Grounding What Users Care About: Multi-Aspect Personalized Modeling from Implicit Preferences

Jiin Park[1] **Misuk Kim**[1,2]*
[1] Department of Artificial Intelligence, Hanyang University, Seoul, Republic of Korea
[2] Department of Data Science, Hanyang University, Seoul, Republic of Korea
{jiinpark, misukkim}@hanyang.ac.kr

Abstract

Large Language Model (LLM) personalization is challenging because user preferences are rarely stated explicitly; instead, they emerge as implicit, multifaceted signals in users' past interaction histories. Existing methods often compress user histories into a single representation or retrieve and inject past records without explicitly linking them to supporting evidence, which can lead to hallucinated or overly generic outputs. In this work, we propose **M**ulti-**A**spect **P**erson**a****L**ized Modeling with **E**vidence (MAPLE). MAPLE structures implicit user preferences through a MULTI-ASPECT PREFERENCE PROFILE and connects them to target-specific grounding evidence through a PERSONALIZED EVIDENCE PLAN. By selecting only generation points that are both relevant to the user and supported by available evidence, MAPLE improves personalization and grounding faithfulness simultaneously. Experiments on the `Games` and `Books` domains with two backbone LLMs show that MAPLE consistently outperforms strong baselines and generalizes robustly under limited-history and unseen-user settings.

1 Introduction

User preferences are generally not explicitly given, so models must infer them from indirect signals such as reviews or interaction histories (Zhang et al., 2025b; Xu et al., 2025). However, real user preferences do not appear as a single, consistent tendency. An aspect refers to a fine-grained perspective that users rely on when evaluating an object. For example, even for the same object, one user may focus on functionality, while another may evaluate it based on price, design, or other perspectives. Thus, the same user may have strong criteria for some aspects, while responding more flexibly to others depending on the context. This implicit and multi-aspect nature makes LLM personalization inherently challenging (Kim et al., 2026; Li et al., 2024).

Existing personalization methods have mainly developed in two directions. The first is memory-based approaches that retrieve or summarize a user's past history and inject it into the prompt (Salemi et al., 2024; Richardson et al., 2023; Zhuang et al., 2024; Mysore et al., 2024). The second is approaches that learn user information as compressed representations, such as embeddings, adapters, or activation vectors (Tan et al., 2024; Liu et al., 2025b; Zhang et al., 2025a). Although these approaches improve personalization performance, they still have a key limitation: they do not explicitly connect information that is important to the user with information that is actually grounded in the current context. As a result, models may generate unsupported personalized content or, conversely, collapse into overly generic responses (Wallat et al., 2025; Schimanski et al., 2024; Sun et al., 2026). Moreover, many existing methods compress users into a single representation, but real user preferences cannot be fully explained by a single vector. A user may show strong preferences for certain aspects while simultaneously holding different or even opposing attitudes toward others. Prior work in recommendation and personalization has repeatedly shown that decomposing user preferences into multiple facets or aspects is effective (Qian et al., 2025; Ye et al., 2025; Du and Li, 2025). Nevertheless, current LLM personalization research still has limitations in structurally modeling users' implicit multi-aspect preferences (Choi et al., 2025; Shi et al., 2025; Sun et al., 2025; Qiu et al., 2025b,a).

To address these limitations, we propose **M**ulti-**A**spect **P**erson**a****L**ized Modeling with **E**vidence (MAPLE), which constructs a MULTI-ASPECT PREFERENCE PROFILE from implicit user histories and connects it with grounding evidence.

*Corresponding author.

Proceedings of the 15th International Joint Conference on Natural Language Processing and the 5th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics, pages 2063–2089
November 6-10, 2026 ©2026 Association for Computational Linguistics

First, we extract preference-related information from user reviews and aggregate it by aspect to construct a MULTI-ASPECT PREFERENCE PROFILE. This structured profile captures which aspects the user cares about and what sentiment tendency the user shows for each aspect, providing interpretable aspect-level personalization signals instead of compressing the user into a single representation. Next, we construct an EVIDENCE BANK by aggregating content repeatedly mentioned by other users about the target item at the aspect level, and build a PERSONALIZED EVIDENCE PLAN that connects user preferences with actual grounding evidence. This allows the generator to produce text based on key information that is both relevant to the user and supported by evidence, rather than directly interpreting the entire user history. Finally, we train the generator to produce the target output conditioned on the MULTI-ASPECT PREFERENCE PROFILE and the PERSONALIZED EVIDENCE PLAN, and further align it with direct preference optimization (DPO) so that it better satisfies both user preferences and grounding evidence. To evaluate the proposed framework, we conduct experiments on the Games and Books domains using Qwen- and Llama-family models. We also evaluate the framework not only in the standard setting where sufficient user history is available, but also under limited-history and unseen-user settings. Experimental results show that MAPLE consistently improves the balance between personalization and grounding faithfulness, and generalizes robustly even when only limited user information is observed.

Our contributions are as follows.

- We propose MAPLE, which structures implicit user histories into aspect-level preferences instead of compressing them into a single representation. This transforms users' past behavioral histories into interpretable aspect-level personalization signals.
- We introduce PERSONALIZED EVIDENCE PLANNING, which aligns user preferences with target item evidence. This enables personalized generation to focus on content that is both important to the user and supported by available evidence, thereby improving both personalization and faithfulness.
- We validate the generalizability and robustness of the proposed method through experiments across multiple domains, models, and limited-history and unseen-user settings.

2 Related Work

Personalized and Difference-aware LLMs. Research on LLM personalization has increasingly focused on leveraging users' interaction histories to reflect individual preferences and intents (Salemi et al., 2024; Chen et al., 2024; Zhang et al., 2025b; Xu et al., 2025). Existing methods are mainly based on memory-based personalization, which retrieves or summarizes past histories and injects them into prompts (Salemi et al., 2024; Richardson et al., 2023; Zhuang et al., 2024; Mysore et al., 2024), or representation-based personalization, which compresses user information into embeddings, adapters, or activation vectors (Tan et al., 2024; Liu et al., 2025b; Zhang et al., 2025a). Recent work has also explored using information from similar users or incorporating collaborative filtering (Choi et al., 2025; Shi et al., 2025; Sun et al., 2025), as well as modeling inter-user differences in natural language or latent space (Qiu et al., 2025b,a). However, these methods generally summarize a user into a single profile or representation, or focus on differences between users, and thus do not explicitly structure the implicit multi-aspect preferences mixed within an individual user.

Multi-aspect Faithful Personalization. In recommendation and federated learning, prior work has explored modeling users with multiple facets rather than a single vector (Qian et al., 2025; Ye et al., 2025; Du and Li, 2025). This work extends this perspective to LLM-based personalized generation by constructing a MULTI-ASPECT PREFERENCE PROFILE from implicit user histories and connecting it with grounding evidence available in the current input. Beyond merely reflecting user preferences, our goal is faithful personalized generation in which user preferences and supporting evidence are jointly aligned.

3 Method

Figure 1 summarizes the overall architecture of MAPLE. Our method constructs a MULTI-ASPECT PREFERENCE PROFILE P_u from the user history and an EVIDENCE BANK E_i from item review histories, then connects the two through a PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$, which is used as input to the generator.

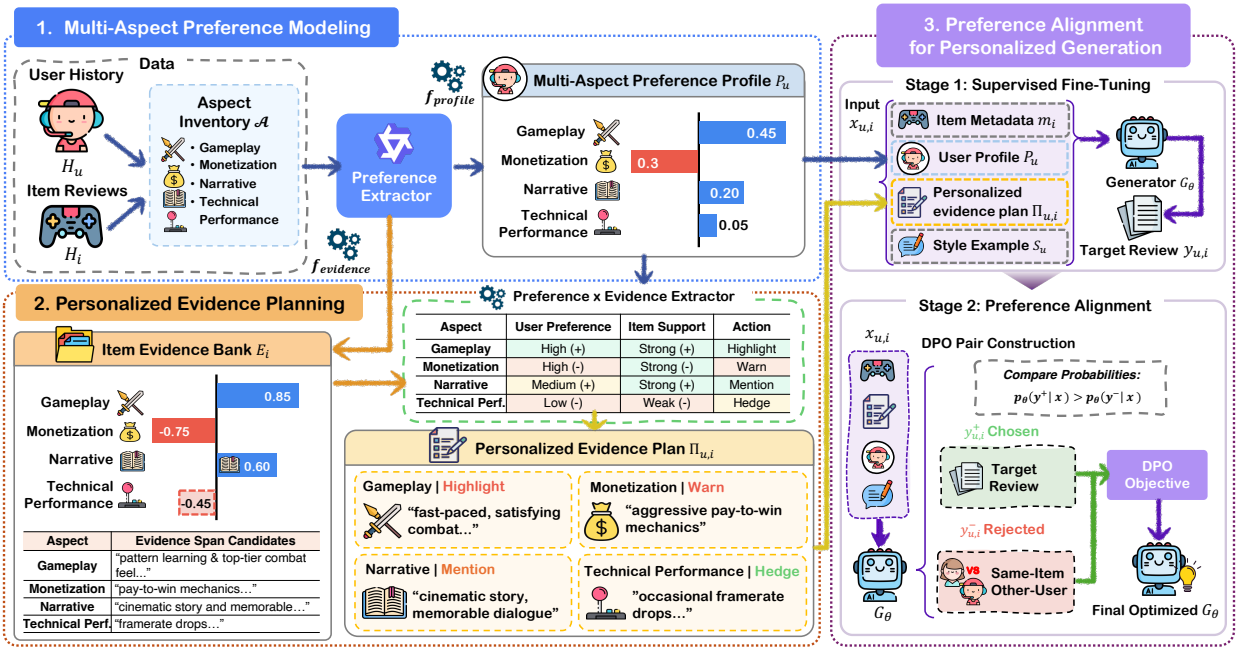


Figure 1: Overview of MAPLE. We extract aspect-level preference tuples from user and item review histories, aggregate them into P_u and E_i, and match them to construct $\Pi_{u,i}$. Bars in P_u indicate aspect importance and user sentiment; bars in E_i indicate mention/support rate and item sentiment, with dashed bars denoting high variance. The generator conditions on ITEM METADATA, P_u, $\Pi_{u,i}$, and style examples, and is trained with SFT followed by grounded DPO.

3.1 Problem Formulation

This work addresses personalized generation in the item review generation setting. Given a user's past history H_u and a target item i, the goal is to generate text $\hat{y}_{u,i}$ that reflects the user's preferences while being grounded in target item evidence. The full dataset consists of a set of user–item interactions D, where each instance is represented as $(u, i, y_{u,i})$, and $y_{u,i}$ denotes the review text written by user u for item i. We follow the leave-one-out protocol widely used in recommender systems (He et al., 2017). For each target interaction $(u, i, y_{u,i})$, we treat $y_{u,i}$ as the generation target and use the user's previous reviews, excluding $y_{u,i}$, as H_u. We also define H_i as the set of reviews observed in the training split for the target item i. To prevent label leakage, the target review $y_{u,i}$ is excluded not only from H_i, but also from the construction of the profile, EVIDENCE BANK, and plan.

The key idea of this work is to structure and connect user preferences with target item evidence, rather than directly using raw user histories. We first construct a MULTI-ASPECT PREFERENCE PROFILE P_u from the user history H_u (see Section 3.2) and build an EVIDENCE BANK E_i that aggregates aspect-level evidence for the target item i. Because this work focuses on review generation, we instantiate E_i as an EVIDENCE BANK constructed from the review history of each target item. However, this concept is not limited to a specific task and can be generalized to any information that supports the current generation, such as document content, retrieved context, dialogue context, or domain knowledge. We then combine P_u and E_i to construct a PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$ (see Section 3.3), which the generator conditions on to produce the final text (see Section 3.4).

3.2 Multi-Aspect Preference Modeling

The goal of this stage is to extract fine-grained preferences implicitly contained in the user history H_u and structure them into a MULTI-ASPECT PREFERENCE PROFILE P_u. To connect P_u and the EVIDENCE BANK E_i in the same aspect space, we first construct a domain-specific aspect set $\mathcal{A}$ from item metadata and reviews in the training split (see Appendix D.2). We then apply a separately trained preference extractor to each review sentence and extract zero or more preference tuples $\ell = (a, d, s, h)$. Here, $a \in \mathcal{A}$ denotes the aspect, d is a descriptor phrase that specifies the aspect, $s \in \{-1, 0, 1\}$ is the sentiment polarity, and h is the original span in the user history where the preference is observed. Through this process, raw histories are converted

into structured preference signals, and a single sentence may yield zero or more preference tuples. Details of the extractor training procedure and output format are provided in Appendix D.3.

Next, we construct P_u by grouping all preference tuples extracted from the history H_u of user u by aspect. This allows us to compute the user-specific aspect importance $p_{u,a}$ and aspect sentiment $q_{u,a}$, which captures the overall sentiment tendency for that aspect, as follows:

$$p_{u,a} = \frac{c_{u,a}}{\sum_{a' \in \mathcal{A}} c_{u,a'}}, \tag{1}$$

$$q_{u,a} = \frac{1}{|R_{u,a}|} \sum_{s \in R_{u,a}} s. \tag{2}$$

where $c_{u,a}$ denotes the number of times user u mentions aspect a, and $R_{u,a}$ denotes the set of sentiment labels observed for that aspect. The final MULTI-ASPECT PREFERENCE PROFILE is then defined as

$$P_u = \{(a, p_{u,a}, q_{u,a})\}_{a \in \mathcal{A}}. \tag{3}$$

3.3 Personalized Evidence Planning

Personalized evidence planning selects the key content that the generator should reflect by matching aspects important to the user with content that is actually supported for the target item. The planner excludes content that is important to the user but lacks sufficient evidence, as well as content that is well supported but less relevant to the user, and includes in $\Pi_{u,i}$ only the content that satisfies both conditions. The EVIDENCE BANK E_i is a structure that aggregates repeatedly observed content from the item review history H_i by aspect. To construct it, we aggregate the preference tuples (a, d, s, h) extracted in Section 3.2 by aspect, and store each aspect's mention rate, mean sentiment, sentiment variance, and candidate descriptors and evidence spans.

Given a target user u and a target item i, the planner selects only aspect-level points that are both important to the user and sufficiently supported by the item evidence. Specifically, it filters out aspects with low user importance or insufficient evidence support, and selects the top K remaining candidates with high user relevance and evidence strength. The final PERSONALIZED EVIDENCE PLAN is represented as follows:

$$\Pi_{u,i} = \{\pi_1, \ldots, \pi_K\}, \qquad \pi = (a, \delta, d, e), \tag{4}$$

where a is the selected aspect and d is the descriptor to be used for generation. The evidence span e is selected by the planner from the candidate spans stored in the EVIDENCE BANK E_i, and δ is an action label that specifies how the selected evidence should be expressed. Here, $\delta \in \{\texttt{highlight}, \texttt{warn}, \texttt{hedge}, \texttt{mention}\}$ determines whether the selected evidence is expressed as a strength, a caution, uncertain information, or auxiliary information. Detailed procedures for evidence selection, ranking, and action assignment are provided in Appendix F.

3.4 Preference Alignment for Personalized Generation

In this stage, we use a two-stage training strategy, applying DPO after SFT, to train the generator in MAPLE. First, the generator is trained to produce natural personalized reviews from inputs that structure user preferences and target item evidence. Then, in DPO, we use real reviews written by other users for the same item as counterfactual negatives, aligning the generator to prefer responses that better match the target user. Using the MULTI-ASPECT PREFERENCE PROFILE P_u and the PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$ defined in Section 3.2 and Section 3.3, respectively, we construct the input to the generator G_θ as follows:

$$x_{u,i} = [m_i; P_u; \Pi_{u,i}; S_u], \tag{5}$$

$$\hat{y}_{u,i} = G_\theta(x_{u,i}), \tag{6}$$

where m_i denotes the cleaned item metadata, and S_u denotes the STYLE EXAMPLE used to reflect the user's writing style and narrative patterns.

Stage 1. Supervised Fine-Tuning. We first train the generator with SFT using training pairs $(x_{u,i}, y_{u,i})$. We use the standard causal language modeling objective:

$$\mathcal{L}_{\text{SFT}} = -\sum_{(u,i)} \sum_{t=1}^{|y_{u,i}|} \log p_\theta(y_{u,i,t} \mid y_{u,i,<t}, x_{u,i}). \tag{7}$$

This stage teaches the generator the basic ability to generate the target review conditioned on P_u, $\Pi_{u,i}$, and S_u.

Stage 2. Preference Alignment. Next, we apply DPO to strengthen preference alignment for the target user. For each input $x_{u,i}$, the chosen response $y^+_{u,i}$ is set to the review $y_{u,i}$ actually written by

the target user. The rejected response $y^-_{u,i}$ is constructed from a real review written by another user $u' \neq u$ for the same item i:

$$y^+_{u,i} = y_{u,i}, \qquad y^-_{u,i} = y_{u',i}, \quad u' \neq u. \tag{8}$$

Because this rejected response is written for the same target item, it provides a stronger comparison than a random negative while still potentially differing from the target user in aspect emphasis, sentiment tendency, or writing style. Thus, DPO trains the generator to prefer reviews that better match the target user, rather than merely producing reviews that fit the item. When multiple candidates are available, we select as a hard negative the review written by the user whose profile is most different from P_u. The DPO objective is defined as follows:

$$\mathcal{L}_{\text{DPO}} = -\log \sigma\Big(\beta\big[r_\theta(y^+_{u,i}, x_{u,i}) - r_\theta(y^-_{u,i}, x_{u,i})\big]\Big). \tag{9}$$

Here, β is a hyperparameter that controls the strength of the preference margin, and $r_\theta(y, x) = \log \frac{p_\theta(y|x)}{p_{\text{ref}}(y|x)}$ denotes the implicit reward defined as the log-probability ratio between the current policy model and the SFT reference model p_{ref}.

As a result, the generator learns to produce reviews for the target item using P_u, $\Pi_{u,i}$, and S_u, while being aligned to prefer responses that better match the target user even for the same item. Details of the pair construction procedure are provided in Appendix H.

4 Experiments

4.1 Experimental Setup

In this section, we aim to answer the following research questions:

- **RQ1:** How effectively does the proposed framework improve personalized and grounded generation across domains and backbone models?
- **RQ2:** How does each component of the proposed framework contribute to personalized and grounded generation?
- **RQ3:** How robust is the proposed method under limited-history and unseen-user settings?
- **RQ4:** Are the generated reviews better grounded in item evidence and more aligned with user preferences?

Datasets. We use the Games and Books domains from Amazon Reviews 2023 [1] to evaluate item review generation (Ni et al., 2019; Peng et al., 2024; Au et al., 2025; Qiu et al., 2025a,b). For each user, we split interactions chronologically: the last interaction is used for test, the immediately preceding interaction for validation, and the remaining interactions for training. Data curation criteria and split statistics are provided in Appendix A.

Baselines. We compare our method against a diverse set of baselines to verify its effectiveness. First, Non-Perso, a non-personalized baseline, generates reviews using only target item information, without incorporating any user-specific information. As prompt-based baseline models, we use RAG (Salemi et al., 2024), which directly includes the user's historical reviews in the prompt; PAG (Richardson et al., 2023), which summarizes the user's history into a profile; and DPL (Qiu et al., 2025b), which constructs a natural-language profile based on the differences between the target user and representative users. As representation-based comparison models, we use OPPU (Tan et al., 2024), which trains user-specific adapters; DEP (Qiu et al., 2025a), which injects user embeddings and inter-user difference embeddings as soft prompts; and PPlug (Liu et al., 2025b), which projects dense user embeddings into the LLM input space. The input formats and implementation details of each baseline are provided in Appendix B.

Evaluation Metrics. We use ROUGE-1/2/L, METEOR, and sacreBLEU to measure the similarity between generated reviews and reference reviews. These lexical-overlap-based metrics have been widely used in prior work on personalized review generation. We separately analyze how well the generated reviews reflect target item evidence and user preferences in Section 4.5.

Implementation Details. For all methods in the main experiments, we use Qwen2.5-7B-Instruct (Yang et al., 2024) and Llama-3.1-8B-Instruct (Grattafiori et al., 2024) as backbone LLMs.[2] Both backbones are compared under the same data construction, input format, training setup, and evaluation protocol. For SFT, we train for 2 epochs with a learning

[1] https://amazon-reviews-2023.github.io/

[2] We use the official Hugging Face checkpoints: Qwen/Qwen2.5-7B-Instruct and meta-llama/Llama-3.1-8B-Instruct.

Table 1: Main results on Games and Books. Scores are displayed to three decimal places using standard rounding. Best results are shown in bold and second-best results are underlined, with ranking determined by the original unrounded values.

Backbone	Domain (→)	Games					Books				
	Method (↓)	R-1	R-2	R-L	BLEU	MET.	R-1	R-2	R-L	BLEU	MET.
Qwen2.5-7B -Instruct	Non-Perso	0.212	0.033	0.118	0.151	0.108	0.218	0.033	0.126	0.726	0.126
	RAG (Salemi et al., 2024)	0.239	0.039	0.128	0.431	0.127	0.243	0.044	0.145	1.235	0.143
	PAG (Richardson et al., 2023)	0.229	0.035	0.125	0.418	0.122	0.232	0.038	0.135	1.160	0.142
	DPL (Qiu et al., 2025b)	0.229	0.033	0.121	0.575	0.131	0.231	0.032	0.128	1.347	0.150
	DEP (Qiu et al., 2025a)	0.269	0.057	0.153	0.588	0.139	0.285	0.077	**0.195**	2.235	0.161
	PPlug (Liu et al., 2025b)	0.269	0.049	0.142	0.352	0.131	0.235	0.034	0.118	1.210	**0.171**
	OPPU (Tan et al., 2024)	0.260	0.055	0.149	0.788	0.127	0.281	0.068	0.161	1.771	0.171
	MAPLE-SFT (w/o DPO)	0.258	0.051	0.144	0.812	0.125	0.275	0.061	0.171	2.109	0.148
	MAPLE	**0.322**	**0.066**	**0.162**	**2.000**	**0.175**	**0.317**	**0.084**	0.188	**3.440**	0.168
Llama-3.1-8B -Instruct	Non-Perso	0.253	0.038	0.130	0.606	0.142	0.246	0.036	0.132	1.447	0.160
	RAG (Salemi et al., 2024)	0.259	0.042	0.133	0.630	0.141	0.248	0.039	0.138	1.682	0.162
	PAG (Richardson et al., 2023)	0.257	0.040	0.131	0.810	0.148	0.240	0.036	0.133	1.714	0.169
	DPL (Qiu et al., 2025b)	0.261	0.040	0.130	0.914	0.153	0.242	0.035	0.131	1.520	0.168
	DEP (Qiu et al., 2025a)	0.264	0.055	0.148	1.042	0.145	0.299	0.087	0.203	2.156	0.173
	PPlug (Liu et al., 2025b)	0.253	0.036	0.116	1.247	**0.164**	0.206	0.027	0.104	0.911	0.170
	OPPU (Tan et al., 2024)	0.273	0.054	0.149	0.955	0.137	0.296	0.080	0.194	2.223	0.161
	MAPLE-SFT (w/o DPO)	0.278	0.053	0.142	0.699	0.136	0.287	0.069	0.181	2.116	0.157
	MAPLE	**0.316**	**0.063**	**0.157**	**1.639**	0.157	**0.326**	**0.104**	**0.210**	**3.645**	**0.196**

rate of 2×10^{-5} and gradient accumulation of 16; the LoRA configuration is $r = 16$, $\alpha = 32$, and dropout 0.05. DPO is initialized from the SFT adapter and trained for 1 epoch with a learning rate of 5×10^{-6} and $\beta = 0.1$. Inference is performed with deterministic decoding, and the default number of style examples is set to 5. Detailed settings are provided in Appendix C.

4.2 Main Results (RQ1)

Table 1 reports the performance of MAPLE and the comparison methods across the two domains. MAPLE achieves the best performance on most metrics and consistently outperforms strong personalization baselines such as DEP, PPlug, and OPPU. This suggests that structuring the generation condition with a MULTI-ASPECT PREFERENCE PROFILE and a PERSONALIZED EVIDENCE PLAN is more effective for producing outputs that are similar to the target review than simply retrieving user histories or compressing them into a single profile or embedding.

Looking at each baseline, Non-Perso shows the lowest performance because it does not use user-specific information, while RAG and PAG achieve some improvements by incorporating the user's historical interactions. DPL, OPPU, DEP, and PPlug also show competitive performance, but MAPLE achieves more stable overall performance by organizing user preferences at the aspect level and linking them to the PERSONALIZED EVIDENCE PLAN.

Table 2: DPO-stage ablation results on the **Games** domain with Qwen2.5-7B-Instruct. Scores are displayed to three decimal places. Best and second-best results are marked **in bold** and underlined.

Variant	R-1	R-L	MET.
MAPLE	**0.322**	**0.162**	**0.175**
MAPLE-SFT (w/o DPO)	0.258	0.144	0.125
w/o style examples	0.302	0.147	0.157
w/o evidence plan	0.304	0.161	0.166
w/o multi-aspect profile	0.281	0.159	0.166
w/o user modeling	0.261	0.147	0.128
w/o evidence bank	0.308	0.161	0.164

At the same time, lexical-overlap metrics alone are insufficient to fully determine whether a generated review faithfully reflects item evidence and user preferences. Therefore, in Section 4.5, we separately analyze how often generated reviews include content not supported by the item EVIDENCE BANK, how well they cover the key points in the PERSONALIZED EVIDENCE PLAN, and how closely they align with the user's major aspects of interest.

4.3 Ablation Studies (RQ2)

This section analyzes the effect of each component in the proposed framework, with the results shown in Table 2. MAPLE-SFT retains the full MAPLE conditioning input and removes only the DPO stage, while w/o style examples removes the STYLE EXAMPLES S_u. w/o evidence plan removes the PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$ from

Table 3: Unseen-user results on Games and Books. Scores are displayed to three decimal places using standard rounding. Best results are shown in bold and second-best results are underlined, with ranking determined by the original unrounded values.

Backbone	Domain (→)	Games					Books				
	Method (↓)	R-1	R-2	R-L	BLEU	MET.	R-1	R-2	R-L	BLEU	MET.
Qwen2.5-7B -Instruct	Non-Perso	0.226	0.038	0.123	0.154	0.113	0.225	0.031	0.122	0.471	0.124
	RAG (Salemi et al., 2024)	0.227	0.043	0.132	0.241	0.124	0.230	0.026	0.124	0.565	0.128
	PAG (Richardson et al., 2023)	0.218	0.038	0.124	0.100	0.104	0.208	0.026	0.123	0.209	0.104
	DPL (Qiu et al., 2025b)	0.260	0.045	0.135	0.420	0.136	0.235	0.023	0.120	0.715	0.121
	DEP (Qiu et al., 2025a)	0.243	0.050	0.141	0.327	0.115	0.245	**0.082**	0.176	1.709	0.130
	PPlug (Liu et al., 2025b)	0.259	0.049	0.136	0.231	0.131	0.234	0.042	0.132	0.146	0.123
	OPPU (Tan et al., 2024)	0.263	0.057	0.147	**0.716**	0.133	0.249	0.042	0.154	1.381	0.129
	MAPLE	**0.278**	**0.058**	**0.153**	0.631	**0.147**	**0.253**	0.070	**0.180**	**2.043**	**0.136**
Llama-3.1-8B -Instruct	Non-Perso	0.248	0.039	0.131	0.574	0.152	0.256	0.035	0.134	1.328	0.161
	RAG (Salemi et al., 2024)	0.233	0.035	0.126	0.348	0.129	0.241	0.030	0.130	0.798	0.146
	PAG (Richardson et al., 2023)	0.239	0.041	0.133	0.341	0.141	0.246	0.031	0.131	1.034	0.144
	DPL (Qiu et al., 2025b)	0.261	0.040	0.131	0.686	**0.154**	0.227	0.029	0.119	1.110	0.149
	DEP (Qiu et al., 2025a)	0.271	**0.072**	**0.164**	0.606	0.146	0.287	0.072	0.178	1.230	0.143
	PPlug (Liu et al., 2025b)	0.253	0.041	0.142	0.354	0.123	0.263	0.044	0.151	0.819	0.123
	OPPU (Tan et al., 2024)	0.273	0.054	0.149	0.702	0.141	0.253	0.039	0.148	**1.664**	0.128
	MAPLE	**0.274**	0.059	0.156	**0.709**	0.148	**0.326**	**0.117**	**0.218**	1.312	**0.197**

the generator input, and w/o evidence bank does not use the target item EVIDENCE BANK E_i when constructing $\Pi_{u,i}$. w/o multi-aspect profile removes only the explicit preference profile P_u, whereas w/o user modeling removes user-specific conditioning (P_u and S_u) and uses a non-personalized plan constructed from E_i. Detailed comparisons of these variants are provided in Appendix I.4.

MAPLE achieves the highest performance across all evaluation metrics. In particular, the performance drops observed for MAPLE-SFT and w/o user modeling indicate that both DPO-based alignment and user-specific conditioning contribute to performance. The performance decrease in w/o style examples suggests that S_u is effective in capturing the user's writing style and narrative patterns. Meanwhile, w/o multi-aspect profile performs worse than MAPLE but maintains higher performance than w/o user modeling. This indicates that user preference information can be partially reflected through the EVIDENCE PLAN or STYLE EXAMPLES, while also showing that providing the explicit P_u remains an important personalization signal. Finally, the performance drops in w/o evidence plan and w/o evidence bank suggest that connecting P_u and E_i through $\Pi_{u,i}$ contributes to improving generation quality.

4.4 Robustness under Limited-History and Unseen-User Settings (RQ3)

RQ3 evaluates generalization under limited user histories and unseen-user settings. In the few-shot setting, we reconstruct P_u and $\Pi_{u,i}$ using only the most recent K reviews before the target review. In the unseen-user setting, personalization is performed using only the observable historical reviews of users who do not appear during training.

Figure 2 shows the few-shot results as K varies. In both domains, the largest performance gain is observed when increasing K from 1 to 4, after which the improvement becomes more gradual. This suggests that a small number of recent reviews is often sufficient to estimate the user's major preference aspects with reasonable stability. However, more historical reviews do not always lead to better performance, as older reviews may contain preferences that are less relevant to the current target item.

The proposed method also maintains stable performance in the unseen-user setting. This indicates that the model does not simply memorize the preferences of specific users, but instead learns the procedure of constructing P_u from user histories and connecting it with E_i. Overall, MAPLE demonstrates that personalized generation remains feasible even under limited-history and unseen-user scenarios.

4.5 Grounding and Qualitative Analyses (RQ4)

RQ4 analyzes whether generated reviews faithfully reflect target item evidence while also capturing user preferences. To this end, we use three metrics: Aspect UCR measures how often the generated text includes aspect-level content not supported by E_i; Coverage Any indicates whether

Case context. *The Missing Ones* (Detective Lottie Parker #1; rating=4.0). **User profile:** from 213 train reviews, the user places dominant importance on **Content Quality** ($w = 0.897$, sentiment $= 0.363$), followed by **Emotional Impact** ($w = 0.101$, sentiment $= 0.634$); other aspects have near-zero importance. **Target evidence:** well-structured whodunit, attractive characters, fast-paced plot, and a **somewhat guessable ending**.

Source	Review excerpt
Gold review	"For a debut novel, the author accomplished a lot: a **well-structured whodunit**, attractive characters, a deep mystery without boring police procedurals, and a **fast-paced plot**. Although the ending looks unpredictable, by the way the author played with the characters **I guessed**."
MAPLE	**"This is a riveting first novel ... a whodunit ... a maze of clues and false trails ... suspense high till the end, although the ending is somewhat predictable ... a good start for a promising career."**

Table 4: Main-text case study comparing the gold review and the review generated by MAPLE. The gold review is shown only for analysis and is not provided to the model. MAPLE preserves the target-specific content-quality signals while retaining the ending-related caveat.

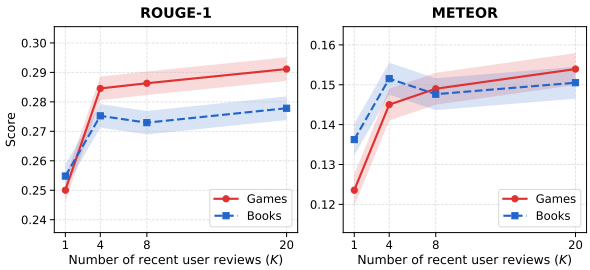


Figure 2: Few-shot performance with varying numbers of recent user reviews K. Performance improves substantially from $K = 1$ to $K = 4$, then tends to saturate or vary mildly. Full results are reported in Table 16.

the key points in $\Pi_{u,i}$ are reflected in the generated text; and `User Align.` evaluates how well the generated text aligns with the user's major preference aspects. The results are presented in Table 5. MAPLE achieves the highest performance on `Coverage Any` and `User Align.`, while keeping `Aspect UCR` low. `RAG` obtains the lowest UCR, but its lower `Coverage Any` and `User Align.` indicate that it tends to generate relatively conservative and less personalized responses. Compared with `DPL`, MAPLE achieves higher `Coverage Any` and lower `Aspect UCR` while maintaining a comparable level of `User Align.`. This demonstrates that $\Pi_{u,i}$ effectively connects user preferences with target item evidence. Table 4 qualitatively illustrates this trend. The user places high importance on *Content Quality* and *Emotional Impact*, while the target evidence includes a whodunit structure, fast pacing, and a predictable ending. MAPLE reflects the positive aspects of the mystery structure and pacing while preserving the caveat that the ending is predictable. Additional metrics and case studies are provided in Appendix I.5 and I.6.

Table 5: Grounding and personalization results on the `Games` domain with Qwen2.5-7B-Instruct. We report three primary metrics in the main text: Aspect UCR, which measures unsupported aspect content (micro, lower is better), coverage-any (micro, higher is better), and user aspect alignment (higher is better). Full results across domains and seen/unseen-user settings are provided in Appendix I.

Model	Aspect UCR ↓	Coverage Any ↑	User Align. ↑
`MAPLE`	0.195	**0.451**	**0.742**
`RAG`	**0.190**	0.389	0.726
`PAG`	0.199	0.389	0.710
`DPL`	0.212	0.412	0.730
`DEP`	0.217	0.374	0.668
`PPlug`	0.227	0.378	0.671
`OPPU`	0.222	0.364	0.665

5 Conclusion

This work proposed MAPLE, a framework that constructs a MULTI-ASPECT PREFERENCE PROFILE from implicit user histories and connects it with target item evidence. Rather than compressing user preferences into a single representation, MAPLE structures them at the aspect level and uses a PERSONALIZED EVIDENCE PLAN to select content that is both important to the user and supported by evidence. Through this design, we showed that faithful personalized generation can jointly account for personalization and grounding faithfulness. Experimental results across multiple domains, backbone models, limited-history settings, and unseen-user scenarios demonstrate that the proposed framework is effective for stable and reliable LLM personalization. We expect this work to serve as a foundation for future personalized generation approaches that integrate aspect-level user modeling with evidence-based generation.

Limitations

This work explicitly connects user preferences with grounding evidence to improve the reliability of personalized generation, but it has several limitations. First, our experiments focus on item review generation. Although we define evidence not as a concept limited to a specific item, but as grounding information that supports the generated content for the current input, our experiments instantiate it as an item-level EVIDENCE BANK. Therefore, for other personalized generation tasks such as document-grounded generation, dialogue response generation, and retrieval-based QA, the EVIDENCE BANK needs to be reconstructed according to the input structure of each task. Second, faithfulness in this work focuses on whether the generated text is supported by the EVIDENCE BANK. Thus, the proposed method aims to reduce content that is unsupported by the input or the review set, but it does not guarantee that the EVIDENCE BANK itself always aligns with real-world facts. For example, if many reviews contain biased or inaccurate information, evidence-grounded generation may also inherit some of these limitations. This can be mitigated by incorporating external knowledge verification or reliable source documents. Third, because the proposed method structures user histories by aspect, the resulting profile may be incomplete when sufficiently diverse user behaviors are not observed. We evaluate this issue under limited-history and unseen-user settings and find that the method maintains stable performance even with only a small number of recent reviews. Nevertheless, in extreme cold-start scenarios where no user history is available, additional signals such as global profiles or explicit user feedback may be required. Finally, this work evaluates personalization performance and grounding through automatic metrics and case studies. While these evaluations are useful for consistently comparing multiple models, they may not fully capture real users' perceived personalization quality, trustworthiness, or satisfaction. Future work should further validate the effectiveness of the proposed method through user studies or long-term interaction settings.

Ethical Statements

This work models user-level preferences from users' historical review data and incorporates them into personalized text generation. While this approach can improve generation quality and personalization performance, it requires careful attention to privacy because users' behavioral patterns or preferences may be revealed. We use only the publicly released Amazon Reviews 2023 dataset for research purposes. Because user-generated reviews may contain inaccurate or offensive content, model outputs may also inherit these limitations. Therefore, any real-world deployment of similar methods should include additional safeguards, such as user consent, data minimization, and anonymization. We used AI-assisted tools only for English proofreading, language polishing, and coding support. All research ideas, experimental design, analyses, and final decisions were conducted and verified by the author.

Acknowledgments

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2020-II201373, Artificial Intelligence Graduate School Program (Hanyang University)); the Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency, funded by the Ministry of Culture, Sports and Tourism in 2026 (No. RS-2026-25525376); and the Korea Planning & Evaluation Institute of Industrial Technology (KEIT), funded by the Ministry of Trade, Industry and Energy (MOTIE), Korea (No. RS-2026-25509027, Development and Demonstration of an Ontology- and On-Device AI-Based Autonomous Operation Agent System for Semiconductor Wafer Manufacturing Processes).

References

Steven Au, Cameron Dimacali, Ojasmitha Pedirappagari, Namyong Park, Franck Dernoncourt, Yu Wang, Nikos Kanakaris, Hanieh Deilamsalehy, Ryan A. Rossi, and Nesreen K. Ahmed. 2025. Personalized graph-based retrieval for large language models. In *Proceedings of the 39th Pacific Asia Conference on Language, Information and Computation*, pages 678–698, Hanoi, Vietnam. Association for Computational Linguistics.

Ruizhe Chen, Xiaotian Zhang, Meng Luo, Wenhao Chai, and Zuozhu Liu. 2024. Pad: Personalized alignment at decoding-time. In *The Thirteenth International Conference on Learning Representations*.

Youngbin Choi, Seunghyuk Cho, Minjong Lee, Moonjeong Park, Yesong Ko, Jungseul Ok, and Dongwoo

Kim. 2025. CoPL: Collaborative preference learning for personalizing LLMs. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 12875–12893, Suzhou, China. Association for Computational Linguistics.

Haizhou Du and Pengfei Li. 2025. Decoupling feature entanglement for personalized federated learning via neural collapse. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, CIKM '25, page 615–624, New York, NY, USA. Association for Computing Machinery.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, and 1 others. 2024. The Llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. *Preprint*, arXiv:2002.02126.

Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. *Preprint*, arXiv:1708.05031.

Yingzhi He, Xiaohao Liu, An Zhang, Yunshan Ma, and Tat-Seng Chua. 2025. Llm2rec: Large language models are powerful embedding models for sequential recommendation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, KDD '25, page 896–907, New York, NY, USA. Association for Computing Machinery.

Yupeng Hou, Jiacheng Li, Xiangjun Fu, Zhankui He, An Yan, Xiusi Chen, and Julian McAuley. 2026. Bridging language and items for retrieval and recommendation: Benchmarking llms as semantic encoders. *Preprint*, arXiv:2403.03952.

Jieyong Kim, Tongyoung Kim, Soojin Yoon, Jaehyung Kim, and Dongha Lee. 2026. RPM: Reasoning-level personalization for black-box large language models. In *International Conference on Learning Representations*.

Xinyu Li, Ruiyang Zhou, Zachary C. Lipton, and Liu Leqi. 2024. Personalized language modeling from personalized human feedback. *Preprint*, arXiv:2402.05133.

Enze Liu, Bowen Zheng, Wayne Xin Zhao, and Ji-Rong Wen. 2025a. Bridging textual-collaborative gap through semantic codes for sequential recommendation. *Preprint*, arXiv:2503.12183.

Jiongnan Liu, Yutao Zhu, Shuting Wang, Xiaochi Wei, Erxue Min, Yu Lu, Shuaiqiang Wang, Dawei Yin, and Zhicheng Dou. 2025b. LLMs + persona-plug = personalized LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9373–9385, Vienna, Austria. Association for Computational Linguistics.

Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton van den Hengel. 2015. Image-based recommendations on styles and substitutes. *Preprint*, arXiv:1506.04757.

Sheshera Mysore, Zhuoran Lu, Mengting Wan, Longqi Yang, Bahareh Sarrafzadeh, Steve Menezes, Tina Baghaee, Emmanuel Barajas Gonzalez, Jennifer Neville, and Tara Safavi. 2024. Pearl: Personalizing large language model writing assistants with generation-calibrated retrievers. In *Proceedings of the 1st Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual (CustomNLP4U)*, pages 198–219, Miami, Florida, USA. Association for Computational Linguistics.

Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 188–197, Hong Kong, China. Association for Computational Linguistics.

OpenAI, Aaron Hurst, Adam Lerer, Adam P. Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, Aleksander Mądry, Alex Baker-Whitcomb, Alex Beutel, Alex Borzunov, Alex Carney, Alex Chow, Alex Kirillov, Alex Nichol, and 400 others. 2024. Gpt-4o system card. *Preprint*, arXiv:2410.21276.

Qiyao Peng, Hongtao Liu, Hongyan Xu, Qing Yang, Minglai Shao, and Wenjun Wang. 2024. Review-llm: Harnessing large language models for personalized review generation. *Preprint*, arXiv:2407.07487.

Jiaming Qian, Xinting Liao, Xiangmou Qu, Zhihui Fu, Xingyu Lou, Changwang Zhang, Pengyang Zhou, Zijun Zhou, Jun Wang, and Chaochao Chen. 2025. Personalized federated recommendation with multi-faceted user representation and global consistent prototype. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, CIKM '25, page 2399–2408, New York, NY, USA. Association for Computing Machinery.

Yilun Qiu, Tianhao Shi, Xiaoyan Zhao, Fengbin Zhu, Yang Zhang, and Fuli Feng. 2025a. Latent inter-user difference modeling for LLM personalization. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 10599–10617, Suzhou, China. Association for Computational Linguistics.

Yilun Qiu, Xiaoyan Zhao, Yang Zhang, Yimeng Bai, Wenjie Wang, Hong Cheng, Fuli Feng, and Tat-Seng Chua. 2025b. Measuring what makes you unique: Difference-aware user modeling for enhancing LLM personalization. In *Findings of the Association for*

Computational Linguistics: ACL 2025, pages 21258–21277, Vienna, Austria. Association for Computational Linguistics.

Chris Richardson, Yao Zhang, Kellen Gillespie, Sudipta Kar, Arshdeep Singh, Zeynab Raeesy, Omar Zia Khan, and Abhinav Sethy. 2023. Integrating summarization and retrieval for enhanced personalization via large language models. *Preprint*, arXiv:2310.20081.

Stephen E. Robertson, Steve Walker, Susan Jones, Micheline Hancock-Beaulieu, and Mike Gatford. 1994. Okapi at TREC-3. In *Proceedings of The Third Text REtrieval Conference, TREC 1994, Gaithersburg, Maryland, USA, November 2–4, 1994*, volume 500-225 of *NIST Special Publication*, pages 109–126. National Institute of Standards and Technology.

Alireza Salemi, Sheshera Mysore, Michael Bendersky, and Hamed Zamani. 2024. LaMP: When large language models meet personalization. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7370–7392, Bangkok, Thailand. Association for Computational Linguistics.

Tobias Schimanski, Jingwei Ni, Mathias Kraus, Elliott Ash, and Markus Leippold. 2024. Towards faithful and robust llm specialists for evidence-based question-answering. *Preprint*, arXiv:2402.08277.

Teng Shi, Jun Xu, Xiao Zhang, Xiaoxue Zang, Kai Zheng, Yang Song, and Han Li. 2025. Retrieval augmented generation with collaborative filtering for personalized text generation. *Preprint*, arXiv:2504.05731.

Chenkai Sun, Ke Yang, Revanth Gangi Reddy, Yi Fung, Hou Pong Chan, Kevin Small, ChengXiang Zhai, and Heng Ji. 2025. Persona-DB: Efficient large language model personalization for response prediction with collaborative data refinement. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 281–296, Abu Dhabi, UAE. Association for Computational Linguistics.

Zhongxiang Sun, Yi Zhan, Chenglei Shen, Weijie Yu, Xiao Zhang, Ming He, and Jun Xu. 2026. When personalization misleads: Understanding and mitigating hallucinations in personalized llms. *Preprint*, arXiv:2601.11000.

Zhaoxuan Tan, Qingkai Zeng, Yijun Tian, Zheyuan Liu, Bing Yin, and Meng Jiang. 2024. Democratizing large language models via personalized parameter-efficient fine-tuning. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 6476–6491, Miami, Florida, USA. Association for Computational Linguistics.

Jonas Wallat, Maria Heuss, Maarten de Rijke, and Avishek Anand. 2025. Correctness is not faithfulness in retrieval augmented generation attributions. In *Proceedings of the 2025 International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval (ICTIR)*, ICTIR '25, page 22–32, New York, NY, USA. Association for Computing Machinery.

Wenjie Wang, Honghui Bao, Xinyu Lin, Jizhi Zhang, Yongqi Li, Fuli Feng, See-Kiong Ng, and Tat-Seng Chua. 2025. Learnable item tokenization for generative recommendation. *Preprint*, arXiv:2405.07314.

Ye Wang, Jiahao Xun, Minjie Hong, Jieming Zhu, Tao Jin, Wang Lin, Haoyuan Li, Linjun Li, Yan Xia, Zhou Zhao, and Zhenhua Dong. 2024. Eager: Two-stream generative recommender with behavior-semantic collaboration. *Preprint*, arXiv:2406.14017.

Yiyan Xu, Jinghao Zhang, Alireza Salemi, Xinting Hu, Wenjie Wang, Fuli Feng, Hamed Zamani, Xiangnan He, and Tat-Seng Chua. 2025. Personalized generation in large model era: A survey. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 24607–24649, Vienna, Austria. Association for Computational Linguistics.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, and 1 others. 2024. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*.

Tiandi Ye, Wenyan Liu, Kai Yao, Lichun Li, Shangchao Su, Cen Chen, Xiang Li, Shan Yin, and Ming Gao. 2025. Towards instance-wise personalized federated learning via semi-implicit bayesian prompt tuning. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*, CIKM '25, page 3877–3887, New York, NY, USA. Association for Computing Machinery.

Jinghao Zhang, Yuting Liu, Wenjie Wang, Qiang Liu, Shu Wu, Liang Wang, and Tat-Seng Chua. 2025a. Personalized text generation with contrastive activation steering. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7128–7141, Vienna, Austria. Association for Computational Linguistics.

Zhehao Zhang, Ryan A. Rossi, Branislav Kveton, Yijia Shao, Diyi Yang, Hamed Zamani, Franck Dernoncourt, Joe Barrow, Tong Yu, Sungchul Kim, Ruiyi Zhang, Jiuxiang Gu, Tyler Derr, Hongjie Chen, Junda Wu, Xiang Chen, Zichao Wang, Subrata Mitra, Nedim Lipka, and 2 others. 2025b. Personalization of large language models: A survey. *Preprint*, arXiv:2411.00027.

Yuchen Zhuang, Haotian Sun, Yue Yu, Rushi Qiang, Qifan Wang, Chao Zhang, and Bo Dai. 2024. Hydra: Model factorization framework for black-box llm personalization. *Preprint*, arXiv:2406.02888.

Table 6: Statistics of the final processed dataset. #Data denotes the number of target interactions in each split. Evidence Len. is the average word length of the EVIDENCE BANK E_i, including ITEM METADATA and aspect-level evidence. Output Len. is the average word length of the target review $y_{u,i}$.

Dataset	Split	#Data	Evidence Len.	Output Len.
Games (7,122)	Train	6,508	887.06±452.11	311.26±361.05
	Val	307	650.34±411.00	351.69±420.24
	Test	307	670.14±465.21	322.88±397.94
Books (29,801)	Train	29,127	738.60±503.21	193.35±208.58
	Val	337	744.08±343.05	146.36±194.97
	Test	337	770.99±501.94	147.18±202.60

Table 7: Comparison of baseline methods and MAPLE. "Current User Data" indicates whether the method uses the target user's history H_u. "Peer User Data" indicates whether the method explicitly uses other users' reviews or representations as a personalization signal. "Explicit Item Evidence" indicates whether the method constructs target-side grounding evidence and uses it during generation.

Method	Training-Free	Current User Data	Peer User Data	Difference-Aware	Explicit Item Evidence
Non-Perso	✓	✗	✗	✗	✗
RAG	✓	✓	✗	✗	✗
PAG	✓	✓	✗	✗	✗
DPL	✓	✓	✓	✓	✗
OPPU	✗	✓	✗	✗	✗
DEP	✗	✓	✓	✓	✗
PPlug	✗	✓	✗	✗	✗
MAPLE	✗	✓	✗	✓	✓

A Dataset Details

Dataset Source and Task. We use the Games and Books domains from the Amazon Reviews 2023 dataset[3] (Hou et al., 2026) to evaluate personalized item review generation. Each instance is a user–item interaction consisting of a user u, an item i, and a review $y_{u,i}$. Following the formulation in the main text, the task is to generate personalized review text conditioned on the user history H_u and the target item i. This task has been widely used in LLM personalization research and is well suited for jointly evaluating the reflection of user preferences and item-grounded generation (Ni et al., 2019; Qiu et al., 2025a,b).

Data Curation. We first removed interactions with missing user IDs, item IDs, ratings, timestamps, review text, or ITEM METADATA. When multiple reviews existed for the same user–item pair, we retained only the most recent review. We then applied iterative 5-core filtering, retaining only users and items with at least five interactions (McAuley et al., 2015; He et al., 2020). To secure sufficient user histories for personalization while reducing heavy-user bias, we excluded users with too few interactions as well as those with more than 500 interactions. Finally, we selected more than 300 users from each domain, and the dataset statistics are reported in Table 6.

Temporal Splits and Leakage Prevention. For each user, we sorted reviews by timestamp and used the most recent interaction for testing, the immediately preceding interaction for validation, and all earlier interactions for training. This follows the leave-one-out protocol widely used in recommender systems (He et al., 2025; Wang et al., 2024, 2025; Liu et al., 2025a). For each target instance $(u, i, y_{u,i})$, the target review $y_{u,i}$ to be generated was excluded when constructing the MULTI-ASPECT PREFERENCE PROFILE P_u, the EVIDENCE BANK E_i, and the PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$. In particular, the text and extractor labels of validation/test target reviews were not used to construct either the MULTI-ASPECT PREFERENCE PROFILE P_u or the EVIDENCE BANK E_i. For the target item i, we built E_i using only train reviews excluding the target review, thereby preventing label leakage from directly referencing the reference review.

[3] https://amazon-reviews-2023.github.io/

Method	Trainable	History Signal	Epochs	LR
Non-Perso	No	–	–	–
RAG	No	recent 5 reviews	–	–
PAG	No	profile + recent 5 reviews	–	–
DPL	No	5 reviews + 8 peers	–	–
OPPU	Yes	per-user LoRA adapter	3 + user-adapt	1×10^{-4}
DEP	Yes	latent user/difference repr.	3	1×10^{-5}
PPlug	Yes	dense user embedding	3	1×10^{-4} / 5×10^{-5}

Table 8: Baseline implementation summary. PPlug uses learning rate 1×10^{-4} for Games and 5×10^{-5} for Books. OPPU follows a task LoRA followed by per-user LoRA adaptation. Prompt-based baselines do not update generator parameters.

B Baselines Details

This appendix summarizes how each baseline is implemented under the experimental setting of this work. All methods use the same train/validation/test split, backbone LLM, target ITEM METADATA m_i, and output condition. Prompt-based baselines use the same generator, differing only in how the personalization signal is constructed, while trainable baselines are compared under training and inference conditions that are as consistent as possible while preserving the original personalization mechanism of each method. In addition, for all baselines, the target review $y_{u,i}$ is not used for retrieval, profile construction, or personalization signal construction. Table 7 summarizes the main properties of the baselines and MAPLE.

B.1 Baselines

Non-Perso. Non-Perso is a non-personalized baseline that does not use user-specific information. Its input includes only the target ITEM METADATA m_i, and it serves as a lower bound without personalization signals.

RAG. RAG (Salemi et al., 2024) retrieves a fixed number of reviews from the target user's historical review set H_u and directly includes them in the generator prompt. It uses the user history in raw-text form, and the target review $y_{u,i}$ is excluded from the retrieval pool.

PAG. PAG (Richardson et al., 2023) summarizes the target user's historical reviews into a compact textual profile and includes it in the generator input. Unlike RAG, which uses individual reviews as they are, PAG compresses the user history into a single natural-language profile.

DPL. DPL (Qiu et al., 2025b) is a difference-aware prompt-based baseline that compares the target user with representative peer users and summarizes inter-user differences into a natural-language profile. We provide the DPL profile to the generator together with the target user's recent reviews.

OPPU. OPPU (Tan et al., 2024) is a baseline that personalizes the generator by training parameter-efficient adapters for each user. Each user's historical reviews are reflected in the parameters of a user-specific LoRA adapter.

DEP. DEP (Qiu et al., 2025a) is a baseline that injects personalization signals as latent soft prompts. It encodes the target user's history and inter-user difference signals into continuous representations and projects them into the LLM input space to guide generation.

PPlug. PPlug (Liu et al., 2025b) is a plug-and-play personalization baseline that encodes the target user's historical reviews into a dense user embedding and projects it into the LLM input space. This enables user information to be incorporated without directly concatenating long user reviews into the prompt.

B.2 Baseline Implementation Details

This section summarizes the main settings used to implement the baselines. All baselines use the same train/validation/test split, target ITEM METADATA m_i and the same decoding configuration, backbone LLM, and decoding budget. The implementation summary is presented in Table 8.

Prompt-based Baselines. Non-Perso, RAG, PAG, and DPL are implemented as prompt-based baselines that do not update the generator parameters. Non-Perso uses only the target ITEM METADATA. RAG includes up to five recent reviews from the target user's train history in the prompt, using them only as signals for capturing the user's writing

style and broad preferences. PAG uses an offline-generated compact user profile together with up to five recent train reviews.

DPL-style Baseline. DPL is implemented by summarizing the differences between the target user and representative peer users into a natural-language profile. For each target instance, we use the target item context as a query and select up to five reviews from the target user's train history using BM25 (Robertson et al., 1994). For inter-user comparison, we consider as candidates the history reviews for which reviews of the same item by other users are available, and select up to eight peer users for each candidate. The default peer selector is kmeans, and we use BAAI/bge-m3 for both user and peer review embeddings. When embedding-based selection is not available, we fall back to recency-based selection. We then use an instruction-tuned LLM to summarize the user–peer differences and use the resulting summary as the final DPL profile.

Trainable Baselines. OPPU, DEP, and PPlug are baselines that use trainable personalization modules. OPPU is implemented as a per-user PEFT baseline that first trains a task-level LoRA and then uses it as initialization to train user-specific LoRA adapters. The task LoRA uses the target item context and up to five user history reviews selected by BM25, excluding reviews of the same item. DEP injects a user-specific representation and an inter-user difference representation as latent soft prompts. PPlug uses precomputed user history embeddings and query embeddings to project a dense user embedding into the LLM input space.

C Experimental Details

This appendix summarizes the training and inference settings of MAPLE described in the main text. The prompt construction and DPO pair construction are provided in Appendix G and Appendix H, respectively.

C.1 Common Training Configuration

Table 9 presents the common settings used for generator training. MAPLE adopts a two-stage training procedure consisting of LoRA-based SFT followed by grounded DPO. Both stages use the same backbone and generation input $x_{u,i}$, but SFT is trained to directly predict the target review $y_{u,i}$, whereas DPO performs preference optimization using chosen/rejected response pairs.

Aspect Discovery Prompt Template

Domain: {domain}

Input. We provide 50 actual product metadata entries and 50 user reviews from the target domain.

Task. Extract high-quality *Aspect Categories* that will serve as the baseline for multi-aspect sentiment analysis.

Step 1: Derive 6–10 core aspects from the data. Do not rely on predefined categories. Derive aspects by analyzing the provided metadata and reviews to identify what consumers evaluate and what factors drive purchasing decisions and satisfaction. Each aspect must represent a distinct and independent dimension, such as user experience, functional behavior, physical characteristics, or cognitive/emotional response.

Step 2: Apply naming conventions. Each aspect name must be a structural and descriptive noun phrase of 2–4 words. The aspect should be abstract enough to serve as a professional evaluation dimension. Avoid overly broad names such as *Quality*, *Design*, *Usability*, *Features*, or *Price*.

Output. Return JSON only:

```
{"aspects": [
    {"aspect_name": "...",
     "description": "..."}
]}
```

Figure 3: Abridged prompt template used to construct the domain-specific aspect inventory $\mathcal{A}$. The prompt asks the model to derive domain-relevant aspect categories from sampled metadata and reviews. The full prompt is included in the supplementary artifact.

C.2 Inference Configuration

All main results and ablation results are obtained under the same deterministic decoding setting. Table 10 summarizes the common inference configuration.

C.3 Model Selection, Seeds, and Hardware

Checkpoints are selected based on performance on the validation split at each training stage. We use fixed random seeds for data splitting, profile construction, EVIDENCE BANK construction, and personalized evidence planning. All generator training and inference are performed on a single NVIDIA H200 GPU with bf16 mixed precision. Additional implementation details and resources are available in our artifact repository.[4]

D Domain-Specific Aspect Inventory and Preference Extraction

This appendix describes the construction of the domain-specific aspect inventory $\mathcal{A}$ and the preference extractor used in MAPLE. $\mathcal{A}$ serves as the basis for connecting P_u and E_i within the same aspect space, while the preference extractor is used to extract preference tuples (a, d, s, h) from each review sentence.

[4] https://github.com/jiin124/maple_artifacts

Setting	SFT stage	Grounded DPO stage
Backbone	Qwen2.5-7B-Instruct / Llama-3.1-8B-Instruct	Qwen2.5-7B-Instruct / Llama-3.1-8B-Instruct
Parameterization	LoRA	LoRA (initialized from SFT adapter)
Training objective	assistant-only token loss	preference optimization with chosen/rejected pairs
Epochs	2	1
Learning rate	2×10^{-5}	5×10^{-6}
Per-device batch size	1	1
Gradient accumulation	16	16
LoRA rank r	16	16
LoRA alpha	32	32
LoRA dropout	0.05	0.05
Max sequence length	4096	4096
Max prompt length	–	2800
Preference coefficient β	–	0.1
Precision	bf16	bf16

Table 9: Common settings for generator training. SFT computes the token-level loss only on the assistant response, while grounded DPO performs preference optimization using chosen/rejected response pairs for the same input.

Setting	Value
Max new tokens	512
Decoding strategy	deterministic
Repetition penalty	1.12
No-repeat n-gram size	4
Style example count	5

Table 10: Common inference settings used in the main experiments. The effect of the number of style examples is analyzed separately in the ablation study.

D.1 Preference Tuple Schema

Given a review sentence x, the preference extractor extracts zero or more preference tuples:

$$\ell = (a, d, s, h). \tag{10}$$

Here, $a \in \mathcal{A}$ denotes an aspect category, d is a descriptor phrase that specifies the aspect, $s \in \{-1, 0, 1\}$ denotes the sentiment polarity, and h is the source span in which the expression is observed in the original text. Since a single sentence may mention multiple aspects simultaneously, it can be associated with multiple tuples.

The same tuple schema is applied to both the user history H_u and the item review history H_i. Tuples extracted from H_u are used to construct P_u, while tuples extracted from H_i are used to build E_i. In this process, the source span h observed in H_i is stored as an evidence candidate and can later be selected as the final evidence span e in $\Pi_{u,i}$.

D.2 Domain-Specific Aspect Inventory

This work uses a domain-specific aspect set $\mathcal{A}$ for each domain. $\mathcal{A}$ serves as the label space of the preference extractor and provides the basis for connecting P_u, E_i, and $\Pi_{u,i}$ within the same aspect space. We perform LLM-based aspect discovery using 50 randomly sampled ITEM METADATA entries and 50 reviews from the training split of each domain.

For aspect discovery, we use `gpt-4o-mini`[5]. Figure 3 shows the prompt used for aspect discovery, and Table 11 presents the final aspect sets for Games and Books.

D.3 Teacher Supervision and Local Preference Extractor

To assign preference tuples to the full corpus, we use a two-stage labeling pipeline that combines teacher supervision with a local preference extractor. First, based on the domain-specific aspect inventory $\mathcal{A}$, we use the teacher LLM `gpt-4o` (OpenAI et al., 2024) to generate teacher labels for 600 sentences from each of Games and Books. Each label is represented in the form of (a, d, s, h).

Using this teacher-labeled data, we LoRA fine-tune a local preference extractor based on `Qwen2.5-3B-Instruct`. The trained extractor predicts zero or more (a, d, s, h) tuples for each sentence and is then applied to the full corpus to construct P_u and E_i.

[5] `gpt-4o-mini-2024-07-18`

Domain	Aspect	Description
Games	Gameplay mechanics	Combat, exploration, puzzles, and progression mechanics
	Control responsiveness	Input response, control feel, camera, and movement
	Visual presentation	Graphics, art style, animation, and visual presentation
	Sound design	Music, sound effects, voice acting, and audio immersion
	Content variety and progression	Content variety, game modes, progression, and unlock structure
	UI/UX and accessibility	Menus, HUD, readability, accessibility, and ease of use
	Narrative and immersion	Story, characters, world-building, and immersion
	Technical performance and stability	Frame rate, bugs, crashes, loading, and server stability
	Replayability and long-term engagement	Replay value, endgame content, and long-term engagement
	Monetization and value	Price, DLC, monetization, and value for money
Books	Content Quality	Plot, writing style, characters, informativeness, and overall quality
	Emotional Impact	Immersion, emotional resonance, and affective response
	Value Cost	Value for money and purchase satisfaction
	Readability Accessibility	Readability, ease of understanding, and clarity of explanation
	Organization Structure	Chapter structure, narrative flow, and structural coherence
	Cultural Contextual Relevance	Cultural context, social relevance, and contextual fit

Table 11: Final domain-specific aspect inventory used for Games and Books. The inventory defines the shared aspect space for constructing P_u, E_i, and $\Pi_{u,i}$.


```
{"user_id": "A1XXXX...",
"profile_source_splits": ["train"],
"num_reviews": 12,
"num_labeled_mentions": 47,
"avg_rating": 4.25,
"shrinkage_alpha": 0.706,
"aspect_importance_shrunk": {"Gameplay mechanics": 0.31, "Visual presentation": 0.18, ...},
"aspect_sentiment_shrunk": {"Gameplay mechanics": 0.74, "Visual presentation": 0.21, ...},
"aspect_difference_from_global": {"Gameplay mechanics": 0.09, "Visual presentation": -0.02, ...}
}
```


Figure 4: Example of one row in the MULTI-ASPECT PREFERENCE PROFILE JSONL file. Each row corresponds to one user and stores the source split, the number of reviews and labeled mentions used to construct the profile, the shrinkage coefficient, and the final aspect-level importance and sentiment values.

D.4 Examples of Extracted Preference Tuples

Table 12 shows representative extracted preference tuples from the Games and Books domains. The examples illustrate that a sentence can contain zero, one, or multiple aspect-level tuples.

E Multi-Aspect Preference Profile Format and Evaluation-Time Variants

This appendix summarizes the smoothing procedure, storage format, and evaluation-time variants of the MULTI-ASPECT PREFERENCE PROFILE P_u used in MAPLE. The domain-specific extractor adapters are released on Hugging Face.[6]

E.1 Profile Smoothing

When user histories are sparse, estimates of aspect importance and sentiment may depend too heavily on a small number of reviews. To mitigate this issue, we interpolate each user-specific profile with a global aspect profile computed from the training users. Although the main text denotes the profile as $P_u = \{(a, p_{u,a}, q_{u,a})\}$ for notational simplicity, the actual implementation applies shrinkage smoothing with the global aspect profile for all users and uses the smoothed values $\tilde{p}_{u,a}$ and $\tilde{q}_{u,a}$ as the downstream profile values. The less user history is available, the more weight is assigned to the global profile; as sufficient history becomes available, the user-specific estimates are reflected more strongly.

Let the global profile be

$$P_{\text{global}} = \{(a, p_a^{(g)}, q_a^{(g)})\}_{a \in \mathcal{A}} \tag{11}$$

where $p_a^{(g)}$ and $q_a^{(g)}$ denote the average importance and sentiment of aspect a, respectively, computed over all training users. Let n_u be the number of reviews used for the profile of user u, and let c be the shrinkage constant. The final profile is computed

[6] games-asde-qwen25-3b-lora and books-asde-qwen25-3b-lora.

Domain	Sentence	Aspect a	Descriptor d	s	Source span h
Games	It is worth the money.	Monetization and value	worth the money	$+1$	worth the money
Games	The explosions are great as well.	Visual presentation	explosions are great	$+1$	explosions are great
Games	The screen size is a concern for older gamers, so if you have trouble seeing small screens, than this is a portable you should skip.	UI/UX and accessibility	screen size is a concern for older gamers	-1	screen size is a concern for older gamers
Games	Despite this, there's no denying that this is definitely an addicting and immersive game with relatively decent replay value.	Replayability and long-term engagement	relatively decent replay value	$+1$	relatively decent replay value
Books	The ending leaves a message of hope, related to the will each individual possess in its more intimate and spiritual values.	Emotional Impact	leaves a message of hope	$+1$	leaves a message of hope
Books	The chapters about the son added no depth or texture to the story.	Content Quality	added no depth or texture to the story	-1	added no depth or texture to the story
Books	His writing is often quite beautiful and lyrical.	Content Quality	often quite beautiful and lyrical	$+1$	often quite beautiful and lyrical
Books	I found it to be well told but not terribly exciting.	Content Quality	well told but not terribly exciting	0	well told but not terribly exciting

Table 12: Examples of preference tuples extracted from the Games and Books domains. Each tuple contains an aspect a, descriptor d, sentiment polarity s, and source span h observed in the original sentence.

Field	Description
user_id	User identifier
profile_source_splits	Source split used to construct the profile (e.g., train, support)
num_reviews	Number of reviews used to construct the profile, n_u
num_labeled_mentions	Total number of sentence-level aspect labels
avg_rating	Average rating over the available user history
shrinkage_alpha	Shrinkage coefficient α_u
aspect_importance	Raw aspect importance before shrinkage, $p_{u,a}$
aspect_sentiment	Raw aspect sentiment before shrinkage, $q_{u,a}$
aspect_importance_shrunk	Final aspect importance after shrinkage, $\tilde{p}_{u,a}$
aspect_sentiment_shrunk	Final aspect sentiment after shrinkage, $\tilde{q}_{u,a}$
aspect_difference_from_global	Difference from global aspect importance, $\tilde{p}_{u,a} - p_a^{(g)}$

Table 13: Main fields in the user profile file. aspect_importance_shrunk and aspect_sentiment_shrunk are the final profile values directly used in downstream personalized evidence planning, while raw estimates are retained for analysis and ablation.

as follows:

$$\alpha_u = \frac{n_u}{n_u + c}, \tag{12}$$

$$\tilde{p}_{u,a} = \alpha_u p_{u,a} + (1 - \alpha_u) p_a^{(g)}, \tag{13}$$

$$\tilde{q}_{u,a} = \alpha_u q_{u,a} + (1 - \alpha_u) q_a^{(g)}. \tag{14}$$

Downstream modules use

$$P_u = \{(a, \tilde{p}_{u,a}, \tilde{q}_{u,a})\}_{a \in \mathcal{A}} \tag{15}$$

as the final profile. We use $c = 5$ in this work.

E.2 Stored Profile Format

The multi-aspect preference profile P_u is stored in JSONL format, where each row corresponds to a single user. Figure 4 shows an example of the storage format, and Table 13 summarizes the main fields. Each profile includes the final importance and sentiment values used for downstream planning, along with raw estimates for analysis and deviations from the global profile.

E.3 Limited-History Profiles

In limited-history evaluation, we recompute the profile using only the most recent K reviews before the target review.

$$H_u^{(K)} = \mathrm{RecentK}(H_u, K), \tag{16}$$

$$P_u^{(K)} = f_{\mathrm{profile}}(H_u^{(K)}). \tag{17}$$

That is, the profile format remains unchanged, while only the amount of user history available

```
{"item_id": "B000XXXX",
"item_title": "Example Item",
"description": "Cleaned item description ...",
"categories": ["Category A", "Category B"],
"review_count": 27,
"bank_source_splits": ["train"],
"aspects": {
"Gameplay mechanics": {
"support_reviews": 9,
"support_labels": 14,
"mention_rate": 0.333,
"mean_sentiment": 0.571,
"sentiment_variance": 0.245,
"positive_descriptors": [{"descriptor": "fun combat", "count": 3, ...}],
"negative_descriptors": [{"descriptor": "repetitive missions", "count": 2, ...}],
"neutral_descriptors": [{"descriptor": "skill tree", "count": 1, ...}]
},
"Visual presentation": {...}
}
}
```


Figure 5: An example row from the item EVIDENCE BANK. Each item row is constructed from train-set reviews and stores which aspects are frequently mentioned, whether they are generally discussed positively or negatively, and what expressions are actually used in the reviews. Personalized evidence planning uses this information to determine "what can be said about this item."

for profile construction is restricted. We then use $P_u^{(K)}$ to reconstruct the personalized evidence plan $\Pi_{u,i}^{(K)}$ in the same manner.

E.4 Unseen-User Profiles

In the unseen-user setting, we separate training users from evaluation users and construct profiles using only the observable historical reviews of users who do not appear during training. Let the observable history be $H_u^{(\mathrm{obs})}$. We define

$$P_u^{(\mathrm{obs})} = f_{\mathrm{profile}}(H_u^{(\mathrm{obs})}) \tag{18}$$

The aspect inventory, preference extractor, and smoothing rule are kept the same as in the default setting. This setting evaluates whether MAPLE can construct P_u from the history of a new user without relying on specific user IDs observed during training.

F Evidence Bank and Personalized Evidence Planning

This appendix describes the implementation details of the EVIDENCE BANK E_i and the personalized evidence plan $\Pi_{u,i}$ used in MAPLE. Since the main text conceptually describes the roles of the two modules, we focus here on the information stored in E_i and the procedure by which the rule-based planner selects evidence-supported points.

F.1 Evidence Bank Construction

The evidence bank E_i is target item evidence constructed by aggregating preference tuples (a, d, s, h) extracted from the item review history H_i at the aspect level. For each aspect, E_i stores the mention rate, mean sentiment, sentiment variance, number of supporting reviews, number of supporting labels, and descriptor–source span candidates. The source span h observed in H_i is used as a candidate that can later be selected as an evidence span e in the PERSONALIZED EVIDENCE PLAN. Figure 5 and Table 14 show an example of the EVIDENCE BANK storage format and its main fields, respectively.

F.2 Planning Candidate Selection and Action Assignment

Given a target user u and a target item i, the planner selects the content to include in $\Pi_{u,i}$ by jointly using P_u and E_i. It first identifies aspects that are important to the user based on the aspect importance values in P_u, and then checks whether those aspects are sufficiently supported in E_i. Aspects that are not important to the user or lack sufficient evidence in E_i are excluded.

For the remaining aspects, the planner selects descriptors and evidence span candidates stored in E_i that are consistent with the sentiment direction of each aspect. This sentiment direction is determined based on the mean sentiment computed by aggregating the polarity labels $s \in \{-1, 0, 1\}$ of

```
{"plan_id": "R123__plan",
 "review_id": "R123",
 "user_id": "A1XXXX",
 "item_id": "B000XXXX",
 "split": "test",
 "item_title": "Example Item",
 "top_user_aspects": [
{"aspect": "Gameplay mechanics", "importance": 0.31, "sentiment": 0.74},
{"aspect": "Monetization and value", "importance": 0.18, "sentiment": -0.42}
 ],
 "personalized_evidence_plan": [
{"aspect": "Gameplay mechanics", "action": "highlight",
 "descriptor": "fun combat", "polarity": 1,
 "evidence": "the combat feels fun and satisfying",
 "item_mention_rate": 0.33, "item_mean_sentiment": 0.57,
 "support_reviews": 9, "support_labels": 14},
{"aspect": "Monetization and value", "action": "warn",
 "descriptor": "overpriced DLC", "polarity": -1,
 "evidence": "the DLC feels overpriced for the amount of content",
 "item_mention_rate": 0.19, "item_mean_sentiment": -0.61,
 "support_reviews": 5, "support_labels": 8}
 ],
 "num_claims": 2
 }
```


Figure 6: Example JSONL format of a personalized evidence plan $\Pi_{u,i}$. Each row corresponds to a user–item instance and links user-important aspects with evidence stored in E_i. Each plan item contains the selected aspect, action label, descriptor, evidence span, and support statistics from E_i.

the preference tuples for each aspect. Specifically, if the mean sentiment of an aspect in E_i is positive, the planner prioritizes positive descriptors and evidence spans; if it is negative, it prioritizes negative descriptors and evidence spans. Each selected point is represented as $\pi = (a, \delta, d, e)$.

The action label δ is determined based on the mean sentiment and sentiment variance aggregated in E_i. If the mean sentiment of an aspect is sufficiently positive and its variance is low, the planner assigns `highlight`, encouraging the generator to express the point as a strength. Conversely, if the mean sentiment is sufficiently negative and its variance is low, the planner assigns `warn`, encouraging the generator to express the point as a caution or weakness. If the absolute value of the mean sentiment is small or the sentiment variance is high, the planner assigns `hedge` to avoid overly assertive statements. Finally, when evidence exists but the point is not strong enough to emphasize, the planner assigns `mention` so that it is reflected briefly as supporting information.

F.3 Ranking and Final Plan Construction

Candidate points with action labels are ranked based on the aspect importance in P_u and the support statistics in E_i. The ranking score reflects aspect importance, mention rate, the magnitude of mean sentiment, and the number of supporting reviews/labels. Thus, points are prioritized when they correspond to aspects that are important to the user and are frequently mentioned and sufficiently supported in the target item's review history.

Finally, the top K points are selected to construct

$$\Pi_{u,i} = \{\pi_1, \ldots, \pi_K\}, \qquad \pi = (a, \delta, d, e)$$

Figure 6 shows an example of the JSONL storage format of $\Pi_{u,i}$. When there are not enough candidates, $\Pi_{u,i}$ is left empty; in this case, the generator is conditioned only on the item metadata m_i, P_u, and style examples S_u. Algorithm 1 summarizes the overall procedure.

G Generator Inputs and Prompt Template

This appendix summarizes the generator input construction, the selection of STYLE EXAMPLES, and the prompt template used in MAPLE.

G.1 Input Components and Training Formats

The generator input consists of the ITEM METADATA m_i, the MULTI-ASPECT PREFERENCE PROFILE P_u, the PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$, and the STYLE EXAMPLES S_u:

$$x_{u,i} = [m_i; P_u; \Pi_{u,i}; S_u]. \tag{19}$$

Field	Description
item_id	Item identifier
item_title	Item title
description	Cleaned item description
categories	Item category information
review_count	Number of reviews used to construct the evidence bank
bank_source_splits	Source split used to construct the evidence bank
aspects	Aspect-level summary information
aspects[a].support_reviews	Number of reviews mentioning aspect a
aspects[a].support_labels	Number of sentence-level labels assigned to aspect a
aspects[a].mention_rate	Proportion of reviews that mention aspect a
aspects[a].mean_sentiment	Mean sentiment for aspect a
aspects[a].sentiment_variance	Sentiment variance for aspect a
aspects[a].positive_descriptors	Positive descriptors and examples
aspects[a].negative_descriptors	Negative descriptors and examples
aspects[a].neutral_descriptors	Neutral descriptors and examples

Table 14: Main fields in the item EVIDENCE BANK. For each aspect, the EVIDENCE BANK stores how frequently the aspect is mentioned, whether it is discussed positively or negatively, and which expressions appear in the reviews. This structure allows the PERSONALIZED EVIDENCE PLAN to select information that is relevant to the user and supported by E_i, rather than relying only on frequent aspects.

Stage	Sys. Msg.	User Msg.	Gold Rev.	Cho./Rej.	Rev. Loss	Direct Gen.
SFT	✓	✓	✓	✗	✓	✗
DPO	✓	✓	✗	✓	✗	✗
Inference	✓	✓	✗	✗	✗	✓

Table 15: Input and supervision format across SFT, DPO, and inference in MAPLE. All stages share the same system message and user prompt. Gold Rev. denotes gold target review supervision; Cho./Rej. denotes chosen/rejected response pairs; Rev. Loss denotes token-level review generation loss; Direct Gen. denotes direct generation at inference time.

Here, S_u denotes STYLE EXAMPLES selected from the same user's historical reviews and is not used as factual evidence about the target item.

SFT, DPO, and inference use the same system message and user prompt, differing only in the form of supervision. SFT uses the target review $y_{u,i}$ as the assistant response, while DPO uses chosen/rejected response pairs for the same prompt. During inference, the generator produces the review directly without access to the reference response. Table 15 summarizes the differences among the three stages.

G.2 Style Example Selection

The STYLE EXAMPLES S_u are examples selected from the same user's historical reviews and are used to capture the user's writing style, length, and expression patterns. Prior work on LLM personalization has also used a small number of a user's historical records as prompts or personalization signals to reflect individual writing styles and preferences (Salemi et al., 2024; Qiu et al., 2025a,b; Mysore et al., 2024). Candidates are selected from the train history of user u, excluding the target review $y_{u,i}$ and any historical reviews of the same item as the target item. Each candidate review is scored by jointly considering its aspect overlap with the `personalized evidence plan`, its similarity to the user's average review length, and its recency. In this work, we assign weights of 0.70, 0.25, and 0.05 to aspect relevance, length fit, and recency, respectively, and select the top five reviews as S_u.

G.3 Prompt Template

Figure 7 shows the generator prompt used in MAPLE. The prompt instructs the model to use the user's historical review examples S_u only as references for writing style, length, and expression patterns. In contrast, the generated content is grounded in the item metadata m_i and the personalized evidence plan $\Pi_{u,i}$. DPO reuses the same prompt as SFT, and the construction of rejected responses is described in Appendix H.

H Counterfactual DPO Pair Construction

This appendix describes how MAPLE constructs DPO pairs. DPO uses the same input $x_{u,i}$ as SFT,

Generator Prompt Template

```
[System Instruction]
You are an evidence-grounded personalized review writer.
Use historical reviews only for tone, verbosity, and phrasing.
Do not copy target-item facts from historical examples.
All target-specific content must be supported by the item metadata or the personalized evidence plan.

[Target Item]
- title: <item title>
- categories: <item categories>
- description: <item description>

[Target User Multi-Aspect Preference]
- <aspect 1>: importance=<...>, user_sentiment=<...>, difference_from_global=<...>
...

[User Writing Examples]
Example 1 | past rating=<...> | past review title=<...>
<past review text>
...

[Personalized Evidence Plan]
1. action=<...> | aspect=<...> | polarity=<...> | descriptor=<...> | evidence=<...>
...

[Output Conditions]
Output only the review text.
```

Figure 7: Generator prompt skeleton used in MAPLE. The prompt separates historical reviews used for writing style from target item information used for content generation. Target-specific content is instructed to rely on item metadata or the personalized evidence plan, while style examples are used only for tone, verbosity, and phrasing.

Algorithm 1 Rule-Based Personalized Evidence Planning in MAPLE

Require: Multi-aspect preference profile P_u, evidence bank E_i, plan size K
Ensure: Personalized evidence plan $\Pi_{u,i}$

$\mathcal{C} \leftarrow \emptyset$ ▷ candidate items
for each aspect $a \in \mathcal{A}$ **do**
  read aspect importance $\tilde{p}_{u,a}$ from P_u
  read mention rate, mean sentiment, variance, and support size from $E_i[a]$
  if $\mathrm{Support}(E_i[a])$ is insufficient **then**
    continue
  end if
  $\rho_a \leftarrow \mathrm{Polarity}(\mathrm{MeanSent}(E_i[a]))$
  $\omega_a \leftarrow \mathrm{Uncertainty}(\mathrm{Var}(E_i[a]))$
  $\delta \leftarrow \mathrm{AssignAction}(\rho_a, \omega_a)$
  $(d, e) \leftarrow \mathrm{SelectEvidence}(E_i[a], \rho_a)$
  $r_a \leftarrow \mathrm{RankScore}(\tilde{p}_{u,a}, \mathrm{MentionRate}(E_i[a]), |\mathrm{MeanSent}(E_i[a])|, \mathrm{SupportSize}(E_i[a]))$
  $\mathcal{C} \leftarrow \mathcal{C} \cup \{(r_a, a, \delta, d, e)\}$
end for
$\Pi_{u,i} \leftarrow$ top-K items in $\mathcal{C}$ ranked by r_a
return $\Pi_{u,i}$

while treating the target user's actual review $y_{u,i}$ as the chosen response $y^+_{u,i}$ and an actual review $y_{u',i}$ written by another user $u' \neq u$ for the same item i as the rejected response $y^-_{u,i}$.

The rejected response preserves relevance to the target item, but may differ from the target user's aspect-level interests, sentiment tendencies, and writing style, thereby providing a stronger comparison target than a random negative. Accordingly, DPO aligns the generator to prefer not merely reviews that fit the item, but reviews that better match the target user even for the same item.

H.1 Hard Negative Selection

For each target user u and item i, we first construct a candidate user set $\mathcal{C}_{u,i}$ from which the rejected response is selected. $\mathcal{C}_{u,i}$ consists of users who wrote a review for the same item i, are different from the target user u, have a valid MULTI-ASPECT PREFERENCE PROFILE, and whose review satisfies the minimum length requirement. When multiple candidates are available, we select the review from the user whose profile is most different from the target user's MULTI-ASPECT PREFERENCE PROFILE as a hard negative.

Let $\tilde{p}_{u,a}$ and $\tilde{q}_{u,a}$ denote the aspect importance and aspect sentiment of user u for aspect a, respectively. We define the profile distance between two users u and v as follows:

$$d(u, v) = \sum_{a \in \mathcal{A}} \lambda_p \left| \tilde{p}_{u,a} - \tilde{p}_{v,a} \right| + \lambda_q \left| \tilde{q}_{u,a} - \tilde{q}_{v,a} \right|. \tag{20}$$

Here, λ_p and λ_q control the relative weights of the aspect-importance difference and the aspect-sentiment difference, respectively. In this work, we use $\lambda_p = 0.7$ and $\lambda_q = 0.3$.

Algorithm 2 Same-Item Counterfactual DPO Pair Construction

Require: SFT-format rows $\mathcal{R}$, review pool $\mathcal{H}$, user profiles $\mathcal{P}$, minimum length threshold τ
Ensure: DPO pairs $\mathcal{D}_{\text{DPO}}$

$\mathcal{D}_{\text{DPO}} \leftarrow \emptyset$
for each row $r \in \mathcal{R}$ **do**
  $u \leftarrow r.\text{user_id}$, $i \leftarrow r.\text{item_id}$
  $x \leftarrow r.\text{prompt}$
  $y^+ \leftarrow r.\text{target_review}$
  if $|y^+| < \tau$ or $P_u \notin \mathcal{P}$ **then**
    continue
  end if
  $\mathcal{C}_{u,i} \leftarrow$ reviews in $\mathcal{H}$ written for item i by users $u' \neq u$
  remove candidates whose review length is $< \tau$ or whose profile is unavailable
  if $\mathcal{C}_{u,i} = \emptyset$ **then**
    continue
  end if
  $h^* \leftarrow \arg\max_{h \in \mathcal{C}_{u,i}} d(u, h.\text{user_id})$
  $y^- \leftarrow h^*.\text{review_text}$
  add $(x, y^+, y^-, \textit{same-item counterfactual})$ to $\mathcal{D}_{\text{DPO}}$
end for
return $\mathcal{D}_{\text{DPO}}$

Finally, we select the candidate user u^* with the largest profile distance and use the review written by that user for item i as the rejected response:

$$u^* = \arg\max_{v \in \mathcal{C}_{u,i}} d(u, v), \qquad y^-_{u,i} = y_{u^*,i}. \quad (21)$$

H.2 Pair Construction Procedure

Algorithm 2 summarizes the DPO pair construction procedure. Since each SFT row already contains $x_{u,i}$, the target review $y_{u,i}$, P_u, $\Pi_{u,i}$, and S_u, DPO pair construction reuses the same prompt and only changes the supervision format to a chosen/rejected pair. The chosen response is the target user's actual review, while the rejected response is an actual review written by another user with the largest profile distance for the same item. In addition, pairs are excluded when either the chosen or rejected response is too short. In this work, we use a pair only when both responses contain at least τ characters, and set $\tau = 80$ in our experiments.

I Additional Results and Analysis

This section reports the complete results summarized in the main text, along with additional ablation results, LLM-judge evaluations, and qualitative analyses.

I.1 Full Limited-History Results

Figure 2 in the main text visualizes the main trends under the limited-history setting as K varies, focusing on ROUGE-1 and METEOR. Table 16 presents the corresponding full results, reporting all automatic evaluation metrics, including ROUGE-2, ROUGE-L, and BLEU. Overall, in both domains, the largest performance gain is observed when increasing K from 1 to 4, after which the improvement becomes smaller or varies slightly across metrics.

Table 16: Full limited-history results with Qwen2.5-7B-Instruct across Games and Books. K denotes the number of recent user reviews used to construct the MULTI-ASPECT PREFERENCE PROFILE and PERSONALIZED EVIDENCE PLAN.

	K	R-1	R-2	R-L	BLEU	MET.
Games	1	0.250	0.050	0.139	0.254	0.124
	4	0.285	0.061	**0.157**	0.801	0.145
	8	0.286	**0.062**	0.156	0.898	0.149
	20	**0.291**	0.061	0.156	**0.913**	**0.154**
Books	1	0.255	0.068	0.166	**2.351**	0.136
	4	0.275	**0.080**	**0.181**	2.283	**0.152**
	8	0.273	0.075	0.177	2.263	0.148
	20	**0.278**	0.076	0.181	1.960	0.151

Table 17: Ablation on the number of STYLE EXAMPLE in MAPLE with Qwen2.5-7B-Instruct.

Domain	#Ex.	R-1	R-2	R-L	BLEU	MET.
Games	0	0.302	0.056	0.147	1.443	0.157
	1	0.308	0.062	0.152	1.717	0.159
	3	0.316	0.064	0.157	1.157	0.157
	5	**0.322**	0.066	0.162	2.000	**0.175**
	8	0.318	**0.070**	**0.164**	**2.155**	0.171
Books	0	0.266	0.047	0.149	1.762	0.142
	1	0.288	0.064	0.170	2.884	0.157
	3	0.282	0.070	0.172	2.878	0.148
	5	**0.317**	**0.084**	**0.188**	**3.440**	**0.168**
	8	0.283	0.074	0.181	3.147	0.150

I.2 Effect of Style Examples

Figure 8 (a,b) visualizes changes in ROUGE-1 and METEOR according to the number of style examples, while Table 17 presents the full metric values. In both domains, performance generally improves when increasing the number of style examples from 0 to 5. This shows that a small number of the user's historical reviews helps reflect the target user's writing style, length, and expression patterns. In contrast, increasing the number to 8 does not yield consistent additional improvements and even decreases some metrics. Therefore, MAPLE uses five STYLE EXAMPLES as the default setting, balancing performance and input length.

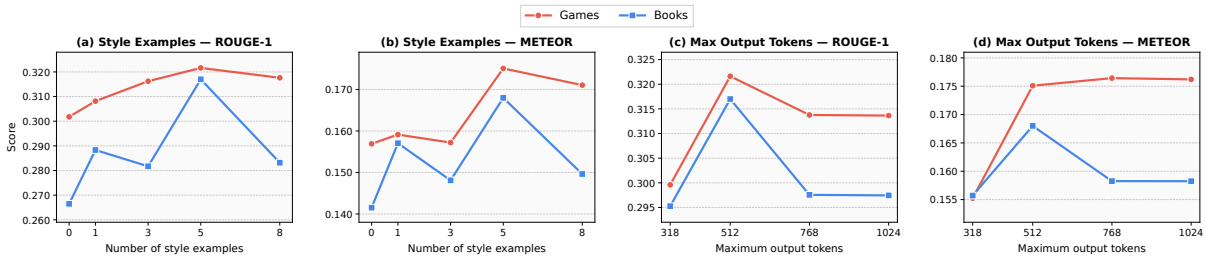


Figure 8: Sensitivity analysis for STYLE EXAMPLE and output length in MAPLE. Panels (a,b) show ROUGE-1 and METEOR trends as the number of STYLE EXAMPLE changes, while panels (c,d) show trends under different maximum output token budgets. Full metric values are reported in Tables 17 and 18.

Table 18: Ablation on the maximum number of new tokens in MAPLE with `Qwen2.5-7B-Instruct`.

Domain	Tokens	R-1	R-2	R-L	BLEU	MET.
Games	318	0.300	0.055	0.146	1.329	0.155
	512	**0.322**	**0.066**	**0.162**	2.000	0.175
	768	0.314	0.064	0.157	**2.234**	**0.176**
	1024	0.314	0.064	0.157	2.201	0.176
Books	318	0.295	0.083	0.188	3.111	0.156
	512	**0.317**	0.084	0.188	3.440	**0.168**
	768	0.298	**0.085**	**0.188**	3.544	0.158
	1024	0.297	0.084	0.188	**3.578**	0.158

I.3 Effect of Output Length

Figure 8 (c,d) visualizes changes in ROUGE-1 and METEOR according to the maximum output token budget, while Table 18 presents the full metric values. Overall, performance improves when increasing the budget from 318 to 512 tokens, but gains beyond 512 tokens are limited. In `Games`, 512 tokens achieves the best performance on the ROUGE metrics, while 768 or 1024 tokens show slight advantages in BLEU and METEOR. In `Books`, 512 tokens yields the highest ROUGE-1 and METEOR scores, whereas larger budgets improve BLEU but reduce ROUGE-1 and METEOR. Therefore, this work uses 512 as the default output budget, balancing generation quality and decoding cost.

I.4 Ablation Variant Details

Section 4.3 analyzes performance changes when key components are removed. Table 19 summarizes which components are used or removed in each ablation variant.

I.5 Additional Grounding and Personalization Metrics

Section 4.5 reports the main results in the `Games` seen-user setting to summarize the balance between grounding and personalization. Table 20 extends the same metrics to both seen-user and unseen-user settings in `Games` and `Books`. Overall, MAPLE maintains a low unsupported claim rate while achieving strong performance in PERSONALIZED EVIDENCE PLAN coverage and user aspect alignment, showing that the trends observed in the main text largely hold across domains and unseen-user settings.

Metric Definitions. The RQ4 metrics are computed based on preference tuples $\ell = (a, d, s, h)$ extracted by reapplying the preference extractor to the generated reviews. `Aspect UCR` is the proportion of generated aspect mentions whose aspects are not supported by the EVIDENCE BANK E_i, where lower values are better. `Coverage Any` measures the proportion of evidence points in the PERSONALIZED EVIDENCE PLAN $\Pi_{u,i}$ that are reflected in the generated text; a point is considered covered if it satisfies exact match, aspect–polarity match, or lexical overlap. `User Align.` is the cosine similarity between the aspect distribution of the generated review and the aspect-importance vector in the user profile P_u.

I.6 Qualitative Case Study

Tables 22 and 23 present the full case studies for the `Games` and `Books` domains, respectively. These cases supplement the qualitative analysis in Section 4.5 and are intended to compare whether each model jointly reflects the user's MULTI-ASPECT PREFERENCE PROFILE and the target item evidence. In both domains, some baselines reflect the overall item sentiment or parts of the user's preferences, but tend to miss target-specific caveats or generate unsupported details. In contrast, the proposed method more reliably preserves evidence-supported target details while reflecting the user's major preference aspects.

Variant	P_u	E_i	$\Pi_{u,i}$	S_u
MAPLE (Full)	✓	✓	✓	✓
w/o evidence plan	✓	–	–	✓
w/o evidence bank	✓	–	✓*	✓
w/o multi-aspect profile	–	✓	✓	✓
w/o user modeling	–	✓	non-personalized	–
w/o style examples	✓	✓	✓	–
MAPLE-SFT (w/o DPO)	✓	✓	✓	✓

Table 19: Component configurations of ablation variants in Section 4.3. P_u denotes the MULTI-ASPECT PREFERENCE PROFILE, E_i the EVIDENCE BANK, $\Pi_{u,i}$ the PERSONALIZED EVIDENCE PLAN, and S_u the STYLE EXAMPLE. ✓* indicates a plan constructed without using the target-side EVIDENCE BANK E_i.

Table 20: Full grounding and personalization results corresponding to Table 5. We report results for both seen-user and unseen-user settings across Games and Books. Lower UCR is better; higher coverage and alignment scores are better. Best and second-best results are marked based on unrounded values.

(a) Seen-user setting

Model	Asp. UCR ↓	Cov. Any ↑	U. Align. ↑
Games (seen-user)			
MAPLE	0.195	**0.451**	**0.742**
Non-Perso	0.235	0.356	0.718
RAG	**0.190**	0.389	0.726
PAG	0.199	0.389	0.710
DPL	0.212	0.412	0.730
DEP	0.217	0.374	0.668
PPlug	0.227	0.378	0.671
OPPU	0.222	0.364	0.665
Books (seen-user)			
MAPLE	0.021	**0.687**	**0.947**
Non-Perso	0.031	0.662	0.918
RAG	0.030	0.647	0.889
PAG	0.031	0.655	0.897
DPL	0.031	0.671	0.931
DEP	**0.011**	0.608	0.932
PPlug	0.030	0.685	0.927
OPPU	0.040	0.608	0.657

(b) Unseen-user setting

Model	Asp. UCR ↓	Cov. Any ↑	U. Align. ↑
Games (unseen-user)			
MAPLE	**0.144**	0.410	0.695
Non-Perso	0.181	0.394	0.646
RAG	0.186	0.365	0.648
PAG	0.153	0.365	0.640
DPL	0.146	0.406	**0.696**
DEP	0.205	0.336	0.639
PPlug	0.189	**0.438**	0.689
OPPU	0.202	0.255	0.449
Books (unseen-user)			
MAPLE	0.013	**0.654**	**0.964**
Non-Perso	0.014	0.622	0.948
RAG	0.013	0.617	0.902
PAG	0.013	0.545	0.874
DPL	**0.012**	0.635	0.921
DEP	0.013	0.577	0.936
PPlug	0.016	0.582	0.873
OPPU	0.015	0.563	0.961

I.7 LLM-as-a-Judge Pairwise Comparison

Table 21 presents the results of pairwise preference evaluation using gpt-4o (OpenAI et al., 2024) as the judge model. In each comparison, the outputs of MAPLE and the baseline were presented in random order to reduce position bias. We also provided the judge with shared evaluation context containing the target user's major preference aspects and target item evidence, and instructed it to evaluate personalization, grounding, and factual faithfulness based on this context. Win rate denotes the proportion of comparisons in which MAPLE is preferred over the corresponding baseline, and W/T/L indicates the numbers of wins, ties, and losses, respectively. The evaluation prompt is summarized in Figure 9.

Overall, MAPLE achieves a win rate above 50% against all baselines in both domains. In Games, it shows particularly large margins over Non-Perso and OPPU, while the comparisons with RAG, PAG, and DPL are relatively competitive. In Books, MAPLE also outperforms all baselines, with especially high win rates over DEP, PPlug, and OPPU. The relatively small gap against RAG, especially in Books, appears to stem from the tendency of RAG to generate substantially longer reviews by incorporating plot summaries, named entities, and descriptive details from the retrieved context. Such outputs can be advantageous under the information-specificity criterion of the LLM judge, although they may also increase the risk of hallucination and become less reliable when sufficient user history is unavailable. Nevertheless, MAPLE still outperforms RAG without directly relying on

LLM Judge Prompt Template

Role. You are an expert LLM judge evaluating personalized item reviews. Given two generated reviews, decide which one better aligns with the user's historical preferences and stylistic patterns while remaining faithful to the provided item facts.

Criterion 1: Information Specificity & Grounding. Prefer the response that uses precise item details from the source evidence, such as named entities, exact locations, concrete mechanics, or specific plot elements. Penalize vague summaries or abstract generalizations.

Criterion 2: Stylistic & Persona Alignment. Prefer the response that better reflects the user's historical writing style, tone, directness, enthusiasm, and review structure. Penalize generic review clichés, overly polished language, or marketing-like phrasing.

Criterion 3: Factual Faithfulness. Ensure that the response introduces no unsupported external knowledge. Hallucinated features, unmentioned sequels, or claims not supported by the target item evidence should be penalized.

Output. Return JSON only:

```
{
  "winner": "A" or "B" or "Tie",
  "dimensions": {
    "information_specificity": "A | B | Tie",
    "stylistic_alignment": "A | B | Tie",
    "factual_faithfulness": "A | B | Tie"
  },
  "reason": "Explain the decision in 2-3 sentences."
}
```

Figure 9: Abridged LLM judge prompt used for pairwise evaluation. The judge receives a shared evaluation context containing target user preferences and target item evidence, compares two generated reviews using information specificity and grounding, stylistic/persona alignment, and factual faithfulness, and returns a structured JSON decision.

long retrieved contexts, further supporting the effectiveness of connecting P_u and E_i through $\Pi_{u,i}$ rather than simply increasing contextual coverage.

Table 21: Pairwise LLM judge results on the Games and Books domains. Win rate denotes the percentage of all comparisons in which MAPLE is preferred over each baseline. W/T/L indicate the number of pairwise comparisons where MAPLE wins, ties, or loses against the corresponding baseline.

Domain	Baseline	Win Rate	W	T	L
Games	Non-Perso	76.22%	234	0	73
	RAG	52.44%	161	0	146
	PAG	54.72%	168	0	139
	DPL	55.37%	170	0	137
	DEP	55.37%	170	12	125
	PPlug	56.35%	173	9	125
	OPPU	69.06%	212	24	71
Books	Non-Perso	53.41%	180	3	154
	RAG	51.34%	173	0	164
	PAG	56.38%	190	13	134
	DPL	53.41%	180	23	134
	DEP	63.20%	213	0	124
	PPlug	65.88%	222	0	115
	OPPU	67.95%	229	0	108

Evaluation context	
Product	*Tom Clancy's Splinter Cell Chaos Theory* (Xbox), item_id=B0002SQ0BS, rating=5.0.
Target user profile	• Built from 13 train reviews; average rating = 4.77. • Dominant aspect: **CONTENT VARIETY AND PROGRESSION** ($w = 0.530$, sentiment $= 0.670$; difference from global $= +0.063$). • Secondary aspects: **CONTROL RESPONSIVENESS** ($w = 0.172$, sentiment $= 0.261$), **VISUAL PRESENTATION** ($w = 0.077$, sentiment $= 0.710$), and **MONETIZATION AND VALUE** ($w = 0.068$, sentiment $= 0.414$). • Although less frequent, **GAMEPLAY MECHANICS** and **SOUND DESIGN** show highly positive sentiment (0.949 and 0.892, respectively).
Gold target review (not provided to models)	"I'm not a big fan for **1-player games** because my brother and I always have friends over. ... Chaos Theory was going to have a **co-op mode** (**I'm a big fan of co-op**). ... They **make you think like a spy**: Which route should I take? Where are the enemies located? What is the best and most quiet way to take them out? ... The graphics are **top-notch**. A key aspect of the game is **sound**, especially online when using a mic."
Target item evidence (excluding target review)	• Cooperative mode is frequently discussed as a salient item feature. • Missions encourage stealth planning: route choice, enemy location, and quiet takedowns. • The game provides many moves, flexible approaches, strong graphics, and sound-based realism.

Method	Generated review (abridged)	Grounding / qualitative finding
Non-Perso	"Finally, a Splinter Cell for me! ... intense stealth missions ... The open-level design allows for strategic play, and **co-op mode** adds a thrilling multiplayer dimension."	Captures co-op and strategic play, but remains generic and does not reflect the user's contrast against one-player games.
RAG	"Finally, a Splinter Cell for me! ... The level design is wide open ... The addition of **cooperative mode** adds a whole new layer of excitement ... Every moment feels like a crucial decision."	Uses relevant item cues, but reads more like a product description and weakly captures the user's social/co-op motivation.
PPlug	"I have been waiting for this game for so long. I am a **huge fan of the first two games** ... I love how you can play as either **Sam or Maya**."	Partially positive, but contradicts the user's stated reason for choosing this title and introduces character details not supported by the target item evidence.
PAG	"Finally, a Splinter Cell for me! ... wide open ... strategic play ... **cooperative mode** ... well-crafted gaming experiences."	Similar to RAG: grounded in broad stealth/co-op evidence, but less personalized to the user's preference for playing with family/friends.
OPPU	"I finally got this game! ... I have always wanted to try out the Splinter Cell series ... I hope they release another one soon."	Too short and generic; misses the key evidence about co-op, tactical route planning, graphics, and sound.
DEP	"I love the graphics and the storyline. I especially like the fact that you can play cooperatively with a friend. **I wish I had a friend to play with though**."	Mentions co-op, but conflicts with the user's social-play context shown in the gold review.
DPL	"Finally, Splinter Cell has arrived ... The ability to choose between stealth and aggression adds a layer of strategy ... I particularly enjoyed the **cooperative mode** ..."	Strongly grounded and competitive, but frames the user as a longtime series fan and is less specific about the user's social gaming motivation.
MAPLE	**"I finally found a Splinter Cell game that I could enjoy ... The game has a lot of replayability since you can play it solo or co-op. The co-op is a blast since you can take down guards together and split up to cover more ground. The game also gives you plenty of tactical options, and the sound design makes the stealth feel more realistic, since every movement and nearby threat feels like something you have to pay attention to ..."**	**Most directly reflects the user's profile-level emphasis on content variety/progression and the target-specific evidence about co-op, tactical freedom, and stealth-oriented gameplay.**

Table 22: Full Games-domain qualitative case study. The gold target review is shown only for analysis and is not provided to the models, and the target item evidence is constructed without using it. The comparison shows whether each model reflects both the user's dominant preference aspects and details supported by the target item evidence.

Evaluation context	
Product	*The Missing Ones: An absolutely gripping thriller with a jaw-dropping twist* (Detective Lottie Parker #1), item_id=1786811138, rating=4.0.
Target user profile	• Built from 213 train reviews; average rating = 3.84. • Dominant aspect: **CONTENT QUALITY** ($w = 0.897$, sentiment $= 0.363$). • Secondary aspect: **EMOTIONAL IMPACT** ($w = 0.101$, sentiment $= 0.634$). • Other aspects have near-zero importance, e.g., Readability/Accessibility ($w = 0.002$).
Gold target review (not provided to models)	"For a debut novel, the author accomplished a lot: a **well-structured whodunit**, attractive characters, a deep mystery without boring police procedurals, and a **fast-paced plot**. Although the ending looks unpredictable, by the way the author played with the characters **I guessed**."
Target item evidence (excluding target review)	• Detective Lottie Parker; debut whodunit. • St Angela's orphanage and recent murders. • Attractive characters. • Fast-paced plot and clue-driven suspense, mystery. • Ending is somewhat guessable.

Method	Generated review (abridged)	Grounding / qualitative finding
Non-Perso	"Riveting first novel! ... the plot keeps you on edge ... **an unexpected twist that left me stunned**."	Captures generic thriller enthusiasm, but contradicts the target-specific caveat that the ending was guessable.
RAG	"... intense narrative and unexpected twists ... while the plot keeps you guessing until the very end ..."	Uses thriller/item cues, but still overstates unpredictability and misses the user's restrained ending judgment.
PPlug	"**Personalized Soft Prompt**: ... The twist at the end was truly jaw-dropping–I never saw it coming!"	Contains prompt-like artifacts and amplifies the opposite of the target review's ending caveat.
PAG	"... plot twists kept me guessing ... **moments that felt a bit predictable** ... character development is strong."	Recovers part of the predictable-ending caveat, but adds less grounded framing such as "well-researched."
OPPU	"... missing persons case involving two families ... the story unfolds slowly ... the ending is satisfying ..."	Introduces unsupported plot details and conflicts with the target evidence that the plot is fast-paced.
DEP	"The story is set in **London** ... DCI Jack Hunter ... all connected to a **cult**."	Mentions a mystery setting but introduces concrete entities and plot details not supported by the target item evidence.
DPL	"... impossible to predict the outcome ... a must-read ... a finale that leaves you breathless."	Over-promotional; ignores the ending-related caveat despite the user's dominant attention to content quality.
MAPLE	**"This is a riveting first novel ... a whodunit ... a maze of clues and false trails ... suspense high till the end, although the ending is somewhat predictable ... a good start for a promising career."**	**Most directly preserves the content-quality signals and retains the predictable-ending caveat rather than over-following the title's promotional framing.**

Table 23: Full Books-domain qualitative case study. The gold target review is shown only for analysis and is not provided to the models, and the target item evidence is constructed without using it. The comparison shows whether each model reflects both the user's dominant preference aspects and details supported by the target item evidence.