TRIPLE: Theory-Driven Integration of Planned and Habitual Behaviors for LLM-based Personalization

Taehyung Noh[1], **Seungwan Jin**[2], **Haein Yeo**[1], **Kyungsik Han**[1, 2*]
[1]Department of Artificial Intelligence, Hanyang University, South Korea
[2]Department of Data Science, Hanyang University, South Korea
{yestaehyung, seungwanjin, haeinyeo, kyungsikhan}@hanyang.ac.kr

Abstract

While large language model (LLM)-based user profiling offers significant potential for personalization, most existing approaches rely on empirical heuristics and lack grounding in the psychological mechanism that drive human behavior. In this paper, we introduce **TRIPLE (Theory-guided Reasoning for Intent and habIt Profiling with LLMs for pErsonalization)**, a novel framework that systematically integrates *dual-process theory* from social psychology into LLM-based user modeling. TRIPLE (1) constructs *a habitual behavior profile* by identifying repeated patterns over time to model automatic responses; (2) builds *an intentional behavior profile* by inferring user attitudes, subjective norms and perceived behavioral control based on the Theory of Planned Behavior (TPB); and (3) generates behavioral rationale that reveal the interaction between habitual and intentional processes to predict user behavior in context-specific situations. We evaluate TRIPLE on five personalization tasks from the LaMP benchmark using multiple open-source LLMs. Results show that TRIPLE consistently outperforms existing in-context learning methods, with especially pronounced gains on complex generative tasks such as headline and title generation.Qualitative analyses further demonstrate that the profiles and reasoning paths generated by TRIPLE provide interpretable and psychologically grounded explanations of user behavior. These findings provide strong evidence that incorporating validated behavioral theories into LLM-based personalization enhances both predictive performance and interpretability, paving a way for theory-driven, socio-cognitively informed user modeling.

Project Page — https://aaai.org/example/code

Introduction

Analyzing user behavioral patterns to derive underlying characteristics is a fundamental component of personalized service design (Koufaris 2002; Tam and Ho 2006; Zhao et al. 2015). Since human behavior typically arises from intentional processes, understanding not only surface-level behaviors but also their intentions is important for achieving a more comprehensive and precise representation of the user (Ajzen 1991). These behavioral intentions are shaped

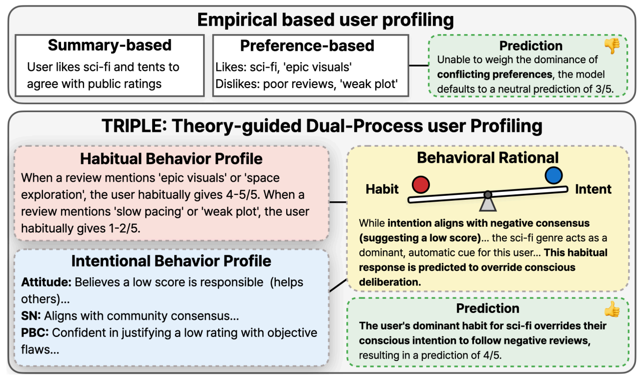

Figure 1: Motivation for TRIPLE. While recent LLM-based profiling methods have advanced transparency compared to traditional black-box models, they often generate unstructured summaries of past behavior without theoretical grounding. This limits their ability to explain the *why* behind user actions. In contrast, TRIPLE integrates Dual-Process Theory, explicitly modeling behavior through a Habitual Profile (automatic actions) and an Intentional Profile (deliberate actions). The Behavioral Rationale then synthesizes these components to generate predictions with interpretable, psychologically-grounded explanations for user behavior.

by individual attitudes, social norms, and perceived behavioral control, while exhibiting latent properties that fluctuate with contextual factors such as time, location, and mood (Ajzen 1985). Thus, recognizing the complex interaction between psychological, social, and contextual dimensions can lead to more accurate behavior prediction and ultimately improve user experiences in personalized systems.

Despite the importance of latent characteristics, effectively capturing them in practical systems remains a significant challenge (De Bra 2017; Tam and Ho 2005). Prior personalization approaches have relied mainly on embedding user behavior data (e.g. reviews, ratings) to infer latent variables such as preferences, emotional states (Zhang, Chen et al. 2020). However, such embedding-based models often suffer from low interpretability, making it difficult to evaluate whether users' cognitive judgment processes are properly reflected (Zhou et al. 2025). This lack of transparency can also hinder the integration of domain knowledge and sociocognitive factors, limiting the flexibility, accountability,

*Corresponding author.

and trustworthiness of the model (McCarthy et al. 2025).

The emergence of large language models (LLMs) is noteworthy as LLMs offer new opportunities to overcome these limitations by enabling more interpretable user modeling. LLMs pre-trained on vast text corpora, encode a broad range of knowledge about human psychology and social norms (Demszky et al. 2023; Kosinski 2024; Park et al. 2023). More importantly, they support natural language-based reasoning, allowing them to generate explicit, interpretable explanations for user states, an advantage over traditional black-box embeddings. Recent studies have explored the use of LLMs for in-context user profiling, where *user profiles* are constructed and used to predict various behaviors, such as product preferences or writing with a certain style. While the specific method of profile construction varies—for example, using raw interaction histories (Salemi et al. 2023), natural language summaries (Liu et al. 2024), or structured preference representations (Bang and Song 2025)—these approaches have shown promise in improving predictive performance and interpretability.

However, these research efforts have largely overlooked deeper socio-technical perspectives. To fully leverage the knowledge embedded in LLMs for understanding users, it is essential to *ground this understanding in validated theoretical frameworks from the social and behavioral sciences*. These frameworks offer structured explanations of human behavior and the interrelations among key constructs, serving as reasoning scaffolds that guide the selection and organization of relevant elements from LLM's knowledge. By incorporating such theory-driven structures into LLM-based modeling, we can achieve more precise, transparent, and psychologically grounded representations of users, which reflects not only their observed behaviors but also the motivations, norms, and cognitive processes that drive them.

In this context, the Dual-Process Theory from social psychology offers an important conceptual foundation (Chaiken and Trope 1999; Evans 2008). This theory distinguishes between two distinct type of human behavior—*habitual* and *intentional*—each governed by different cognitive mechanisms (Strack and Deutsch 2004). Habitual behavior arises from repeated actions in stable contexts and lead to automatic responses that can predict future behavior independently of conscious intention (Wood and Neal 2007). In contrast, intentional behavior is based on behavioral intentions, shaped by attitudes, subjective norms, and perceived behavioral control, providing explanatory power for short-term, goal-directed actions (Ajzen 1991). These two behavioral processes function complementarily, with their influence varying according to contextual factors (Ouellette and Wood 1998). Therefore, modeling both processes in LLM input contexts allows for a more nuanced and accurate understanding of user behavior.

In this paper, we propose **TRIPLE (Theory-guided Reasoning for Intent and habit Profiling with LLMs for pErsonalization)**, a novel LLM-based profiling framework that systematically integrates dual-process theory into user modeling. TRIPLE generates two distinct but complementary profiles. First, the intentional behavior profile models goal-directed actions by inferring behavioral intentions and their components—attitudes, subjective norms, and perceived behavioral control—based on the Theory of Planned Behavior (TPB) framework. Second, the habitual behavior profile captures automatic, routine behaviors by identifying repeated behavioral patterns in user history. This profile models unconscious tendencies that TPB fails to capture. Additionally, TRIPLE generates a behavioral rationale, which explicitly articulates, in natural language, how and why each profile contributes to the prediction for a given situation. This rationale captures the interaction between the two profiles, providing interpretable insights into how different behavioral processes apply in different contexts.

To evaluate the effectiveness of TRIPLE, we conducted experiments on five personalized tasks (i.e., news categorization, movie tagging, product rating, news headline generation, and scholarly title generation) using the LaMP dataset (Salemi et al. 2023), a widely used benchmark in user modeling research. Results across multiple open-source LLMs showed that TRIPLE outperformed existing in-context learning profiling methods, with particularly notable improvements on the generative tasks. This highlights the value of incorporating social science theories into LLM-based user modeling. Furthermore, a qualitative evaluation confirmed that TRIPLE-generated profiles offered clearer and more comprehensible explanations of user behavior compared to existing methods.

This study contributes not only to improved prediction performance but also to theoretical transparency in user modeling. By making explicit both the outcomes and reasoning processes behind user behaviors, TRIPLE enhances interpretability and user understanding—offering a practical and extensible framework for theory-grounded personalization. As such, TRIPLE serves as a leading example of socio-technical AI modeling that promotes conceptual and algorithmic extensions in the user modeling literature and offers a generalizable methodology for personalization across diverse domains.

Background

Theories of Behavioral Intention and Habit

Traditional research in social psychology has long interpreted human behavior as the result of conscious and deliberative reasoning. The Theory of Reasoned Action (Fishbein and Ajzen 1975) and the Theory of Planned Behavior (Ajzen 1991) propose that behavioral intention is formed through the interplay of attitudes, subjective norms, and perceived behavioral control, and intention is assumed to predict behavior. These theories have been validated across a wide range of behavioral domains (McEachan et al. 2011).

However, these intention-based models are limited in their ability to explain automatic and habitual behaviors, resulting in what is commonly referred to as the intention-behavior gap (Sheeran 2002). Ouellette et al. defined habit as an automatic behavioral response learned through repeated goal-directed actions (Ouellette and Wood 1998), often triggered by environmental cues and carried out with minimal conscious deliberation. Habitual behaviors, therefore, arise through psychological mechanisms that are fundamentally

Table 1: Comparison of TRIPLE with representative empirical methods. While existing approaches primarily rely on empirical heuristics to capture behavioral patterns, TRIPLE is the first framework to incorprate psychological theories to model the underlying cognitive mechanisms that drive user behavior.

Method	Profile Content and Generation Method	Profile Representation	Theoretical Foundation
PAG (Richardson et al. 2023)	Summarizes entire user history offline	Descriptive summary of overall preferences	Empirical
ONCE (Liu et al. 2024)	Infers high-level interests and attributes from user history	List of inferred interests and attributes	Empirical
GPG (Zhang 2024)	Describes specific behavioral styles	Descriptive text of specific user styles	Empirical
PURE (Bang and Song 2025)	Extracts explicit preferences directly	Structured lists of likes/dislikes/features	Empirical
TRIPLE (Ours)	**Analyzes psychological drivers and mechanisms**	**Habits and Intentions**	**Dual processing theory, Habit Theory, Theory of Planned Behavior**

distinct from those governing intentional behaviors.

To unify these perspectives, *dual-process theory* integrates intention-centered and habit-based models, proposing that intentional and habitual processes function complementarily in shaping human behavior (Ouellette and Wood 1998). This theory has been validated across multiple domains (Evans and Stanovich 2013), improving prediction accuracy in consumer (Samson and Voyer 2012) and health behavior (Brown et al. 2021) by differentiating deliberative from habitual processes. Recently applied to human-AI interaction, dual-process modeling has enhanced LLM agents' performance in collaborative tasks (Zhang 2025).

Building on this foundation, our study proposes a novel framework that integrates dual-process theory with LLM in-context learning to support deeper and more interpretable user modeling. We model intentional and habitual behaviors as distinct profiles and analyze their context-dependent contributions through explicit behavioral rationales. By combining the reasoning capabilities of LLMs with established social-psychological theories, we aim to advance interpretable and personalized behavior prediction in a theoretically grounded and computationally practical manner.

LLM-based In-Context Personalization

The emergence of LLMs has opened new opportunities for user profiling. Early studies use the direct inclusion of users' past behaviors in prompts. For example, PAG (Richardson et al. 2023) extracts keywords from user behaviors to generate concise profiles. However, this approach is limited to simple keyword enumeration, failing to capture relationships between behaviors or contextual meanings. ONCE (Liu et al. 2024) expresses user interests through natural language summarization but remains at static profiles that cannot reflect temporal changes or situational differences. PURE (Bang and Song 2025) model user preferences through structured representations of likes and dislikes. While it offers limited update mechanisms over time, it remains focused on surface-level preferences and lacks the capacity to explain underlying psychological drivers or habitual behavior patterns. Table 1 summarizes the key conceptual differences between these approaches and the proposed TRIPLE framework. Prior studies on LLM-based profiling, while advancing performance and interpretability, remain largely empirical, focusing on behavioral pattern extraction rather than established theories of human behavior.

Our study integrates the Theory of Planned Behavior and habit theory validated in social psychology into LLMs in-context learning. This enables the model to not merely describe user behavior but also explain why it occurs through underlying psychological processes (e.g., intentions, attitudes, social norms, and habitual cues). To our knowledge, this is the first study to systematically integrate dual-process behavioral theories into an LLM-based personalization framework.

Method

We propose a novel user profiling method, TRIPLE, based on Dual-Process Theory (DPT) (Chaiken and Trope 1999; Evans 2008) to enhance the prediction and interpretation of user behavior. DPT distinguishes between two fundamental cognitive systems: habitual processes, which are automatic and shaped by repeated experiences, and intentional processes, which are driven by conscious deliberation. Building on this framework, TRIPLE integrates DPT with LLM personalization by modeling user behavior through two complementary profiles: a habitual profile $P_u^H(t)$ and an intentional profile $P_u^I(t)$. For each user u, these profiles are constructed using the most recent k query-behavior pairs $h_t = \{x_t, y_t\}$ drawn from the user histories $H_u = \{h_0, h_1, ..., h_t\}$ within a sliding window up to time t. Additionally, TRIPLE generates a behavioral rationale $R_u^{\text{history}}(h_t)$ that combines both profiles to identify which process is more dominant in a given context. This approach enhances behavioral prediction accuracy and provides interpretable insights into the underlying cognitive mechanisms of personalized user behavior.

User Profiling with Dual-Process Theory

Habitual Profile According to habit theory (Wood and Neal 2007), habits are automatic behavioral patterns triggered by specific situational cues and shaped through long-term experience. We therefore define a user's habitual behavior profile $P_u^H(t)$ as the set of automatic response patterns that surface consistently in behavioral data.

To capture users' stable, largely unconscious behavioral foundations, we generate $P_u^H(t)$ with a sliding window $W_t = \{h_{t-k+1}, ..., h_t\}$ of size k that contains the most recent k histories. The window excludes future information yet updates over time, ensuring a balance between two properties: (1) habits evolve over time, so outdated patterns

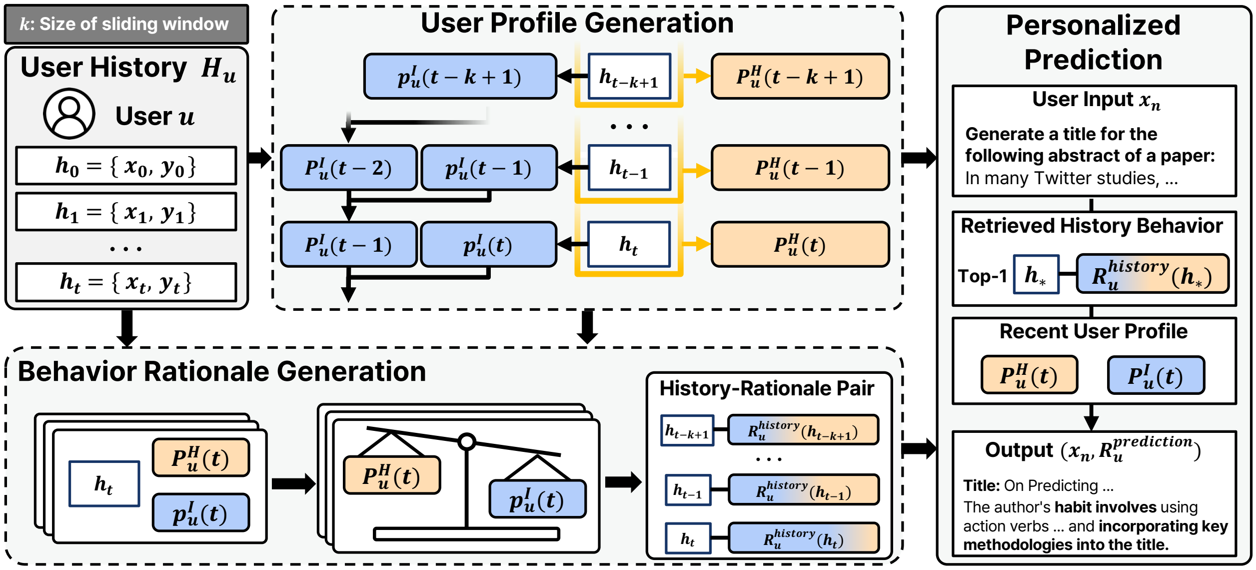

Figure 2: Overview of TRIPLE framework, showing how user history is processed into habitual and intentional profiles, which are then synthesized through behavioral rationale generation to produce personalized predictions.

should fade, and (2) recently reinforced habits best predict future actions. Thus, the window preserves only the most relevant and currently active habitual patterns.

Analyzing the window W_t at once lets the LLM detect recurring cure-response links more directly than step-wise incremental updates. We generate the habitual profile $P_u^H(t)$ using function $\mathcal{L}_{\text{habit}}$, as follows:

$$P_u^H(t) = \mathcal{L}_{\text{habit}}(\mathcal{P}_{\text{habit}}, W_t) \tag{1}$$

where $\mathcal{L}_{\text{habit}}$ denotes the LLM model, and the prompt $\mathcal{P}_{\text{habit}}$ is a prompt that explicitly instructs the LLM to extract situational cues and their repetitive behavioral patterns in a structured format. The resulting habitual profile distills the user's routine automatic behaviors.

Intentional Profile The Theory of Planned Behavior (TPB) (Ajzen 1991) suggests that intentional behavior emerges from deliberative processes involving attitudes, subjective norms, and perceived behavioral control. Based on TPB, we model the intentional profile $P_u^I(t)$ as a dynamic belief system that reflects how users' behavioral intentions change over time, by analyzing situational factors and behaviors within a sliding window of size k ending at time t. This profile comprises the following components:

- **Attitude**: An individual's positive or negative evaluation of specific behaviors.
- **Subjective Norm**: Perception of social pressure from others and expectations regarding behavior performance.
- **Perceived Behavioral Control**: Beliefs about how easy or difficult it is perceived to perform specific behaviors.

TRIPLE manages $P_u^I(t)$ using an incremental update mechanism, a design choice that directly reflects the conscious deliberation described in the TPB. Unlike habits, which are derived from repeated behavioral patterns, intentions emerge in response to specific situations and are then integrated into the user's evolving belief system. For each behavioral history $h_t = \{x_t, y_t\}$, TRIPLE generates a short-term intentional profile $p_u^I(t)$ by inferring how the situation x_t triggered the behavior y_t, guided by the TPB's three components (i.e., attitude, subjective norm, and perceived behavioral control). $p_u^I(t)$ is computed as follows:

$$p_u^I(t) = \mathcal{L}_{\text{intent}}(\mathcal{P}_{\text{TPB}}, h_t) \tag{2}$$

where $\mathcal{L}_{\text{intent}}$ denotes the LLM model, and $\mathcal{P}_{\text{TPB}}$ represents a prompt that instructs the LLM to construct a profile consisting of the three TPB components based on the history h_t. $p_u^I(t)$ captures the user's intention at time t and serves as the basic unit for updating the long-term intentional profile. $p_u^I(t)$ is then integratd with the prior intentional profile $P_u^I(t-1)$ to form the new intentional profile $P_u^I(t)$:

$$P_u^I(t) = \mathcal{L}_{\text{update}}(\mathcal{P}_{\text{update}}, \{P_u^I(t-1), p_u^I(t)\}) \tag{3}$$

where $\mathcal{L}_{\text{update}}$ denotes the LLM model, and $\mathcal{P}_{\text{update}}$ compares $p_u^I(t)$ with the previous profile $P_u^I(t-1)$, reinforcing consistent beliefs, refining conflicting ones, or mirroring how conscious thought gradually stabilizes while filtering out fleeting intentions. This incremental process simulates how deliberate beliefs evolve over time.

In summary, the habitual profile $P_u^H(t)$ captures basic behavioral patterns in a long-term, stable, and automatic way, while the intentional profile $P_u^I(t)$ reflects situational adjustments in a short-term, changing, and conscious way.

Behavioral Rationale Generation To identify which system dominated a specific action, we perform a post-hoc analysis of h_t using $P_u^H(t)$ and $P_u^I(t)$. This yields a personalized rationale $R_u^{\text{history}}(h_t)$, as follows:

$$R_u^{\text{history}}(h_t) = \mathcal{L}_{\text{rationale}}(\mathcal{P}_{\text{rationale}}, \{h_t, P_u^H(t), P_u^I(t)\}) \tag{4}$$

where $\mathcal{L}_{\text{rationale}}$ denotes the LLM model, and $\mathcal{P}_{\text{rationale}}$ guides the LLM through four steps:

1. **Habit-based initial prediction:** Identify habitual cues based on the given query x_t and habit profile $P_u^H(t)$, and predict habitual behavior.
2. **Intention-based review:** Compare initial predictions with $P_u^I(t)$ to analyze points of agreement or conflict between the profiles.
3. **Self-verification:** Classify the consistency between the analysis results of the two profiles as match, partial match, partial mismatch, or complete mismatch, and explain the differences.
4. **Integrated reasoning:** Based on the analyzed consistency, ultimately infer why the actual behavior y_t occurred. In cases of mismatch, specifically describe which intentional elements suppressed habitual responses.

$R_u^{\text{history}}(h_t)$ clarifies individual DPT dynamics, enhancing interpretability by showing when users rely on habitual shortcuts versus conscious deliberation.

Personalized Prediction

To predict future user behavior, we integrate three components: habitual profiles, intentional profiles, and the most relevant history-rationale pair. Given a new query x_n, TRIPLE identifies the most relevant history h_* from W_t by using BM25 (Robertson, Zaragoza et al. 2009), and retrieves the mapped $R_u^{\text{history}}(h_*)$. This in-context example shows how the user previously balanced habits and intentions, guiding the LLM to mirror that decision style. TRIPLE then integrates x_n, both profiles, and the retrieved example, predicting the user behavior $\hat{y}_n$ and rationale $R_u^{\text{predict}}(x_n)$, as follows:

$$(\hat{y}_{\text{n}}, R_u^{\text{predict}}(x_{\text{n}})) = \mathcal{L}_{\text{predict}}(\mathcal{P}_{\text{predict}}, \{x_{\text{n}}, P_u^H(t), P_u^I(t), h_*, R_u^{\text{history}}(h_*)\}) \quad (5)$$

where $\mathcal{L}_{\text{predict}}$ denotes the LLM model, and the prompt $\mathcal{P}_{\text{predict}}$ instructs the LLM to predict user behavior and generate a behavioral rationale following step-by-step guidelines. This DPT-based reasoning learns from past dynamic decision-making, yielding higher accuracy and personalization than relying on static profiles only.

Experiments

Experimental Setup

Dataset and Tasks We conduct experiments using LaMP (Salemi et al. 2023), a LLM personalization benchmark. Following prior research (Zhuang et al. 2024), We evaluate TRIPLE on five tasks spanning both classification and generation tasks: (1) personalized news categorization LaMP-2N, (2) personalized movie tagging LaMP-2M, (3) personalized rating prediction LaMP-3, (4) personalized news headline generation LaMP-4, and (5) personalized scholarly title generation LaMP-5. The data follows LaMP's time-based split, where profiles are created from historical data and evaluated on future data, reflecting real-world scenarios. Following the experimental setup of prior work (Zhuang et al. 2024), we randomly select 50 users for LaMP-2N/M and 100 users for LaMP-3/4/5. Each user profile contains up to 10 most recent data points, with evaluation on 4 test samples per user for LaMP-2N/M and 10 test samples per user for LaMP-3/4/5.

Implementation Details We compare TRIPLE against existing in-context learning-based profiling methods for personalization. These methods explicitly construct user profiles and incorporate them into prompts, enabling LLMs to generate outputs without fine-tuning. Our evaluation includes four profiling approaches: PAG (Richardson et al. 2023), ONCE (Liu et al. 2024), GPG (Zhang 2024), and PURE (Bang and Song 2025), along with a zero-shot baseline. To implement TRIPLE, we employed three open-source LLMs of varying sizes: LLaMA-3.1-70B-Instruct, LLaMA-3.1-8B-Instruct, and LLaMA-3.2-3B-Instruct (Grattafiori et al. 2024). See the supplementary materials for the prompts and additional settings used.

Main Results

To evaluate the effectiveness of our proposed framework, we conducted a comprehensive set of experiments on five distinct personalization tasks from the LaMP benchmark, using three different open-source Large Language Models. Comparative results against established in-context learning profiling methods are detailed in Table 2.

Empirical results reveal that TRIPLE consistently achieves superior or highly competitive performance across all tasks. Its advantages are especially prominent in the two generative tasks: news headline generation (LaMP-4) and scholarly title generation (LaMP-5). On these tasks, which demand a deeper understanding of a user's nuanced creative and stylistic preferences, TRIPLE achieves state-of-the-art performance by a significant margin across all evaluation metrics (ROUGE-1, ROUGE-L, and BLEU).

This significant improvement on generative tasks can be attributed to TRIPLE's unique architecture. By explicitly modeling both a user's stable Habitual Behavior Profile and their dynamic Intentional Behavior Profile, TRIPLE captures a more comprehensive and psychologically-grounded representation of the user. The Behavioral Rationale further guides the model to understand how and why a user arbitrates between habit and intention in a given context, a crucial factor for replicating generative styles.

Furthermore, TRIPLE also secures top-tier performance on classification (LaMP-2N, LaMP-2M) and regression (LaMP-3) tasks, underscoring the versatility and robustness of our theory-grounded approach. These findings validate our central hypothesis: integrating validated behavioral science theories into LLM-based user modeling yields substantial gains in both predictive and generative personalization.

Ablation Studies

To evaluate the contribution of each component, we conducted an ablation study. Table 3 presents the results. The findings first underscore the distinct roles of the two profiles. The Habitual Profile alone establishes a remarkably strong baseline, even outperforming the full model on LaMP-2N.

Table 2: Main experiment results on the LaMP benchmark. R-1 and R-L represent ROUGE-1 and ROUGE-L. ↑ denotes that higher values are better, while ↓ implies that lower values are preferred. The best score for each task is highlighted in **bold**.

Dataset (→)		LaMP-2N		LaMP-2M		LaMP-3		LaMP-4			LaMP-5		
Model (↓)	Method (↓)	Acc ↑	F-1 ↑	Acc ↑	F-1 ↑	MAE ↓	RSME ↓	R-1 ↑	R-L ↑	BLEU ↑	R-1 ↑	R-L ↑	BLEU ↑
LLaMA-3.1-70B-Instruct	Zero-shot	0.690	0.447	0.310	0.282	0.349	0.679	0.131	0.117	0.802	0.402	0.341	4.048
	PAG (Richardson et al. 2023)	0.720	0.486	0.585	0.448	0.316	0.643	0.098	0.086	0.475	0.144	0.128	1.200
	ONCE (Liu et al. 2024)	**0.795**	0.564	0.535	0.363	0.358	0.769	0.098	0.087	0.492	0.139	0.120	1.380
	GPG (Zhang 2024)	0.735	0.478	0.580	0.432	0.322	0.666	0.119	0.106	1.733	0.407	0.341	7.341
	PURE (Bang and Song 2025)	0.725	0.462	0.325	0.247	0.405	0.856	0.120	0.106	0.511	0.416	0.344	6.627
	TRIPLE (Ours)	**0.795**	**0.573**	**0.610**	**0.487**	**0.280**	**0.603**	**0.197**	**0.180**	**4.994**	**0.493**	**0.430**	**11.467**
LLaMA-3.1-8B-Instruct	Zero-shot	0.655	0.466	0.265	0.205	0.443	0.766	0.091	0.080	0.256	0.132	0.120	1.421
	PAG (Richardson et al. 2023)	0.785	0.528	0.455	**0.337**	0.426	0.877	0.051	0.044	0.234	0.128	0.114	0.842
	ONCE (Liu et al. 2024)	**0.805**	0.613	0.445	0.278	0.464	0.962	0.070	0.061	0.364	0.129	0.108	1.230
	GPG (Zhang 2024)	0.780	**0.617**	0.465	0.301	0.352	0.691	0.108	0.095	0.961	0.403	0.336	6.772
	PURE (Bang and Song 2025)	0.635	0.375	0.220	0.126	0.941	1.469	0.074	0.067	0.255	0.111	0.096	1.158
	TRIPLE (Ours)	0.800	0.613	**0.470**	0.313	**0.302**	**0.660**	**0.163**	**0.149**	**3.868**	**0.465**	**0.400**	**10.515**
LLaMA-3.2-3B-Instruct	Zero-shot	0.630	0.392	0.280	0.144	0.541	0.852	0.113	0.099	0.418	0.137	0.124	1.357
	PAG (Richardson et al. 2023)	0.760	0.503	0.435	**0.331**	0.378	0.731	0.121	0.108	1.473	0.427	0.355	7.620
	ONCE (Liu et al. 2024)	0.730	0.541	0.355	0.173	0.544	0.958	0.068	0.059	0.287	0.135	0.113	1.249
	GPG (Zhang 2024)	0.635	0.360	0.420	0.166	0.362	**0.721**	0.112	0.098	1.174	0.376	0.308	5.160
	PURE (Bang and Song 2025)	0.480	0.188	0.150	0.024	0.572	1.071	0.107	0.092	0.348	0.376	0.301	4.896
	TRIPLE (Ours)	**0.770**	**0.585**	**0.475**	0.290	**0.361**	0.745	**0.144**	**0.130**	**2.873**	**0.445**	**0.377**	**8.758**

Table 3: Ablation study of TRIPLE's core components on the LaMP dataset, conducted with the Llama-3.1-70B-Instruct model. The baseline (first row) uses a standard in-context learning approach with the 10 most recent raw user history records. The highlighted row represents the full TRIPLE model with all components.

Method	Components			LaMP-2N		LaMP-2M		LaMP-3		LaMP-4			LaMP-5		
	Habit	Intention	Rationale	Acc ↑	F-1 ↑	Acc ↑	F-1 ↑	MAE ↓	RSME ↓	R-1 ↑	R-L ↑	BLEU ↑	R-1 ↑	R-L ↑	BLEU ↑
Baseline	✗	✗	✗	0.755	0.529	0.425	0.339	0.308	0.641	0.137	0.123	2.947	0.372	0.304	4.048
Ours	✓	✗	✗	**0.805**	0.540	0.545	0.389	0.294	0.608	0.161	0.153	4.394	0.334	0.288	10.178
	✗	✓	✗	0.780	0.533	0.535	0.425	0.328	0.654	0.152	0.137	2.656	0.459	0.394	9.049
	✓	✓	✗	0.790	0.534	0.648	0.430	0.290	0.653	0.168	0.154	3.810	0.473	0.407	10.484
	✓	✓	✓	0.795	**0.573**	**0.610**	**0.487**	**0.280**	**0.603**	**0.197**	**0.180**	**4.994**	**0.493**	**0.430**	**11.467**

This suggests that stable, long-term behavioral patterns provide a powerful foundation for personalization. The Intentional Profile is less effective in isolation but captures a different dimension of user behavior. The most critical insight emerges when both profiles are included without the Behavioral Rationale. In this configuration, performance degrades on several tasks compared to using the Habit profile alone. This indicates that simply combining two profiles can introduce conflicting signals, potentially "confusing" the LLM due to an absence of a mechanism to arbitrate between habitual tendencies and situational intentions.

The Behavioral Rationale addresses this challenge by serving as an explicit reasoning scaffold that synthesizes the two profiles into a coherent interpretation. When included, the full TRIPLE exhibits significant performance gain across tasks, particularly in generative settings. These results underscore that the rationale is a core component that enables the model to reconcile and dynamically weigh the influence of habit and intention. This confirms the effectiveness of our DPT-guided approach and highlights the important role of interpretability in various user modeling scenarios.

Performance Analysis with Varying Context Size

Figure 3 shows how each method performs with varying k used for personalized prediction, evaluating how effectively additional user history improves performance. Overall, most methods show performance improvements as k increases. This suggests that LLMs benefit from additional behavioral context when modeling user preference. However, the extent of performance gains varies significantly across methods. Notably, TRIPLE achieves superior initial performance across most tasks at k=1. At this level, repetitive behavioral patterns are difficult to detect, so performance largely depends on the model's ability to infer a robust intentional profile from a single behavior. While the baselines show gradual performance improvements with increasing k, their performance stagnates beyond k=4. This suggests that simply summarizing past behavioral data appears limited in its ability to extract deeper insights from larger behavioral windows.

In contrast, TRIPLE achieves the best performance across all k values and demonstrates excellent scalability. As k increases, TRIPLE shows more noticeable improvements than other methods, leading to a widening performance gap with the baselines. This trend is particularly pronounced on generation tasks (LaMP-4, LaMP-5), where TRIPLE's BLEU score increase steadily, while baseline models stagnate (Figure 3). This demonstrates that TRIPLE's superior capacity to capture and replicate nuanced, personalized generative styles based on rich behavioral histories.

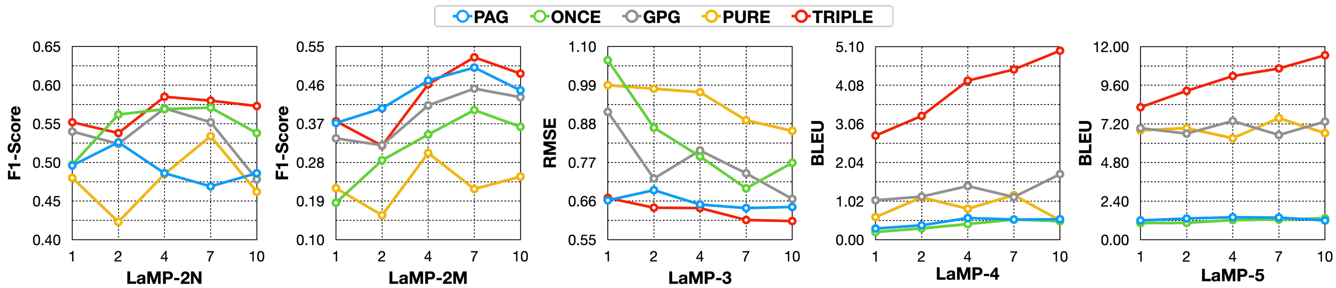

Figure 3: Performance variation of different profiles as a function of k values.

Paper Abstract	In many Twitter studies, it is important to know where a tweet came from in order to use the tweet content to study regional user behavior. However, researchers using Twitter to understand user behavior often lack sufficient geo-tagged data. Given the huge volume of Twitter data there is a need for accurate automated geolocating solutions. Herein, we present a new method to predict a Twitter user's location based on the information in a single tweet. We integrate text and user profile meta-data into a single model using a convolutional neural network. Our experiments demonstrate that our neural model substantially outperforms baseline methods, achieving 52.8% accuracy and 92.1% accuracy on city-level and country-level prediction respectively.
Title	On Predicting Geolocation of Tweets using Convolutional Neural Networks.
Habitual Profile	The author's habitual title patterns include **using action verbs like 'Exploring', 'Understanding', and 'Extracting'** to describe the research process, and **incorporating key concepts or methodologies like 'agent-based dynamic network simulation', 'social media analytics', and 'probabilistic geographical aspect-opinion model' to highlight the study's focus.** When presenting findings related to social media analysis, they structure titles as …
Intentional Profile	A. **Attitude**: The user has a positive evaluation of titles that **balance creativity, accuracy, informativeness, innovation, thoroughness, and real-world impact**, aligning with their values of effective academic title writing, academic rigor, and practical applications. **They believe such titles can lead to valuable outcomes, such as increased academic impact, engagement, and suitability for publication in reputable journals.** B. **Subjective norm**: The user **feels pressure to conform to established academic norms that prioritize clarity, concision, accuracy, informativeness**, … driven by their desire to meet established academic standards and expectations in their field. They are also influenced by the expectation that academic titles should … **adhere to a formal and objective tone.** C. **Perceived behavioral control:** The user **feels confident in their ability to generate effective titles that accurately reflect the content, driven by their perceived expertise in the field, understanding of methodologies and approaches**, … This confidence is reinforced by their ability to apply their knowledge and skills to create effective titles, summarize the main contributions of papers, and explain complex models and their evaluation results.
Behavior Rationale	The researcher chose the title 'On Predicting Geolocation of Tweets using Convolutional Neural Networks' due to **a combination of habitual and deliberate cognitive processes.** Their habitual writing patterns, … influenced the use of action verbs … **but the TPB factors, particularly the subjective norm and perceived behavioral control, led to a more concise and formal title that prioritizes clarity and accuracy. The TPB factors overrode the habitual response** to create a title that **meets established academic standards and expectations.** The final title reflects a balance between creativity and informativeness, aligning with the researcher's values and goals.

Figure 4: Examples of the profiles and the Behavior Rationale generated by TRIPLE on the LaMP-5.

Qualitative study

We conduct a qualitative analysis to assess the interpretability and psychological grounding of the profiles and rationales generated by TRIPLE. A key limitation of existing methods lies in their inability to provide transparent explanations for user behavior. In contrast, our analysis demonstrates how TRIPLE overcomes this by explicitly modeling the dual-process nature of human cognition—capturing the interaction between habitual and intentional behaviors—and articulates this dynamic through clear, natural language rationales. For instance, in a LaMP-5, TRIPLE generates the following rationale for a user's final title choice: "The title was selected due to the alignment of their habitual writing patterns and TPB factors." This rationale clarifies that the user's automatic tendency to write concise, descriptive titles was consistent with their conscious intention to communicate research effectively, as defined by their TPB profile. This showcases TRIPLE's ability to predict and explain decisions with cognitive transparency.

TRIPLE also demonstrates its unique strength in situations where a user's habitual tendencies and intentional goals are in conflict. As can be seen in the habitual profile, a user habitually began titles with action verbs (e.g.,"using phrases like 'Exploring', 'Understanding', or 'Extracting' to describe the research ...), but the final title adopted a more formal academic structure ("On Predicting..."). The generated rationale stated "The TPB factors overrode the habitual response to create a title that meets established academic standards." This reasoning reveals a cognitive conflict and its resolution, illustrating how subjective norms—the perceived pressure to conform to academic conventions—can override automatic tendencies. By identifying and explaining the psychological mechanism driving the behavioral shift, TRIPLE offers a nuanced account of user decision-making. TRIPLE demonstrates a depth of psychological fidelity that surpasses existing empirical approaches, which often lack the structure to capture or interpret such interactions. See the supplementary file for this user's past behavior and more examples.

Conclusion

This paper addresses a key limitation in current LLM-based personalization: a reliance on empirical heuristics that overlook the psychological mechanisms driving human behavior. We introduced TRIPLE, a novel framework that models users by integrating established behavioral theories from social psychology. Grounded in Dual-Process Theory, TRIPLE constructs two distinct profiles: a Habitual Profile, derived from repetitive behavioral patterns, and an Intentional Profile, based on the Theory of Planned Behavior (TPB). TRIPLE generates a Behavioral Rationale, an explicit explanation of how these two profiles interact to produce a final action. Experiments on the LaMP benchmark demonstrated that TRIPLE consistently outperformed existing in-context learning methods across multiple open-source LLMs. Ablation studies further confirmed that each component contributes to its overall performance. Overall, the primary contribution of this work lies in the successful integration of behavioral theories into LLM-based personalization, supported by a generalizable framework that can be extended to other theoretical models and user-centric applications. By modeling not only what users do but also why they do it, TRIPLE advances a new, theoretically grounded direction for user modeling. This research serves as a foundational example of theory-driven AI that can deepen human-AI interaction in personalization and beyond.

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) and the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea government (MOTIE) (No. RS-2025-02633871, Development of automated data pre-processing technology for industrial AI).

References

Ajzen, I. 1985. From intentions to actions: A theory of planned behavior. *Action control: From cognition to behavior/Springer*.

Ajzen, I. 1991. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2): 179–211.

Bang, S.; and Song, H. 2025. LLM-based User Profile Management for Recommender System. *arXiv preprint arXiv:2502.14541*.

Brown, D. J.; Charlesworth, J.; Hagger, M. S.; and Hamilton, K. 2021. A Dual-Process Model Applied to Two Health-Promoting Nutrition Behaviours. *Behavioral Sciences*, 11(12): 170.

Chaiken, S.; and Trope, Y. 1999. *Dual-process theories in social psychology*. Guilford Press.

De Bra, P. 2017. Challenges in user modeling and personalization. *IEEE Intelligent Systems*, 32(5): 76–80.

Demszky, D.; Yang, D.; Yeager, D. S.; Bryan, C. J.; Clapper, M.; Chandhok, S.; Eichstaedt, J. C.; Hecht, C.; Jamieson, J.; Johnson, M.; et al. 2023. Using large language models in psychology. *Nature Reviews Psychology*, 2(11): 688–701.

Evans, J. S. B. 2008. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, 59(1): 255–278.

Evans, J. S. B.; and Stanovich, K. E. 2013. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3): 223–241.

Fishbein, M.; and Ajzen, I. 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Addison-Wesley.

Grattafiori, A.; Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Vaughan, A.; et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Kosinski, M. 2024. Evaluating large language models in theory of mind tasks. *Proceedings of the National Academy of Sciences*, 121(45): e2405460121.

Koufaris, M. 2002. Applying the technology acceptance model and flow theory to online consumer behavior. *Information systems research*, 13(2): 205–223.

Liu, Q.; Chen, N.; Sakai, T.; and Wu, X.-M. 2024. Once: Boosting content-based recommendation with both open- and closed-source large language models. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, 452–461.

McCarthy, S.; O'Raghallaigh, P.; Kelleher, C.; and Adam, F. 2025. A socio-cognitive perspective of knowledge integration in digital innovation networks. *The Journal of Strategic Information Systems*, 34(1): 101871.

McEachan, R. R. C.; Conner, M.; Taylor, N. J.; and Lawton, R. J. 2011. Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, 5(2): 97–144.

Ouellette, J. A.; and Wood, W. 1998. Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124(1): 54.

Park, J. S.; O'Brien, J.; Cai, C. J.; Morris, M. R.; Liang, P.; and Bernstein, M. S. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, 1–22.

Richardson, C.; Zhang, Y.; Gillespie, K.; Kar, S.; Singh, A.; Raeesy, Z.; Khan, O. Z.; and Sethy, A. 2023. Integrating summarization and retrieval for enhanced personalization via large language models. *arXiv preprint arXiv:2310.20081*.

Robertson, S.; Zaragoza, H.; et al. 2009. The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends® in Information Retrieval*, 3(4): 333–389.

Salemi, A.; Mysore, S.; Bendersky, M.; and Zamani, H. 2023. Lamp: When large language models meet personalization. *arXiv preprint arXiv:2304.11406*.

Samson, A.; and Voyer, B. G. 2012. Two minds, three ways: dual system and dual process models in consumer psychology. *AMS review*, 2: 48–71.

Sheeran, P. 2002. Intention-behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1): 1–36.

Strack, F.; and Deutsch, R. 2004. Reflective and impulsive determinants of social behavior. *Personality and social psychology review*, 8(3): 220–247.

Tam, K. Y.; and Ho, S. Y. 2005. Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Information systems research*, 16(3): 271–291.

Tam, K. Y.; and Ho, S. Y. 2006. Understanding the impact of web personalization on user information processing and decision outcomes. *MIS quarterly*, 865–890.

Wood, W.; and Neal, D. T. 2007. A new look at habits and the habit-goal interface. *Psychological review*, 114(4): 843.

Zhang, J. 2024. Guided profile generation improves personalization with llms. *arXiv preprint arXiv:2409.13093*.

Zhang, S. 2025. Leveraging Dual Process Theory in Language Agent Framework for Real-time Simultaneous Human-AI Collaboration. *arXiv preprint arXiv:2502.11882*. Accepted by ACL 2025 Main.

Zhang, Y.; Chen, X.; et al. 2020. Explainable recommendation: A survey and new perspectives. *Foundations and Trends® in Information Retrieval*, 14(1): 1–101.

Zhao, Z.; Cheng, Z.; Hong, L.; and Chi, E. H. 2015. Improving user topic interest profiles by behavior factorization. In *Proceedings of the 24th international conference on world wide web*, 1406–1416.

Zhou, Y.; Wang, H.; He, J.; and Wang, H. 2025. Based Explainable Recommendations: A Transparency Perspective. *ACM Transactions on Recommender Systems*, 3(3): 1–20.

Zhuang, Y.; Sun, H.; Yu, Y.; Qiang, R.; Wang, Q.; Zhang, C.; and Dai, B. 2024. Hydra: Model factorization framework for black-box llm personalization. *Advances in Neural Information Processing Systems*, 37: 100783–100815.