R-HIBR: RELIABILITY-CONSTRAINED EVIDENCE REVISION FOR SEQUENTIAL RECOMMENDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

When should auxiliary evidence be allowed to revise a contextual recommendation, and by how much? We study this question as reliability-constrained evidence revision. R-HIBR retains a semantic-initialized contextual anchor and constructs transition and historical-semantic evidence whose score contributions reflect support and behavioral relevance. Bayesian posterior-predictive lift measures successor evidence beyond catalog prevalence, while deterministic support attenuation controls its multi-horizon contribution. Training transition discrimination calibrates historical semantics, and a persistence probe informs bounded contextual routing. The resulting analysis links evidence estimation to a score-space intervention budget: sufficiently separated rankings are preserved, a pairwise evidence advantage can overcome an anchor deficit, and jointly vanishing corrections recover the anchor. Calibration perturbations are bounded through both revision paths. Under matched full-catalog evaluation, R-HIBR leads all 12 Amazon HR/NDCG metrics and all four OnlineRetail metrics. Five-dataset ablations, 45 sensitivity evaluations, and a four-encoder transfer study characterize how evidence construction and intervention control contribute to the final ranking.

1 INTRODUCTION

Sequential recommendation uses ordered interactions to form a contextual ranking of the next item. Beyond this contextual prediction, observed transitions identify recurring successors and language representations expose semantic relations between items. These sources can refine a ranking, but frequency and similarity alone do not establish how much influence they should have. A sparse transition may be accidental, and a close semantic neighbor may be an unlikely successor. We therefore ask: *when should auxiliary evidence be allowed to revise a contextual ranking, and by how much?*

We formulate this question as *reliability-constrained evidence revision*, with three challenges. **C1: Evidence estimation under sparse support.** Transition evidence must distinguish successor structure from catalog prevalence while accounting for limited observations. **C2: Semantic relevance under behavioral mismatch.** A semantic space must align with the domain's observed successors, and historical intent must remain relevant to the current prefix. **C3: Bounded intervention in ranking space.** Useful evidence must reach the final decision through an explicit, context-dependent revision budget. Together, these challenges separate evidence estimation, relevance assessment, and permission to alter an existing contextual prediction.

We introduce R-HIBR as a realization of this principle. Bayesian-shrunk multi-horizon renewal estimates transition-specific lift and attenuates weakly supported contributions for **C1**. Training transition discrimination and a semantic persistence probe assess historical language evidence for **C2**. A bounded router regulates both corrections around the contextual anchor for **C3**. The theory connects posterior shrinkage to intervention budgets, conditional ranking preservation and correction, and complete-score calibration stability. Five-dataset comparisons and ablations, calibration sensitivity, and semantic-source transfer examine the resulting evidence-to-decision chain.

Our contributions are:

- We formulate reliability-constrained evidence revision, separating contextual prediction from evidence estimation, relevance assessment, and bounded ranking intervention.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

- We construct posterior-predictive transition lift and behaviorally calibrated historical-semantic evidence, with support attenuation and contextual control of their score contributions.
- We connect posterior shrinkage and finite intervention budgets to ranking preservation, sufficient correction conditions, anchor recovery, and complete-score calibration sensitivity.
- We report 20 matched primary HR/NDCG comparisons and characterize evidence construction and regulation through five-dataset ablations, 45 sensitivity evaluations, and four semantic encoders.

2 Related Work

Sequential recommendation. Sequential models use Markov factorization and recurrent dynamics (Rendle et al., 2010; Hidasi et al., 2016), attention (Kang & McAuley, 2018; Sun et al., 2019), or frequency filtering (Zhou et al., 2022). MPT pretrains causal sequence processing on synthetic Markov chains (Xu et al., 2026). These approaches supply contextual predictions; R-HIBR studies how explicit transition and semantic evidence can revise a contextual anchor under support and intervention constraints.

Language representations for recommendation. Universal and text-centric pretraining aligns language with interaction sequences (Hou et al., 2022; Li et al., 2023b); LLaRA adapts language representations for recommendation (Liao et al., 2024), while Sentence-T5 and E5 provide semantic embeddings (Ni et al., 2022; Wang et al., 2022). R-HIBR retains semantic initialization of the anchor and additionally assesses the behavioral relevance and bounded ranking contribution of explicit historical-language evidence.

3 Method

R-HIBR realizes evidence revision in four steps: form a contextual anchor, estimate transition evidence, assess historical-semantic relevance, and bound their contributions to one catalog-wide decision. Reliability enters through posterior shrinkage, deterministic support attenuation, and training-derived semantic calibration. Sections 3.1–3.4 follow this evidence-to-ranking computation.

3.1 Problem setup and contextual anchor

Let $\mathcal{U}$ and $\mathcal{V}$ denote the user and item sets, and let $N = |\mathcal{V}|$ be the catalog size. User $u \in \mathcal{U}$ has a chronological sequence $S_u = (i_1^u, \ldots, i_{L_u}^u)$, where L_u is its length and $i_t^u \in \mathcal{V}$ is the item at event index t. Given the observed prefix $S_{u,1:t}$, the task ranks every $j \in \mathcal{V}$ for the next item i_{t+1}^u. We omit the superscript u when discussing one sequence.

Write $\mathcal{N}(\mathbf{v}) = \mathbf{v}/\max\{\|\mathbf{v}\|_2, \epsilon_{\mathrm{n}}\}$ for normalization, with $\epsilon_{\mathrm{n}} = 10^{-12}$. Each item j has a frozen, normalized language vector $\boldsymbol{s}_j \in \mathbb{R}^{d_s}$, where d_s is its feature dimension. An adaptor $g_\theta : \mathbb{R}^{d_s} \to \mathbb{R}^d$ produces $\boldsymbol{e}_j = g_\theta(\boldsymbol{s}_j)$ in hidden dimension d. A causal encoder F_Θ yields

$$\boldsymbol{h}_t = \mathcal{N}(F_\Theta(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_t})), \qquad a_t(j) = \frac{\langle \boldsymbol{h}_t, \bar{\boldsymbol{e}}_j \rangle}{\tau}, \tag{1}$$

where $\bar{\boldsymbol{e}}_j = \mathcal{N}(\boldsymbol{e}_j)$, $\tau > 0$ is the temperature, and $a_t(j)$ is the contextual anchor score. The causal encoder transforms the ordered prefix into a sequence-dependent state, making the anchor contextual. Its semantic field also supplies the historical-intent correction below; the two uses provide contextual encoding and explicit preference matching, respectively. The adaptor and encoder are initialized from an MPT checkpoint and remain fixed during R-HIBR adaptation.

3.2 Bayesian-shrunk multi-horizon renewal

All transition statistics are estimated from training sequences. Let $c_\ell(i, j)$ count occurrences in which item j appears ℓ events after item i, and let $c_\ell(i) = \sum_{j \in \mathcal{V}} c_\ell(i, j)$. Let $c(j)$ be the training frequency

of item j. The smoothed catalog probability is

$$p_0(j) = \frac{c(j) + \frac{1}{2}}{\sum_{v \in \mathcal{V}} c(v) + \frac{1}{2}N}. \tag{2}$$

Let D be the maximum transition horizon. For lag $\ell \in \{1, \ldots, D\}$, source i, and prior mass $\alpha > 0$, a Dirichlet prior centered on the fixed training estimate p_0 gives the posterior predictive probability

$$\widehat{p}_\ell(j \mid i) = \frac{c_\ell(i, j) + \alpha p_0(j)}{c_\ell(i) + \alpha}. \tag{3}$$

For support budget K, the set $\mathcal{C}_\ell(i)$ contains at most K observed successors, ordered by decreasing transition count with item-ID tie breaking. For $j \in \mathcal{C}_\ell(i)$, define its posterior information lift

$$e_\ell(i, j) = \mathrm{clip}_{[0, B_{\max}]}(\log \widehat{p}_\ell(j \mid i) - \log p_0(j)), \tag{4}$$

where $B_{\max} > 0$ caps the correction. The raw ratio compares source-conditioned posterior prediction with catalog prevalence; clipping retains positive, bounded transition-specific lift rather than rewarding successor frequency alone.

Equation 3 is the Bayesian estimation step. After this posterior is formed, deterministic support factors regulate its contribution to the ranking score:

$$r_\ell^{\mathrm{src}}(i) = \frac{c_\ell(i)}{c_\ell(i) + \lambda_{\mathrm{src}}}, \qquad r_\ell^{\mathrm{pair}}(i, j) = \frac{c_\ell(i, j)}{c_\ell(i, j) + \lambda_{\mathrm{pair}}}, \tag{5}$$

where $\lambda_{\mathrm{src}}, \lambda_{\mathrm{pair}} > 0$ set the support scales for score attenuation. They regulate the influence of the lift after probability estimation. Let $q_{\ell,i}(j) \in \{1, \ldots, K\}$ be the position of j in $\mathcal{C}_\ell(i)$ and define the rank-survival bonus

$$v_{\ell,i}(j) = \lambda_{\mathrm{rs}} \log \frac{K + \frac{1}{2}}{q_{\ell,i}(j) - \frac{1}{2}}, \tag{6}$$

where $\lambda_{\mathrm{rs}} \geq 0$ controls this ranking bias toward frequently supported successors. With $D_t = \min(D, t)$ available history items and indicator $\mathbb{I}[\cdot]$, the renewal energy is

$$b_t(j) = \sum_{m=0}^{D_t - 1} \eta^m \mathbb{I}[j \in \mathcal{C}_{m+1}(i_{t-m})]\, r_{m+1}^{\mathrm{src}}(i_{t-m}) r_{m+1}^{\mathrm{pair}}(i_{t-m}, j) \cdot \left[e_{m+1}(i_{t-m}, j) + v_{m+1, i_{t-m}}(j)\right], \tag{7}$$

where $\eta \in (0, 1)$ discounts longer horizons. This addresses **C1** by separating posterior evidence estimation from support-constrained score intervention; the rank bonus favors high-support successors within each lag.

3.3 Behaviorally calibrated semantic evidence

The semantic history intent is the normalized recency-weighted average

$$\boldsymbol{\mu}_t = \mathcal{N}\left(\sum_{m=0}^{D_t - 1} \beta^m \boldsymbol{s}_{i_{t-m}}\right), \qquad m_t(j) = \langle \boldsymbol{\mu}_t, \boldsymbol{s}_j \rangle, \tag{8}$$

where $\beta \in (0, 1)$ is the recency factor and $m_t(j) \in [-1, 1]$ is the historical-intent score. An auxiliary head $g_\psi : \mathbb{R}^d \to \mathbb{R}^{d_s}$ forms a *semantic persistence probe*, $\boldsymbol{\nu}_t = \mathcal{N}(g_\psi(\boldsymbol{h}_t))$. Next-item semantic supervision gives it a sequence-implied future direction. Its agreement with $\boldsymbol{\mu}_t$ provides a diagnostic routing feature for the historical correction; candidate scores use $m_t(j)$.

Global semantic credibility is estimated only from training transitions. Let P be the number of sampled adjacent transitions, let (x_p, y_p) be the p-th realized source–successor pair, and let $z_{p,r}$ be its r-th deterministic negative for $r \in \{1, \ldots, R\}$. The contrastive semantic gap is

$$\Delta_{\mathrm{sem}} = \frac{1}{P} \sum_{p=1}^{P} \left[\langle \boldsymbol{s}_{x_p}, \boldsymbol{s}_{y_p} \rangle - \frac{1}{R} \sum_{r=1}^{R} \langle \boldsymbol{s}_{x_p}, \boldsymbol{s}_{z_{p,r}} \rangle \right]. \tag{9}$$

Let $\sigma(x) = (1 + e^{-x})^{-1}$. For calibration center $c \in \mathbb{R}$ and sharpness $\kappa > 0$, the global semantic reliability is

$$\gamma = \sigma(\kappa(\Delta_{\text{sem}} - c)) \in (0, 1). \tag{10}$$

The margin measures semantic discrimination of observed training successors, grounding credibility in domain-specific behavioral alignment. The encoder-transfer study uses the order-based calibration in Appendix B.1, substituted for γ throughout the same scoring and routing equations.

For a prefix, define history coverage $x_t^{\text{hist}} = D_t/D$, semantic consistency $x_t^{\text{sem}} = D_t^{-1} \sum_{m=0}^{D_t-1} \langle \boldsymbol{s}_{i_{t-m}}, \boldsymbol{\mu}_t \rangle$, and a normalized transition-support statistic

$$x_t^{\text{sup}} = \tanh\left(\frac{\log\left(1 + \sum_{m=0}^{D_t-1} \eta^m c_{m+1}(i_{t-m})\right)}{\chi} \right), \tag{11}$$

where $\chi = 5$ is the log-support normalization scale. This deterministic statistic summarizes discounted transition support along the prefix, separately from posterior probability estimation. Probe agreement is $x_t^{\text{agree}} = \langle \boldsymbol{\nu}_t, \boldsymbol{\mu}_t \rangle$. The router receives

$$\boldsymbol{x}_t = [x_t^{\text{hist}},\ \gamma x_t^{\text{sem}},\ x_t^{\text{sup}},\ \gamma x_t^{\text{agree}}]^\top \in \mathbb{R}^4. \tag{12}$$

3.4 Bounded ranking revision and learning

Let $W_r \in \mathbb{R}^{2\times 4}$ and $\boldsymbol{b}_r \in \mathbb{R}^2$ denote the active renewal and semantic rows of the residual router. For support center $q_0 \in (0, 1)$ and slope $\zeta > 0$, its outputs are

$$\begin{bmatrix} z_t^{\text{B}} \\ z_t^{\text{S}} \end{bmatrix} = \tanh\left(W_r \boldsymbol{x}_t + \boldsymbol{b}_r + \begin{bmatrix} 0 \\ \zeta\gamma(q_0 - x_t^{\text{sup}}) \end{bmatrix} \right). \tag{13}$$

The analytic component contributes a positive semantic logit below q_0 and a negative one above it. The learned residual jointly adjusts both coefficients using all four features, including support. Together they form a support-aware, bounded contextual modulator. Its final catalog energy is

$$\ell_t(j) = a_t(j) + \lambda_{\text{B}}(1 + \delta z_t^{\text{B}}) b_t(j) + \gamma\lambda_{\text{S}}(1 + \delta z_t^{\text{S}}) m_t(j), \tag{14}$$

where $\lambda_{\text{B}}, \lambda_{\text{S}} > 0$ are base scales and $\delta \in [0, 1)$ is the routing range. This is *multi-stage global semantic calibration*: γ scales semantic-related router features, the analytic semantic bias, and the explicit semantic contribution. One training-derived statistic therefore couples both contextual routes and semantic score magnitude to the usefulness of the semantic space.

For ground-truth successor $y_t = i_{t+1}$, the training objective is

$$\mathcal{L} = -\log \frac{\exp(\ell_t(y_t))}{\sum_{j\in\mathcal{V}} \exp(\ell_t(j))} + \lambda_{\text{I}} \left(1 - \langle \boldsymbol{\nu}_t, \boldsymbol{s}_{y_t} \rangle\right), \tag{15}$$

where $\lambda_{\text{I}} \geq 0$ controls probe alignment. We use a *two-stage adaptation–calibration protocol*. First, the probe and residual router are adapted with $\gamma = 1$, keeping semantic vectors, adaptor, encoder, and base scales fixed. After checkpoint selection, the training-only statistic determines γ for evaluation of that fixed checkpoint. This separates contextual adaptation from dataset-level calibration. Equation 14 produces one full-catalog ranking.

4 Reliability and Ranking Revision Guarantees

We connect evidence reliability to bounded intervention, conditional correction, and calibration stability. Let $\mathcal{E}_t \subseteq \mathcal{V}$ be the common eligible-item set after evaluation masking. For $1 \leq k < |\mathcal{E}_t|$, write $\text{Top}_k(f)$ for the top-k item set under score f and $\text{gap}_k(f)$ for its k-th minus $(k+1)$-st largest score, using item-ID tie breaking. Proofs appear in Appendix A.

4.1 Support-controlled transition evidence

Proposition 1 (Posterior shrinkage and information lift). *For fixed p_0 and $c_\ell(i) > 0$, define the empirical distribution $\widetilde{p}_\ell(j \mid i) = c_\ell(i, j)/c_\ell(i)$ and shrinkage coefficient $\omega_\ell(i) = c_\ell(i)/(c_\ell(i) + \alpha)$.*

Then

$$\widehat{p}_\ell(j \mid i) = (1 - \omega_\ell(i))p_0(j) + \omega_\ell(i)\widetilde{p}_\ell(j \mid i),$$
$$\log \frac{\widehat{p}_\ell(j \mid i)}{p_0(j)} = \log\left[1 + \omega_\ell(i)\left(\frac{\widetilde{p}_\ell(j \mid i)}{p_0(j)} - 1\right)\right]. \tag{16}$$

At a fixed empirical ratio above one, raw lift increases with $\omega_\ell(i)$ and tends to zero as $\omega_\ell(i) \to 0$; clipping preserves nondecreasing dependence. At zero source count, $\widehat{p}_\ell = p_0$ and this source–horizon pair contributes no renewal evidence.

The mixture coefficient controls departures from catalog prevalence during estimation; the source–pair factors then attenuate their score contribution.

Lemma 1 (Finite renewal evidence)**.** *For every prefix and candidate, Equation 7 satisfies*

$$0 \le b_t(j) \le B_{\mathrm{R}} := (B_{\max} + \lambda_{\mathrm{rs}} \log(2K+1)) \frac{1-\eta^D}{1-\eta}. \tag{17}$$

4.2 Bounded intervention, preservation, and correction

Proposition 2 (Anchor and semantic correction budgets)**.** *Fix a prefix and its realized routes. Set $\Lambda_t = \lambda_{\mathrm{B}}(1 + \delta z_t^{\mathrm{B}})$, $\epsilon_t = \gamma\lambda_{\mathrm{S}}(1 + \delta z_t^{\mathrm{S}})$, and $Q_t(j) = a_t(j) + \Lambda_t b_t(j)$. Let b_t^- and b_t^+ be the minimum and maximum renewal scores over $\mathcal{E}_t$. Then, for $j \in \mathcal{E}_t$,*

$$\Lambda_t b_t^- - \epsilon_t \le \ell_t(j) - a_t(j) \le \Lambda_t b_t^+ + \epsilon_t,$$
$$|\ell_t(j) - Q_t(j)| \le \epsilon_t. \tag{18}$$

In particular, the total correction is finite:

$$\max_{j \in \mathcal{E}_t} |\ell_t(j) - a_t(j)| \le \Lambda_t b_t^+ + \epsilon_t \le (1+\delta)(\lambda_{\mathrm{B}} B_{\mathrm{R}} + \gamma\lambda_{\mathrm{S}}). \tag{19}$$

If $\mathrm{gap}_k(a_t) > \Lambda_t(b_t^+ - b_t^-) + 2\epsilon_t$, then $\mathrm{Top}_k(\ell_t) = \mathrm{Top}_k(a_t)$. If $\mathrm{gap}_k(Q_t) > 2\epsilon_t$, then $\mathrm{Top}_k(\ell_t) = \mathrm{Top}_k(Q_t)$.

Equation 19 defines a prefix-dependent intervention neighborhood around the anchor. If $b_t^+ \to 0$ and $\epsilon_t \to 0$, the complete score recovers the anchor uniformly. The reference Q_t retains the contextual anchor and realized renewal route.

Proposition 3 (Evidence-supported ranking correction)**.** *For distinct eligible items y, j, define their reliability-weighted evidence advantage*

$$\mathcal{A}_t(y, j) = \Lambda_t[b_t(y) - b_t(j)] + \epsilon_t[m_t(y) - m_t(j)]. \tag{20}$$

The condition $\mathcal{A}_t(y, j) > a_t(j) - a_t(y)$ is equivalent to $\ell_t(y) > \ell_t(j)$, correcting the pair when next item y lies below j under the anchor. Appendix A.4 gives a catalog-wide top-k entry condition.

Preservation and correction share one budget: the realized evidence advantage determines whether a particular anchor deficit can be crossed.

4.3 Calibration sensitivity of the complete score

Proposition 4 (Deterministic calibration perturbation)**.** *Fix the prefix, semantic field, transition statistics, and adapted parameters. Write $\ell_t(j;\gamma)$ for Equation 14 with all occurrences of γ varied together. Let $w_{12}, w_{14}, w_{22}, w_{24}$ be the indicated entries of W_r, and define*

$$L_t^{\mathrm{B}} = |w_{12}x_t^{\mathrm{sem}} + w_{14}x_t^{\mathrm{agree}}|,$$
$$L_t^{\mathrm{S}} = |w_{22}x_t^{\mathrm{sem}} + w_{24}x_t^{\mathrm{agree}} + \zeta(q_0 - x_t^{\mathrm{sup}})|. \tag{21}$$

For any $\gamma, \gamma' \in [0,1]$,

$$\max_{j \in \mathcal{E}_t} |\ell_t(j;\gamma) - \ell_t(j;\gamma')| \le C_t|\gamma - \gamma'|,$$
$$C_t = \lambda_{\mathrm{B}}\delta b_t^+ L_t^{\mathrm{B}} + \lambda_{\mathrm{S}}(1 + \delta + \delta L_t^{\mathrm{S}}). \tag{22}$$

Thus $\mathrm{gap}_k(\ell_t(\cdot;\gamma')) > 2C_t|\gamma - \gamma'|$ suffices to preserve the top-k set under the calibration change.

The first term of C_t accounts for the changing renewal route; the second covers semantic routing and its outer scale. For main-study gap values Δ_{sem} and Δ'_{sem} with fixed c, κ, their calibrated values satisfy $|\gamma - \gamma'| \le \min\{1, \kappa|\Delta_{\mathrm{sem}} - \Delta'_{\mathrm{sem}}|/4\}$. This connects a change in the measured semantic gap to both corrections and the final ranking.

5 EXPERIMENTS

Our evaluation follows the evidence-revision principle: same-setting comparisons assess the final ranking, sensitivity and semantic-source transfer examine its operating conditions, and ablations distinguish evidence construction from intervention control. Ranking-depth profiles connect these studies to one full-catalog decision.

5.1 EXPERIMENTAL SETUP

Datasets and protocol. We evaluate Amazon Beauty, Sports, and Toys, together with Yelp2018-10core and OnlineRetail. All datasets use chronological leave-one-out: the final interaction is test, the penultimate interaction is validation, and earlier interactions form the training prefix. Evaluation ranks the full catalog with observed-history items masked.

Table 1: Dataset statistics under the shared evaluation protocol.

Dataset	Users	Items	Interactions	Avg. length
Beauty	22,363	12,101	198,502	8.88
Sports	35,598	18,357	296,337	8.32
Toys	19,412	11,924	167,597	8.63
Yelp2018-10core	77,277	45,638	2,103,896	27.23
OnlineRetail	16,520	3,469	519,906	31.47

Table 1 contrasts short Amazon histories with longer Yelp and OnlineRetail sequences. The variation in catalog size and interaction volume provides distinct settings for examining how transition support and semantic context shape the two corrections.

Baselines. The baselines are GRU4Rec (Hidasi et al., 2016), SASRec (Kang & McAuley, 2018), FMLPRec (Zhou et al., 2022), HSTU (Zhai et al., 2024), SASRec+ (Klenitskiy & Vasilev, 2023), UniSRec (Hou et al., 2022), Recformer (Li et al., 2023a), E4SRec (Li et al., 2023b), Qwen2.5 (Yang et al., 2024), MPT-Adaptor (Xu et al., 2026), and MPT-LoRA (Xu et al., 2026). All comparisons share the data splits, candidate construction, history masking, and metric definitions.

Evaluation. We report full-catalog Hit Rate (HR) and normalized discounted cumulative gain (NDCG) at cutoffs 10 and 20. Ranking profiles also include cutoff 5 and mean reciprocal rank (MRR); other studies report NDCG@10. Sentence-T5 provides the default frozen semantics; implementation details appear in Appendix B.2.

5.2 SAME-PROTOCOL BASELINE COMPARISON

Table 2: Full-catalog comparison on three Amazon datasets under the shared MPT setting (Xu et al., 2026). H/N denote HR/NDCG; red/blue bold indicate best/second best. Improve reports gains over the strongest baseline; −− denotes no gain.

	Beauty				Sports				Toys			
Method	H@10	N@10	H@20	N@20	H@10	N@10	H@20	N@20	H@10	N@10	H@20	N@20
GRU4Rec	0.0524	0.0278	0.0788	0.0344	0.0259	0.0135	0.0407	0.0172	0.0490	0.0282	0.0704	0.0335
SASRec	0.0753	0.0417	0.1060	0.0495	0.0405	0.0212	0.0607	0.0263	0.0846	0.0493	0.1115	0.0561
FMLPRec	0.0775	0.0454	0.1072	0.0529	0.0413	0.0238	0.0593	0.0283	0.0855	0.0521	0.1124	0.0589
HSTU	0.0902	0.0535	0.1246	0.0621	0.0550	0.0317	0.0788	0.0377	0.0981	0.0610	0.1288	0.0687
SASRec+	0.0982	0.0591	0.1348	0.0683	0.0576	0.0337	0.0824	0.0399	0.1040	0.0645	0.1393	0.0733
UniSRec	0.0803	0.0382	0.1233	0.0490	0.0446	0.0221	0.0699	0.0285	0.0829	0.0380	0.1235	0.0483
Recformer	0.0676	0.0333	0.1017	0.0419	0.0345	0.0167	0.0555	0.0220	0.0842	0.0405	0.1221	0.0501
E4SRec	0.0773	0.0469	0.1046	0.0538	0.0387	0.0218	0.0563	0.0262	0.0798	0.0492	0.1061	0.0558
Qwen2.5	0.0826	0.0464	0.1200	0.0558	0.0473	0.0256	0.0712	0.0316	0.0784	0.0433	0.1114	0.0515
MPT-Adaptor	**0.1118**	**0.0673**	**0.1531**	**0.0777**	**0.0651**	**0.0373**	0.0924	**0.0442**	**0.1202**	**0.0705**	**0.1617**	**0.0809**
MPT-LoRA	0.1100	0.0658	0.1514	0.0763	0.0644	0.0365	**0.0927**	0.0436	0.1192	0.0701	0.1611	0.0806
R-HIBR	**0.1174**	**0.0717**	**0.1587**	**0.0822**	**0.0715**	**0.0417**	**0.1007**	**0.0491**	**0.1337**	**0.0824**	**0.1763**	**0.0931**
Improve	+5.01%	+6.54%	+3.66%	+5.79%	+9.83%	+11.80%	+8.63%	+11.09%	+11.23%	+16.88%	+9.03%	+15.08%

Table 2 places R-HIBR first on all 12 matched HR/NDCG metrics. Gains in both hit-based and position-sensitive metrics show that the revised ranking benefits retrieval and discounted order across the three product domains. This is consistent with combining transition-specific lift beyond catalog

prevalence with behaviorally calibrated historical intent. The comparison establishes the effectiveness of the complete revision rule; the component study subsequently distinguishes the contribution of evidence construction from its regulation.

Table 3: Full-catalog comparison on OnlineRetail and Yelp under the same setting. Red/blue bold indicate best/second best. Improve reports gains over the strongest baseline; −− denotes no gain.

	OnlineRetail				Yelp			
Method	HR@10	NDCG@10	HR@20	NDCG@20	HR@10	NDCG@10	HR@20	NDCG@20
GRU4Rec	0.1175	0.0561	0.1931	0.0751	0.0246	0.0116	0.0448	0.0166
SASRec	0.1245	0.0524	0.2014	0.0718	0.0298	0.0123	0.0539	0.0184
FMLPRec	0.1333	0.0548	0.2237	0.0775	0.0334	0.0138	0.0606	0.0207
HSTU	0.1452	0.0656	0.2304	0.0870	0.0453	0.0215	0.0773	0.0295
SASRec+	0.1589	0.0719	**0.2575**	0.0968	**0.0466**	0.0209	**0.0796**	0.0292
UniSRec	0.1424	0.0662	0.2275	0.0877	0.0343	0.0170	0.0567	0.0226
Recformer	0.1407	0.0683	0.2194	0.0881	0.0247	0.0121	0.0429	0.0166
E4SRec	0.1551	**0.0749**	0.2528	**0.0995**	0.0291	0.0139	0.0513	0.0195
Qwen2.5	0.1503	0.0731	0.2356	0.0945	0.0411	0.0204	0.0682	0.0272
MPT-Adaptor	**0.1608**	0.0741	0.2572	0.0982	0.0461	0.0224	0.0783	0.0305
MPT-LoRA	0.1568	0.0746	0.2502	0.0980	**0.0495**	**0.0241**	**0.0832**	**0.0325**
R-HIBR	**0.1748**	**0.0897**	**0.2645**	**0.1123**	0.0465	**0.0239**	0.0763	**0.0314**
Improve	+8.71%	+19.76%	+2.72%	+12.86%	−−	−−	−−	−−

Table 3 extends evidence revision to transaction and review histories. R-HIBR leads all four OnlineRetail metrics and achieves the second-highest Yelp NDCG at both cutoffs. Renewal removal in Table 8 reduces NDCG@10 by 20.44% and 9.96%, respectively. These changes are consistent with **C1**: explicitly estimated successor evidence contributes useful ranking revisions across distinct interaction regimes, with posterior lift and support attenuation acting together.

5.3 Robustness and Semantic-Source Transfer

Table 4: Complete test metrics for R-HIBR with Sentence-T5 semantics.

Dataset	HR@5	NDCG@5	HR@10	NDCG@10	HR@20	NDCG@20	MRR
OnlineRetail	0.1093	0.0686	0.1748	0.0897	0.2645	0.1123	0.0783
Beauty	0.0865	0.0618	0.1174	0.0717	0.1587	0.0822	0.0650
Sports	0.0504	0.0349	0.0715	0.0417	0.1007	0.0491	0.0382
Toys	0.0991	0.0712	0.1337	0.0824	0.1763	0.0931	0.0736
Yelp2018-10core	0.0279	0.0179	0.0465	0.0239	0.0763	0.0314	0.0238

Table 4 describes the same revised ranking at three cutoffs and through MRR. Beauty and Toys concentrate more hits early, while OnlineRetail gains broader coverage at larger cutoffs. The fixed per-prefix score connects candidate inclusion and discounted order to one decision, complementing the baseline comparison with a complete account of where relevant items fall after revision.

We next vary the center c and sharpness κ defined in Equation 10, one at a time around $c = 0.11$ and $\kappa = 150$. Nine distinct parameter settings yield 45 evaluations across the five test sets, using the corresponding adapted checkpoints. Tables 5 and 6 report every numerical result.

Figure 1 shows a broad stable region: centers from 0.07 to 0.13 retain at least 99.5% of reference NDCG@10 on every dataset, and sharpness values of 100 or larger give closely aligned results. Calibration changes both contextual routes and the explicit semantic coefficient while retaining the transition statistics. The plateau characterizes the empirical operating region of the complete reliability-to-score mapping analyzed in Proposition 4.

Table 5: Test NDCG@10 for the center sweep with $\kappa = 150$.

c	OnlineRetail	Beauty	Sports	Toys	Yelp
0.07	0.089666	0.071743	0.041704	0.082376	0.023912
0.09	0.089663	0.071743	0.041705	0.082376	0.023934
0.11	0.089663	0.071743	0.041732	0.082374	0.023927
0.13	0.089663	0.071743	0.041551	0.082222	0.023932
0.15	0.089663	0.071477	0.040601	0.080542	0.023932

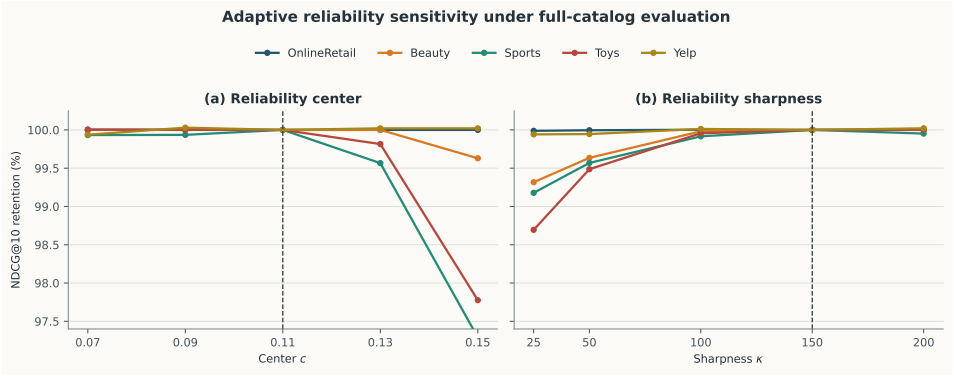


Figure 1: Sensitivity of adaptive reliability over 45 full-catalog test evaluations. Each curve reports NDCG@10 as a percentage of the dataset-specific reference at $c = 0.11$ and $\kappa = 150$; dashed lines mark the reference values.

Table 5 examines the boundary for admitting semantic influence. Raising c to 0.15 lowers Sports and Toys NDCG@10, while OnlineRetail changes little. The stronger response in domains benefiting from explicit semantics is consistent with their training transition alignment and ablation profiles. The common reference lies within the stable interval across all five datasets, characterizing a useful balance between contextual prediction and calibrated revision.

Table 6: Test NDCG@10 for the sharpness sweep with $c = 0.11$; the reference $\kappa = 150$ row is included for completeness.

κ	OnlineRetail	Beauty	Sports	Toys	Yelp
25	0.089652	0.071253	0.041389	0.081299	0.023913
50	0.089659	0.071481	0.041552	0.081951	0.023914
100	0.089665	0.071730	0.041697	0.082338	0.023930
150	0.089663	0.071743	0.041732	0.082374	0.023927
200	0.089663	0.071743	0.041712	0.082376	0.023932

Table 6 shows Amazon results approaching a plateau as κ increases, while OnlineRetail and Yelp vary little. Sharper calibration distinguishes training margins above and below the center more decisively, affecting both routes and the semantic scale. The small changes beyond $\kappa = 100$ identify a stable tested region for translating behavioral alignment into score intervention.

Semantic-source transfer examines evidence revision across representation geometries under the order-based calibration in Appendix B.1. Renewal settings, router structure, data splits, and evaluation remain shared within the study. Each frozen semantic field supplies its training-derived credibility and adapted checkpoint, changing both the contextual anchor and historical-semantic evidence.

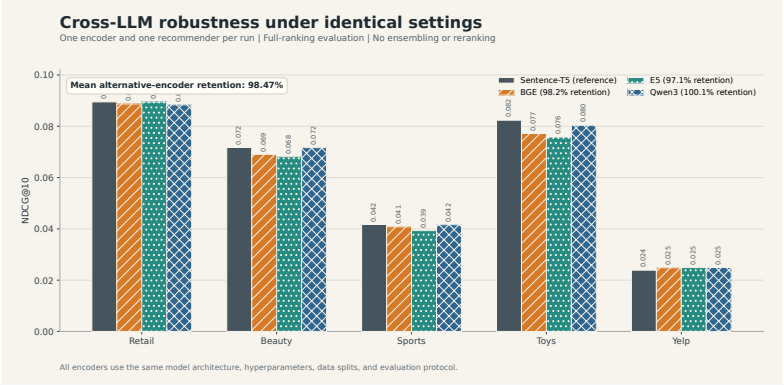


Figure 2: Full-catalog encoder transfer with the shared order-based calibration. Each bar uses one frozen encoder and its corresponding adapted recommender.

Figure 2 shows 98.47% average NDCG@10 retention for alternative encoders relative to this study's Sentence-T5 reference. The complete revision architecture remains effective across these semantic fields, with shared renewal settings and each field's credibility grounded in its measured training transition discrimination.

Table 7: Encoder-transfer NDCG@10 under order-based semantic calibration.

Encoder	OnlineRetail	Beauty	Sports	Toys	Yelp
Sentence-T5	**0.089520**	**0.071732**	**0.041705**	**0.082373**	0.023892
BGE	0.088724	0.069053	0.040822	0.077204	0.024851
E5	0.089500	0.068144	0.039321	0.075814	0.024909
Qwen3	0.088394	0.071681	0.041625	0.080406	**0.024939**

Table 7 shows Sentence-T5 leading on Amazon and OnlineRetail, and Qwen3 leading on Yelp. This domain-dependent ordering is consistent with assessing semantic geometry through behavioral alignment. The common renewal estimator and bounded composition support the complete scorer across the substituted anchor representations and historical-semantic fields under the specified transfer calibration.

5.4 Mechanism attribution: renewal and semantic correction

Table 8: Complete NDCG@10 ablation on all five datasets under the shared full-ranking protocol. Bold marks the best result in each column.

Variant	OnlineRetail	Beauty	Sports	Toys	Yelp
w/o global calibration	0.089056	0.071643	0.041704	0.082276	0.023892
w/o explicit semantic score	0.089563	0.068325	0.040129	0.077389	0.023832
w/o Bayesian renewal	0.071332	0.069817	0.039137	0.076630	0.021545
w/o adaptive router	0.089401	0.071692	0.041590	0.082245	0.023827
w/o reliability shrinkage	0.088787	0.066936	0.039829	0.079501	0.021289
w/o rank survival	0.085884	0.071671	0.041518	0.082110	0.023711
Full R-HIBR	**0.089663**	**0.071743**	**0.041732**	**0.082374**	**0.023927**

Evidence construction versus intervention control. Table 8 places the full model first in every column. Renewal removal causes the largest loss on OnlineRetail and Toys; source–pair shrinkage is particularly influential on Beauty and Yelp, consistent with support-aware transition evidence in **C1**. Explicit-semantic removal reduces Amazon NDCG@10 by 3.84–6.05%, matching its historical-relevance role in **C2**. Because this removal retains the semantic-initialized anchor, the loss concerns direct history-to-candidate revision beyond contextual encoding. Router and global-calibration removals yield smaller changes, consistent with intervention control in **C3**, while rank survival matters most on OnlineRetail. These effects distinguish the directions supplied by transition and semantic evidence from the support and contextual controls governing their contribution.

6 Conclusion

R-HIBR views sequential recommendation as reliability-constrained evidence revision: a contextual ranking is refined according to the support and behavioral relevance of auxiliary evidence. Posterior transition lift and calibrated historical semantics provide revision directions, while support attenuation and bounded routing regulate their influence. The analysis connects these operations to both preservation and correction: an intervention budget protects sufficiently separated rankings, and an evidence advantage exceeding an anchor deficit reverses the corresponding pairwise order. Jointly vanishing corrections recover the anchor, and calibration changes induce bounded complete-score perturbations. Matched comparisons favor R-HIBR on all reported Amazon and OnlineRetail metrics; ablations, sensitivity, and semantic-source transfer characterize the empirical roles of evidence and its regulation. The resulting principle separates the relevance of auxiliary information from its authority to alter a contextual decision. Statistical support and behavioral alignment govern evidence construction and scaling; the realized advantage relative to the anchor deficit then characterizes whether a specific ordering is revised.

AI USE STATEMENT

Generative AI tools assisted language polishing, manuscript organization, and reasoning about the methodological exposition and mathematical arguments. The authors are responsible for the final manuscript, including its methods, theoretical claims, experimental results, and references.

REPRODUCIBILITY STATEMENT

We provide the implementation code for R-HIBR. Appendix B.2, Table 9, documents the model parameters, optimization settings, random seed, and evaluation protocol. Appendix B.1 specifies the encoder-transfer calibration, distinguishing it from the main-study configuration.

REFERENCES

Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. Session-based recommendations with recurrent neural networks. In *International Conference on Learning Representations*, 2016.

Yupeng Hou, Shanlei Mu, Wayne Xin Zhao, Yaliang Li, Bolin Ding, and Ji-Rong Wen. Towards universal sequence representation learning for recommender systems. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 585–593, 2022. doi: 10.1145/3534678.3539381.

Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *IEEE International Conference on Data Mining*, pp. 197–206, 2018. doi: 10.1109/ICDM.2018.00035.

Anton Klenitskiy and Alexey Vasilev. Turning dross into gold loss: Is BERT4Rec really better than SASRec? In *Proceedings of the 17th ACM Conference on Recommender Systems*, pp. 1120–1125, 2023. doi: 10.1145/3604915.3610644.

Jiacheng Li, Ming Wang, Jin Li, Jinmiao Fu, Xin Shen, Jingbo Shang, and Julian McAuley. Text is all you need: Learning language representations for sequential recommendation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1258–1267, 2023a. doi: 10.1145/3580305.3599519.

Xinhang Li, Chong Chen, Xiangyu Zhao, Yong Zhang, and Chunxiao Xing. E4SRec: An elegant effective efficient extensible solution of large language models for sequential recommendation. *arXiv preprint arXiv:2312.02443*, 2023b.

Jiayi Liao, Sihang Li, Zhengyi Yang, Jiancan Wu, Yancheng Yuan, Xiang Wang, and Xiangnan He. LLaRA: Large language-recommendation assistant. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1785–1795, 2024. doi: 10.1145/3626772.3657690.

Jianmo Ni, Gustavo Hernández Ábrego, Noah Constant, Ji Ma, Keith B. Hall, Daniel Cer, and Yinfei Yang. Sentence-T5: Scalable sentence encoders from pre-trained text-to-text models. In *Findings of the Association for Computational Linguistics: ACL 2022*, pp. 1864–1874. Association for Computational Linguistics, 2022. doi: 10.18653/v1/2022.findings-acl.146. URL https://aclanthology.org/2022.findings-acl.146/.

Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized Markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, pp. 811–820, 2010. doi: 10.1145/1772690.1772773.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 1441–1450, 2019. doi: 10.1145/3357384.3357895.

Liang Wang, Nan Yang, Xiaolong Huang, Binxing Jiao, Linjun Yang, Daxin Jiang, Rangan Majumder, and Furu Wei. Text embeddings by weakly-supervised contrastive pre-training. *arXiv preprint arXiv:2212.03533*, 2022.

Cong Xu, Guoliang Li, Jun Wang, and Wei Zhang. Markovian pre-trained transformer for next-item recommendation. *arXiv preprint arXiv:2601.08275*, 2026. URL https://arxiv.org/abs/2601.08275.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. Qwen2.5 Technical Report. *arXiv preprint arXiv:2412.15115*, 2024.

Jiaqi Zhai, Lucy Liao, Xing Liu, Yueming Wang, Rui Li, Xuan Cao, Leon Gao, Zhaojie Gong, Fangda Gu, Jiayuan He, Yinghai Lu, and Yu Shi. Actions speak louder than words: Trillion-parameter sequential transducers for generative recommendations. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 58484–58509. PMLR, 2024. URL https://proceedings.mlr.press/v235/zhai24a.html.

Kun Zhou, Hui Yu, Wayne Xin Zhao, and Ji-Rong Wen. Filter-enhanced MLP is all you need for sequential recommendation. In *Proceedings of the ACM Web Conference 2022*, pp. 2388–2399, 2022. doi: 10.1145/3485447.3512111.

A PROOFS

A.1 PROOF OF PROPOSITION 1

Proof. For $c_\ell(i) > 0$, substituting $c_\ell(i, j) = c_\ell(i)\widetilde{p}_\ell(j \mid i)$ into Equation 3 gives

$$\begin{aligned}\widehat{p}_\ell(j \mid i) &= \frac{c_\ell(i)\widetilde{p}_\ell(j \mid i) + \alpha p_0(j)}{c_\ell(i) + \alpha} \\ &= (1 - \omega_\ell(i))p_0(j) + \omega_\ell(i)\widetilde{p}_\ell(j \mid i).\end{aligned}$$

Since $p_0(j) > 0$, division by $p_0(j)$ followed by the logarithm proves Equation 16. The difference from the catalog prior contracts exactly:

$$\widehat{p}_\ell(j \mid i) - p_0(j) = \omega_\ell(i)\big(\widetilde{p}_\ell(j \mid i) - p_0(j)\big).$$

Viewing source support as continuous while keeping the empirical proportions and prior fixed,

$$\frac{\partial \omega_\ell(i)}{\partial c_\ell(i)} = \frac{\alpha}{(c_\ell(i) + \alpha)^2} > 0, \qquad \frac{\partial}{\partial \omega_\ell(i)} \log \frac{\widehat{p}_\ell(j \mid i)}{p_0(j)} = \frac{\widetilde{p}_\ell(j \mid i)/p_0(j) - 1}{1 + \omega_\ell(i)(\widetilde{p}_\ell(j \mid i)/p_0(j) - 1)}.$$

The derivative of the raw lift is positive when the empirical ratio exceeds one. The raw lift tends to zero as $\omega_\ell(i) \to 0$, and clipping is nondecreasing. At $c_\ell(i) = 0$, Equation 3 gives $p_0(j)$ directly, the support set is empty, and that source–horizon pair contributes zero to Equation 7.

For a dense-support regime, consider proportional growth of positive counts while keeping p_0 and the selected count ordering fixed. For a selected pair, $\omega_\ell(i)$, $r_\ell^{\mathrm{src}}(i)$, and $r_\ell^{\mathrm{pair}}(i, j)$ tend to one, and $\widehat{p}_\ell(j \mid i) \to \widetilde{p}_\ell(j \mid i)$. Hence its renewal contribution before horizon discounting satisfies

$$r_\ell^{\mathrm{src}}(i) r_\ell^{\mathrm{pair}}(i, j)\big[e_\ell(i, j) + v_{\ell,i}(j)\big] \longrightarrow \mathrm{clip}_{[0, B_{\max}]}\left(\log \frac{\widetilde{p}_\ell(j \mid i)}{p_0(j)}\right) + v_{\ell,i}(j).$$

This limit retains the implemented clipping, selected support, and rank bonus; Equation 7 subsequently applies the same horizon weights. □

A.2 PROOF OF LEMMA 1

Proof. Equation 4 gives $0 \le e_\ell(i, j) \le B_{\max}$. Both factors in Equation 5 lie in $[0, 1]$. For every selected position $q_{\ell,i}(j) \in \{1, \dots, K\}$, Equation 6 is nonnegative and is maximized at position one:

$$0 \le v_{\ell,i}(j) \le \lambda_{\mathrm{rs}} \log \frac{K + \frac{1}{2}}{\frac{1}{2}} = \lambda_{\mathrm{rs}} \log(2K + 1).$$

The indicator in Equation 7 also lies in $\{0, 1\}$. Bounding each summand and extending D_t to D yields

$$\begin{aligned}0 \le b_t(j) &\le \sum_{m=0}^{D-1} \eta^m \left(B_{\max} + \lambda_{\mathrm{rs}} \log(2K + 1)\right) \\ &= \left(B_{\max} + \lambda_{\mathrm{rs}} \log(2K + 1)\right) \frac{1 - \eta^D}{1 - \eta},\end{aligned}$$

which is finite because $\eta \in (0, 1)$. □

A.3 PROOF OF PROPOSITION 2

Proof. Equation 13 gives $z_t^{\mathrm{B}}, z_t^{\mathrm{S}} \in [-1, 1]$. Since $\delta \in [0, 1)$, the realized coefficients obey

$$0 < \Lambda_t \le \lambda_{\mathrm{B}}(1 + \delta), \qquad 0 \le \epsilon_t \le \gamma \lambda_{\mathrm{S}}(1 + \delta).$$

Normalization in Equation 8 ensures $|m_t(j)| \le 1$. Therefore

$$\ell_t(j) - a_t(j) = \Lambda_t b_t(j) + \epsilon_t m_t(j) \in [\Lambda_t b_t^- - \epsilon_t, \Lambda_t b_t^+ + \epsilon_t], \qquad |\ell_t(j) - Q_t(j)| = \epsilon_t |m_t(j)| \le \epsilon_t.$$

Lemma 1 gives $0 \le b_t^- \le b_t^+ \le B_{\mathrm{R}}$, yielding Equation 19. Since Λ_t is bounded, $b_t^+ \to 0$ together with $\epsilon_t \to 0$ implies $\max_{j \in \mathcal{E}_t} |\ell_t(j) - a_t(j)| \to 0$, establishing uniform anchor recovery when both revision contributions vanish. For $j \in \mathrm{Top}_k(a_t)$ and $j' \notin \mathrm{Top}_k(a_t)$ with both items eligible,

$$\ell_t(j) - \ell_t(j') \ge a_t(j) - a_t(j') - \Lambda_t(b_t^+ - b_t^-) - 2\epsilon_t \ge \mathrm{gap}_k(a_t) - \Lambda_t(b_t^+ - b_t^-) - 2\epsilon_t.$$

The first margin condition makes this positive for every inside–outside pair. No outside item can enter the top-k set. Repeating the argument around Q_t gives $\ell_t(j) - \ell_t(j') \ge \mathrm{gap}_k(Q_t) - 2\epsilon_t > 0$ for its inside–outside pairs. Both conclusions compare scores on the same eligible set; Q_t retains the contextual anchor and realized renewal route. □

A.4 Proof of Proposition 3

Proof. Fix the prefix and its realized routes as in Proposition 2. Subtracting the two instances of Equation 14 gives

$$\ell_t(y) - \ell_t(j) = a_t(y) - a_t(j) + \Lambda_t[b_t(y) - b_t(j)] + \epsilon_t[m_t(y) - m_t(j)] = a_t(y) - a_t(j) + \mathcal{A}_t(y, j).$$

Thus the corrected difference is positive exactly when the evidence advantage exceeds the anchor deficit. The same budgets also give

$$\mathcal{A}_t(y, j) \le \Lambda_t(b_t^+ - b_t^-) + 2\epsilon_t.$$

An anchor deficit at least this large cannot be reversed, connecting the correction condition to the preservation bound.

For a catalog-wide sufficient condition, let y be the eligible next item and fix a set $\mathcal{J}_t \subseteq \mathcal{E}_t \setminus \{y\}$ of size $k - 1$. If

$$\mathcal{A}_t(y, j) > a_t(j) - a_t(y) \qquad \text{for every } j \in \mathcal{E}_t \setminus (\mathcal{J}_t \cup \{y\}),$$

then y outranks every item outside $\mathcal{J}_t \cup \{y\}$. At most $k - 1$ items can precede it, so $y \in \mathrm{Top}_k(\ell_t)$. When $y \notin \mathrm{Top}_k(a_t)$, choosing $\mathcal{J}_t$ as the anchor's first $k - 1$ items supplies a sufficient condition for entering the corrected top-k set. For $k = 1$, $\mathcal{J}_t$ is empty. The condition covers all remaining eligible competitors under the same realized revision rule. □

A.5 Proof of Proposition 4

Proof. Under the fixed quantities in the proposition, the pre-activation of each router row is affine in γ. Its coefficient is the signed quantity inside the corresponding absolute value in Equation 21. Since the derivative of $\tanh$ has absolute value at most one,

$$|z_t^{\mathrm{B}}(\gamma) - z_t^{\mathrm{B}}(\gamma')| \le L_t^{\mathrm{B}}|\gamma - \gamma'|, \qquad |z_t^{\mathrm{S}}(\gamma) - z_t^{\mathrm{S}}(\gamma')| \le L_t^{\mathrm{S}}|\gamma - \gamma'|.$$

Here $z_t^{\mathrm{B}}(\gamma)$ and $z_t^{\mathrm{S}}(\gamma)$ denote the same router in Equation 13 evaluated at γ. The renewal contribution changes by at most $\lambda_{\mathrm{B}}\delta b_t^+ L_t^{\mathrm{B}}|\gamma - \gamma'|$. For the semantic coefficient, add and subtract $\gamma'(1 + \delta z_t^{\mathrm{S}}(\gamma))$:

$$\begin{aligned}
&|\gamma(1 + \delta z_t^{\mathrm{S}}(\gamma)) - \gamma'(1 + \delta z_t^{\mathrm{S}}(\gamma'))| \\
&\quad \le |\gamma - \gamma'|\,|1 + \delta z_t^{\mathrm{S}}(\gamma)| + \gamma'\delta|z_t^{\mathrm{S}}(\gamma) - z_t^{\mathrm{S}}(\gamma')| \\
&\quad \le (1 + \delta + \delta L_t^{\mathrm{S}})|\gamma - \gamma'|,
\end{aligned}$$

using $\gamma' \le 1$. Multiplication by $\lambda_{\mathrm{S}}|m_t(j)| \le \lambda_{\mathrm{S}}$, followed by addition of the renewal bound, proves Equation 22. The contextual anchor itself stays fixed. For an inside–outside pair under $\ell_t(\cdot; \gamma')$, the score difference under γ is at least

$$\mathrm{gap}_k(\ell_t(\cdot; \gamma')) - 2C_t|\gamma - \gamma'|.$$

Strict positivity proves the stated top-k preservation. Finally, the sigmoid derivative satisfies $\sigma'(x) = \sigma(x)(1 - \sigma(x)) \le 1/4$. Applying the mean value theorem to Equation 10 at the two gap values, and using the range of the sigmoid, gives

$$|\gamma - \gamma'| \le \min\Big\{1, \frac{\kappa}{4}|\Delta_{\mathrm{sem}} - \Delta'_{\mathrm{sem}}|\Big\}.$$

This deterministic inequality applies with the same calibration center and sharpness and the same adapted checkpoint. □

B Additional Experimental Details

B.1 Encoder-transfer calibration

The main comparisons, ablations, and sensitivity study use the contrastive-gap calibration in Equation 10. The encoder-transfer study instead uses pair ordering on the same kind of training transitions and deterministic negatives. With P, R, (x_p, y_p), and $z_{p,r}$ defined in Section 3.3, set $d_p^+ = \langle \boldsymbol{s}_{x_p}, \boldsymbol{s}_{y_p} \rangle$ and $d_{p,r}^- = \langle \boldsymbol{s}_{x_p}, \boldsymbol{s}_{z_{p,r}} \rangle$. Their probability-of-superiority statistic is

$$A_{\mathrm{pair}} = \frac{1}{PR} \sum_{p=1}^{P} \sum_{r=1}^{R} \left(\mathbb{I}[d_p^+ > d_{p,r}^-] + \tfrac{1}{2} \mathbb{I}[d_p^+ = d_{p,r}^-] \right). \tag{23}$$

Define the clipped rank effect $\rho = \mathrm{clip}_{[\varepsilon_A, 1-\varepsilon_A]}(2A_{\mathrm{pair}} - 1)$, with $\varepsilon_A = 10^{-6}$, and $\mathrm{logit}(x) = \log(x/(1-x))$ for $x \in (0,1)$. The order calibrator is

$$\gamma_{\mathrm{ord}} = \sigma(\kappa_{\mathrm{ord}}[\mathrm{logit}(\rho) - \mathrm{logit}(\rho_\star)]), \qquad \rho_\star = 0.20, \quad \kappa_{\mathrm{ord}} = 6. \tag{24}$$

The same $\rho_\star$ and κ_{ord} are used for every dataset and encoder. Substituting γ_{ord} for γ in Equations 12, 13, and 14 specifies the transfer configuration. Pair ordering is invariant to a common strictly increasing score transformation; replacing an encoder can change that ordering and both semantic-dependent scoring paths. Table 7 reports the corresponding encoder-specific checkpoint series, including its Sentence-T5 reference.

B.2 Reproducibility configuration

Table 9: Core configuration and study-specific semantic calibration.

Component	Setting
Sequential anchor	4 Transformer layers, 2 heads, hidden size 256
Sequence regularization	maximum length 50, dropout 0.2
Renewal statistics	$D = 12$, $K = 16$, $\alpha = 0.1$
Support modulation	$\lambda_{\mathrm{src}} = 4$, $\lambda_{\mathrm{pair}} = 1$, $\lambda_{\mathrm{rs}} = 0.82$
Evidence scales	$\eta = 0.8$, $\beta = 0.4$, $\lambda_{\mathrm{B}} = 0.36$, $\lambda_{\mathrm{S}} = 1.8$
Local router	$q_0 = 0.67$, $\zeta = 20$, $\delta = 0.3$
Main-study calibration	$c = 0.11$, $\kappa = 150$, $R = 16$ negatives
Calibration sample budget	at most 200,000 adjacent training transitions
Encoder-transfer calibration	$\rho_\star = 0.20$, $\kappa_{\mathrm{ord}} = 6$
Numerical and support scales	$B_{\max} = 6$, $\chi = 5$, $\epsilon_{\mathrm{n}} = 10^{-12}$
Score and objective	$\tau = 0.07$, $\lambda_{\mathrm{I}} = 0.08$
Optimization	AdamW, learning rate 10^{-3}, weight decay 0.1
Training	batch size 512, seed 42
Evaluation	chronological leave-one-out, full-catalog ranking

Analysis. Table 9 distinguishes shared score-construction settings from the two semantic calibration configurations. Main comparisons and sensitivity use cosine gaps, while transfer uses pair ordering to accommodate different semantic geometries. The common architecture, support budget, and catalog evaluator keep the studies tied to the same renewal mechanism, with each reported result associated with its specified calibration and checkpoint.