PRIME: Language Model Personalization with Cognitive Memory and Thought Processes

Anonymous ACL submission

Abstract

Large language model (LLM) personalization aims to align model outputs with individuals' unique preferences and opinions. While recent efforts have implemented various personalization methods, a unified theoretical framework that can systematically understand the drivers of effective personalization is still lacking. In this work, we integrate the well-established cognitive dual-memory model into LLM personalization, by mirroring episodic memory to historical user engagements and semantic memory to long-term, evolving user beliefs. Specifically, we systematically investigate memory instantiations and introduce a unified framework, PRIME, using episodic and semantic memory mechanisms. We further augment PRIME with a novel *personalized thinking* capability inspired by the slow thinking strategy. Moreover, recognizing the absence of suitable benchmarks, we introduce CMV dataset specifically designed to evaluate long-context personalization. Extensive experiments validate PRIME's effectiveness across both long- and short-context scenarios. Further analysis confirms that PRIME effectively captures dynamic personalization beyond mere popularity biases.

1 Introduction

Personalization (Schafer et al., 2001; Berkovsky et al., 2005) aims to tailor model outputs to individual users' needs, preferences and beliefs, moving beyond generic responses (Zhang et al., 2018; Huang et al., 2022; Tseng et al., 2024). While Large language models (LLMs) excel at diverse NLP tasks, users' demand for personalized LLMs that reflect their unique histories and preferences has grown (Salemi et al., 2024; Liu et al., 2025). For instance, we have seen personalization adopted into commercial applications, such as OpenAI's customizable GPTs,[1] which are essential for building trust and reducing interaction friction (Castells et al., 2015). In this work, we formally define a *personalized LLM* as one adapted to align with the individual preferences, characteristics, and beliefs, by utilizing user-specific attributes, past engagements, and context the user was exposed to (Zhang et al., 2024d). Various techniques have been explored for LLM personalization, including prompt engineering (Petrov and Macdonald, 2023; Kang et al., 2023), retrieval-augmented generation (Salemi et al., 2024; Mysore et al., 2024), efficient fine-tuning (Tan et al., 2024; Zhang et al., 2024b), and reinforcement learning from human feedback (Li et al., 2024). Yet these piecemeal approaches lack a unified framework for systematically identifying what makes personalization effective. We posit that drawing inspiration from established cognitive models of human memory (Atkinson and Shiffrin, 1968b) offers a principled way to understand and advance LLM personalization. Specifically, we propose a **dual-memory model** (Tulving et al., 1972; Tulving, 1985; Schacter et al., 2009) with **episodic memory** (specific personal experiences) and **semantic memory** (abstract knowledge and beliefs) that parallels existing LLM personalization techniques.

Based on the cognitive model, we begin by examining memory instantiations to understand their strengths and weaknesses. Next, we present a unified framework, dubbed **PRIME** (**P**ersonalized **R**easoning with **I**ntegrated **ME**mory), to integrate both memory mechanisms in principle. Such integration facilitates a holistic understanding of user queries and histories, enabling the model to generate responses that are both contextually relevant and aligned with the user's long-term beliefs. Furthermore, within PRIME, we introduce the generation of chain-of-thoughts (CoTs) using **personalized thinking**, which draws on the slow thinking strategy (Muennighoff et al., 2025; Chen et al., 2025). Yet, we find that generic CoT reasoning can hinder performance on tasks that require personal-

[1] https://openai.com/index/introducing-gpts/

ized perspectives (Guo et al., 2025). In contrast, by adapting the self-distillation strategy (Zhang et al., 2019; Pham et al., 2022; Wang et al., 2023), we unlock LLM's *personalized thinking capability*. This ability guides the model to perform customized reasoning, yielding more accurate, user-aligned responses and richer reasoning traces that reflect the user's history and traits.

Meanwhile, benchmarking LLM personalization capabilities is hindered by a lack of suitable datasets (Tseng et al., 2024). Most datasets focus on short-context queries and surface-level imitation (e.g., stylistic mimicry; Wu et al., 2020; Salemi et al., 2024), neglecting genuine personalization—*users' latent beliefs and perspectives*—which requires modeling deeper, long-term preferences and traits. To this end, we introduce a novel dataset derived from the *Change My View* (CMV) Reddit forum,[2] which comprises 133 challenging evaluation posts by 41 active authors, along with their $7,514$ historical engagements. CMV discussions feature *extended dialogues* where participants seek to change the original poster's (OP's) opinion on varied topics. We cast the interactions into a ranking-based recommendation task, where the objective is to identify the response that effectively alters the OP's point of view, as acknowledged by the OP.

We conduct extensive empirical experiments on both our curated CMV data and an existing LLM personalization benchmark—LaMP (Salemi et al., 2024). Results show that 1) semantic memory model behaves generally more robust than episodic memory model; 2) our proposed PRIME is compatible with models of different families and sizes, yielding better results; 3) the novel *personalized thinking* plays a pivotal role in improving personalization. Our analysis also demonstrates that personalized thinking can be enabled in training-free settings, offering flexibility in handling users with limited history which is often framed as the "cold-start" challenge (Zhang et al., 2025b). To assess how well our models capture user-specific characteristics, we inject other users' histories and measure the resulting performance drop, confirming that our method captures dynamic personalization rather than bandwagon biases.

In summary, our contributions are threefold:[3]

- We propose PRIME, a cognitively inspired unified framework for LLM personalization, further augmented with *personalized thinking*.
- We introduce a challenging dataset, derived from the CMV forum, with nuanced user beliefs and preferences in long-context setting.
- Experiments showcase the effectiveness and fidelity of PRIME, as well as the pivotal role of personalized thinking.

[2] https://www.reddit.com/r/changemyview/

[3] Code and data will be released. We will create a project page for our arxiv version.

2 Related Work

2.1 LLM Personalization

Methods for Personalization. Early personalization in NLP relied on explicit user models, i.e., a structured representation of user traits, to tailor system outputs (Amato and Straccia, 1999; Purificato et al., 2024). They rely on static demographic features (e.g., age, gender, location) and use handcrafted rules to adapt outputs (Gou et al., 2014; Kim et al., 2013; Gao et al., 2013). Latent-factor techniques like matrix factorization (Koren et al., 2009; Jiang et al., 2014) decompose the user-item interaction matrix into low-dimensional embeddings. The Transformer architecture (Vaswani et al., 2017) enables the learnable user embedding approach (Qiu et al., 2021; Deng et al., 2023). However, they overlook *unstructured* user-written content and fail to generalize across tasks, yielding shallow, brittle personalization and underscoring the need for more robust methods.

With LLMs, three major paradigms have emerged: prompt engineering, retrieval-augmented generation, and training-based parameterization. Prompt-based approaches prepend user context, such as profile summaries (Richardson et al., 2023) or past interactions (Liu et al., 2023; Petrov and Macdonald, 2023; Kang et al., 2023), to the model input, but this method suffers from LLMs' limited context window. An improved version relies on retrievers like BM25 (Robertson and Zaragoza, 2009) and FAISS (Douze et al., 2024) to fetch relevant user history, which is then included in the model input (Madaan et al., 2022; Salemi et al., 2024; Mysore et al., 2024). However, noisy or irrelevant retrieval limits their ability to capture fine-grained user preferences. To address these challenges, recent studies have proposed to parameterize the historical engagement through training, by learning embeddings (Doddapaneni et al., 2024; Ning et al., 2024), by fine-tuning light-weight adapters (Tan et al., 2024; Zhang et al., 2024b), or by employing RLHF to align with individuals' preferences (Christiano et al., 2017; Ouyang et al., 2022; Li et al.,

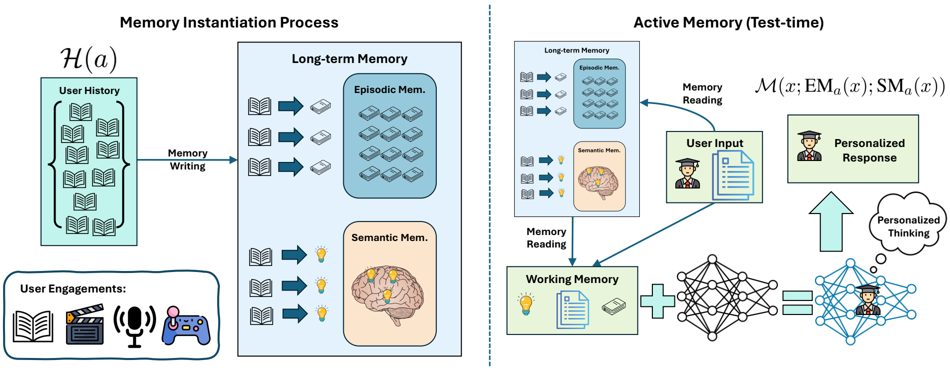

Figure 1: Overview of our unified framework, PRIME, inspired by dual-memory model (Tulving et al., 1972). PRIME is further augmented with *personalized thinking*, yielding more accurate, user-aligned responses.

2024). While these piecemeal approaches improve personalization on its own, there lacks a unified framework to bring them together. In this work, we bridge the gap by mirroring the dual-memory model in the human cognition process.

Datasets. The advance of personalized LLMs has been hampered by a shortage of comprehensive benchmarks (Tseng et al., 2024). Existing ones predominantly target short-context queries (Li et al., 2020; Salemi et al., 2024), and some even contain no user-generated context at all but just user-level metadata (Harper and Konstan, 2016; Wu et al., 2020). These tasks, while useful, assess personalization in a rather shallow way, such as simple rating prediction for short movie reviews (Ni et al., 2019) or capturing surface-level stylistic pattern in writing (Salemi et al., 2024). They overlook subtle dimensions of personalization, such as users' latent stance and evolving preferences during extended interactions. We also refer readers to Appendix B for discussions on the personalization evaluation.

2.2 Memory Mechanism for LLM

Decades of psychological research have converged on the following human memory components: sensory register, short-term memory, and long-term memory (Atkinson and Shiffrin, 1968a). Regarding the durable long-term memory, further distinction has been made between *episodic* and *semantic* memory (Tulving et al., 1972; Tulving, 1985). Episodic memory refers to autobiographical events we can re-experience (Tulving, 2002; Clayton et al., 2007), e.g., recalling a specific conversation that happened last night. Semantic memory, on the other hand, refers to general facts and knowledge we have accumulated (Saumier and Chertkow, 2002; McRae and Jones, 2013), such as knowing that NLP stands for Natural Language Processing. In this work, we posit that **the dual structure—episodic vs. semantic memories—is especially pertinent to LLM personalization**, as it mirrors the difference between remembering what happened in a particular interaction (*episodes*), and knowing what is true about the users' opinions, beliefs, and preferences (*semantics*).

Integrating memory into LLM-based systems quickly becomes a research frontier, as it holds the key to extending LLMs beyond fixed context windows, especially critical for LLM agents (Zhang et al., 2024a,c). A standard implementation of episodic memory is retrieval-based: past interactions (Park et al., 2023) and external facts (Yao et al., 2023) are indexed in a database and fetched on demand. In contrast, semantic memory is mostly realized parametrically: model's parameters are updated by training on user data to embed user-level knowledge (Zhang et al., 2024b; Magister et al., 2024). Recent hybrid approaches attempt to combine these two by merely concatenating textual summaries with retrieved experiences (Tan et al., 2024; Zhong et al., 2024; Gupta et al., 2024), resulting in only superficial fusion. Recognizing the isolated usage and the shallow integration, we formulate a more principled approach that enables deep information flow between episodic and semantic memories, which enables the successful use of the newly proposed *personalized thinking*.

3 CMV Dataset Construction

Change My View (CMV) is a Reddit forum (`r/ChangeMyView`) where participants discuss to understand different viewpoints on various topics. CMV has been widely used for studies on argumentation (Ji et al., 2018; Lin et al., 2024) and framing (Peguero and Watanabe, 2024). To our knowledge, we are the first to use CMV for LLM personalization, *defining personalization as recommending the most persuasive reply for a given OP* *(original post)*. An evaluation example from our dataset is shown in Figure A9.

CMV Dataset Curation. We obtained the raw CMV data (OPs, comments, and reply threads) from Academic Torrents.[4] We split the data chronologically: interactions from 2013–2022 form the *historical engagement set* and those from 2023–2024 form the *evaluation set*. The 2023 cutoff mitigates the data contamination issue—evaluation data have been part of the training corpus—since many open-weight models used in this study have the knowledge cutoff in 2023 (Dong et al., 2024a). We restructure each interaction by flattening the original multi-branch structures into linear threads of (OP, direct reply, follow-ups). We discard any thread containing deleted contents or authors, marked with "[deleted]" or "[removed]", since they offer no helpful personalization signal.

To convert conversations into a recommendation task, we exploit CMV's *delta* mechanism[5] to label replies: A direct reply that receives a delta becomes a *positive* example; all other direct replies under the same OP form the *negative* pool. For the sake of simplicity, we only consider single-turn conversations and truncate all follow-ups.

User Selection and Query Construction We restrict to *active users* who awarded at least 10 deltas in the historical engagement set (2013–2022) and granted at least one delta in 2023–2024. This yields 56 authors. Each evaluation query contains an OP and one of its delta-awarded replies with non-delta replies to the same OP as negatives. Our initial evaluation set comprises 327 queries from 56 OP authors. We further filter data based on their difficulty level, with details in Appendix C to mitigate popularity heuristics (Ji et al., 2020).

Statistics. Our final evaluation set includes 133 queries by 41 OP authors, supported by $7,514$ historical conversations published from 2013 to 2022. For the evaluation set (2023-2024), OP posts average 409 tokens; positive and negative replies average 200.2 and 105.8 tokens, respectively. Each positive reply is paired with 47.5 negatives on average ($6,317$ negatives total). In the historical engagement set, active authors have on average 28.1 positive and 155.1 negative conversations each.

[4]https://academictorrents.com/details/20520c420c6c846f555523babc8c059e9daa8fc5/

[5]OP authors award a "Δ" to replies that change their view.

4 Memory Instantiation

Inspired by cognitive theories of memory (Tulving et al., 1972), we investigate how different instantiations of episodic and semantic memories affect the LLM personalization. More specifically, we are interested in instantiating the memory-writing mechanism, i.e., how experiences are *encoded* into memory, and the memory-reading mechanism, i.e., how that information is *utilized* at test time. This study aims to provide insights into the **strengths and limitations** of various memory configurations.

Personalization with Dual-Memory. We adopt the dual-memory architecture, comprising episodic memory (EM) and semantic memory (SM), to define a **personalized LLM**, denoted as $\tilde{\mathcal{M}}$. The model processes an input query x from user a as follows:

$$\tilde{\mathcal{M}}(x) = \mathcal{M}(x; \mathrm{EM}_a(x); \mathrm{SM}_a(x)) \tag{1}$$
$$= \mathcal{M}(x; \phi(x, \mathcal{H}(a)); \theta \oplus \Delta_{\mathcal{H}(a)}\theta) \tag{2}$$

$\mathcal{M}$ represents the base LLM with parameters θ, $\mathcal{H}(a)$ denotes the historical engagements of user a, ϕ is the recall function for episodic memory, and $\Delta_{\mathcal{H}(a)}$ signifies the user-specific preference encoded in the personalized semantic memory. The operator $\oplus$ indicates the fusion of base LLM parameters with personalized adjustments.

For this set of preliminary experiments, we utilize LLAMA-3.1-8B (Dubey et al., 2024) and QWEN2.5-7B (Yang et al., 2024), for their representativeness. We conduct experiments on CMV data, and see Figure A9 for evaluation query.

Episodic Memory Instantiation. The writing mechanism typically involves storing raw interaction data for efficiency and completeness. We thus focus on the reading mechanism, exploring several recall strategies, $\phi(\cdot)$: 1) recall *complete history* (i.a., Shinn et al., 2023), 2) recall *most recent history* (i.a., Wang et al., 2024), and 3) recall *relevant history* (i.a., Park et al., 2023). Since full-history recall is intractable for long-context conversations, we focus our experiments on both recent and relevant recall. Additionally, we experiment with augmenting episodic memory using profile summaries derived from semantic memory (Richardson et al., 2023).

Semantic Memory Instantiation. We first explore different instantiations of the memory-writing function, specifically focusing on deriving $\Delta_{\mathcal{H}(a)}$ by *internalizing* information from user history

	Non-P	Recent	Relevant
Llama-3.1-8B	26.58	26.88 (26.67)	25.68 (25.96)
Qwen2.5-7B	27.89	25.51 (25.51)	25.66 (26.18)

Table 1: Aggregated results on **episodic memory** instantiation (10 runs). Complete results refer to Table A2. Parenthesized numbers represent textual-summary augmentation, which is usually beneficial.

$\mathcal{H}(a)$, i.e., encoding abstract concepts (e.g., preferences) into semantic memory. There are two forms of personalized semantic memory, $\Delta_{\mathcal{H}(a)}$: parametric and textual forms. We provide a brief summary and Table A4 presents the input–output mappings for each instantiation.

Parametric form, $\Delta_{\mathcal{H}(a)}\theta$, encodes user preferences into the model's parameters. We examine several training objectives:

- **Input-Only Training**: Suitable when human-written personalized outputs are unavailable (Tan et al., 2024). Objectives include *next token prediction (NTP)* and *conditional input generation (CIG)*, e.g., generate a post based on the title.
- **Fine-Tuning (FT)**: The most common practice to personalize model parameters (Zhang et al., 2024b; Magister et al., 2024; Tan et al., 2024), and we have two variants: *output-oriented FT (O-FT)* and *task-oriented FT (T-FT)*, depending on whether end task information is handy.
- **Preference Tuning**: Alternative to RLHF, employs methods like *DPO* (Rafailov et al., 2023) and *SIMPO* (Meng et al., 2024), an efficient variant without the need for the reference model, to align model outputs with user preferences. Although RLHF has been used to learn user preferences (Li et al., 2024), its simpler alternative, preference tuning, remains largely unexplored for LLM personalization.

Textual form represents user preferences as text, usually in the summary form. We explore:

- *Hierarchical Summarization(HSumm)*: hierarchically aggregates current interactions into concise summaries (Zhong et al., 2024).
- *Parametric Knowledge Reification (PKR)*: a novel method that leverages a model, trained on a user's engagement history but not for summarization, to generate a concise profile summary.

During the **memory reading** process, as shown in Equation (2), if semantic memory is in parametric form, the model parameters are adjusted as $\theta + \Delta_{\mathcal{H}(a)}\theta$; if in the textual form, $\oplus$ is implemented as prefixing the generated profile summary to the input query q.

For instantiations that involve training, we utilize LoRA (Hu et al., 2022) for its efficiency and interpretability, allowing $\Delta_{\mathcal{H}(a)}\theta$ as an abstract state to represent user-specific preferences and beliefs.

Discussion and Analysis. Tables 1 and 2 present comprehensive results for episodic and semantic memory instantiations. Table A2 also provides insights into the efficiency aspect.

Our experiments reveal that episodic memory grounded in simple recency often outperforms a semantic-similarity retrieval strategy—both in accuracy and speed—because the most recent interactions tend to be the strongest predictors of immediate user behavior. In contrast, semantic memory allows us to infer user preferences and latent traits even without task-specific labels, as validated by the improved performances achieved through *input-only training*. The best performance is reached by the *task fine-tuning (T-FT)*, which directly learns the mapping from the input query to the final desired outcome. Surprisingly, preference-tuning methods underperform here, which deserves more investigation in the future. Overall, using semantic memory (SM) alone generally leads to higher performance compared to using episodic memory (EM) alone. This suggests that semantic abstraction of user preferences and history might be more effective for personalization than simply recalling specific interactions.

It is important to emphasize that most of these memory instantiations have been examined individually in prior work, but never evaluated together on a common benchmark. To our knowledge, this study delivers the first comprehensive, head-to-head assessment of their personalization performance on long-context queries under a unified evaluation framework.

5 Framework: PRIME

5.1 Unified Framework

Section 4 offers insights into the instantiation of episodic and semantic memory separately, which is a common practice in the literature. Only a few works attempt to combine the two, and those mostly operate in the textual space only (Richardson et al., 2023; Zhong et al., 2024). To this end, we introduce our PRIME (illustrated in Figure 1) to unify both memory types, *so that the model can leverage detailed event histories alongside generalized user profiles*. This framework draws inspiration

	Non-P	Input Only		Fine Tuning		Preference Tuning		Textual	
		NTP	CIG	O-FT	T-FT	DPO	SIMPO	HSumm	PKF (ours)
Llama-3.1-8B	26.58	29.22	29.79	25.47	31.24	26.33	24.45	27.07	26.62
Qwen2.5-7B	27.89	28.11	28.41	28.01	30.20	28.04	17.37	26.83	27.02

Table 2: Average results of Hit@1, Hit@3, MRR and DCG@3 on semantic memory configuration (10 runs) Complete results refer to Table A2 where we additionally analyze the **time efficiency**. Non-P is a non-personalized baseline. Overall, the best configuration is to instantiate *parametric semantic memory with task-oriented fine-tuning*, if the task information is available. Parametric semantic memory generally outperforms its textual counterpart, whereas the preference-tuning approach delivers suboptimal results and thus deserves further investigation.

from the well-established cognitive theory of the dual-memory model (Tulving et al., 1972).

To maintain efficiency during training and inference, we implement episodic memory via *recency-based recall* and semantic memory via *task-oriented fine-tuning*. Importantly, our PRIME is virtually compatible with all valid instantiation approaches, as confirmed in section 4. Once instantiated, we freeze both memories.

At test time (right part of Figure 1), we process each input query x from an arbitrary user a following Equation (2). That is, we activate the corresponding LoRA matrices trained for the user a to enable personalized semantic memory through parameters merging, $\theta + \Delta_{\mathcal{H}(a)}\theta$. Next, we retrieve the most recent experiences for user a from the episodic memory to form a context-aware input query, $x \bigoplus \phi(x, \mathcal{H}(a))$, where $\bigoplus$ denotes text concatenation. We further augment our PRIME by *personalized thinking* (section 5.2), which jointly leverages these memories to generate more faithful, user-aligned responses and exhibit richer personalized reasoning traces for improved interpretability.

5.2 Personalized Thinking

Slow thinking, demonstrated by long CoT methods like DeepSeek-R1 (Guo et al., 2025) and s1 (Muennighoff et al., 2025), has shown promise, but its use in personalization is still in its infant stage. We are thus motivated to apply the slow thinking strategy to unlock personalized thinking

However, due to the fast thinking training paradigm (i.e., direct mapping from input to output), we find that fine-tuned LLMs have been turned into a specialist model and overfitted to the target space, i.e., losing the generalist capability of generating meaningful intermediate thoughts when prompted. A common error is repetition of tokens. To this end, we decide to unlock personalized thinking capabilities through training on **synthesized personalized thoughts**.

Capitalizing on the recent success of self-distillation (Zhang et al., 2019; Pham et al., 2022; Wang et al., 2023), we design the following algorithm to produce intermediate thoughts and feed them back to the model itself for learning the personalized thinking process. We start by an LLM with instantiated parametric semantic memory, i.e., $\tilde{\mathcal{M}}_{\mathrm{SE}_\alpha}(\cdot) = \mathcal{M}(\cdot; \emptyset; \mathrm{SM}_a(\cdot))$

- Step 1 (Profile Generation): We prompt $\tilde{\mathcal{M}}_{\mathrm{SE}_\alpha}(x)$ to generate a summary for a user, derived from the training on that user's history, following the same *Parametric Knowledge Reification* approach as described in Section 4.[6]
- Step 2 (Review History Engagement): We convert each historical engagement into a query as in the evaluation format (Figure A9), and we prompt the $\tilde{\mathcal{M}}_{\mathrm{SE}_\alpha}(x)$ to revisit all past engagements, and produce answers on them.
- Step 3 (Fast-thinking Filtering): We compare the produced answer with the ground truth answer, and then apply rejection sampling (Zhu et al., 2023; Yuan et al., 2023) to keep the queries that the model is able to get right.
- Step 4 (Proxy LLM Initialization & Reasoning): We follow the textual semantic memory reading process in section 4 and use the summary generated by $\tilde{\mathcal{M}}_{\mathrm{SE}_\alpha}(\cdot)$ to instantiate a proxy model, $\tilde{\mathcal{M}}'_{\mathrm{SE}_\alpha}(\cdot)$, of $\tilde{\mathcal{M}}_{\mathrm{SE}_\alpha}(\cdot)$. Next, we perform reverse engineering by feeding into the $\tilde{\mathcal{M}}'_{\mathrm{SE}_\alpha}(x)$ the input query x and answer, and prompt it to generate meaningful intermediate thoughts that could lead to the final personalized answer.
- Step 5 (Slow-thinking Filtering): After obtaining the intermediate thoughts and predicted answer produced by the proxy LLM, we perform another round of rejection sampling to keep the ones where the final answer matches the ground truth.

After obtaining the synthesized personalized thoughts, we perform standard fine-tuning with cross entropy, where the input is still a plain query q, but the model is expected to generate both personalized thinking trace and the final answer.

[6]Despite the model fails to generate meaningful thoughts, we find it still capable to generate meaningful summaries.

Our work also draws a clear distinction from concurrent work (Tang et al., 2025; Zhang et al., 2025a) on eliciting slow-thinking for LLM personalization. First, they only focus on the recommendation task, which is just one sub-task of LLM personalization, while we focus on various LLM personalization sub-tasks. Second, they use virtual tokens, i.e., a sequence of vectors, as intermediate steps while we are producing real tokens for the intermediate reasoning step, so users get insights into the reasoning trace. Third, the study in Zhang et al. (2025a) is limited to small-scale models (50M parameters) with its efficacy for larger models remain unverified, while we experiment with a diverse array of LLMs ranging from 3B to 14B.

6 Experiment

Datasets and Tasks. We conduct a holistic evaluation of LLM personalization across four task types (ranking, classification, regression, and generation), and on both short- and long-context queries. To this end, we benchmark models on our curated CMV to specifically probe long-context understanding. We also include a public LLM personalization benchmark, LaMP (Salemi et al., 2024), offering a testbed for all aforementioned tasks except for the ranking task. Specifically, we include their LaMP-1 to LaMP-5, and remove LaMP-6 and 7. We exclude LaMP-6 because it relies on a private dataset to which we have no access, and LaMP-7 because its GPT-3.5–generated labels may not faithfully represent real user behavior. Dataset statistics are included in Table A1, and Figure A9 shows an evaluation example from CMV dataset. Evaluation metrics for each task are in Appendix A.2,

Setup. We include recent, strong LLMs, showing promising results on various benchmarks. Specifically, we cover a diverse array of LLMs, ranging from mini LLMs (3B) to medium LLMs (14B), as shown in Table 3 and discussed in Appendix A.1.

On LaMP benchmark, for fair comparison with the SOTA approach (Tan et al., 2024), which is built upon LLAMA2-7B (Touvron et al., 2023), we only report performances based on LLAMA-3.1-8B.

Baselines and PRIME Variants. On both benchmarks, we compare our proposed PRIME with the non-personalized baseline, and generic reasoners like R1-DISTILL-LLAMA. We also compare our approach with the SOTA system on LaMP tasks, OPPU (Tan et al., 2024), which uses 100× more data by training on vast non-target users' history before fine-tuning on a target user's history.

Meanwhile, we compare PRIME with several variants: episodic memory only (EM), semantic memory only (SM), and PRIME with no personalized thinking (DUAL). For PRIME variants, we instantiate their memory in the same way as PRIME.

7 Results and Analysis

7.1 Main results

Major results are included in Table 3, and the full results (across all 5 metrics) can be found in Table A3. Below are our major findings.

1) *Generic Reasoning is not All We Need*: Enabling generic chain-of-thought often underperforms the non-thinking baseline (see Table A3). The uncustomized reasoning trace merely scratches the surface, seeking broad answers rather than to-the-point, user-specific responses. A detailed case study appears in Appendix D.

2) *Semantic Memory (*SM*) Beats Episodic Memory (*EM*)*: Consistent with our major finding in Section 4, SM alone generally outperforms EM alone, regardless of the model size or family.

3) DUAL *Often Underperforms* SM *Alone*: Surprisingly, integrating both memory types without personalized thinking (DUAL) occasionally yields lower or comparable results than SM along. This suggests that potential conflicts between episodic and semantic memories could backfire if not properly mediated.

4) *Model-agnostic Effectiveness*: PRIME consistently enhances performance across all base models at different scales, illustrating that our PRIME framework is robust and model-agnostic.

5) *Personalized Thinking Is Crucial*: By augmenting DUAL with personalized thinking, PRIME achieves superior performance over nearly all variants. This highlights the pivotal role of customized reasoning in improving personalization.

Results on LaMP. LaMP is a public benchmark of short-context queries that mainly tests surface-level personalization (e.g., imitating writing style). Although PRIME is designed to capture latent, evolving preferences, we are also interested in PRIME's ability of handling short, simple queries.

As shown in Table 4, the trends mirror those on CMV: SM outperforms EM, and DUAL sometimes trails SM due to potential memory conflicts. Crucially, personalized thinking in PRIME helps

	Non-P		EM		SM		DUAL		PRIME	
	Hit@3	Avg	Hit@3	Avg	Hit@3	Avg	Hit@3	Avg	Hit@3	Avg
Llama-3.2-3B	38.65	26.44	38.42	26.76	**43.61**	**30.25**	41.95	28.87	42.93	29.95
Llama-3.1-8B	36.77	26.58	44.14	31.43	43.01	31.24	44.59	32.24	**45.79**	**34.13**
Ministral-8B	36.77	25.60	39.92	27.94	**40.83**	27.97	**40.83**	28.39	40.75	**28.99**
Qwen2.5-7B	39.10	27.89	42.33	28.10	43.38	30.20	41.58	28.71	**45.19**	**32.29**
Qwen2.5-14B	41.28	30.24	44.96	30.81	51.35	37.22	52.03	37.68	**52.03**	**38.15**
Phi-4	41.50	29.63	44.89	31.71	42.63	31.09	43.98	32.61	**47.29**	**35.15**

Table 3: Results on CMV evaluation set (average of 10 runs). Avg is the aggregated metric of Hit@1, Hit@3, DCG@3, and MRR. Refer to Table A3 for breakdown results. Best results for *each* base model are **bold**. Non-P is a non-personalized baseline. PRIME performs better across the board, and is compatible with various base models.

Task (Metric)	Non-P	R1	EM	SM	DUAL	PRIME	SOTA
LaMP1 (Acc) ↑	44.7	47.2	49.6	46.3	52.8	**54.5**	79.7
LaMP1 (F1) ↑	30.9	46.9	45.7	31.7	52.9	**54.5**	79.4
LaMP2 (Acc) ↑	33.6	29.6	43.6	53.3	50.0	**54.3**	64.8
LaMP2 (F1) ↑	28.2	26.5	34.4	40.5	39.1	**42.7**	54.0
LaMP3 (MAE) ↓	.313	.366	.268	.214	**.188**	.223	.143
LaMP3 (RMSE) ↓	.605	.620	.567	.482	**.453**	.491	.378
LaMP4 (R-1) ↑	12.4	11.5	13.8	16.9	18.6	**18.8**	19.4
LaMP4 (R-L) ↑	11.0	10.1	12.4	15.2	16.7	**16.8**	17.5
LaMP5 (R-1) ↑	44.8	40.7	47.0	50.1	**52.2**	47.3	52.5
LaMP5 (R-L) ↑	34.7	32.7	38.5	44.8	**47.3**	40.6	47.3

Table 4: Results on LaMP benchmark. Non-P is the non-personalized baseline, R1 denotes R1-DISTILL-LLAMA. SOTA results, OPPU (Tan et al., 2024), use 100× more data of non-target users for training. Best performance among non-OPPUs is **bold**. In general, personalized thinking in PRIME leads to better results while the DUAL variant is a competitive baseline.

yield better results while DUAL is a competitive baseline for surface-level personalization tasks. This is inline with recent findings that overthinking might harm simple tasks (Sui et al., 2025). While PRIME surpasses all non-SOTA baselines, it remains behind the OPPU (Tan et al., 2024), which is trained on 100× more data—including other users' histories. This cross-user training clearly violates privacy constraints in reality by exposing private data (Kim et al., 2025). Given PRIME's use of only each user's own history, we deem the remaining gap acceptable.

7.2 Further Analysis

Train-free Personalized Thinking. Cold-start—performing personalized tasks with minimal history (e.g., $\leq$5 engagements)—remains challenging (Zhang et al., 2025b). We thus decide to approach this challenge with our proposed *personalized thinking* but under the training-free setting. Specifically, we prompt EM (Figure A13), and compare the training-free thinking with the standard non-thinking prompting.

As shown in Figure A2, personalized thinking boosts all metrics except Hit@1 (Figure A4). Despite trailing trained PRIME, it always outperforms other baselines including the generic reasoner (e.g., R1-distill LLMs), highlighting personalized thinking's promise even without training.

Profile Replacement. To evaluate the extent to which PRIME faithfully leverages and ingest a user's unique history, we perform a controlled "profile-replacement" experiment: for each test query, we substitute the target user's engagement history, for both episodic and semantic memory, with that of (i) the most similar user, (ii) a random user, or (iii) a maximally dissimilar user. There is also a "Self" baseline.

We report Hit@1 and average performance in Figure A1 and detailed breakdown in Figure A7. For both evaluated models, performance is generally at peak under the Self condition and degrades as the replacement profile diverges, and reaches the lowest when the profile is dramatically different. This consistent decline confirms that PRIME's reasoning and predictions depend critically on correct user history, and cannot be explained by simple bandwagon patterns or popularity heuristics (Ji et al., 2020). This further demonstrates that PRIME faithfully captures dynamic, user-specific preferences.

8 Conclusion

Inspired by the cognitive dual-memory model, we first systematically study different memory instantiations and then propose PRIME, a unified framework that integrates episodic and semantic memory mechanisms. We further augment PRIME with a novel *personalized thinking* capability, yielding more accurate, user-aligned responses and richer reasoning traces. To assess long-context personalization, we introduce the CMV dataset and conduct extensive experiments, which demonstrate the effectiveness of both PRIME and personalized thinking. Finally, our further analysis confirms PRIME 's fidelity to each user's unique history.

9 Limitations

Evaluation benchmarks. In this work, we have included two evaluation benchmarks, aiming to cover a diverse array of tasks, genres and domains. Yet, these two benchmarks cannot comprehensively represent the entire spectrum. For example, recent research efforts venture into long-form personalization (Kumar et al., 2024). In future research, we plan to extend PRIME to more applications, and examine its true generalizability in the wild.

GPU resources. The base LLMs used in this work are of 3 to 14 billions parameters. It is thus more time-consuming than traditionally small models like BERT (Devlin et al., 2019) at inference time, which in turn results in a higher carbon footprint. Specifically, we run each base LM on 1 single NVIDIA A40 or NVIDIA L40 with significant CPU and memory resources. The combined inference time for each LM on the three benchmarks ranges from 10 to 20 hours, depending on the configurations.

References

Marah I Abdin, Jyoti Aneja, Harkirat S. Behl, Sébastien Bubeck, Ronen Eldan, Suriya Gunasekar, Michael Harrison, Russell J. Hewett, Mojan Javaheripi, Piero Kauffmann, James R. Lee, Yin Tat Lee, Yuanzhi Li, Weishung Liu, Caio C. T. Mendes, Anh Nguyen, Eric Price, Gustavo de Rosa, Olli Saarikivi, and 8 others. 2024. Phi-4 technical report. *CoRR*, abs/2412.08905.

Giuseppe Amato and Umberto Straccia. 1999. User profile modeling and applications to digital libraries. In *Research and Advanced Technology for Digital Libraries, Third European Conference, ECDL'99, Paris, France, September 22-24, 1999, Proceedings*, volume 1696 of *Lecture Notes in Computer Science*, pages 184–197. Springer.

R. C. Atkinson and R. M. Shiffrin. 1968a. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence, editors, *The Psychology of Learning and Motivation*, volume 2, pages 89–195. Academic Press, New York.

R.C. Atkinson and R.M. Shiffrin. 1968b. Human memory: A proposed system and its control processes. volume 2 of *Psychology of Learning and Motivation*, pages 89–195. Academic Press.

Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. 2005. Entertainment personalization mechanism through cross-domain user modeling. In *Intelligent Technologies for Interactive Entertainment, First International Conference, INTETAIN 2005, Madonna di Campiglio, Italy, November 30 - December 2, 2005, Proceedings*, volume 3814 of *Lecture Notes in Computer Science*, pages 215–219. Springer.

Pablo Castells, Neil J. Hurley, and Saúl Vargas. 2015. Novelty and diversity in recommender systems. In Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors, *Recommender Systems Handbook*, pages 881–918. Springer US.

Zhipeng Chen, Yingqian Min, Beichen Zhang, Jie Chen, Jinhao Jiang, Daixuan Cheng, Wayne Xin Zhao, Zheng Liu, Xu Miao, Yang Lu, and 1 others. 2025. An empirical study on eliciting and improving r1-like reasoning models. *arXiv preprint arXiv:2503.04548*.

Paul F. Christiano, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 4299–4307.

Nicola S. Clayton, Lucie H. Salwiczek, and Anthony Dickinson. 2007. Episodic memory. *Current Biology*, 17(6):R189–R191.

Naihao Deng, Xinliang Zhang, Siyang Liu, Winston Wu, Lu Wang, and Rada Mihalcea. 2023. You are what you annotate: Towards better models through annotator representations. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12475–12498, Singapore. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Sumanth Doddapaneni, Krishna Sayana, Ambarish Jash, Sukhdeep Sodhi, and Dima Kuzmin. 2024. User embedding model for personalized language prompting. In *Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024)*, pages 124–131, St. Julians, Malta. Association for Computational Linguistics.

Yihong Dong, Xue Jiang, Huanyu Liu, Zhi Jin, Bin Gu, Mengfei Yang, and Ge Li. 2024a. Generalization or memorization: Data contamination and trustworthy evaluation for large language models. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 12039–12050, Bangkok, Thailand. Association for Computational Linguistics.

Yijiang River Dong, Tiancheng Hu, and Nigel Collier. 2024b. Can LLM be a personalized judge? In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 10126–10141, Miami, Florida, USA. Association for Computational Linguistics.

Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2024. The faiss library.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, and 82 others. 2024. The llama 3 herd of models. *CoRR*, abs/2407.21783.

Rui Gao, Bibo Hao, Shuotian Bai, Lin Li, Ang Li, and Tingshao Zhu. 2013. Improving user profile with personality traits predicted from social media content. In *Seventh ACM Conference on Recommender Systems, RecSys '13, Hong Kong, China, October 12-16, 2013*, pages 355–358. ACM.

Liang Gou, Michelle X. Zhou, and Huahai Yang. 2014. Knowme and shareme: understanding automatically discovered personality traits from social media and user sharing preferences. In *CHI Conference on Human Factors in Computing Systems, CHI'14, Toronto, ON, Canada - April 26 - May 01, 2014*, pages 955–964. ACM.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Priyanshu Gupta, Shashank Kirtania, Ananya Singha, Sumit Gulwani, Arjun Radhakrishna, Gustavo Soares, and Sherry Shi. 2024. MetaReflection: Learning instructions for language agents using past reflections. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 8369–8385, Miami, Florida, USA. Association for Computational Linguistics.

F. Maxwell Harper and Joseph A. Konstan. 2016. The movielens datasets: History and context. *ACM Trans. Interact. Intell. Syst.*, 5(4):19:1–19:19.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, and 1 others. 2022. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2):3.

Xiaolei Huang, Lucie Flek, Franck Dernoncourt, Charles Welch, Silvio Amir, Ramit Sawhney, and Diyi Yang. 2022. Usernlp'22: 2022 international workshop on user-centered natural language processing. In *Companion of The Web Conference 2022, Virtual Event / Lyon, France, April 25 - 29, 2022*, pages 1176–1177. ACM.

Lu Ji, Zhongyu Wei, Xiangkun Hu, Yang Liu, Qi Zhang, and Xuanjing Huang. 2018. Incorporating argument-level interactions for persuasion comments evaluation using co-attention model. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3703–3714, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Yitong Ji, Aixin Sun, Jie Zhang, and Chenliang Li. 2020. A re-visit of the popularity baseline in recommender systems. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, pages 1749–1752. ACM.

Guangyuan Jiang, Manjie Xu, Song-Chun Zhu, Wenjuan Han, Chi Zhang, and Yixin Zhu. 2023. Evaluating and inducing personality in pre-trained language models. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Meng Jiang, Peng Cui, Fei Wang, Wenwu Zhu, and Shiqiang Yang. 2014. Scalable recommendation with social contextual information. *IEEE Trans. Knowl. Data Eng.*, 26(11):2789–2802.

Wang-Cheng Kang, Jianmo Ni, Nikhil Mehta, Maheswaran Sathiamoorthy, Lichan Hong, Ed H. Chi, and Derek Zhiyuan Cheng. 2023. Do llms understand user preferences? evaluating llms on user rating prediction. *CoRR*, abs/2305.06474.

Jieun Kim, Ahreum Lee, and Hokyoung Ryu. 2013. Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model. *International Journal of Industrial Ergonomics*, 43(5):450–461.

Kyuyoung Kim, Jinwoo Shin, and Jaehyung Kim. 2025. Personalized language models via privacy-preserving evolutionary model merging. *arXiv preprint arXiv:2503.18008*.

Hannah Rose Kirk, Alexander Whitefield, Paul Röttger, Andrew M. Bean, Katerina Margatina, Juan Ciro, Rafael Mosquera, Max Bartolo, Adina Williams, He He, Bertie Vidgen, and Scott A. Hale. 2024. The PRISM alignment project: What participatory, representative and individualised human feedback reveals about the subjective and multicultural alignment of large language models. *CoRR*, abs/2404.16019.

Yehuda Koren, Robert M. Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37.

Ishita Kumar, Snigdha Viswanathan, Sushrita Yerra, Alireza Salemi, Ryan A. Rossi, Franck Dernoncourt, Hanieh Deilamsalehy, Xiang Chen, Ruiyi Zhang, Shubham Agarwal, Nedim Lipka, and Hamed Zamani. 2024. Longlamp: A benchmark for personalized long-form text generation. *CoRR*, abs/2407.11016.

Lei Li, Yongfeng Zhang, and Li Chen. 2020. Generate neural template explanations for recommendation. In *CIKM '20: The 29th ACM International Conference on Information and Knowledge Management, Virtual*

Event, Ireland, October 19-23, 2020, pages 755–764. ACM.

Xinyu Li, Zachary C. Lipton, and Liu Leqi. 2024. Personalized language modeling from personalized human feedback. *CoRR*, abs/2402.05133.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Jiayu Lin, Guanrong Chen, Bojun Jin, Chenyang Li, Shutong Jia, Wancong Lin, Yang Sun, Yuhang He, Caihua Yang, Jianzhu Bao, Jipeng Wu, Wen Su, Jinglu Chen, Xinyi Li, Tianyu Chen, Mingjie Han, Shuaiwen Du, Zijian Wang, Jiyin Li, and 10 others. 2024. Overview of ai-debater 2023: The challenges of argument generation tasks. *CoRR*, abs/2407.14829.

Jiahong Liu, Zexuan Qiu, Zhongyang Li, Quanyu Dai, Jieming Zhu, Minda Hu, Menglin Yang, and Irwin King. 2025. A survey of personalized large language models: Progress and future directions. *arXiv preprint arXiv:2502.11528*.

Junling Liu, Chao Liu, Renjie Lv, Kang Zhou, and Yan Zhang. 2023. Is chatgpt a good recommender? A preliminary study. *CoRR*, abs/2304.10149.

Aman Madaan, Niket Tandon, Peter Clark, and Yiming Yang. 2022. Memory-assisted prompt editing to improve GPT-3 after deployment. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 2833–2861, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Lucie Charlotte Magister, Katherine Metcalf, Yizhe Zhang, and Maartje ter Hoeve. 2024. On the way to LLM personalization: Learning to remember user conversations. *CoRR*, abs/2411.13405.

Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, UK.

Ken McRae and Michael N. Jones. 2013. 206 semantic memory. In *The Oxford Handbook of Cognitive Psychology*. Oxford University Press.

Yu Meng, Mengzhou Xia, and Danqi Chen. 2024. Simpo: Simple preference optimization with a reference-free reward. *Advances in Neural Information Processing Systems*, 37:124198–124235.

Meta AI Team. 2024. Llama 3.2: Revolutionizing edge ai and vision with open, customizable models for edge and mobile devices. `https://ai.meta.com/blog/llama-3-2-connect-2024-vision-edge-mobile-devices/`.

Mistral AI team. 2024. Un ministral, des ministraux: Introducing the world's best edge models. `https://mistral.ai/news/ministraux`.

Niklas Muennighoff, Zitong Yang, Weijia Shi, Xiang Lisa Li, Li Fei-Fei, Hannaneh Hajishirzi, Luke Zettlemoyer, Percy Liang, Emmanuel Candès, and Tatsunori Hashimoto. 2025. s1: Simple test-time scaling. *arXiv preprint arXiv:2501.19393*.

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.

Lin Ning, Luyang Liu, Jiaxing Wu, Neo Wu, Devora Berlowitz, Sushant Prakash, Bradley Green, Shawn O'Banion, and Jun Xie. 2024. User-llm: Efficient LLM contextualization with user embeddings. *CoRR*, abs/2402.13598.

Damilola Omitaomu, Shabnam Tafreshi, Tingting Liu, Sven Buechel, Chris Callison-Burch, Johannes C. Eichstaedt, Lyle H. Ungar, and João Sedoc. 2022. Empathic conversations: A multi-level dataset of contextualized conversations. *CoRR*, abs/2205.12698.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F. Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.

Joon Sung Park, Joseph C. O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, UIST 2023, San Francisco, CA, USA, 29 October 2023- 1 November 2023*, pages 2:1–2:22. ACM.

Arturo Martínez Peguero and Taro Watanabe. 2024. Change my frame: Reframing in the wild in r/changemyview. *CoRR*, abs/2407.02637.

Aleksandr V. Petrov and Craig Macdonald. 2023. Generative sequential recommendation with gptrec. *CoRR*, abs/2306.11114.

Minh Pham, Minsu Cho, Ameya Joshi, and Chinmay Hegde. 2022. Revisiting self-distillation. *arXiv preprint arXiv:2206.08491*.

Erasmo Purificato, Ludovico Boratto, and Ernesto William De Luca. 2024. User modeling and user profiling: A comprehensive survey. *CoRR*, abs/2402.09660.

Zhaopeng Qiu, Xian Wu, Jingyue Gao, and Wei Fan. 2021. U-BERT: pre-training user representations for improved recommendation. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 4320–4327. AAAI Press.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36:53728–53741.

Christopher 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. *CoRR*, abs/2310.20081.

Stephen E. Robertson and Hugo Zaragoza. 2009. The probabilistic relevance framework: BM25 and beyond. *Found. Trends Inf. Retr.*, 3(4):333–389.

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.

Daniel Saumier and Howard Chertkow. 2002. Semantic memory. *Current Neurology and Neuroscience Reports*, 2(6):516–522.

Daniel L. Schacter, Daniel T. Gilbert, and Daniel M. Wegner. 2009. Semantic and episodic memory. In *Psychology*, pages 185–186. Macmillan.

J. Ben Schafer, Joseph A. Konstan, and John Riedl. 2001. E-commerce recommendation applications. *Data Min. Knowl. Discov.*, 5(1/2):115–153.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. 2023. Reflexion: language agents with verbal reinforcement learning. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Yang Sui, Yu-Neng Chuang, Guanchu Wang, Jiamu Zhang, Tianyi Zhang, Jiayi Yuan, Hongyi Liu, Andrew Wen, Shaochen Zhong, Hanjie Chen, and Xia Ben Hu. 2025. Stop overthinking: A survey on efficient reasoning for large language models. *CoRR*, abs/2503.16419.

Annalisa Szymanski, Noah Ziems, Heather A. Eicher-Miller, Toby Jia-Jun Li, Meng Jiang, and Ronald A. Metoyer. 2025. Limitations of the llm-as-a-judge approach for evaluating LLM outputs in expert knowledge tasks. In *Proceedings of the 30th International Conference on Intelligent User Interfaces, IUI 2025, Cagliari, Italy, March 24-27, 2025*, pages 952–966. ACM.

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.

Jiakai Tang, Sunhao Dai, Teng Shi, Jun Xu, Xu Chen, Wen Chen, Wu Jian, and Yuning Jiang. 2025. Think before recommend: Unleashing the latent reasoning power for sequential recommendation. *arXiv preprint arXiv:2503.22675*.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton-Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, and 49 others. 2023. Llama 2: Open foundation and fine-tuned chat models. *CoRR*, abs/2307.09288.

Yu-Min Tseng, Yu-Chao Huang, Teng-Yun Hsiao, Wei-Lin Chen, Chao-Wei Huang, Yu Meng, and Yun-Nung Chen. 2024. Two tales of persona in LLMs: A survey of role-playing and personalization. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 16612–16631, Miami, Florida, USA. Association for Computational Linguistics.

Endel Tulving. 1985. How many memory systems are there? *American psychologist*, 40(4):385.

Endel Tulving. 2002. Episodic memory: From mind to brain. *Annual Review of Psychology*, 53(Volume 53, 2002):1–25.

Endel Tulving and 1 others. 1972. Episodic and semantic memory. *Organization of memory*, 1(381-403):1.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008.

Lei Wang, Jingsen Zhang, Hao Yang, Zhiyuan Chen, Jiakai Tang, Zeyu Zhang, Xu Chen, Yankai Lin, Ruihua Song, Wayne Xin Zhao, Jun Xu, Zhicheng Dou, Jun Wang, and Ji-Rong Wen. 2024. User behavior simulation with large language model-based agents. *Preprint*, arXiv:2306.02552.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. Self-instruct: Aligning language models with self-generated instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian, Danyang Liu, Xing Xie, Jianfeng Gao, Winnie Wu, and Ming Zhou. 2020. MIND: A large-scale dataset for news recommendation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3597–3606, Online. Association for Computational Linguistics.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, and 22 others. 2024. Qwen2.5 technical report. *CoRR*, abs/2412.15115.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R. Narasimhan, and Yuan Cao. 2023. React: Synergizing reasoning and acting in language models. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.

Zheng Yuan, Hongyi Yuan, Chengpeng Li, Guanting Dong, Chuanqi Tan, and Chang Zhou. 2023. Scaling relationship on learning mathematical reasoning with large language models. *CoRR*, abs/2308.01825.

Junjie Zhang, Beichen Zhang, Wenqi Sun, Hongyu Lu, Wayne Xin Zhao, Yu Chen, and Ji-Rong Wen. 2025a. Slow thinking for sequential recommendation. *arXiv preprint arXiv:2504.09627*.

Kai Zhang, Yangyang Kang, Fubang Zhao, and Xiaozhong Liu. 2024a. LLM-based medical assistant personalization with short- and long-term memory coordination. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 2386–2398, Mexico City, Mexico. Association for Computational Linguistics.

Kai Zhang, Lizhi Qing, Yangyang Kang, and Xiaozhong Liu. 2024b. Personalized LLM response generation with parameterized memory injection. *CoRR*, abs/2404.03565.

Linfeng Zhang, Jiebo Song, Anni Gao, Jingwei Chen, Chenglong Bao, and Kaisheng Ma. 2019. Be your own teacher: Improve the performance of convolutional neural networks via self distillation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3713–3722.

Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing dialogue agents: I have a dog, do you have pets too? In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia. Association for Computational Linguistics.

Weizhi Zhang, Yuanchen Bei, Liangwei Yang, Henry Peng Zou, Peilin Zhou, Aiwei Liu, Yinghui Li, Hao Chen, Jianling Wang, Yu Wang, Feiran Huang, Sheng Zhou, Jiajun Bu, Allen Lin, James Caverlee, Fakhri Karray, Irwin King, and Philip S. Yu. 2025b. Cold-start recommendation towards the era of large language models (llms): A comprehensive survey and roadmap. *CoRR*, abs/2501.01945.

Zeyu Zhang, Xiaohe Bo, Chen Ma, Rui Li, Xu Chen, Quanyu Dai, Jieming Zhu, Zhenhua Dong, and Ji-Rong Wen. 2024c. A survey on the memory mechanism of large language model based agents. *CoRR*, abs/2404.13501.

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, Ju-Ying Wu, Xiang Chen, Zichao Wang, Subrata Mitra, Nedim Lipka, and 2 others. 2024d. Personalization of large language models: A survey. *ArXiv*, abs/2411.00027.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Wanjun Zhong, Lianghong Guo, Qiqi Gao, He Ye, and Yanlin Wang. 2024. Memorybank: Enhancing large language models with long-term memory. In *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pages 19724–19731. AAAI Press.

Xinyu Zhu, Junjie Wang, Lin Zhang, Yuxiang Zhang, Yongfeng Huang, Ruyi Gan, Jiaxing Zhang, and Yujiu Yang. 2023. Solving math word problems via cooperative reasoning induced language models. In

Tasks	#Q	|Q|	#History	Output Format
CMV	133	1561.4	183.2	ranking
LaMP-1	123	29.0	317.5	2-way class
LaMP-2	3,302	48.6	55.6	15-way class
LaMP-3	112	183.9	959.8	[1, 2, 3, 4, 5]
LaMP-4	6,275	18.2	270.1	short generation
LaMP-5	107	161.9	442.9	short generation

Table A1: Basic statistics of evaluation sets. #Q indicates the number of queries. |Q| is token-based input query length, excluding template tokens. #History tells the number of historical engagements per user.

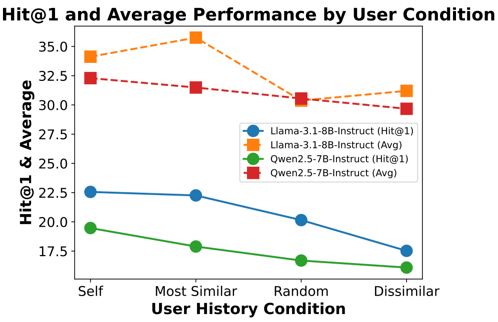

Figure A1: Hit@1 and average performance under four user-profile replacement conditions. Performance decline as the profile diverges, confirming PRIME 's faithfulness to user history.

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 4471–4485, Toronto, Canada. Association for Computational Linguistics.

A Implementation Details

A.1 Models Used for Experiments on CMV

For all base models, we use their instruction-fine-tuned versions for experiments. We have provided model cards in the footnotes.

- **Mini LLM**: LLAMA-3.2-3B (Meta AI Team, 2024)[7]
- **Small LLM**: LLAMA-3.1-8B (Dubey et al., 2024),[8] QWEN2.5-7B (Yang et al., 2024),[9] MINISTRAL-8B (Mistral AI team, 2024)[10].
- **Medium LLM**: QWEN2.5-14B (Yang et al., 2024),[11] PHI-4 (Abdin et al., 2024)[12]

[7] meta-llama/Llama-3.2-3B-Instruct
[8] meta-llama/Llama-3.1-8B-Instruct
[9] Qwen/Qwen2.5-7B-Instruct
[10] mistralai/Ministral-8B-Instruct-2410
[11] Qwen/Qwen2.5-7B-Instruct
[12] microsoft/phi-4

A.2 Evaluation Metrics

For our CMV benchmark, considering it is a ranking task (Manning et al., 2008), we adopt Hit@1, Hit@3, DCG@3 and MRR. For LaMP, we follow the official metrics (Salemi et al., 2024), and use accuracy and F1-score for classification tasks (LaMP-1 and LaMP-2), MAE and RMSE for the ordinal regression task (LaMP-3), and ROUGE-1 and ROUGE-L (Lin, 2004) for text generation tasks (LaMP-4 and LaMP-5). Note that all metrics are the higher the better, except for RMSE and MAE used for the LaMP-3.

B Additional Literature Review on Evaluation Challenge

Although benchmarks like PRISM (Kirk et al., 2024) and Empathetic Conversation (Omitaomu et al., 2022) offer a testbed for long-context query evaluation, their evaluation relies on generic metrics, e.g., ROUGE (Lin, 2004), or uses LLM-as-a-judge (Zheng et al., 2023). The former only measures surface-level similarity, while the latter demands extensive prompt engineering (Dong et al., 2024b; Szymanski et al., 2025) and still falls short of truly *imitating* individual users' preferences (Jiang et al., 2023), as the models do not consistently hold the target user's persona. To address these gaps, we introduce a new CMV-based benchmark focusing on long-form, persuasion-driven recommendation tasks, enabling direct and objective assessment of LLM personalization without relying on proxy judgments

C CMV Data Filtering by Difficulty

To ensure queries demand personalization rather than commonsense reasoning or popularity heuristics (Ji et al., 2020), we apply two instruction-tuned small yet powerful LLMs (LLAMA-3.1-8B (Dubey et al., 2024), QWEN2.5-7B (Yang et al., 2024),) on each query[13] without providing user history. We perform 10 runs per model, computing Hit@1 and Hit@3. We retain queries with Hit@1 ≤ 0.3 and Hit@3 ≤ 0.5, removing 15 authors and 194 queries to focus on challenging personalization items.

[13] We only consider 9 sampled negatives to form a query, the same setting as in section 4.

Model	Instantiation	Hit@1	Hit@3	DCG@3	MRR	Avg	W. Efficiency	R. Efficiency
				No Personalization				
Llama-3.1-8B	Baseline	16.32	36.77	28.10	25.11	26.58	N/A	N/A
Qwen2.5-7B		16.91	39.10	29.43	26.13	27.89		
				Episodic Memory (EM)				
Llama-3.1-8B	Recent	**16.62**	37.22	**28.36**	**25.33**	**26.88**	Fastest	Slow
Qwen2.5-7B		13.91	37.47	27.10	23.57	25.51		
Llama-3.1-8B	Relevant	16.17	35.41	27.00	24.12	25.68	Fastest	Slower
Qwen2.5-7B		13.23	38.50	27.36	23.56	25.66		
Llama-3.1-8B	Recent+PKR	**16.62**	36.84	28.10	25.10	26.67	Medium	Slower
Qwen2.5-7B		14.29	37.07	27.05	23.62	25.51		
Llama-3.1-8B	Relevant+PKR	15.64	36.32	27.45	24.41	25.96	Medium	Slowest
Qwen2.5-7B		13.76	**39.02**	27.88	24.07	26.18		
				Semantic Memory (SM)				
Llama-3.1-8B	NTP	17.44	41.20	30.93	27.31	29.22	Fast	Fast
Qwen2.5-7B		16.84	39.55	29.71	26.34	28.11		
Llama-3.1-8B	CIG	17.74	41.95	31.56	27.92	29.79	Fast	Fast
Qwen2.5-7B		16.77	40.23	30.05	26.57	28.41		
Llama-3.1-8B	Output FT	14.66	36.47	26.99	23.75	25.47	Medium-Fast	Fast
Qwen2.5-7B		16.54	39.85	29.58	26.08	28.01		
Llama-3.1-8B	Task FT	**19.62**	43.01	**32.96**	**29.36**	**31.24**	Medium	Fast
Qwen2.5-7B		16.99	**43.38**	32.15	28.28	30.20		
Llama-3.1-8B	DPO	15.41	37.37	27.89	24.64	26.33	Slowest	Fast
Qwen2.5-7B		16.77	39.55	29.61	26.22	28.04		
Llama-3.1-8B	SIMPO	14.21	34.81	25.89	22.88	24.45	Slow	Fast
Qwen2.5-7B		10.08	24.66	18.44	16.30	17.37		
Llama-3.1-8B	HSumm	16.32	37.89	28.62	25.44	27.07	Slowest	Medium
Qwen2.5-7B		15.04	38.80	28.50	24.97	26.83		
Llama-3.1-8B	PKR	16.69	36.39	28.12	25.26	26.62	Medium	Medium
Qwen2.5-7B		15.34	39.02	28.63	25.08	27.02		

Table A2: Complete results of the preliminary study where we study the strengths and limitations of various memory configurations. Results are the average of 10 runs. Recent/Relevant+PKR are effectively hybrid approaches using both episodic and textual semantic memories. W. Efficiency refers to memory writing or memory instantiation efficiency. For episodic memories, the writing time is the index creation time cost, which is extremely fast, compared to semantic memory writing. For semantic memory, we determine the efficiency label based on the train flops. For example, given a history of 15 engagements, the train flops of NSP/CIG is around 1e+16, while that of DPO is almost 1e+17. R. Efficiency measures the time overhead of both memory reading and the subsequent inference step. This overhead grows linearly with the number of retrieved past interactions—and increases further if a textual profile summary is prepended. In contrast,parametric semantic memories incur minimal inference cost, since all it needs to process is the input query without worrying about the past interaction retrieval.

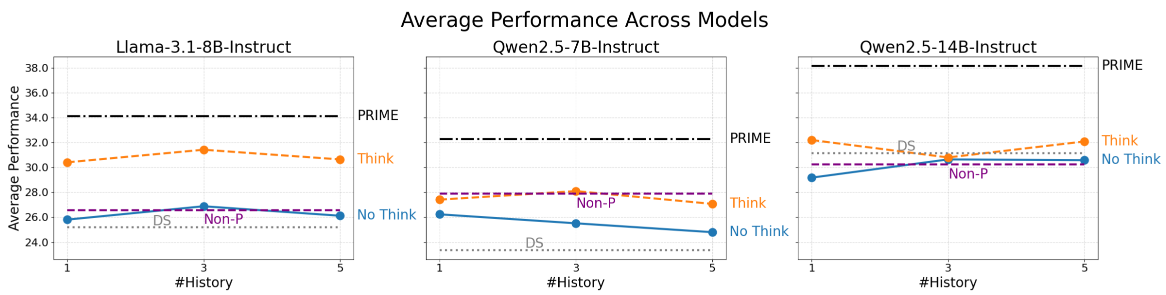

Figure A2: Average performance for Profile Replacement study.

Model	Hit@1	Hit@3	DCG@3	MRR	Avg
		No Personalization			
Llama-3.2-3B	14.51	38.65	28.09	24.49	26.44
Llama-3.1-8B	16.32	36.77	28.10	25.11	26.58
Ministral-8B	14.36	36.77	27.27	24.00	25.60
Qwen2.5-7B	16.91	39.10	29.43	26.13	27.89
Qwen2.5-14B	19.40	41.28	31.77	28.51	30.24
Phi-4	17.97	41.50	31.27	27.77	29.63
DeepSeek-Llama-3.1-8B	13.61	36.77	26.68	23.64	25.18
DeepSeek-Qwen2.5-7B	13.08	33.76	24.70	21.91	23.36
DeepSeek-Qwen2.5-14B	17.97	44.66	32.96	28.99	31.15
		EM			
Llama-3.2-3B	15.41	38.42	28.36	24.84	26.76
Llama-3.1-8B	19.10	44.14	33.11	29.35	31.43
Ministral-8B	16.09	39.92	29.63	26.10	27.94
Qwen2.5-7B	13.98	42.33	30.14	25.95	28.10
Qwen2.5-14B	16.92	44.96	32.76	28.58	30.81
Phi-4	18.57	44.89	33.63	29.76	31.71
		SM			
Llama-3.2-3B	17.22	**43.61**	**32.25**	**27.91**	**30.25**
Llama-3.1-8B	19.62	43.01	32.96	29.36	31.24
Ministral-8B	15.34	**40.83**	29.78	25.94	27.97
Qwen2.5-7B	16.99	43.38	32.15	28.28	30.20
Qwen2.5-14B	23.38	51.35	39.16	34.99	37.22
Phi-4	19.85	42.63	32.65	29.24	31.09
		DUAL			
Llama-3.2-3B	16.09	41.95	30.78	26.65	28.87
Llama-3.1-8B	20.15	44.59	34.06	30.16	32.24
Ministral-8B	16.24	**40.83**	30.08	26.40	28.39
Qwen2.5-7B	15.71	41.58	30.66	26.90	28.71
Qwen2.5-14B	23.76	52.03	39.59	35.34	37.68
Phi-4	21.58	43.98	34.13	30.76	32.61
		PRIME			
Llama-3.2-3B	**17.29**	42.93	31.81	27.78	29.95
Llama-3.1-8B	**22.56**	**45.79**	**35.87**	**32.28**	**34.13**
Ministral-8B	**17.14**	40.75	**30.81**	**27.26**	**28.99**
Qwen2.5-7B	**19.47**	**45.19**	**34.16**	**30.35**	**32.29**
Qwen2.5-14B	**24.29**	**52.03**	**40.17**	**36.09**	**38.15**
Phi-4	**23.01**	**47.29**	**36.93**	**33.37**	**35.15**

Table A3: Full results on CMV. Best results for *each* base model are **bold**.

	Input	Output
NTP	"author": {author}. "title": {title}. "body": {body}<EOS>	author": {author}. "title": {title}. "body": {body}
CIG	For the topic "{title}", the author "{author}" states:	{body}
Output-FT (O-FT)	The author, "{author}", has engaged with users on the Change-My-View subreddit across various original posts (OPs). Based on the author"s preference and engagement patterns, generate a persuasive response that is highly likely to change their viewpoint on the following post. "title": {title}. "body": {body}	"reply": {positive_reply}
Task-FT (T-FT)	The author, "{author}", has engaged with users on the Change-My-View subreddit across various original posts (OPs) and is seeking alternative opinions to alter their viewpoint. Currently, the author is creating a new OP titled "{title}", with the content: ***{body}.*** From the "candidate replies" JSON file below, select the best reply (using "option ID") that best challenges the author"s view. {candidates}	["{option ID}"]
Preference Tuning	"author": {author}. "title": {title} "body": {body}	"reply": {positive_reply} / "reply": {negative_reply}

Table A4: Input-output mapping for each parametric semantic memory instantiation. In the context of a single-turn CMV conversation, there are always the following fields: *title*, *body*, *author* and *reply*. If a reply receives a Δ, then it is a positive reply; otherwise, it is a negative reply. Such pair can directly support the preference tuning. However, in reality, it is not always the case we have the access to aforementioned items. If the reply is unavailable, one may resort to **NTP** or **CIG**. For **NTP**, the output is essentially the left-shifted input. If output is available, one may utilize fine-tuning paradigm such as **O-FT** or **T-FT**, where the latter will be preferred if we are able to know the task information. In this study, we can convert raw replies into the desired task format following the prompt shown in Figure A8.

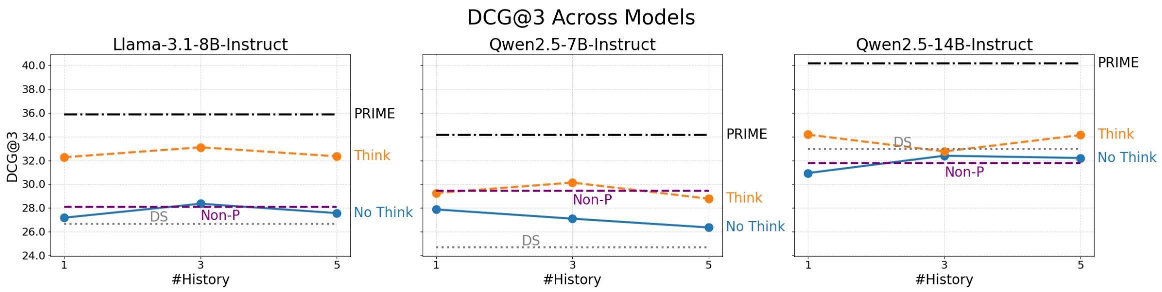

Figure A3: DCG@3 metric for Profile Replacement study.

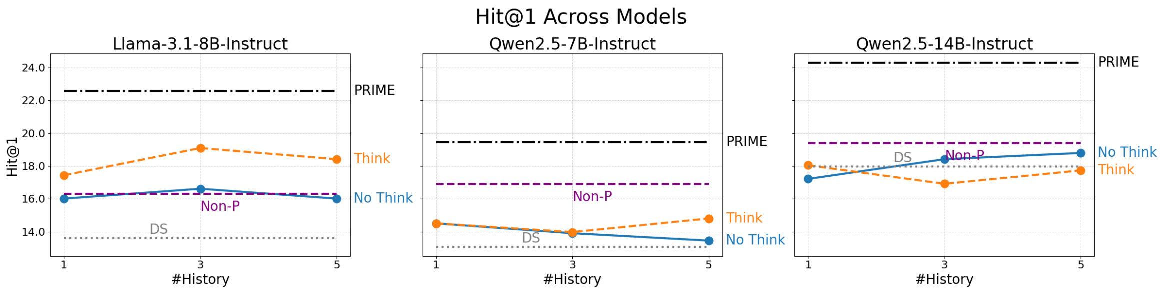

Figure A4: Hit@1 metric for Profile Replacement study.

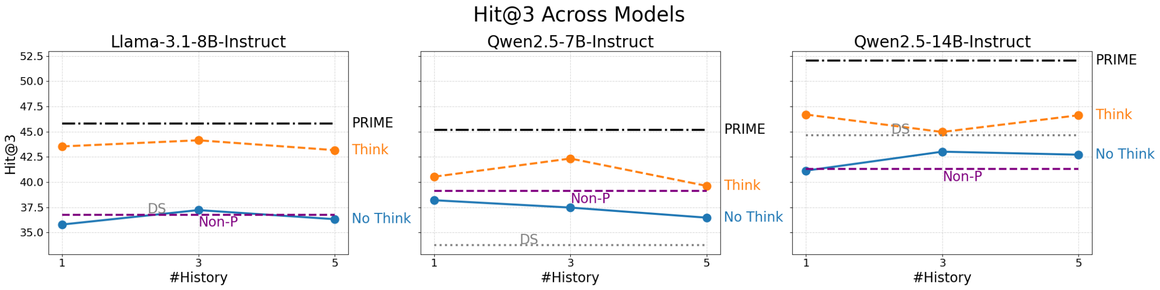

Figure A5: Hit@3 metric for Profile Replacement study.

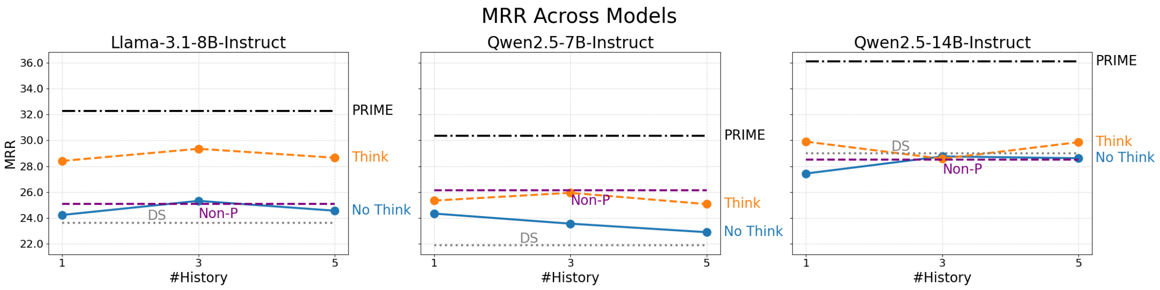

Figure A6: MRR metric for Profile Replacement study.

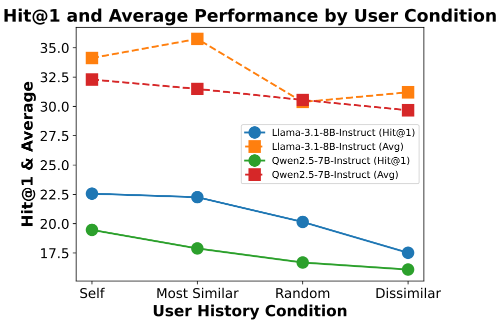

Figure A7: Hit@1, Hit@3, DCG@3 and MRR performances under four user-profile replacement conditions–self, most similar, random and dissimilar. In general, all metrics decide as the profile diverges, confirming PRIME 's sensitivity to user history, showing that PRIME indeed captures the dynamic personalization rather than just bandwagon biases.

Evaluation Query

The author, `{AUTHOR}`, has engaged with users on the Change-My-View subreddit across various original posts (OPs) and is seeking alternative opinions to alter their viewpoint.
Currently, the author is creating a new OP titled

{OP TITLE}

with the following content:

{OP CONTENT}

From the `candidate replies` JSON file below, select the top 3 replies (using `option ID`) that best challenge the author's view. Rank them from most to least compelling.

```
[
  { 'option ID': '...','challenger':'...','reply':'...'},
  { 'option ID': '...','challenger':'...','reply':'...'},
]
```

Output format: Output a valid JSON array of "`option ID`" strings representing the selected replies. Each element must be a double-quoted string. The response should contain nothing but the JSON array and end with "#END".

Figure A8: Standard prompt for CMV evaluation query.

Evaluation Query

The author, `kingpatzer`, has engaged with users on the Change-My-View subreddit across various original posts (OPs) and is seeking alternative opinions to alter their viewpoint.
Currently, the author is creating a new OP titled

> "CMV: Those who attribute gun ownership rates as the cause of the problem of gun violence in terms of criminal gun deaths are not merely mistaken; they are disingenuous"

with the following content:

> The data has been clear for a very long time: the relationship between guns and gun homicides doesn't show any strong correlation.
>
> I have personally taken the cause-of-death data from https://wonder.cdc.gov/, grouping results by year and state, and selecting *Homicide, Firearm* as the cause of death. I then matched that data to the per-capita gun-ownership statistics by state from the ATF, as reported by Hunting Mark (https://huntingmark.com/gun-ownership-stats/).
>
> A standard correlation analysis between firearm homicide rates per 100,000 and per-capita gun ownership yields an r^2 of 0.079 (no meaningful correlation). A similar global analysis by nation gives an r^2 of 0.02...
>
> The only way to associate gun ownership with gun violence is by including suicides by firearm, which I argue is disingenuous. We don't count suicide by hanging as "rope violence" when discussing strangulation, nor overdoses as "drug violence," etc.

From the `candidate replies` JSON file below, select the top 3 replies (using `option ID`) that best challenge the author's view. Rank them from most to least compelling.

```
[
  { 'option ID': '...','challenger':'...','reply':'...'},
  { 'option ID': '...','challenger':'...','reply':'...'},
]
```

Output format: Output a valid JSON array of "`option ID`" strings representing the selected replies. Each element must be a double-quoted string. The response should contain nothing but the JSON array and end with "#END".

Figure A9: A formatted evaluation query from our CMV dataset. We have also provided a candidate reply in Figure A10

Positive Candidate Reply (option F)

- **Option ID**: `option F`
- **Challenger**: `An-Okay-Alternative`
- **Reply**:

  Your data for gun-ownership is weapons per capita, not individual gun owners per capita. If we're testing access to guns vs. gun homicides, someone who owns 30 guns wouldn't be 30× as likely to commit a homicide.

  Secondly, the ATF list includes only specially regulated firearms (e.g. fully automatic weapons, short-barreled shotguns) and NFA items bought by law enforcement.

  Thus, the data omits common firearms, is biased by collectors and hobbyists, and is confounded by law-enforcement purchases.

  Overall, this calculation is weak evidence against a gun–homicide correlation and doesn't imply critics are dishonest.

Figure A10: Ground-truth reply to the evaluation query shown in Figure A9, which receives Δ awarded by the OP author.

Personalized Thinking

User Experience: The author, `kingpatzer`, has engaged with users on the Change-My-View subreddit across various original posts (OPs), seeking alternative opinions to alter their viewpoint. In past discussions, they've shown openness to different perspectives and a willingness to consider data and evidence about gun ownership and gun violence. However, they can become defensive when challenged and sometimes dismiss opposing viewpoints.

OP Summary: The author is creating a new OP titled

> CMV: Those who attribute gun ownership rates as the cause of the problem of gun violence in terms of criminal gun deaths are not merely mistaken; they are disingenuous

They argue that available data show no strong correlation between gun ownership rates and gun homicides, and that including firearm suicides is disingenuous—comparing it to "rope violence" or "drug violence" when discussing strangulation or overdoses.

Sketch Outline:
To challenge the author's view, we need to address the following points:
1. The author's interpretation of the data: We need to examine the author's methodology and data analysis to see if it is sound and if it supports their conclusion. 2. The inclusion of suicide by guns: We need to discuss the relevance of including suicide by guns in the data set and why it is not disingenuous to do so. 3. The comparison to other causes of death: We need to compare the author's approach to other causes of death, such as drug overdoses, to see if it is consistent and fair. 4. The author's tone and language: We need to address the author's tone and language, which some may perceive as dismissive and condescending.
Key questions to ask:
* What is the author's definition of "disingenuous" and how does it apply to including suicide by guns in the data set? * Is the author's methodology sound and does it support their conclusion? * How does the author's approach to gun ownership and gun violence compare to other causes of death? * Why is the author using a tone and language that some may perceive as dismissive and condescending?

Analysis (option F):
This reply challenges the author's view by pointing out that the data used to support their conclusion is flawed, and that the author's methodology is not sound. The reply also highlights the importance of considering the social and cultural context in which gun ownership and gun violence occur.

Reflection:
Based on the analysis, the top 3 replies that best challenge the author's view are:
1. Option F: This reply challenges the author's view by pointing out that the data used to support their conclusion is flawed, and that the author's methodology is not sound. The reply also highlights the importance of considering the social and cultural context in which gun ownership and gun violence occur. 2. Option D: This reply challenges the author's view by pointing out that there is a strong link between gun ownership and impulsive suicide, and that including suicide by guns in the data set is not disingenuous. The reply also highlights the importance of considering the cultural and social context in which gun ownership and gun violence occur. 3. Option C: This reply provides a detailed and well-researched argument that challenges the author's view, pointing out that there is a significant correlation between gun ownership and gun violence. The reply also highlights the importance of considering the social and cultural context in which gun ownership and gun violence occur.
These replies are the most compelling because they engage with the author's perspective, provide reasoned and respectful arguments, and challenge the author's view in a thoughtful and nuanced way.

Answer:

$$\left[\text{"option F"},\ \text{"option D"},\ \text{"option C"}\right]$$

Figure A11: Personalized thinking trace generated by our PRIME to the evaluation query shown in Figure A9. Due to space limit, we only show the analysis for option F, but we keep all the other components intact.

D Case study of Reasoning Trace

This section provides a brief overview on the reasoning traces by PRIME (Figure A11), i.e., personalized thinking, and by R1-DISTILL-LLAMA (Figure A12), which is a generic reasoner.

Specifically, our personalized thinking explicitly emphasizes on the *author's historical sensitivity to methodological rigor*, grounding its selection of option F ("data used to support their conclusion is flaw") in a nuanced understanding of the author's past engagements. This showcases how our thinking traces implements personalization. In contrast, the generic reasoner (R1-DISTILL-LLAMA) superficially mentions methodological issues in option F but ultimately prioritizes other responses (A, C, and D) based on *broader counter-evidence and direct contradictions*. Thus, R1-DISTILL-LLAMA's reasoning is weakened by a lack of personalization, failing to adequately weigh methodological precision aligned with the author's long-term behavior and patterns.

Generic Reasoning by R1-DISTILL-LLAMA

Think:
Alright, I need to figure out which of the provided replies best challenge the author's viewpoint. The author argues that there's no strong correlation between gun ownership rates and gun homicides, and that including suicides in the data is disingenuous.
Looking at the replies, I should focus on those that either provide counterarguments, point out flaws in the author's methodology, or present evidence contradicting their claims.
Option A: This reply points out that ignoring suicides is misleading because many suicides could be prevented by restricting gun access. It also mentions that while suicides by other means are controlled, there's no such regulation for guns. This directly challenges the author's exclusion of suicides, making it a strong point.
Option B: This uses an analogy about other causes of death to question the author's approach. It's a clever way to highlight that the author's reasoning might be inconsistent, suggesting the argument is flawed. This is compelling because it questions the validity of the author's methodology.
Option C: This reply cites a study that found a positive correlation between gun ownership and firearm homicides. It's strong evidence against the author's claim, making it a solid challenge.
Option D: This argues that there's a correlation between gun ownership and gun deaths, including suicides, and criticizes the author's exclusion of suicides. It's a direct counterargument, supporting the idea that the author is being disingenuous.
Option E: This reply references a global study showing a link between gun ownership and gun violence, which directly contradicts the author's data. It's a good point because it uses external evidence to challenge the author.
Option F: This critique points out methodological issues with the author's data, such as not measuring individual gun owners and including restricted firearms. It's detailed and highlights the limitations of the author's approach, making it a strong challenge.
Option G: This reply calls the author's statement absurd, stating that gun homicides are clearly correlated with gun presence. It's a strong, straightforward counterargument.
Option H: This argues that while the author excludes suicides, measures are taken for other causes, but not for gun suicides. It's relevant because it highlights the inconsistency in excluding suicides, supporting the idea of disingenuousness.
Option I: This defends the author's character, saying disagreement doesn't make someone disingenuous. It's more of a defense rather than a direct challenge, so it's less compelling.
Option J: This points out that the global study the author cited was flawed and incorrect. It's a good challenge because it questions the validity of the author's global data.
After evaluating, the top three replies that most effectively challenge the author's viewpoint are A, C, and D. They each present strong arguments, either by providing counter-evidence, highlighting methodological flaws, or pointing out inconsistencies in the author's reasoning.

Answer:

$$\left[\text{"option A"},\ \text{"option C"},\ \text{"option D"}\right]$$

Figure A12: Generic reasoning trace generated by R1-DISTILL-LLAMA to the evaluation query shown in Figure A9.

Personalized Thinking Prompt

You are an AI assistant designed with **bionics-inspired episodic memory** capabilities. Your objective is to approach the task by emulating human-like episodic recall, drawing from past interactions, contextual understanding, and nuanced reasoning to deliver informed and thoughtful judgments.
Treat the author's past conversations as episodic memories that guide your reasoning and decision-making throughout the process. Prioritize capturing the user's values and patterns from past interactions and integrating these into your reasoning.

Instructions

1. **Ingest Author History:**
   - Leveraging your trained semantic memory, extract and synthesize insights from the author's prior interactions. Summarize the author's past patterns, preferences, values, and beliefs to establish an episodic memory.
   - Write a concise summary of history conversations within `<user experience>` tags. This summary should serve as your episodic memory for later steps.
2. **Summarize the New OP:**
   - Review the content of the new OP carefully, identifying salient events, major arguments, core themes, and the author's explicit and implicit viewpoints.
   - Write a concise summary of the new OP within `<OP summary>` tags.
3. **Sketch an Outline:**
   - Combine insights from your episodic memory (`<user experience>`) with the context from the new OP (`<OP summary>`). Conduct reasoning that incorporates the author's past preferences and patterns to create a strategic outline for how to challenge or respond to the author's viewpoint.
   - Highlight the most important points or questions for challenging the author's view and encapsulate these in a concise outline within `<sketch outline>` tags.
4. **Evaluate Candidate Replies:**
   - Analyze each candidate reply from the provided JSON file in terms of strength, relevance, and weaknesses. Base your evaluations on your episodic memory, OP reasoning, and the outlined strategy.
   - Present this evaluation as a dictionary within `<analysis>` tags, e.g., `{'option A': [analysis], 'option B': [analysis], . . . }`.
5. **Reflect and Rank Top Replies:**
   - Reflect on and integrate all insights to determine the most compelling replies—those engaging the author's view and providing reasoned, respectful, and novel insights.
   - Identify the top three replies by option ID, ranking them from highest to lowest compellingness. Include your concise reflection within `<reflection>` tags.
6. **Answer and Conclude:**
   - Output your selection as a valid JSON array of strings within `<answer>` tags, e.g., `["option ID", "option ID", "option ID"]`.
   - End your response immediately with #END.

Output Format:

```
<user experience>[concise user experience summary]</user experience>
<OP summary>[concise OP summary]</OP summary>
<sketch outline>[concise sketched outline]</sketch outline>
<analysis>{'option A': [analysis], 'option B': [analysis], ...}</analysis>
<reflection>[concise reflection]</reflection>
<answer>["option ID","option ID","option ID"]</answer>
#END
```

Figure A13: Our personalized thinking prompt, designed for CMV evaluation query.