qquad d = 1, \dots, D. \tag{5.5}$$

Let $\mathcal{D}^+(M) = \{d : C_d(M) > 0\}$ be the set of dimensions that appear at least once in the finite pool, and let $L(M) = \min\{K, |\mathcal{D}^+(M)|\}$. Define

$$\mathcal{S}^K(M) \subset \mathcal{D}^+(M), \qquad |\mathcal{S}^K(M)| = L(M), \tag{5.6}$$

to be the set of indices corresponding to the $L(M)$ largest positive values of $C_d(M)$, with ties broken by a fixed rule known to the consumer. When no confusion arises, I suppress the dependence on M and write $\mathcal{S}^K$. Upon entry, the consumer observes the effective summary vector

$$a = (a_d)_{d\in\mathcal{S}^K}, \qquad a_d \in \{0,1\}, \tag{5.7}$$

where a_d is a compressed signal about dimension d. A natural benchmark is that each a_d is the majority-rule overview of the non-missing dimension-d signals:

$$\mathbb{P}(a_d = 1 \mid R, M) = \begin{cases} 1, & \sum_{n:\, m_{nd}=1} r_{nd} > C_d(M)/2, \\ \tau_d, & \sum_{n:\, m_{nd}=1} r_{nd} = C_d(M)/2, \\ 0, & \sum_{n:\, m_{nd}=1} r_{nd} < C_d(M)/2, \end{cases} \qquad d \in \mathcal{S}^K, \tag{5.8}$$

where the benchmark fixes the neutral dimension-level tie-breaking probability $\tau_d = 1/2$, known to the consumer. Thus every reported summary component has positive realized prevalence. If the finite pool contains at least K positive-prevalence dimensions, then $|\mathcal{S}^K(M)| = K$; if it contains fewer, the unreported zero-prevalence dimensions are uninformative null components and are dropped from the effective summary. The consumer knows the aggregation rule, knows that at most the K most prevalent dimensions are summarized, and knows that any posterior about the remaining unopened reviews must be consistent with both of those facts.

Under this neutral componentwise majority benchmark, an exact mixed sentiment state can arise only through a tie in the non-missing signals for a displayed dimension. Hence it has positive probability only when $C_d(M)$ is even. If $C_d(M)$ is odd, majority rule maps every realized dimension-level signal pool into either a favorable or unfavorable summary. A platform can still report a mixed category for odd coverage counts, but doing so corresponds to the wider three-cell threshold rule in Online Appendix C.4.1, not to a pure majority tie.

This specification is intentionally asymmetric. Covered dimensions are summarized for free; uncovered dimensions are not. The summary is therefore informative about common dimensions of discussion, not necessarily about the dimensions a particular consumer values.

For the closed-form analysis, I now specialize the common kernel to top-K prevalence selection and componentwise majority. These rules make it possible to separate what the consumer learns about quality from what she learns about topic arrival. They are analytical benchmarks rather than claims about a deployed platform. Section 6 returns to general known kernels, recovered rules, folding, and three-cell sentiment bands after the benchmark mechanisms have been established.

5.3 What remains to be learned?

After seeing a finite-slot summary, the consumer faces two distinct unknowns. She may still be uncertain about quality on a dimension, and she may be uncertain whether another review will discuss that dimension at all. The posterior must track both. Let h_t denote the history after t opened reviews: it records which reviews have been opened and the full sparse vectors observed on those reviews. Write $\omega = (\mathcal{S}^{\phi}, a)$ for the realized summary output, where $\mathcal{S}^{\phi}$ is the set of dimensions selected for display by rule ϕ and a collects their reported sentiment cells. In the majority benchmark $\mathcal{S}^{\phi} = \mathcal{S}^{K}$; below, a is shorthand for the full observed output whenever its displayed set is clear from context. For any covered dimension $d \in \mathcal{S}^{\phi}$, let

$$\mu_{d,t}(a, h_t) = \mathbb{P}(\theta_d = 1 \mid a, h_t) \tag{5.9}$$

denote the posterior on latent quality dimension d after observing the summary and the first t opened reviews. At entry, before any sparse reviews are opened, write $\mu_{d,0}(a) \equiv \mu_{d,0}(a, \emptyset)$. For uncovered dimensions, the summary does not directly report a compressed sign, but it can still carry indirect information through their nonselection. Under the majority benchmark this is the fact that those dimensions failed to make the top-K set. Formally, the consumer updates jointly over (θ, M) using the integrated likelihood of the observed summary and opened reviews:

$$\mathbb{P}(\theta, M \mid \omega, h_t, z; \phi) \propto \mathbb{P}(\theta)\, G(M)\, \mathcal{L}^{\phi}(\omega, h_t \mid \theta, M, z), \tag{5.10}$$

where

$$\mathcal{L}^{\phi}(\omega, h_t \mid \theta, M, z) = \sum_{R \in \mathcal{R}(M; h_t)} \underbrace{\mathbb{P}(R \mid \theta, M)}_{\text{signal-matrix probability}} \underbrace{Q_{\phi}(\omega \mid R, M, z)}_{\text{summary compatibility}}. \tag{5.11}$$

Here $\mathcal{R}(M; h_t)$ is the finite set of full sparse signal matrices that are consistent with coverage matrix M and the opened-review history h_t. Equation (5.11) is the multi-dimensional analogue of the scalar H-function: it integrates over the unobserved signs in the finite review pool rather than conditioning on an unobserved full signal matrix. Equation (5.10) is therefore the precise updating object behind both the benchmark and the recovered-rule formulation. Under the top-K majority benchmark, Q_{ϕ} imposes prevalence selection and componentwise majority; for a recovered rule it imposes the estimated joint mapping from the corpus to the displayed set and cells. Online Appendix D.1 derives (5.10), its factorized special case, and the related updating objects step by step.

The economic content of this integration is straightforward. The consumer asks which latent qualities and coverage patterns could have produced both the displayed summary and the reviews already opened. She then averages over the unseen corpora that remain possible. This prevents the model from treating an unobserved signal matrix as if the consumer had seen it.

Marginalizing (5.10) over M gives the quality posterior in (5.9).

All subsequent posterior-predictive objects condition on the realized output and the rule; those arguments are suppressed to keep the dynamic program readable. The consumer's expected utility at history (a, h_t) is

$$\delta(a, h_t) = \mathbf{x}'\beta - \alpha P + \sum_{d=1}^{D} \gamma_d \mu_{d,t}(a, h_t). \tag{5.12}$$

Inspection technology is passive: opening one more review means drawing one review uniformly from the unopened finite pool. Let

$$\eta_{d,t}(a, h_t) = \mathbb{P}\big(m_{n_{t+1},d} = 1 \mid a, h_t\big), \tag{5.13}$$

where n_{t+1} denotes a uniformly chosen unopened review. Because the consumer knows the summary rule and the finite review count, she can compute this probability from the joint posterior:

$$\eta_{d,t}(a, h_t) = \mathbb{E}\left[\frac{C_d(M) - c_{d,t}}{N - t} \;\middle|\; a, h_t\right], \tag{5.14}$$

where $c_{d,t}$ is the number of opened reviews up to time t that covered dimension d. This is the exact formal version of the intuition that the consumer keeps reading when there is still a meaningful chance that the next unopened review will speak to an uncovered dimension she cares about. The first main multidimensional theorem, Theorem 2 below, applies this object to top-K selection and shows that omission itself can lower the expected topic-arrival probability.

Define the exact predictive probability that the next covered signal on dimension d is favorable by

$$\pi^{+}_{d,t}(a, h_t) = \mathbb{P}(r_{n_{t+1},d} = 1 \mid m_{n_{t+1},d} = 1, a, h_t). \tag{5.15}$$

This is a posterior-predictive object computed from (5.10); it is not generally equal to $\mu_{d,t}\rho_d + (1 - \mu_{d,t})(1 - \rho_d)$. For a displayed dimension, the summary cell restricts the total number of favorable signs in the finite pool, so the same scalar H-ratio logic must be used. The simple expression

$$\mu_{d,t}(a, h_t)\rho_d + (1 - \mu_{d,t}(a, h_t))(1 - \rho_d)$$

is valid only in the special case in which the event that the next unopened review covers d is not itself informative about θ_d and the summary imposes no additional count restriction on the unobserved signs of dimension d.

The unconditional predictive probability that the next unopened review delivers a favorable signal on dimension d is therefore

$$\psi^{+}_{d,t}(a, h_t) = \underbrace{\eta_{d,t}(a, h_t)}_{\text{chance the topic appears}} \quad \underbrace{\pi^{+}_{d,t}(a, h_t)}_{\substack{\text{positive-sign chance} \\ \text{conditional on coverage}}}, \tag{5.16}$$

which is simply the product rule using the exact conditional probability in (5.15). If topic incidence is itself informative about latent quality, that information is incorporated inside $\pi^+_{d,t}$ through the conditioning event $m_{n_{t+1},d} = 1$.

Equation (5.16) is the heart of the core model. In the single-dimensional model every next review is automatically payoff-relevant. Here, continuation depends not only on how informative a review would be if opened, but also on the posterior probability that the next unopened review will address a utility-relevant dimension at all. Online Appendix D.1.1 records the corresponding tractability result: the posterior and Bellman problem are exact finite-state objects, while sign-invariant componentwise rules admit a binomial-tail or binomial-interval factorization.

5.4 The decision to continue reading

The consumer does not read simply because uncertainty remains. She reads only when the chance of encountering a relevant topic, the information carried by that topic, and its importance for her decision jointly justify the cost. To connect the updating objects above to search behavior, let the within-option state after the summary be $s_t = (a, h_t)$, where h_t records the first t opened sparse reviews. Let $\bar{M}$ denote the continuation value of leaving the current option for the best outside opportunity. The current stopping payoff is

$$S(s_t;\gamma) = \max\left[0, \delta(a, h_t), \bar{M}\right]. \tag{5.17}$$

The within-option Bellman equation is

$$V(s_t) = \max\left\{S(s_t;\gamma),\ -c^W(t+1) + \mathbb{E}[V(s_{t+1}) \mid s_t]\right\}. \tag{5.18}$$

The stopping value is the best of exit, keeping the current option, and taking the outside opportunity. Continuation is costly because another review must be opened before its sparse content is known.

The expectation in (5.18) is over all sparse next-review realizations consistent with the posterior in (5.10). That is precisely where the finite-pool logic matters: after seeing the summary and some opened reviews, the consumer updates not only latent-quality beliefs but also the probability distribution over what the unopened reviews still discuss.

Define the exact dynamic gross value of starting within-option search at state s_t as

$$\bar{c}^{MD}(s_t;\gamma) = \mathbb{E}[V(s_{t+1}) \mid s_t] - S(s_t;\gamma). \tag{5.19}$$

Thus the exact dynamic cutoff is $c^W(t+1) \le \bar{c}^{MD}(s_t;\gamma)$. For an analytical lower bound, consider the feasible policy that opens one more review and then stops.

Let $\mathcal{X}(s_t)$ denote the finite set of sparse next-review realizations that can arrive at state s_t, and let $s_t(x)$ be the post-review state after realization $x \in \mathcal{X}(s_t)$. The exact primitive one-step value of opening one more review and then stopping is

$$\mathcal{I}^{MD}(s_t;\gamma) = \mathbb{E}[S(s_t(X_{t+1});\gamma) \mid s_t] - S(s_t;\gamma), \tag{5.20}$$

where the expectation is taken under the posterior predictive distribution implied by (5.10).

Online Appendix D.1 gives the dynamic cutoff proof, focal omitted-dimension decomposition, and one-step inspection inequality. Omission generates verification only when the omitted information is taste-relevant, still likely to arrive, and capable of changing the stopping decision.

Coverage and residual learnability. Two scalars organize this economic tradeoff. The first asks how much of what the consumer cares about is already displayed. The second asks how much taste-relevant content can still arrive in another review.

Definition 2 (Coverage share)**.** For consumer preference vector γ and summary coverage set $\mathcal{S}^K$, define

$$\mathrm{Cove}(\gamma; K) = \frac{\overbrace{\sum_{d \in \mathcal{S}^K} \gamma_d}^{\text{taste weight on displayed dimensions}}}{\underbrace{\sum_{d=1}^{D} \gamma_d}_{\text{total taste weight}}} \in [0,1]. \tag{5.21}$$

The coverage share measures how much of the consumer's preference mass lies on dimensions explicitly summarized by AI.

Definition 3 (Residual learnability)**.** For consumer preference vector γ and history (a, h_t), define

$$\mathrm{RL}_t(\gamma; a, h_t) = \sum_{d \notin \mathcal{S}^K} \underbrace{\gamma_d}_{\text{taste weight}} \underbrace{\eta_{d,t}(a, h_t)}_{\text{topic-arrival chance}} . \tag{5.22}$$

Residual learnability is the preference-weighted posterior probability that one additional unopened review will speak to an uncovered dimension. It is the simplest scalar object capturing the new friction in (5.16). A consumer may care about an uncovered dimension, but if the remaining unopened reviews are unlikely to discuss it then the incentive to keep reading is small; conversely, a dimension can be omitted by the summary yet still sustain verification if it is both salient to the consumer and sufficiently likely to appear in the remaining finite review pool.

The next results follow this causal chain. Prevalence selection first changes what omission means. The remaining topic-arrival beliefs then interact with the consumer's tastes and stopping margin to determine whether she reads. Finally, those same forces determine which dimensions a consumer-oriented summary would place in its scarce slots.

5.5 Self-concealing omission

Return to the USB-C hub. If charging slowdown is absent from a top-K summary, the consumer learns two things at once: she receives no free sentiment about slowdown, and she infers that slowdown was discussed less often than the displayed topics. The second inference can make a future review look less promising even though slowdown remains important to her. To isolate that selection effect, the comparison benchmark is an *exogenous-display* environment: dimension d is not displayed, but the nondisplay event is statistically independent of the realized coverage count $C_d(M)$. This keeps the absence of free sentiment information fixed while removing the selection information carried by a top-K prevalence rule.

Theorem 2 (Self-concealing omission). Consider the top-K prevalence-selection benchmark with fixed, known tie-breaking at the entry history $t = 0$. Fix a dimension d, and let

$$O_d = \{d \notin \mathcal{S}^K(M)\}$$

be the event that dimension d is omitted from the summary. Suppose the coverage count $C_d(M)$ is independent of the vector of other coverage counts under G, has nondegenerate support, and $0 < \mathbb{P}(O_d) < 1$. Let

$$g_d(c) = \mathbb{P}(O_d \mid C_d(M) = c)$$

denote the induced omission probability. Assume that g_d is not constant on the support of $C_d(M)$. Then top-K selection and the independence of coverage counts imply that $g_d(c)$ is weakly decreasing in c. Bayes' rule therefore gives, at every support point,

$$\frac{Prb(C_d(M) = c \mid O_d)}{Prb(C_d(M) = c)} = \frac{g_d(c)}{\mathbb{P}(O_d)}, \tag{5.23}$$

which is weakly decreasing and nonconstant in c. Thus the posterior coverage count conditional on omission is dominated by the unconditional count in the monotone-likelihood-ratio order:

$$C_d(M) \mid O_d \preceq_{\text{MLR}} C_d(M),$$

where the displayed ordering uses the convention that the conditional-to- unconditional probability-mass ratio in (5.23) is decreasing. Consequently,

$$\mathbb{P}(C_d(M) \geq c \mid O_d) \leq \mathbb{P}(C_d(M) \geq c) \qquad \text{for every } c, \tag{5.24}$$

with strict inequality for at least one cutoff. More generally, for every weakly increasing function h,

$$\mathbb{E}[h(C_d(M)) \mid O_d] \leq \mathbb{E}[h(C_d(M))], \tag{5.25}$$

and the inequality is strict whenever h is strictly increasing on the support. Taking $h(c) = c/N$ yields the original topic-arrival implication:

$$\mathbb{E}\big[\eta_{d,0} \mid O_d\big] = \frac{1}{N}\,\mathbb{E}[C_d(M) \mid O_d] < \frac{1}{N}\,\mathbb{E}[C_d(M)]. \tag{5.26}$$

Thus omission is not merely absence of sentiment information: under a prevalence-based display rule, omission is itself evidence that the omitted dimension has a uniformly lower posterior distribution of latent coverage and is less likely to appear in the remaining review pool.

In the exogenous-display benchmark, nondisplay is independent of coverage, so the conditional distribution of $C_d(M)$ equals its unconditional distribution. Hence (5.25) orders every omitted-dimension verification gain that is increasing in latent coverage, not only its mean. In particular, let $B_d \geq 0$ denote a fixed gross belief-unit gain per unit of taste, conditional on the next review covering dimension d. When the focal one-step gain takes the linear form $\gamma_d \eta_{d,0} B_d$, top-K omission weakly lowers its conditional expectation relative to exogenous display, and strictly lowers it if $B_d > 0$. Moreover, whenever omitted positive-coverage states with $\eta_{d,0} < N^{-1}\,\mathbb{E}[C_d(M)]$ occur with positive probability, there is a positive-probability set of realized omitted states and inspection costs for which the consumer verifies under exogenous display but stops under top-K omission even though the omitted dimension can still be present in the finite review pool.

Proof in Online Appendix D.1.

Why omission shifts the coverage distribution. Holding the other dimensions' counts fixed, increasing $C_d(M)$ can only move dimension d upward in the top-K ranking. Independence keeps the distribution of those competing counts fixed as $C_d(M)$ varies, so the omission probability $g_d(c)$ decreases with c. Bayes' rule then makes the conditional-to-unconditional mass ratio decreasing, which gives MLR dominance, FOSD, and all increasing-function inequalities. The topic-arrival mean is the special case $h(c) = c/N$.

Theorem 2 is the selection counterpart to the scalar compression-consistency theorem. A finite-slot summary does not just choose which dimensions receive free sentiment. Because the chosen dimensions are selected from the same corpus the consumer may later inspect, the omitted set changes beliefs about the usefulness of further review reading. A prevalence-based summary can therefore rationally discourage search on omitted dimensions through a topic-arrival channel, even when the platform is truthful and even when omitted reviews would sometimes contain decision-relevant information.

When the interface reveals displayed coverage counts, the selection result also has a directly measurable implication. If all K slots are filled and c_* is the smallest displayed count, every omitted dimension satisfies

$$C_d(M) \leq c_*, \qquad \eta_{d,t} \leq \frac{c_* - c_{d,t}}{N - t}. \tag{5.27}$$

The inequality is weak because a dimension tied at the last slot can lose the fixed tie-break. If the summary is under-filled, omission instead implies $C_d = 0$. This is an order-statistic support restriction rather than a new behavioral mechanism, so its formal statement and step-by-step proof are in Online Appendix Corollary 9. Empirical use requires exact counts and a common eligible pool; rounded or capped counts yield an interval version of (5.27).

5.6 Who continues reading?

Consumers who see the same truthful summary need not make the same verification choice. The main interior case is the one in which some preference mass lies on covered dimensions and some lies on uncovered dimensions. Then the summary partially resolves relevant uncertainty but leaves a residual incentive to keep reading. Coverage and residual learnability alone do not generically order verification across all possible taste vectors, because omitted dimensions can differ in posterior uncertainty, signal precision, and stopping-payoff wedges. Thus Cove and residual learnability are not advertised as universal sufficient statistics. They are diagnostic objects. Online Appendix D.1 gives the formal one-step ordering that becomes valid under homogeneous omitted-dimension payoff wedges; outside that qualification the same objects should be read as empirical and numerical predictions.

The same objects distinguish three economically different cases. When $\mathrm{Cove}(\gamma; K) = 0$, the summary covers none of the dimensions entering the consumer's utility. It cannot directly move her payoff-relevant quality beliefs; any effect must come from what nondisplay reveals about topic arrival (Online Appendix Proposition 19). When $0 < \mathrm{Cove}(\gamma; K) < 1$, the summary resolves some needs but not others, so verification persists selectively. Finally, when $\mathrm{Cove}(\gamma; K) = 1$ and $\mathrm{RL}_t(\gamma; a, h_t)$ is close to zero, little relevant content remains discoverable and the consumer is pushed back toward the scalar benchmark.

At the one-step level this ordering becomes a taste region. In the one-topic, local-separability specialization, let $\Omega_{d,t}(s_t)$ denote the expected one-step gross payoff gain per unit of taste weight on dimension d, including the probability that the next review covers that dimension. Online Appendix D.1 derives the focal-dimension gain and the resulting taste hyperplane $\sum_d \gamma_d \Omega_{d,t}(s_t) \geq c^W(t+1)$. Consumers whose weights load on dimensions with high residual gains continue, while consumers whose needs are already covered stop. The empirical design must therefore observe or elicit hard needs or attribute weights rather than impose one propensity to read at a given summary state.

5.7 When is a truthful summary sufficient?

Truthful reporting does not by itself imply that the summary is enough for a decision. What matters is whether any feasible completion of the hidden corpus could reverse the consumer's preferred stopping action. This question can be reduced to one taste-weighted index. Write

$$O = \max\{0, \bar{M}\}, \qquad A = \mathbf{x}'\beta - \alpha P, \qquad b = O - A,$$

and define the decision index

$$\Lambda_t \equiv \sum_{d=1}^{D} \gamma_d \, \mu_{d,t}. \tag{5.28}$$

The stopping payoff can then be written as the scalar kinked function

$$S(s_t; \gamma) = O + (\Lambda_t - b)_+, \tag{5.29}$$

with a single kink at the taste hyperplane $\{\mu : \sum_d \gamma_d \mu_d = b\}$. Since each posterior belief is a martingale, Λ_t is a martingale.

Let $\mathfrak{T}(s_t)$ be the finite set of terminal posterior states reached with positive probability after revealing the remaining pool. Its exact decision-index support is

$$\Lambda_t^{T,\min} = \min_{\bar{s} \in \mathfrak{T}(s_t)} \sum_{d=1}^{D} \gamma_d \mu_d(\bar{s}), \qquad \Lambda_t^{T,\max} = \max_{\bar{s} \in \mathfrak{T}(s_t)} \sum_{d=1}^{D} \gamma_d \mu_d(\bar{s}). \tag{5.30}$$

Theorem 3 (Decision insufficiency and residual verification). Fix any known finite truthful summary rule Q_ϕ and any post-summary state $s_t = (\omega, h_t)$. Under the affine expected-utility structure in (5.12), the following statements hold.

(i) If $b \notin (\Lambda_t^{T,\min}, \Lambda_t^{T,\max})$, every reachable index lies on one affine branch of the stopping payoff. Every inspection policy therefore has zero gross value, and strictly positive costs imply immediate stopping.

(ii) If $\Lambda_t^{T,\min} < b < \Lambda_t^{T,\max}$, full revelation has strictly positive gross value

$$\mathcal{J}^T(s_t;\gamma) = \mathbb{E}[S(\bar{s};\gamma) \mid s_t] - S(s_t;\gamma), \qquad \bar{s} \in \mathfrak{T}(s_t), \tag{5.31}$$

and verification starts whenever

$$\sum_{\ell=t+1}^{N} c^W(\ell) < \mathcal{J}^T(s_t;\gamma),$$

because opening the remaining pool and then stopping is feasible. If the terminal interval is nondegenerate, this strict straddle and positive gross value hold on a nonempty open set of type-margin pairs. A truthful summary that is not decision-sufficient therefore cannot eliminate residual value uniformly over stopping margins.

Proof in Online Appendix D.1.

Truthfulness is therefore weaker than decision sufficiency. Omitted or folded information has positive gross value when it leaves terminal index support on both sides of the consumer's margin. Online Appendix Theorem 4 supplies the terminal-hull equality, absorbing zero region, and dynamic sandwich; Online Appendix Section D.1.2 characterizes states on which the one-step cutoff is exact. Together, these results separate states in which hidden evidence can still change the action from states in which uncertainty is behaviorally irrelevant. The affine-index restriction is substantive: with nonlinear utility, continuation generally depends on the full reachable belief polytope.

For the hub example, suppose every corpus still compatible with the summary leaves the hub preferable to the outside option. Uncertainty may remain, but it cannot affect the action, so reading has no value. If compatible corpora imply different choices because charging performance is unresolved, the summary is truthful but decision-insufficient. The consumer then reads when the available improvement is large enough relative to cost.

5.8 Which dimensions deserve scarce slots?

The hyperplane representation also clarifies what a summary should cover if the objective is to reduce costly review reading. A platform that displays the most frequently discussed dimensions is not generally solving the consumer's within-product information problem. Frequency matters only through the posterior probability that a future review will discuss a dimension; the consumer's verification motive also depends on taste weight and on whether learning that dimension can change the stopping action.

For example, compatibility may be discussed in many hub reviews but already be settled for the marginal consumer. Charging slowdown may be mentioned less often yet determine whether she buys. Giving the next slot to compatibility follows prevalence; giving it to charging performance follows consumer relevance.

Let F denote the population distribution of consumer taste vectors γ. In the one-topic, local-separability environment, suppose that covering a dimension changes a fixed componentwise per-unit gain from $\Omega^U_{d,t}$ to $\Omega^C_{d,t}$, with $0 \le \Omega^C_{d,t} \le \Omega^U_{d,t}$. The formal display-set problem is then additive. Online Appendix D.1 states the exact minimization problem and proof. The resulting consumer-relevance index is

$$\mathcal{R}_{d,t} = \underbrace{\mathbb{E}_F[\gamma_d]}_{\text{mean taste weight}} \underbrace{\left(\Omega^U_{d,t}(s_t) - \Omega^C_{d,t}(s_t)\right)}_{\text{verification removed by display}} . \tag{5.32}$$

The K slots that minimize expected one-step inspection pressure are the dimensions with the largest values of this index. A top-K prevalence rule is therefore inspection-pressure optimal only when prevalence ranks dimensions the same way as $\mathcal{R}_{d,t}$.

For the comparison, let $\mathrm{Prev}_d > 0$ denote the prevalence score used to rank dimension d. In the finite-pool benchmark this can be $C_d(M)$ or its share; in the large-pool limit it can be λ_d. Define consumer relevance per unit of prevalence by

$$w_d = \frac{\mathcal{R}_{d,t}}{\mathrm{Prev}_d}. \tag{5.33}$$

Theorem 4 (Sharp performance bound for prevalence ranking). Consider the additive one-topic, local-separability environment just described. Fix $K < D$. Let S^P be any size-K set maximizing total prevalence, and let S^R be any size-K set maximizing total consumer relevance:

$$S^P \in \arg\max_{|S|=K} \sum_{d\in S} \mathrm{Prev}_d, \qquad S^R \in \arg\max_{|S|=K} \sum_{d\in S} \mathcal{R}_{d,t}.$$

(i) (Approximation guarantee.) If relevance per unit of prevalence is bounded as

$$0 < \underline{w} \le w_d \le \overline{w} < \infty \qquad \text{for every } d,$$

then the prevalence-selected set captures at least the fraction $\underline{w}/\overline{w}$ of the maximal inspection-pressure reduction:

$$\frac{\sum_{d\in S^P}\mathcal{R}_{d,t}}{\sum_{d\in S^R}\mathcal{R}_{d,t}} \geq \frac{\underline{w}}{\overline{w}}. \tag{5.34}$$

In particular, if w_d is constant across dimensions, prevalence ranking is exactly optimal for this objective.

(ii) (Sharpness.) The factor $\underline{w}/\overline{w}$ is sharp as a dimension-free guarantee. Already with $K = 1$ and $D = 2$, the achieved ratio can be made arbitrarily close to $\underline{w}/\overline{w}$ from above.

(iii) (No-alignment converse.) Over the unrestricted additive class, in which prevalence imposes no positive lower bound on $\min_d w_d/\max_d w_d$, prevalence has no positive uniform approximation guarantee. For every $K < D$ and $\varepsilon > 0$, there exist positive prevalence scores, a normalized taste distribution, and componentwise covered and uncovered gains satisfying $0 \leq \Omega^C_{d,t} \leq \Omega^U_{d,t}$ such that the top-K prevalence rule uniquely selects S^P, yet

$$\frac{\sum_{d\in S^P}\mathcal{R}_{d,t}}{\max_{|S|=K}\sum_{d\in S}\mathcal{R}_{d,t}} < \varepsilon. \tag{5.35}$$

An omitted dimension can then have arbitrarily greater consumer relevance than a displayed one, so replacing the displayed low-relevance dimension strictly reduces expected inspection pressure.

Proof in Online Appendix D.1.

The theorem identifies exactly what disciplines the use of commonness as a design heuristic. If consumer relevance delivered per unit of prevalence does not vary much across dimensions, prevalence is provably near-optimal. If that alignment ratio is unrestricted, the guarantee collapses to zero: a frequently mentioned dimension can be irrelevant for the marginal consumer's decision, while a less frequent dimension carries the residual verification motive. The failure result is therefore the converse of a quantitative performance bound, not merely an unrestricted counterexample.

This design implication is deliberately not a welfare theorem. It ranks summary slots by their ability to reduce review-opening pressure in the one-step within-product problem. Welfare design also depends on choice quality and on the value of information that remains available through later costly inspection. The important implication is still sharp: prevalence alone is not the primitive consumer-relevance object. The primitive object is prevalence weighted by taste and by the decision-changing residual uncertainty that a summary slot removes.

5.9 Why evidence abundance does not eliminate slot scarcity

One might expect the finite-slot problem to disappear when a product has many reviews. That intuition is correct for a displayed scalar dimension: majority aggregation becomes nearly decisive. It is false for a taste-relevant dimension that never receives a slot. More

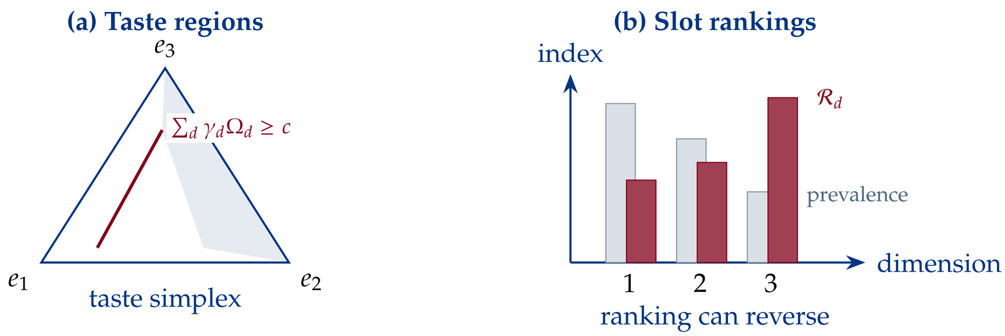

Figure 4: Taste-dependent verification and slot ranking. The left panel shows the taste hyperplane that separates one-step inspection from stopping. The right panel illustrates why raw prevalence need not rank summary slots in the same way as the consumer-relevance index $\mathcal{R}_d$.

evidence improves the precision of what is displayed, but it does not expand display capacity. The next theorem separates these two limits.

Theorem 5 (Slot scarcity in large review pools). Fix D, $K < D$, priors $\mu_{d,0} \in (0, 1)$, and precisions $\rho_d \in (1/2, 1)$. Consider a sequence of products with review-pool size $N \to \infty$. Suppose the top-K summary rule is sign-invariant in dimension selection, coverage is independent of latent quality, and the coverage shares satisfy

$$\frac{C_d(M)}{N} \longrightarrow \lambda_d \in [0, 1]$$

in probability for every d. Assume that at least $K + 1$ limiting shares are positive, that the K-th and $(K + 1)$-st largest positive values of $(\lambda_1, \ldots, \lambda_D)$ are distinct, and let $S^*(\lambda)$ be the corresponding limiting top-K set. Finally, assume the consumer's coverage prior G is correctly specified for this sequence, in the sense that it coincides with the data-generating coverage law used to form the limiting shares. All convergence statements below are under the data-generating law, at the post-summary entry state, and along the high-probability event on which the displayed set is $S^*(\lambda)$ and displayed majority cells equal their latent dimension states.

Then the following hold.

(i) (Displayed dimensions are asymptotically resolved.) With probability approaching one, $\mathcal{S}^K(M) = S^*(\lambda)$. For every $d \in S^*(\lambda)$ with $\lambda_d > 0$, the displayed dimension-level majority posterior converges in probability to the true dimension state:

$$\mu_{d,0}(a) \longrightarrow \theta_d.$$

Consequently, for any fixed number of additional inspections, the gross value of further scalar verification about already displayed dimensions converges to zero.

(ii) (Omitted dimensions retain topic-arrival content.) For every $d \notin S^*(\lambda)$,

$$\eta_{d,0}(a, \emptyset) \to \lambda_d \quad \text{in probability.}$$

Under sign-invariant selection and prior independence across dimensions, omission does not directly reveal the sign of θ_d; it affects the verification problem through topic-arrival probabilities.

(iii) (Pivotal omitted dimensions sustain verification pressure.) Fix an omitted dimension $d^* \notin S^*(\lambda)$ with $\lambda_{d^*} > 0$ and $\gamma_{d^*} > 0$. Suppose that, after taking the large-pool limit for displayed dimensions on the high-probability event described above, the dimension-specific margin

$$\widehat{b}_{d^*} \equiv \frac{b - \sum_{r \in S^*(\lambda)} \gamma_r \theta_r - \sum_{\substack{q \notin S^*(\lambda) \\ q \neq d^*}} \gamma_q \mu_{q,0}}{\gamma_{d^*}} \tag{5.36}$$

lies between the two one-signal posteriors generated by a covered signal on dimension d^*. That is, with $\mu^+_{d^*} = \mathbb{P}(\theta_{d^*} = 1 \mid r_{nd^*} = 1, m_{nd^*} = 1)$ and $\mu^-_{d^*} = \mathbb{P}(\theta_{d^*} = 1 \mid r_{nd^*} = 0, m_{nd^*} = 1)$,

$$\mu^-_{d^*} < \widehat{b}_{d^*} < \mu^+_{d^*}. \tag{5.37}$$

Suppose additionally that the limiting next-review experiment is *focal-separable* for d^*: conditional on covering d^*, it changes no other payoff-relevant posterior component, and conditional on not covering d^*, its expected stopping payoff is weakly at least the current stopping payoff. Then the limiting one-step gross value of a passive review is at least $\lambda_{d^*} J^\infty_{d^*} > 0$, where $J^\infty_{d^*}$ is the covered-d^* Jensen gap derived in the proof. In particular, if the first review-opening cost is strictly below this lower bound, the one-step policy of opening one more review and then stopping strictly dominates immediate stopping for all sufficiently large N; hence verification starts in the full dynamic problem as well.

Moreover, the gross value of allocating one additional *targeted* free summary slot to d^* is bounded away from zero whenever the weaker full-information pivotality condition $0 < \widehat{b}_{d^*} < 1$ holds. Thus large review pools eliminate residual scalar uncertainty on displayed dimensions, but they do not eliminate the economic value of scarce summary slots when omitted dimensions are both payoff-relevant and pivotal.

Proof in Online Appendix D.1.

Why the large-pool limit differs from scalar learning. As N grows, the law of large numbers makes displayed prevalence shares and displayed majority cells converge to their population limits. Conditional on the displayed set, majority cells for displayed dimensions become nearly decisive, so scalar residual verification on those dimensions vanishes. But omitted dimensions are not summarized; their topic-arrival probabilities converge to positive prevalence limits when those dimensions are present in the review corpus. If such an omitted dimension is taste-relevant and pivotal, a focal-separable review that might cover it retains a positive Jensen gap even at large scale.

Theorem 5 is the large-pool version of the paper's main thesis. If the summary were scalar and unlimited, evidence abundance would make post-summary verification disappear. With finite slots, evidence abundance resolves displayed dimensions but does not resolve omitted ones. The relevant scarcity is therefore not the number of reviews; it is the number of summary slots allocated to dimensions consumers may care about. The theorem is also deliberately conditional: if an omitted dimension is not pivotal for the stopping decision, then its residual uncertainty has little behavioral bite. The harmed consumers are those whose tastes load on omitted dimensions that remain both uncertain and decision-changing.

Online Appendix Corollary 14 gives the corresponding capacity threshold. Uniform large-pool elimination follows once capacity eventually covers every positive-share dimension; with a fixed slot shortage, an open set of taste-margin pairs continues to verify at sufficiently low positive costs. The formal statement is kept online because it is the uniform-capacity corollary of Theorem 5, not a separate mechanism.

Two boundary cases locate the finite-slot mechanism. Online Appendix D.4.1 states and proves both. If the summary covers none of the consumer's utility-relevant dimensions, it cannot directly update those quality beliefs; any effect must operate through coverage inference and topic-arrival probabilities. At the opposite knife edge, if one covered dimension is the only payoff-relevant dimension and every review covers it, the multidimensional problem collapses exactly to the scalar benchmark. Between these endpoints lies the paper's main region: summaries partially cover relevant dimensions, and residual topic arrival governs whether verification persists.

6 Beyond the Benchmark Rule

Majority and top-K prevalence selection make the mechanisms transparent; they are not required to define the consumer's problem. Under any known finite truthful kernel Q_ϕ, the integrated likelihood in (2.5) and its sparse-review counterpart (5.11) deliver the exact posterior and predictive transition, and the Bellman problem remains (2.8). Theorem 3 is the rule-independent behavioral result: residual evidence has zero gross value exactly when its terminal decision-index support stays on one side of the stopping margin. Online Appendix Proposition 2 gives the formal scope statement and Online Appendix Theorem 4 gives the corresponding dynamic support geometry.

The additional benchmark structure determines which sharper results follow:

Restriction	Result it supports
Known finite Q_ϕ	Integrated posterior, predictive transition, Bellman cutoff, martingale updating, and decision sufficiency.
Sign-invariant componentwise rule	Dimensionwise likelihood calculations and rule-specific H-objects.
Top-K prevalence selection	Self-concealing omission and observable coverage-support bounds.
Componentwise majority	Binomial closed forms, shared-pool attenuation, and large-pool resolution of displayed dimensions.
Three-cell threshold band	Persistent mixed-cell uncertainty when a latent signal share lies inside the middle band.

Plugging in a recovered rule. If an empirical stage recovers a componentwise count kernel $\widehat{q}_\phi(a \mid Y, N, z)$, define

$$\widehat{H}_a^\phi(m, y; p, z) = \sum_{u=0}^{N-m} \binom{N-m}{u} p^u (1-p)^{N-m-u} \widehat{q}_\phi(a \mid y+u, N, z). \tag{6.1}$$

Replacing H_a^τ with $\widehat{H}_a^\phi$ gives the posterior and transition, after which the Bellman recursion is unchanged. For a joint sparse rule, $\widehat{Q}_\phi$ is inserted directly into (5.11); exactness is preserved, although componentwise binomial factorization can be lost. Online Appendix Section D.1.1 gives the full construction. Estimating Q_ϕ, propagating first-stage uncertainty, and identifying search costs belong to the companion methods paper rather than this theory.

Folding and local robustness. A displayed aspect may pool fine dimensions, such as leakage, temperature, and charging reliability under "functionality." This is a known kernel whose cell is computed from grouped columns of the sparse corpus. The exact integrated posterior remains valid; what is generally lost is componentwise factorization. Online Appendix Proposition 14 shows that, for belief movements small relative to a fixed smoothing scale, local verification value is decision curvature times the taste-weighted posterior variance left by the rule, up to a controlled remainder. Online Appendix Corollary 13 therefore makes folding and nondisplay locally equivalent whenever they leave the same variance in the consumer's within-fold taste direction.

Three-cell contrast. A nondegenerate mixed band replaces the majority tail with a binomial interval. Online Appendix Theorem 3 shows that a displayed dimension is asymptotically resolved when its agreement precision lies beyond the endpoint threshold, but a mixed cell converges to the prior when both state-contingent signal shares lie inside the middle band. Such a cell can therefore sustain verification when it is decision-pivotal.

Under binary majority the middle-band case is absent for $\rho_d > 1/2$. Full finite-pool three-cell likelihoods and the exact-tie caveat are in Online Appendix Section C.4.1.

7 Amazon-Informed Quantitative Illustration

This section asks only whether the mechanisms can be economically nontrivial at empirically plausible evidence volumes. It does not estimate structural precision, tastes, or inspection costs. A preliminary consumer-interface audit contains 8,795 displayed aspect cells on 1,292 products. The median aspect count is 69, the recovered descriptive sentiment thresholds are $\widehat{q}_\ell = 0.2707$ and $\widehat{q}_h = 0.6667$, and the mean modal-sign share is 0.823. The last object is used as an agreement sensitivity input, not as a truth-validated estimate of structural ρ.

Table 2: Scalar verification at representative audit cells

Cell	N	Agreement	Attenuation	$\max \mathcal{I}/\gamma$	Region at $\kappa^W/\gamma = .01$
Negative	71	0.800	77.0%	4.79×10^{-23}	empty
Mixed	73	0.606	—	0.0483	$[0.387, 0.540]$
Positive	66	0.896	48.3%	1.32×10^{-29}	empty

Notes: Agreement is the median modal-sign share within the cell type and is used as a sensitivity calibration. Attenuation is the disconfirming log-odds reduction relative to the raw-signal update. The region is the interval of stopping margins for which one-step verification covers the reported normalized cost under the recovered three-cell thresholds.

The scalar calculation separates decisive and mixed states. At representative positive and negative cells, compression-consistent attenuation is large while the maximum residual scalar verification value is effectively zero. The representative mixed cell instead leaves a maximum normalized value of 0.0483 and a nontrivial decision-margin interval even at normalized cost 0.01. This is a state-contingent implication, not a calibration of population search: the pooled scalar model with one common precision substantially underpredicts the audit's mixed-cell frequency, so structural use requires heterogeneity or a directly controlled laboratory environment.

The finite-slot calculation uses only displayed counts. Among 813 products that fill all eight slots, let c_* be the smallest displayed count and $c_{\max}$ the largest. Exact top-K selection implies $C_d \leq c_*$ for every omitted dimension and the common eligible pool must satisfy $N \geq c_{\max}$, so

$$\eta_{d,0} \leq \frac{c_*}{c_{\max}}.$$

The median product-level bound is 0.164. In the one-topic specialization with omitted-dimension precision 0.70 and taste share 0.50, this gives a median peak upper bound of 0.0082 on verification value normalized by total taste scale. At normalized costs 0.005 and 0.010, respectively, 75.9% and 37.6% of full-slot products are not ruled out by the count-only bound. These are possibility bounds, not estimated opening rates: measuring the eligible review-pool denominator is required to turn them into topic-arrival estimates. Online Appendix Section D.2 derives every formula and records the measurement caveats.

The illustration reinforces the paper's qualitative restriction. Large, decisive scalar cells leave little reason to inspect another binary sign; economically meaningful verification is concentrated in mixed cells and in omitted or folded dimensions that remain likely to arrive and matter for the consumer's decision. Online Appendix Section D.3 records the corresponding laboratory and field predictions without making them part of the present identification task.

8 Conclusion

A consumer facing an AI summary has already received information before she decides whether to read a source. That does not make the source-opening decision trivial. The summary is a compressed view of the same finite evidence she can inspect, so it changes both what she believes about the product and what she expects the hidden evidence to contain. This paper asks when that free, truthful compression is enough for choice and when the consumer still pays to look underneath it.

The scalar benchmark isolates the first answer. Because the summary and raw reviews share an evidence pool, a review cannot be interpreted as though it were independent of the summary. Compression-consistent updating attenuates raw-review belief movements while the summary still restricts the hidden pool; as reviews reveal what was compressed, that informational restriction is depleted. The consumer reads only when the evidence can move her across a relevant stopping margin. Signal precision therefore has two opposing effects: it improves the free summary, reducing residual uncertainty, but it also makes a raw review more useful. Verification can consequently be nonmonotone in precision rather than simply falling as information quality improves. At large review volumes, however, a scalar majority summary becomes nearly decisive. Persistent reading at that scale requires another source of decision-relevant uncertainty.

The multidimensional model provides that source. A finite-slot summary may accurately describe the dimensions it shows and still omit or fold a dimension that matters to a particular consumer. Top-K selection also makes omission informative: a dimension that fails to appear is inferred to be less prevalent and hence less likely to arrive in another review. This self-concealing effect reduces verification, but it need not eliminate it. Reading persists when an omitted or folded dimension remains likely enough to appear, receives enough taste weight, and can change the stopping choice. As the corpus grows with fixed display capacity, majority sentiment on shown dimensions becomes precise while uncertainty about scarce, unshown slots can remain behaviorally relevant. More evidence therefore does not substitute for more summary capacity.

This distinction changes the natural design criterion. Ranking dimensions by prevalence is effective when prevalence, consumer tastes, and residual decision value are sufficiently aligned. Outside that region it can perform poorly, because the most frequently discussed dimension need not be the one whose disclosure saves consumers the most verification.

The consumer-facing object is taste-weighted residual verification value: how much costly, decision-changing learning a slot removes. The quantitative illustration is consistent with this division of labor. Decisive scalar cells leave little residual sign-learning value, whereas mixed cells and omitted dimensions can leave economically nontrivial verification regions. Those calculations discipline magnitudes; they are not estimates of population search costs or claims about a platform's objective.

The sharp comparative statics rely on the transparent majority and top-K benchmarks. The broader machinery requires less: for any known finite truthful compression kernel, integrating over hidden corpora gives the exact posterior, predictive transitions, and within-product Bellman problem. A recovered empirical rule can therefore replace the benchmark kernel, though its behavior need not inherit majority-specific attenuation formulas or top-K omission rankings. The paper's claim is accordingly limited but useful: truthful compression is not equivalent to full disclosure, and its effect on verification depends jointly on shared-pool inference, decision margins, and which dimensions scarce summary capacity resolves.

This within-product theory supplies the demand-side foundation for the companion projects. The numerical paper endogenizes the value of moving to another product; the methods paper studies identification after the deployed compression rule is measured; and the experimental project tests the resulting state-contingent predictions under controlled primitives. Strategic summary design, unknown-rule learning, endogenous reviewing, imperfect authenticity, and fabricated content remain distinct extensions. Keeping them outside the present model preserves the paper's central economic question: when does an honest summary allow a consumer to stop searching, and when does economically important information remain worth finding?

Online Appendix

The separate Online Appendix contains all proofs, step-by-step algebra, aggregation-rule extensions, endpoint cases, operational results, experimental details, and the notation compendium cited throughout the paper.

References

Agarwal, Saharsh, and Ananya Sen. 2026. "Google AI Overviews and Publisher Traffic: Evidence from a Field Experiment." Working paper, SSRN No. 6513059. https://ssrn.com/abstract=6513059.

Bundorf, M. Kate, Maria Polyakova, and Ming Tai-Seale. 2024. "How Do Consumers Interact with Digital Expert Advice? Experimental Evidence from Health Insurance." *Management Science* 70(11): 7617–7643.

Austen-Smith, David, and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review* 90(1): 34–45.

Boland, Philip J. 1989. "Majority Systems and the Condorcet Jury Theorem." *Journal of the Royal Statistical Society: Series D (The Statistician)* 38(3): 181–189.

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105(7): 2183–2203.

Anderson, Simon P., and Régis Renault. 2006. "Advertising Content." *American Economic Review* 96(1): 93–113.

Armstrong, Mark. 2017. "Ordered Consumer Search." *Journal of the European Economic Association* 15(5): 989–1024.

Aybas, Yunus C., and Eray Turkel. 2026. "Persuasion with Coarse Communication." *Economic Theory*, advance online publication.

Benjamin, Daniel J. 2019. "Errors in Probabilistic Reasoning and Judgment Biases." In *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 69–186. Elsevier.

Branco, Fernando, Monic Sun, and J. Miguel Villas-Boas. 2012. "Optimal Search for Product Information." *Management Science* 58(11): 2037–2056.

Bizzotto, Jacopo, Jesper Rüdiger, and Adrien Vigier. 2020. "Testing, Disclosure and Approval." *Journal of Economic Theory* 187: 105002.

Choi, Michael, Anovia Yifan Dai, and Kyungmin Kim. 2018. "Consumer Search and Price Competition." *Econometrica* 86(4): 1257–1281.

Cheng, Lu, Ruocheng Guo, and Huan Liu. 2021. "Estimating Causal Effects of Multi-Aspect Online Reviews with Multi-Modal Proxies." *arXiv:2112.10274*.

Chahrour, Ryan. 2014. "Public Communication and Information Acquisition." *American Economic Journal: Macroeconomics* 6(3): 73–101.

Cheng, Xienan. 2023. "Deterrence, Diversion, and Encouragement: How Freely Provided Information May Distort Learning." Working paper, SSRN No. 4382692.

Dai, Weijia, Ginger Z. Jin, Jungmin Lee, and Michael Luca. 2018. "Aggregation of Consumer Ratings: An Application to Yelp.com." *Quantitative Marketing and Economics* 16(3): 289–339.

Chade, Hector, and Edward Schlee. 2002. "Another Look at the Radner-Stiglitz Nonconcavity in the Value of Information." *Journal of Economic Theory* 107(2): 421–452.

Chevalier, Judith A., and Dina Mayzlin. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research* 43(3): 345–354.

De los Santos, Babur, Ali Hortaçsu, and Matthijs Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review* 102(6): 2955–2980.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. 2015. "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err." *Journal of Experimental Psychology: General* 144(1): 114–126.

Doval, Laura. 2018. "Whether or Not to Open Pandora's Box." *Journal of Economic Theory* 175: 127–158.

Dziuda, Wioletta. 2011. "Strategic Argumentation." *Journal of Economic Theory* 146(4): 1362–1397.

Dye, Ronald A. 1985. "Disclosure of Nonproprietary Information." *Journal of Accounting Research* 23(1): 123–145.

Fang, Lu, Yanyou Chen, Chiara Farronato, Zhe Yuan, and Yitong Wang. 2024. "Platform Information Provision and Consumer Search: A Field Experiment." *NBER Working Paper* 32099.

Gavilan, Diana, Maria Avello, and Gema Martinez-Navarro. 2018. "The Influence of Online Ratings and Reviews on Hotel Booking Consideration." *Tourism Management* 66: 53–61.

Gibbard, Peter. 2022. "A Model of Search with Two Stages of Information Acquisition and Additive Learning." *Management Science* 68(2): 1212–1217.

Giovannoni, Francesco, and Daniel J. Seidmann. 2007. "Secrecy, Two-Sided Bias and the Value of Evidence." *Games and Economic Behavior* 59(2): 296–315.

Greminger, Rafael P. 2022. "Optimal Search and Discovery." *Management Science* 68(5): 3904–3924.

Gu, Chris, and Yike Wang. 2022. "Consumer Online Search with Partially Revealed Information." *Management Science* 68(6): 4215–4235.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24(3): 461–483.

Hauser, John R., and Birger Wernerfelt. 1990. "An Evaluation Cost Model of Consideration Sets." *Journal of Consumer Research* 16(4): 393–408.

Honka, Elisabeth. 2014. "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry." *RAND Journal of Economics* 45(4): 847–884.

Hu, Xingbao (Simon), and Yang Yang. 2020. "Determinants of Consumers' Choices in Hotel Online Searches: A Comparison of Consideration and Booking Stages." *International Journal of Hospitality Management* 86: 102370.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101(6): 2590–2615.

Ke, T. Tony, Zuo-Jun Max Shen, and J. Miguel Villas-Boas. 2016. "Search for Information on Multiple Products." *Management Science* 62(12): 3576–3603.

Ke, T. Tony, and J. Miguel Villas-Boas. 2019. "Optimal Learning Before Choice." *Journal of Economic Theory* 180: 383–437.

Keane, Michael P., and Kenneth I. Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics* 76(4): 648–672.

Kristensen, Dennis, Patrick K. Mogensen, John M. Moon, and Bertel Schjerning. 2021. "Solving Dynamic Discrete Choice Models Using Smoothing and Sieve Methods." *Journal of Econometrics* 223(2): 328–360.

Le Treust, Maël, and Tristan Tomala. 2019. "Persuasion with Limited Communication Capacity." *Journal of Economic Theory* 184: 104940.

Liao, Xiaoye, and Michal Szkup. 2026. "Coordination with Sequential Information Acquisition." *Theoretical Economics* 21(1): 132–166.

Logg, Jennifer M., Julia A. Minson, and Don A. Moore. 2019. "Algorithm Appreciation: People Prefer Algorithmic to Human Judgment." *Organizational Behavior and Human Decision Processes* 151: 90–103.

Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105(1): 272–298.

Matysková, Ludmila, and Alfonso Montes. 2023. "Bayesian Persuasion with Costly Information Acquisition." *Journal of Economic Theory* 211: 105678.

Alavi, Soogand, and Salar Nozari. 2025. "AI Product Review Summaries and Content Homogenization: The Changing Landscape of UGC." Working paper, University of Iowa.

McCall, J. J. 1970. "Economics of Information and Job Search." *Quarterly Journal of Economics* 84(1): 113–126.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12(2): 380–391.

Nitzan, Shmuel, and Jacob Paroush. 1982. "Optimal Decision Rules in Uncertain Dichotomous Choice Situations." *International Economic Review* 23(2): 289–297.

Olszewski, Wojciech, and Richard Weber. 2015. "A More General Pandora Rule?" *Journal of Economic Theory* 160: 429–437.

Saeedi, Maryam, and Ali Shourideh. 2026. "Optimal Rating Design under Moral Hazard." *arXiv:2008.09529*.

Seidmann, Daniel J., and Eyal Winter. 1997. "Strategic Information Transmission with Verifiable Messages." *Econometrica* 65(1): 163–169.

Shin, Hyun Song. 2003. "Disclosures and Asset Returns." *Econometrica* 71(1): 105–133.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy* 69(3): 213–225.

Titova, Maria, and Kun Zhang. 2025. "Persuasion with Verifiable Information." *Journal of Economic Theory* 230: 106102.

Ursu, Raluca M. 2018. "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions." *Marketing Science* 37(4): 530–552.

Ursu, Raluca M., Stephan Seiler, and Elisabeth Honka. 2025. "The Sequential Search Model: A Framework for Empirical Research." *Quantitative Marketing and Economics* 23(1): 165–213.

Vana, Prasad, and Anja Lambrecht. 2021. "The Effect of Individual Online Reviews on Purchase Likelihood." *Marketing Science* 40(4): 708–730.

Wei, Dong. 2021. "Persuasion under Costly Learning." *Journal of Mathematical Economics* 94: 102451.

Yoon, Young-Ro. 2015. "Strategic Behavior in Acquiring and Revealing Costly Private Information." *International Review of Economics & Finance* 39: 133–148.

Weitzman, Martin L. 1979. "Optimal Search for the Best Alternative." *Econometrica* 47(3): 641–654.

Blackwell, David. 1953. "Equivalent Comparisons of Experiments." *Annals of Mathematical Statistics* 24(2): 265–272.

Hopenhayn, Hugo, and Maryam Saeedi. 2026. "Optimal Simple Ratings." *Journal of Industrial Economics*, advance online publication.

Diamond, Peter A. 1971. "A Model of Price Adjustment." *Journal of Economic Theory* 3(2): 156–168.

Nelson, Phillip. 1970. "Information and Consumer Behavior." *Journal of Political Economy* 78(2): 311–329.

Nelson, Phillip. 1974. "Advertising as Information." *Journal of Political Economy* 82(4): 729–754.

Roberts, John H., and James M. Lattin. 1991. "Development and Testing of a Model of Consideration Set Composition." *Journal of Marketing Research* 28(4): 429–440.

Stahl, Dale O. 1989. "Oligopolistic Pricing with Sequential Consumer Search." *American Economic Review* 79(4): 700–712.

Wolinsky, Asher. 1986. "True Monopolistic Competition as a Result of Imperfect Information." *Quarterly Journal of Economics* 101(3): 493–511.