UNIER: A UNIFIED BENCHMARK FOR ITEM-LEVEL AND PATH-LEVEL EXERCISE RECOMMENDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Personalized exercise recommendation dynamically aligns pedagogical resources with individual knowledge mastery, which is crucial for satisfying students' dynamic learning needs in modern education. Existing approaches primarily follow two paradigms: Item-Level Exercise Recommendation (ILER) ranks exercises for immediate recommendation, whereas Path-Level Exercise Recommendation (PLER) plans ordered exercise sequences. Despite sharing the same ultimate objective, disparate evaluation setups have kept these two lines of research isolated, hindering unified benchmarking and fair comparison. To fill the gap, in this paper, we present a **Uni**fied Benchmark for **E**xercise **R**ecommendation (UniER), a comprehensive evaluation framework that unifies ILER and PLER. Specifically, we introduce Weighted Cognitive Gain (WCG) to compare estimated knowledge gains under different instructional objectives and shared interaction budgets. We benchmark 18 representative methods spanning four recommendation categories on nine public datasets. Through analyses of effectiveness, generalizability, robustness, and efficiency, we find that the leading methods change with the interaction budget: ILER methods achieve the highest average scores at one step, whereas PLER methods lead at ten steps. Performance under data sparsity, cold-start, and label noise further varies across methods and datasets, providing evidence for selecting recommendation methods according to the intended learning setting. Furthermore, we provide an open-source codebase of UniER to foster reproducible research and outline potential directions for future investigations[1].

1 INTRODUCTION

Exercise Recommendation (ER) has become a cornerstone of personalized learning, transforming how educational resources are aligned with individual learner needs Cheng et al. (2025); Liu et al. (2019). The fundamental objective of ER is actively providing appropriate exercises to assist students in acquiring knowledge, thereby eliminating the need for self-directed searches. Although ER shares the notion of "recommendation", it comes with significant challenges that are not present in more conventional recommendation domains, such as e-commerce and media recommendation Rajapakse & Jannach (2025); Greenwood et al. (2024); Zhang et al. (2024b). In particular, interaction pattern mining techniques that are effective in conventional recommendations fail to reflect students' knowledge states, and optimizing solely for user preferences is insufficient to promote knowledge mastery. Consequently, effective ER requires paradigms that extend beyond behavioral similarity and preference optimization to directly support knowledge acquisition.

As one of the representative ER paradigms, Item-Level Exercise Recommendation (ILER) follows the traditional item-based recommender system paradigm, standing as a classical research direction Wu et al. (2020); Guan et al. (2023); Cheng et al. (2025). Specifically, ILER formulates the task as a next-step recommendation process, aiming to retrieve an unordered set of exercises tailored to the student's immediate cognitive needs. Existing ILER methods can be categorized into two branches, namely end-to-end approaches and two-stage approaches. End-to-end approaches seamlessly project unmastered knowledge concepts into the exercise space and make recommendations directly. To further maximize the concept diversity within the final recommendation list, two-stage approaches explicitly decouple the process into filtering and re-ranking phases. Nevertheless,

[1] https://anonymous.4open.science/r/UniER-4532/

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

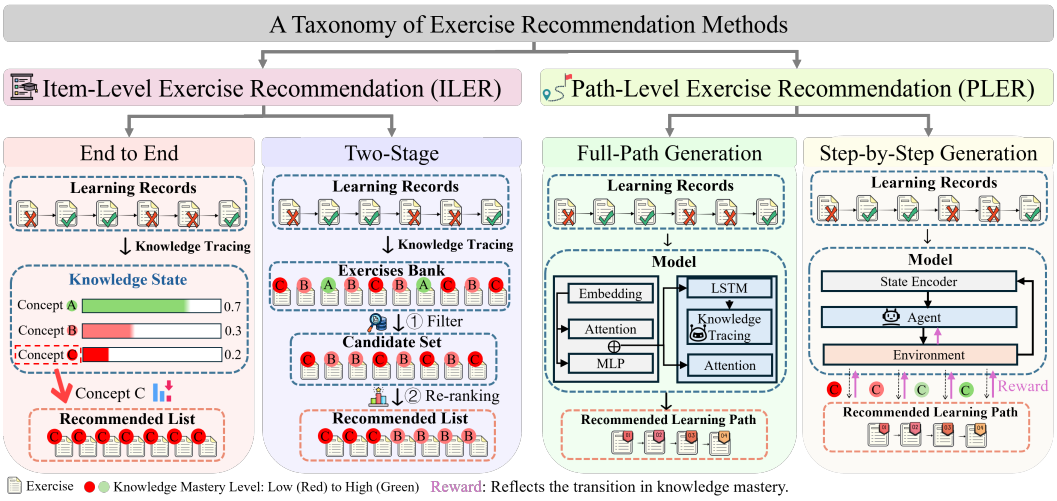

Figure 1: A Taxonomy of Exercise Recommendation Methods.

existing ILER methods treat the recommendation process as isolated steps rather than a coherent learning path, which neglects the continuous nature of learning processes.

Driven by the necessity to construct coherent learning paths, another line of research, Path-Level Exercise Recommendation (PLER), draws increasing attention in recent years Liu et al. (2019); Li et al. (2024); Cheng et al. (2026). PLER aims to generate an ordered sequence of exercises, meaning it is designed to maintain cognitive smoothness, and indicating a primary focus on achieving predefined pedagogical goals. Existing PLER approaches usually follow two paradigms: full-path generation and step-by-step generation. Full-path generation methods employ an encoder-decoder framework to output the entire learning path in a single step by capturing latent dependencies between exercises. In contrast, step-by-step generation methods based on reinforcement learning (RL) model learning path generation as a Markov Decision Process (MDP) Pateria et al. (2021), dynamically generating personalized paths aligned with students' cognitive development patterns. Compared to ILER, PLER places greater emphasis on modeling long-term sequential dependencies and goal-oriented learning planning. However, it often relies on predefined paths, thereby neglecting the pedagogical flexibility and immediate feedback inherent in individual learning processes.

Despite the methodological diversity in current ER research, the evaluation of ER approaches remains inadequate and incomplete, mainly due to two key challenges. ***Challenge 1****: Lack of comprehensive evaluation protocols.* As ILER and PLER paradigms diverge fundamentally in both predictive mechanisms and core objectives, they inherently rely on incompatible evaluation metrics. For instance, ILER predominantly relies on generic recommendation metrics that are fundamentally agnostic to pedagogical applicability. Even when incorporating KT, such evaluations prioritize isolated concept mastery, thereby neglecting the continuous dynamics of the learning process. In contrast, while PLER explicitly models the learning path, it lacks a comprehensive characterization of the overall knowledge framework, heavily relying on aggregated path-level rewards for evaluation. This disjointed evaluation protocol yields one-sided performance characterizations, hindering a comprehensive understanding of practical algorithmic efficacy. ***Challenge 2****: Lack of unified horizontal comparisons*. Current studies report outcomes under disparate datasets and heterogeneous settings, lacking a unified benchmark to facilitate fair cross-method and cross-paradigm comparisons. Given the intrinsic differences between the ILER and PLER paradigms, direct comparison is inherently precluded. Consequently, although these methods have achieved remarkable progress within their specific settings, evaluating their relative strengths and applicable scenarios remains challenging, thereby creating a bottleneck for effective real-world deployment.

To fill the gap, in this paper, we develop a **Uni**fied Benchmark for **E**xercise **R**ecommendation, abbreviated as UniER. To overcome ***Challenge 1***, grounded in the Item Response Theory (IRT) Lord (1952; 2012), we propose Weighted Cognitive Gain (WCG), a unified metric to quantify algorithmic performance for holistic ER evaluation. WCG models cognitive gain through

customizable weights, allowing us to evaluate each algorithm from multiple perspectives by instantiating different evaluation principles via varying weight configurations. To address ***Challenge 2***, building upon WCG, we establish a comprehensive benchmark UniER based on 9 datasets to fairly compare 18 ILER and PLER methods under a unified experimental setting, exploring the performance landscape of state-of-the-art approaches across diverse scenarios and domains. In terms of **effectiveness**, we conduct a cross-paradigm empirical study within a unified benchmark, systematically uncovering the performance trade-offs between item-level accuracy and path-level coherence. To assess **generalizability**, we stress-test existing models under varying degrees of data sparsity and cold-start conditions, thereby revealing the inherent limitations of current paradigms under data scarcity. Furthermore, we examine model **robustness** by introducing stochastic noise into educational datasets, which allows us to quantify the resilience of these methods against data perturbations. Finally, we benchmark **efficiency** by quantifying training/inference latency and peak memory footprint, providing insights into the practical scalability of representative ER approaches. At the end of this paper, we discuss the future directions of this emerging research.

2 PRELIMINARIES

In this section, we formally define two core paradigms of exercise recommendation (ER), namely Item-Level Exercise Recommendation (ILER) and Path-Level Exercise Recommendation (PLER).

Unified Problem Formulation. The primary objective of ER is to recommend a subset of suitable exercises based on a student's historical interaction sequence. Let $\mathcal{S} = \{s_1, s_2, \ldots, s_{|\mathcal{S}|}\}$ denote the set of students, $\mathcal{C} = \{c_1, c_2, \ldots, c_{|\mathcal{C}|}\}$ denote the set of Knowledge Concepts (KCs), and $\mathcal{E} = \{e_1, e_2, \ldots, e_{|\mathcal{E}|}\}$ denote the set of exercises, where $|\mathcal{S}|$, $|\mathcal{C}|$, and $|\mathcal{E}|$ represent their respective total numbers. Furthermore, the relationship between exercises and KCs is represented by a binary Q-matrix $\mathbf{Q} \in \{0,1\}^{|\mathcal{E}|\times|\mathcal{C}|}$, where $Q_{i,j} = 1$ indicates that exercise e_i involves KC c_j, and $Q_{i,j} = 0$ otherwise. The historical interaction sequence of student $s \in \mathcal{S}$ up to time step t is defined as $\mathcal{H}_{s,t} = \{(e^{(1)}, r^{(1)}), (e^{(2)}, r^{(2)}), \ldots, (e^{(t)}, r^{(t)})\}$, where $r^{(i)} = 1$ indicates that the student correctly answered exercise $e^{(i)} \in \mathcal{E}$, and $r^{(i)} = 0$ otherwise.

Item-Level Exercise Recommendation (ILER). The ILER paradigm primarily aims to recommend exercises that target a student's cognitive weaknesses. To achieve this, conventional approaches leverage Knowledge Tracing (KT) techniques Piech et al. (2015); Ma et al. (2025) to diagnose unmastered KCs and retrieve candidate exercises covering these concepts Huang et al. (2019); Ghosh et al. (2020); Liu et al. (2023b;a). Specifically, the KT model estimates the student's mastery over KCs from the historical interaction sequence $\mathcal{H}_{s,t}$, resulting in a continuous knowledge state vector $\mathcal{M}_{\text{KT}}(\mathcal{H}_{s,t}) \in [0,1]^{|\mathcal{C}|}$. The KC-level knowledge gaps are represented as $(\mathbf{1} - \mathcal{M}_{\text{KT}}(\mathcal{H}_{s,t}))$ and mapped to the exercise space via the Q-matrix $\mathbf{Q}$, yielding exercise-level gap scores. The final recommendation set $\mathcal{E}_I$ is obtained by selecting the Top-K exercises with the highest gap scores:

$$\mathcal{E}_I = \text{Top-}K\left(\mathbf{Q} \cdot (\mathbf{1} - \mathcal{M}_{\text{KT}}(\mathcal{H}_{s,t}))\right). \tag{1}$$

However, this formulation exhibits a locally greedy selection strategy, repeatedly recommending exercises associated with the most unmastered KCs, which leads to homogeneous recommendation results. To mitigate this issue, recent works adopt two-stage architectures that decouple candidate generation from ranking to promote recommendation diversity. Building upon KT-based retrieval, these paradigms incorporate a diversified re-ranking stage to cover a broader spectrum of KCs Wu et al. (2020); Ren et al. (2023); Liu et al. (2025). In the *retrieval stage*, the algorithm identifies a candidate pool $\mathcal{E}_{\text{cand}}$ by selecting the Top-N exercises that correspond to the student's estimated knowledge gaps. Subsequently, in the *re-ranking stage*, the system selects an optimal subset $\mathcal{E}_I \subset \mathcal{E}_{\text{cand}}$ of size K ($K < N$) that maximizes the set-level KC diversity:

$$\mathcal{E}_I = \underset{\mathcal{E}' \subset \mathcal{E}_{\text{cand}}, |\mathcal{E}'|=K}{\arg\max} \Omega_{div}(\mathcal{E}', \mathbf{Q}). \tag{2}$$

The objective function $\Omega_{div}(\cdot, \mathbf{Q})$ promotes conceptual diversity by maximizing pairwise dissimilarity in KC coverage among the selected exercises. Both paradigms can be viewed under a unified framework of ILER. The output of an ILER method is a Top-K subset of exercises:

$$\mathcal{E}_I = \{e_1, e_2, \ldots, e_K\} \subset \mathcal{E}. \tag{3}$$

Notably, the elements in $\mathcal{E}_I$ are treated as an *unordered set*, without any explicit sequential relationship among the exercises. This drastically contrasts with path-level exercise recommendations, where the internal ordering and transition of exercises are explicitly modeled.

Path-Level Exercise Recommendation (PLER). The primary objective of PLER is to generate coherent learning paths that guide students toward specific pedagogical goals. As a representative work, SRC Chen et al. (2023) casts this task within a set-to-sequence paradigm, employing an encoder-decoder architecture to facilitate efficient and concept-aware path generation. Formally, given the student's historical interaction sequence $\mathcal{H}_{s,t}$ and a target goal G, SRC aims to identify an optimal learning path $\mathcal{T}^*$ of length N. This task is framed within a full-path generation framework, where the entire sequence is generated to optimize global coherence:

$$\mathcal{T}^* = \text{Decoder}\left(\text{Encoder}(\mathcal{H}_{s,t}, G)\right) = \langle e_{p_1}, e_{p_2}, \ldots, e_{p_N} \rangle, \tag{4}$$

where $\text{Encoder}(\cdot)$ compresses both the target goals and an unordered set of unmastered knowledge concepts extracted from the historical sequence $\mathcal{H}_{s,t}$ into a latent global representation. Subsequently, $\text{Decoder}(\cdot)$ autoregressively maps this context into a fixed-length learning path within the exercise space $\mathcal{E}^N$. By adopting this full-path generation approach, the system is capable of modeling the underlying inter-exercise sequential dependencies and prerequisite constraints.

Alternatively, another prevalent approach within PLER formulates the task as a Markov Decision Process (MDP), predominantly leveraging Reinforcement Learning (RL) to dynamically generate optimal learning paths Liu et al. (2019); Li et al. (2023); Zhang et al. (2024a); Li et al. (2024); Cheng et al. (2026). Within these RL-based methods Konda & Tsitsiklis (1999); Tang et al. (2019); Kubotani et al. (2021), several advanced approaches explicitly incorporate knowledge structure graphs to model prerequisites among KCs, ensuring that foundational KCs are mastered before advancing to the target goal Liu et al. (2019); Zhang et al. (2024a). Formally, at each sequential step $i \in \{1, \ldots, N\}$, the agent observes the current state, which comprises the target goal G and the real-time knowledge state estimated from the dynamic interaction history $\mathcal{H}_{s,t+i-1}$, and selects the optimal exercise $e^*_{p_i}$ following a parameterized policy π_Θ:

$$e^*_{p_i} = \arg\max_{e \in \mathcal{E}} \pi_\Theta(e \mid \mathcal{H}_{s,t+i-1}, G). \tag{5}$$

Crucially, unlike static generation, the historical context expands dynamically after each intra-path recommendation, such that $\mathcal{H}_{s,t+i} = \mathcal{H}_{s,t+i-1} \cup \{(e^*_{p_i}, r_{p_i})\}$, where r_{p_i} denotes the simulated instantaneous student feedback. Through these sequential state transitions and localized policy decisions, the algorithm synthesizes a strictly *ordered sequence* of length N, designed to guide learners step-by-step:

$$\mathcal{T}^* = \langle e^*_{p_1}, e^*_{p_2}, \ldots, e^*_{p_N} \rangle, \tag{6}$$

where the internal transition from $e^*_{p_i}$ to $e^*_{p_{i+1}}$ is intrinsically driven by the maximization of long-term pedagogical outcomes. While these dynamic PLER methodologies demonstrate profound capabilities in formulating logically coherent paths, they have largely evolved as a siloed research domain. This divergence has resulted in a critical lack of cross-paradigm evaluation methods that can holistically benchmark both ILER and PLER within a unified recommendation framework.

3 Benchmark Design

In this section, we provide a comprehensive overview of UniER covering the Weighted Cognitive Gain metric, algorithms, evaluation datasets, and implementation details.

3.1 Weighted Cognitive Gain Metric

Evaluating the quality of recommended exercises is fundamental to personalized learning systems. However, existing evaluation metrics are highly fragmented and often focus on isolated aspects of performance, making it difficult to fairly compare different methodologies. More importantly, they implicitly assume that all knowledge concepts (KCs) are equally important, which contradicts real-world instructional settings where learning objectives vary across scenarios. In practice, certain KCs may play a more critical role depending on pedagogical goals, such as intensive exam-oriented training versus broad knowledge exploration. To address this limitation, we propose Weighted Cognitive Gain (WCG), a configurable evaluation metric that allows flexible weighting over KCs.

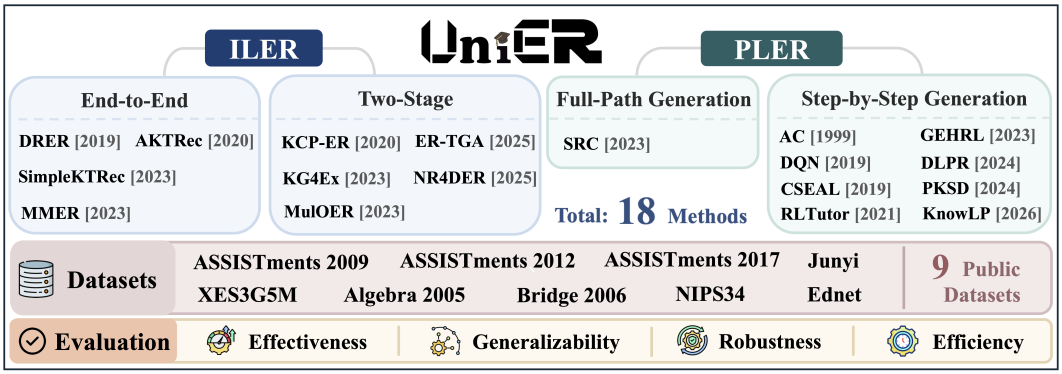

Figure 2: An overview of UniER.

By explicitly modeling concept importance, WCG enables scenario-aware evaluation and provides a unified framework for fair comparison across different recommendation paradigms.

Unified Output. To align evaluation budgets while accommodating different uses of recommendation lists, we represent each method's executed exercises as an ordered sequence $\mathcal{T}_{\text{unified}} = \langle e_1, \ldots, e_N \rangle$ under a shared budget of N interactions. PLER follows its native path-generation procedure, whereas ILER uses one of three sequencing protocols: *(i) DKT-guided adaptive sequencing*, which models practice with feedback: at each step i, DKT selects the easiest-to-master exercise from the current recommendation list $\mathcal{E}_I^{(i)}$ if a candidate satisfies the learnability criterion; otherwise, the highest-ranked unmastered exercise is selected. The resulting response is appended to the interaction history, and the recommendation list is regenerated before the next selection. *(ii) Static score ordering*, which preserves the model's ranking by executing the initial list in descending order of recommendation scores. *(iii) Static random ordering*, which represents list use without following the model's ranking by randomly permuting the initial list once and executing the resulting sequence. The latter two protocols retain the initial list throughout execution, whereas the first updates recommendations after every interaction. We use protocol (i) for the main-text results and compare all three protocols in Appendix D to assess sensitivity to how ILER recommendations are executed.

Unified Evaluation Metric. To establish a rigorously fair benchmark, we formalize a unified metric: WCG. We introduce a task-specific target weight distribution $\mathbf{w} \in [0,1]^{|\mathcal{C}|}$, subject to $\sum_{c \in \mathcal{C}} w_c = 1$. Under this unified evaluation framework, the performance of any generated learning path $\mathcal{T}_{\text{unified}}$ is quantified by measuring its expected weighted knowledge mastery improvement:

$$\text{WCG}(\mathcal{T}_{\text{unified}} \mid \mathbf{w}) = \sum_{c \in \mathcal{C}} w_c \cdot \left(\mathcal{M}_{\text{KT}}^{(c)}(\mathcal{H}_{s,t} \oplus \text{Sim}(\mathcal{T}_{\text{unified}})) - \mathcal{M}_{\text{KT}}^{(c)}(\mathcal{H}_{s,t}) \right), \tag{7}$$

where $\mathcal{M}_{\text{KT}}^{(c)}(\mathcal{H}_{s,t})$ and $\mathcal{M}_{\text{KT}}^{(c)}(\mathcal{H}_{s,t} \oplus \text{Sim}(\mathcal{T}_{\text{unified}}))$ denote the student's estimated mastery level on concept c before and after executing the learning path, respectively. Here, $\oplus$ denotes sequence concatenation, and $\text{Sim}(\cdot)$ represents the simulated response generation process for the recommended path, executed by the underlying knowledge state simulator. The weight configuration $\mathbf{w}$ can be flexibly adjusted to accommodate specific pedagogical requirements, a key advantage that enables UniER to generalize across diverse educational contexts. In this study, we primarily instantiate two representative settings to demonstrate this adaptability in real-world scenarios:

- **Task 1: Targeted Goal Achievement (TGA).** Let $\mathcal{C}_{\text{target}} \subset \mathcal{C}$ denote a specific set of target concepts a student needs to master. We construct a highly sparse weight vector by defining $w_c = 1/|\mathcal{C}_{\text{target}}|$ if $c \in \mathcal{C}_{\text{target}}$, and $w_c = 0$ otherwise. This setting strictly evaluates the model's architectural capacity to synthesize logically progressive paths that fulfill specific pedagogical goals, effectively quantifying the average cognitive gain across the designated targets. In practical, TGA aligns with the pedagogical scenario of targeted exam preparation, where a student must efficiently master a predefined set of crucial KCs.
- **Task 2: Global Proficiency Promotion (GPP).** Let $\mathcal{C}_{\text{unmastered}} \subset \mathcal{C}$ denote the set of all concepts the student has yet to master at time t. We define a uniform distribution over these concepts

Table 1: Comparison in terms of GPP@10 and TGA@10. The best three results are highlighted by 1st, 2nd, and 3rd. "Avg." indicate the average GPP@10 and TGA@10 across all datasets.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M		Avg.
	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	T&G
DRER	-0.208	-0.183	0.055	0.033	-0.003	0.002	-0.487	-0.400	0.083	0.122	0.365	0.286	0.222	0.084	0.023
AKTRec	-0.083	-0.051	0.023	0.025	-0.000	-0.001	-0.183	-0.016	0.090	0.137	0.315	0.188	0.224	0.089	0.057
S-KTRec	-0.087	-0.052	0.030	0.020	-0.001	-0.000	-0.171	-0.020	0.091	0.133	0.314	0.185	0.222	0.091	0.059
MMER	-0.141	-0.126	0.085	0.042	-0.009	-0.005	-0.014	0.044	0.082	0.043	0.338	0.247	0.249	0.032	0.070
KCP-ER	-0.174	-0.121	0.154	0.023	-0.002	0.001	-0.224	0.047	0.104	0.128	0.310	0.187	0.224	0.087	0.070
KG4Ex	-0.182	-0.118	0.116	0.057	-0.002	-0.002	0.061	0.060	0.095	0.136	0.282	0.140	0.236	0.130	0.087
MulOER	-0.083	-0.065	0.065	0.034	-0.006	-0.003	0.030	0.087	0.094	0.131	0.271	0.162	0.228	0.086	0.086
ER-TGA	-0.114	-0.083	0.104	0.052	0.002	0.006	0.015	0.027	0.091	0.142	-0.063	0.003	0.248	0.130	0.009
NR4DER	-0.144	-0.090	0.104	0.047	0.002	0.002	-0.117	0.039	0.095	0.145	0.314	0.183	0.228	0.099	0.074
SRC	0.024	-0.135	0.167	0.082	0.000	0.007	0.306	-0.340	0.102	-0.103	0.116	0.170	-0.026	0.017	0.011
AC	0.283	0.179	0.244	0.096	0.000	0.005	-0.38	0.405	0.457	0.355	0.419	0.485	0.902	-0.246	0.190
DQN	0.164	0.121	0.084	0.126	0.004	0.006	0.092	0.486	0.212	0.284	0.509	0.105	0.877	0.255	0.215
CSEAL	0.132	0.208	0.162	0.054	-0.001	0.005	0.042	0.363	0.419	0.370	0.400	0.481	0.914	-0.118	0.210
RLTutor	0.226	0.246	0.213	0.125	0.005	0.005	-0.181	0.469	0.318	0.378	0.617	0.443	0.888	-0.126	0.259
GEHRL	-0.226	-0.144	0.161	0.034	0.017	0.007	0.329	0.484	-0.277	-0.214	-0.17	0.101	0.918	0.243	0.018
DLPR	0.093	0.140	0.181	0.112	0.006	0.006	-0.109	0.364	0.218	0.275	0.322	0.357	0.755	0.237	0.179
PKSD	0.268	0.290	0.235	0.125	0.001	0.006	0.225	0.417	0.288	0.381	0.644	0.517	0.914	0.152	0.311
KnowLP	-0.031	0.171	0.203	0.123	0.004	0.005	0.107	0.370	0.255	0.355	0.474	0.381	0.803	0.256	0.232

by setting $w_c = 1/|\mathcal{C}_{\text{unmastered}}|$ if $c \in \mathcal{C}_{\text{unmastered}}$, and $w_c = 0$ otherwise. This task evaluates the model's capability to maximize overall knowledge coverage and broad deficit remediation. Practically, this corresponds to the pedagogical scenario of comprehensive end-of-term reviews, where the primary objective is to elevate the student's overall academic proficiency.

3.2 Benchmark Algorithms

To ensure a comprehensive and representative evaluation, UniER incorporates a diverse suite of state-of-the-art ER methods. These algorithms are categorized into two primary paradigms based on their recommendation granularity and structural design: Item-Level Recommendation (ILER) and Path-Level Recommendation (PLER). For a concise summary of the specific model configurations and implementation details, please refer to Figure 2 and Appendix B.

3.3 Datasets and Implementation Details

Datasets. We evaluate UniER on nine widely recognized datasets: ASSISTments (2009, 2012, and 2017) Abdelrahman & Wang (2019), Bridge2006, Junyi, XES3G5M Liu et al. (2023c), Algebra2005 Stamper et al. (2010), Ednet, and NIPS34 Liu et al. (2022), which span a diverse array of educational domains and scales. Comprehensive data statistics and preprocessing specifications are detailed in Appendix A and C.2.

Hyperparameter Search. To obtain the performance upper bounds of various methods on ILER/PLER tasks, we conduct a random search to find the optimal hyperparameters. The search space is detailed in Tables 4 and 5. The random search is conducted 20 times or for a maximum of one day per method per dataset to ensure fairness. For more details related to the experimental setup in UniER, please refer to Appendix C.2.

4 Experimental Results & Discussion

In this section, we present a comprehensive evaluation and in-depth analysis on the UniER benchmark. Specifically, our experiments are designed to answer the following research questions (RQs):

- **RQ1 (Effectiveness):** How do ILER and PLER perform under the unified WCG metric?
- **RQ2 (Generalizability):** How does performance degrade under data sparsity and cold-start?
- **RQ3 (Robustness):** How robust are these representative models against underlying label noise?
- **RQ4 (Efficiency):** What is the trade-off between computational cost and performance?

Table 2: Comparison in terms of GPP@1 and TGA@1. The best three results are highlighted by 1st, 2nd, and 3rd. "Avg." indicate the average GPP@1 and TGA@1 across all datasets.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M		Avg.
	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	T&G
DRER	-0.154	-0.138	0.089	0.043	-0.004	-0.002	0.176	0.270	0.075	0.081	0.038	0.085	0.216	0.071	0.077
AKTRec	0.019	0.044	0.087	0.051	0.001	0.000	0.050	0.064	0.076	0.106	0.333	0.166	0.224	0.069	0.095
S-KTRec	0.013	0.037	0.051	0.041	0.001	0.000	0.043	0.058	0.072	0.100	0.343	0.179	0.218	0.051	0.091
MMER	-0.080	-0.063	-0.207	-0.053	-0.006	0.003	0.013	0.034	0.062	0.066	0.238	0.180	0.254	0.073	0.049
KCP-ER	0.012	0.036	0.065	0.033	0.000	0.000	-0.113	0.034	0.088	0.104	0.245	0.170	0.219	0.075	0.079
KG4Ex	-0.133	-0.097	0.111	0.048	-0.002	-0.002	0.000	0.000	0.068	0.092	0.200	0.068	0.213	0.065	0.062
MulOER	-0.012	0.010	0.071	0.047	-0.006	-0.002	-0.026	0.006	0.069	0.085	0.211	0.153	0.198	0.048	0.074
ER-TGA	0.005	0.026	0.113	0.064	-0.002	-0.001	0.096	0.121	0.117	0.157	-0.120	-0.120	0.030	-0.141	0.009
NR4DER	-0.024	0.010	0.084	0.039	-0.002	0.002	-0.016	0.074	0.063	0.087	0.258	0.143	0.234	0.113	0.084
SRC	0.260	-0.070	0.221	-0.082	0.000	-0.002	0.207	0.056	0.032	-0.254	-0.410	0.181	-0.020	0.003	0.024
AC	-0.113	-0.061	-0.156	-0.027	0.000	0.002	-0.667	-0.216	-0.281	-0.156	-0.165	-0.070	0.400	-0.185	-0.120
DQN	-0.114	-0.048	-0.129	-0.027	0.000	0.001	-0.644	-0.202	-0.316	-0.167	-0.174	-0.051	0.291	-0.231	-0.134
CSEAL	0.075	0.017	-0.111	-0.015	-0.001	0.002	-0.207	-0.007	0.127	-0.133	0.145	0.055	0.436	-0.243	-0.005
RLTutor	-0.091	-0.056	-0.168	-0.024	0.001	0.001	-0.334	-0.181	-0.056	-0.161	0.027	0.040	0.447	-0.164	-0.066
GEHRL	-0.056	-0.067	0.038	0.040	-0.005	-0.006	-0.500	-0.466	-0.439	-0.413	-0.184	-0.181	0.357	0.370	-0.135
DLPR	-0.068	-0.094	0.131	0.075	-0.003	0.002	-0.788	-0.243	-0.247	-0.197	-0.392	0.011	0.614	0.328	-0.046
PKSD	-0.032	-0.031	-0.106	-0.015	0.001	0.002	-0.494	-0.183	0.190	-0.149	0.213	-0.025	0.512	-0.191	-0.004
KnowLP	-0.047	-0.037	0.140	0.085	0.006	0.012	-0.418	-0.205	0.087	-0.152	0.095	0.021	0.523	0.338	0.039

4.1 Overall Performance Comparison (RQ1)

Experiment Design. We compare 18 baseline algorithms across nine benchmark datasets using TGA@N and GPP@N, where $N \in \{1, 10\}$ denotes the number of exercise interactions. Evaluating both single-step and ten-step gains allows us to examine how relative performance varies with the interaction budget.

Experimental Results. Tables 1 and 2 report TGA and GPP under single-step and ten-step interaction budgets, respectively. Following prior studies Liu et al. (2019); Li et al. (2024), we use the same DKT-based student simulator to provide interaction feedback during policy training and to estimate knowledge gains during evaluation. This shared simulation protocol provides a consistent student environment for comparing recommendation methods. Appendix D complements the main evaluation with conventional recommendation metrics and additional WCG@N settings, and examines sensitivity to the evaluation simulator by replacing DKT with SimpleKT while keeping the trained recommendation models fixed. We highlight the following key observations:

Observation ❶: Leading methods change with the interaction budget. At one step, ILER methods achieve the eight highest average scores, led by AKTRec (0.095) and S-KTRec (0.091). By comparison, KnowLP, the strongest PLER method at this budget, achieves an average score of 0.039. At ten steps, PLER methods occupy the seven highest average ranks. PKSD (0.311) and RLTutor (0.259) outperform the strongest ILER method, KG4Ex (0.087). Evaluating both budgets therefore reveals complementary strengths of the two paradigms that a single-horizon comparison would not fully capture. This reversal suggests matching recommendation methods to the time and effort learners can devote to practice.

Observation ❷: Targeted and global gains can diverge. On XES3G5M, AC achieves higher TGA@10 than KnowLP (0.902 vs. 0.803) but lower GPP@10 (-0.246 vs. 0.256). This contrast shows that stronger target-specific gains need not coincide with stronger global estimated gains. Jointly reporting TGA and GPP reveals differences obscured when considering either objective alone. These rankings support method choices for targeted remediation and broad review.

Observation ❸: Method advantages vary across learning environments. Although PKSD achieves the highest average score at ten steps, GEHRL outperforms it on Ednet in both TGA (0.329 vs. 0.225) and GPP (0.484 vs. 0.417). At one step, the PLER method DLPR exceeds the strongest ILER method, MMER, on XES3G5M in TGA@1 (0.614 vs. 0.254). These results reveal complementary strengths across the evaluated learning environments, with no single method consistently performing best. By exposing these differences under shared evaluation protocols, UniER provides a common basis for assessing the relative strengths of methods across learning environments.

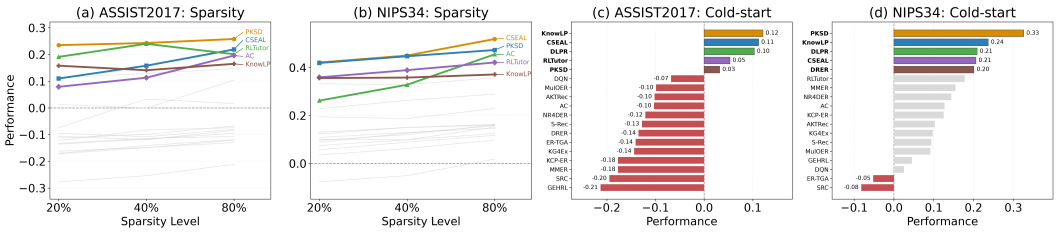

Figure 3: Generalizability analysis under data sparsity and cold-start settings. In subfigures (a) and (b), the top-five methods are highlighted, while the remaining methods are represented by gray lines.

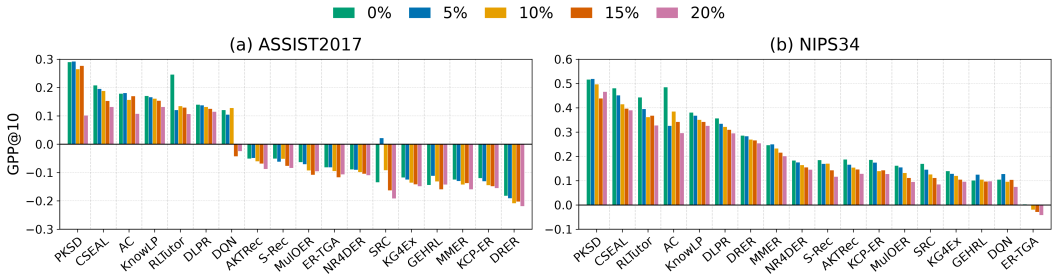

Figure 4: Performance under different perturbation levels.

4.2 Performance under Data Sparsity and Cold-Start (RQ2)

Experiment Design. To assess the generalizability of ILER/PLER methods, we simulate two extreme yet common educational scenarios on the ASSIST2017 and NIPS34 datasets: *data sparsity* and *cold-start*. For data sparsity, we evaluate performance at three levels of data availability (20%, 40%, and 80%). For the cold-start scenario, we test the models on a specialized subset where each student's historical interaction sequence is truncated to only five steps.

Experimental Results. Figure 3 compares GPP@10 under varying levels of data availability (panels a and b) and cold-start conditions (panels c and d). We highlight the following key observations:

Observation ❹: Strong gains can persist under substantial data sparsity. On NIPS34, reducing data availability from 80% to 20% decreases PKSD's GPP@10 from 0.473 to 0.419, retaining 88.6% of its score at 80% availability. In comparison, AC retains 57.6%, declining from 0.455 to 0.262. At 20% availability, eight of nine ILER methods also achieve positive gains on NIPS34, including DRER (0.228), whereas all nine yield negative gains on ASSIST2017. Thus, sparse interaction logs can still support positive estimated gains in both paradigms, with the attainable gains depending on the method and learning environment.

Observation ❺: Data sparsity and cold-start reveal different method strengths. On ASSIST2017, PKSD leads across the evaluated sparsity settings but falls behind KnowLP with five-interaction learner histories (GPP@10: 0.032 vs. 0.121). AC likewise shifts from positive gains across the sparsity settings to -0.103 under cold-start. On NIPS34, however, PKSD remains the strongest cold-start method (0.326), ahead of KnowLP (0.237). These results distinguish tolerance to dataset-level sparsity from effectiveness with limited learner histories, revealing different method strengths under the two forms of data scarcity.

4.3 Stability Against Label Noise (RQ3)

Experiment Design. We assess robustness to label noise on ASSIST2017 and NIPS34 by randomly altering correctness labels in student records at rates of 5%, 10%, 15%, and 20%, using the unperturbed records as the 0% baseline.

Experimental Results. Figure 4 reports GPP@10 across the evaluated noise levels, including the 0% baseline. We highlight the following observation:

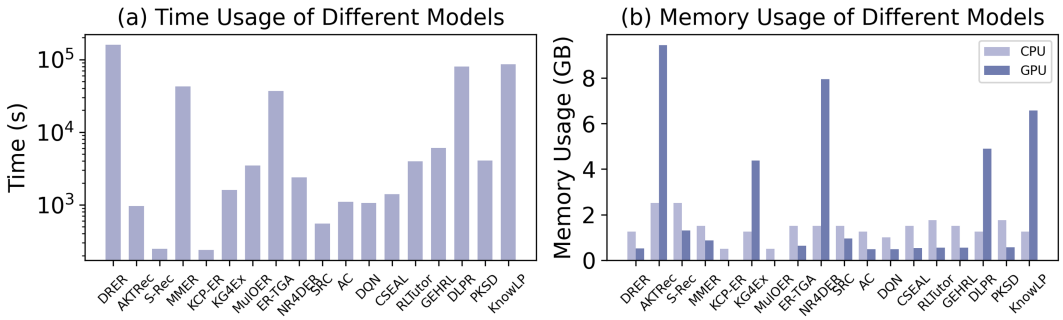

Figure 5: Time and memory usage comparison.

Observation ❻: Noise tolerance varies across methods and datasets. At 20% label noise, PKSD's GPP@10 decreases from 0.290 to 0.102 on ASSIST2017, retaining 35.2% of its baseline score. On NIPS34, it retains 90.1%, decreasing from 0.517 to 0.466. Eight of nine ILER methods also maintain positive gains on NIPS34, with DRER retaining 88.8% of its baseline score (0.286 to 0.254). The high retention achieved by PKSD and DRER on NIPS34 shows that methods from both paradigms can preserve most of their estimated gains under label noise.

4.4 Efficiency Analysis (RQ4)

Experiment Design. We assess the runtime and memory usage of ILER/PLER methods on ASSIST2017 using default hyperparameter settings.

Experimental Results. The computational overhead of each method, quantified by total time consumption and memory usage, is detailed in Figures 5 (a) and (b), respectively. Based on these findings, we highlight a striking observation:

Observation ❼: Extreme intra-paradigm scalability variance. Figure 5, computational overhead varies drastically even among RL-based models. Among the evaluated RL-based methods, DLPR and KnowLP incur substantial computational costs under the default settings on ASSIST2017. Their runtimes reach 79,540 and 86,400 seconds, respectively, equivalent to approximately 22.1 and 24.0 hours. Their GPU memory usage is also comparatively high, reaching 4.90 and 6.57 GB, respectively. Conversely, streamlined RL-based methods (*e.g.*, AC, DQN, and CSEAL) efficiently reach convergence in a fraction of the time ($\sim 10^3$s) while maintaining a lightweight memory footprint (< 1 GB). This reveals that the true computational bottleneck stems from coupling expensive structural graph searches with policy optimization, rather than the intrinsic sampling cost of the RL paradigm itself.

5 Conclusion and Future Directions

We present UniER, a unified benchmark for ILER and PLER that compares estimated knowledge gains under shared interaction budgets and different instructional objectives. Across nine datasets and 18 methods, ILER and PLER achieve the highest average scores at one and ten steps, respectively, while sensitivity to sparsity, cold-start, and label noise varies across methods and datasets. These findings highlight the importance of matching recommendation methods to the intended learning setting and motivate two research directions: **(1) Causal and interactive evaluation.** Future work should validate simulator-estimated gains against measured learning outcomes and develop evaluation protocols that better identify the causal effects of recommendations. **(2) Multi-objective evaluation.** Future benchmarks should explicitly characterize trade-offs among mastery, efficiency, and diversity under clearly specified educational goals.

Limitation. Our benchmark relies on offline evaluation with simulated feedback. It therefore cannot fully capture how models adapt to real-time changes in student behavior in actual tutoring scenarios. As an evolving benchmark, future versions will address these gaps by incorporating interactive simulation protocols, supported by the integration of more diverse, task-specific datasets.

AI USE STATEMENT

Generative AI tools were used to improve the readability and clarity of the manuscript and to provide suggestions on presentation and organization. They were not used to formulate the research hypotheses, design the methodology or experiments, or interpret the experimental results. All AI-assisted edits were reviewed and verified by the authors. The authors take full responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

To facilitate reproducibility, we specify the WCG metric and unified recommendation execution protocols in Section 3.1 and describe the student-simulator setting in Section 4.1. Dataset descriptions and statistics are provided in Appendix A, and the evaluated methods are described in Appendix B. Section 3.3 and Appendix C.2 document the hyperparameter search procedure, implementation details, and hardware specifications, with search spaces listed in Tables 4 and 5. Appendix D provides additional results and sensitivity analyses covering sequence construction protocols, interaction budgets, and evaluation simulators. We provide the UniER codebase, experimental configurations, and usage examples through our anonymous repository, with further information on resource accessibility and documentation summarized in Appendix C.3.

REFERENCES

Ghodai Abdelrahman and Qing Wang. Knowledge Tracing with Sequential Key-Value Memory Networks. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval (SIGIR 2019)*, pp. 175–184, 2019.

Xianyu Chen, Jian Shen, Wei Xia, Jiarui Jin, Yakun Song, Weinan Zhang, Weiwen Liu, Menghui Zhu, Ruiming Tang, Kai Dong, Dingyin Xia, and Yong Yu. Set-to-Sequence Ranking-based Concept-Aware Learning Path Recommendation. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-2023)*, pp. 5027–5035, 2023.

Xinghe Cheng, Xufang Zhou, Liangda Fang, Chaobo He, Yuyu Zhou, Weiqi Luo, Zhiguo Gong, and Quanlong Guan. NR4DER: Neural Re-ranking for Diversified Exercise Recommendation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-2025)*, pp. 1738–1747, 2025.

Xinghe Cheng, Zihan Zhang, Jiapu Wang, Liangda Fang, Chaobo He, Quanlong Guan, Shirui Pan, and Weiqi Luo. GraphRAG-Induced Dual Knowledge Structure Graphs for Personalized Learning Path Recommendation. In *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI-2026)*, pp. 14610–14620, 2026.

Aritra Ghosh, Neil Heffernan, and Andrew S.Lan. Context-Aware Attentive Knowledge Tracing. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD-2020)*, pp. 2330–2339, 2020.

Sophie Greenwood, Sudalakshmee Chiniah, and Nikhil Garg. User-item fairness tradeoffs in recommendations. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*, volume 37, pp. 114236–114288, 2024.

Quanlong Guan, Fang Xiao, Xinghe Cheng, Liangda Fang, Ziliang Chen, Guanliang Chen, and Weiqi Luo. KG4Ex: An Explainable Knowledge Graph-Based Approach for Exercise Recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM-2023)*, pp. 597–607, 2023.

Zhenya Huang, Qi Liu, Chengxiang Zhai, Yu Yin, Enhong Chen, Weibo Gao, and Guoping Hu. Exploring Multi-Objective Exercise Recommendations in Online Education Systems. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM-2019)*, pp. 1261–1270, 2019.

Vijay R. Konda and John N. Tsitsiklis. Actor-Critic Algorithms. In *Advances in Neural Information Processing Systems 12 (NIPS-1999)*, pp. 1008–1014, 1999.

Yoshiki Kubotani, Yoshihiro Fukuhara, and Shigeo Morishima. Rltutor: Reinforcement Learning Based Adaptive Tutoring System by Modeling Virtual Student with Fewer Interactions. *arXiv preprint arXiv:2108.00268*, 2021.

Qingyao Li, Wei Xia, Li'ang Yin, Jian Shen, Renting Rui, Weinan Zhang, Xianyu Chen, Ruiming Tang, and Yong Yu. Graph Enhanced Hierarchical Reinforcement Learning for Goal-oriented Learning Path Recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM-2023)*, pp. 1318–1327, 2023.

Qingyao Li, Wei Xia, Li'ang Yin, Jiarui Jin, and Yong Yu. Privileged Knowledge State Distillation for Reinforcement Learning-based Educational Path Recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD-2024)*, pp. 1621–1630, 2024.

Fei Liu, Xuegang Hu, Shuochen Liu, Chenyang Bu, and Le Wu. Meta Multi-Agent Exercise Recommendation: A Game Application Perspective. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD-2023)*, pp. 1441–1452, 2023a.

Guowei Liu, Meirui Ren, Longjiang Guo, Jin Li, and Miao Ma. Comprehensive Exercise Recommendation with Practicality, Generalizability, and Versatility in AI-Driven Education. *Information Processing & Management*, 62(3):104051, 2025.

Qi Liu, Shiwei Tong, Chuanren Liu, Hongke Zhao, Enhong Chen, Haiping Ma, and Shijin Wang. Exploiting Cognitive Structure for Adaptive Learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD-2019)*, pp. 627–635, 2019.

Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, Jiliang Tang, and Weiqi Luo. pyKT: A Python Library to Benchmark Deep Learning based Knowledge Tracing Models. In *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*, volume 35, pp. 18542–18555, 2022.

Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, and Weiqi Luo. SimpleKT: a Simple but Tough-to-Beat Baseline for Knowledge Tracing. In *Proceedings of the 11th International Conference on Learning Representations (ICLR-2023)*, pp. 20276–20287, 2023b.

Zitao Liu, Qiongqiong Liu, Teng Guo, Jiahao Chen, Shuyan Huang, Xiangyu Zhao, Jiliang Tang, Weiqi Luo, and Jian Weng. XES3G5M: A Knowledge Tracing Benchmark Dataset with Auxiliary Information. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)*, volume 36, pp. 32958–32970, 2023c.

Frederic Lord. A theory of test scores. *Psychometric monographs*, 1952.

Frederic M. Lord. *Applications of Item Response Theory to Practical Testing Problems*. Routledge, 2012.

Haiping Ma, Yi Yin, Ziwen Wang, Changqian Wang, Xiaoshan Yu, Shangshang Yang, and Xingyi Zhang. Learning patterns-guided data generation for knowledge tracing. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2025)*, pp. 2031–2041, 2025.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, volume 32, 2019.

Shubham Pateria, Budhitama Subagdja, Ah hwee Tan, and Chai Quek. Hierarchical Reinforcement Learning: A Comprehensive Survey. *ACM Computing Surveys (CSUR)*, 54(5):1–35, 2021.

Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas Guibas, and Jascha Sohl-Dickstein. Deep Knowledge Tracing. In *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, volume 28, 2015.

Dilina Chandika Rajapakse and Dietmar Jannach. Reassessing the Effectiveness of Reinforcement Learning based Recommender Systems for Sequential Recommendation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2025)*, pp. 3306–3314, 2025.

Yimeng Ren, Kun Liang, Yuhu Shang, and Yiying Zhang. MulOER-SAN: 2-Layer Multi-Objective Framework for Exercise Recommendation with Self-Attention Networks. *Knowledge-Based Systems*, 260:110117, 2023.

J Stamper, A Niculescu-Mizil, S Ritter, G Gordon, and K Koedinger. Algebra I 2005-2006 and Bridge to Algebra 2006-2007. *Development data sets from KDD Cup*, 2010.

Xueying Tang, Yunxiao Chen, Xiaoou Li, Jingchen Liu, and Zhiliang Ying. A Reinforcement Learning Approach to Personalized Learning Recommendation Systems. *British Journal of Mathematical and Statistical Psychology*, 72(1):108–135, 2019.

Zhengyang Wu, Ming Li, Yong Tang, and Qingyu Liang. Exercise Recommendation Based on Knowledge Concept Prediction. *Knowledge-Based Systems*, 210:106481, 2020.

Haotian Zhang, Shuanghong Shen, Bihan Xu, Zhenya Huang, Jinze Wu, Jing Sha, and Shijin Wang. Item-Difficulty-Aware Learning Path Recommendation: From a Real Walking Perspective. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD-2024)*, pp. 4167–4178, 2024a.

Kaike Zhang, Qi Cao, Yunfan Wu, Fei Sun, Huawei Shen, and Xueqi Cheng. Understanding and Improving Adversarial Collaborative Filtering for Robust Recommendation. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*, volume 37, pp. 120381–120417, 2024b.

Appendix

Table of Contents

A Detailed Description of Datasets in UniER **14**

B Detailed Description of Algorithms in UniER **14**

C Detailed Information of UniER **17**

C.1 Other Metrics 17

C.2 Additional Experimental Details 17

C.3 Reproducibility 20

D Additional Experimental Results and Analysis **21**

APPENDIX

A Detailed Description of Datasets in UniER

The description of datasets in UniER in given as follows:

- **ASSISTments 2009 / 2012 / 2017.** These are benchmark knowledge-tracing datasets collected from the ASSISTments online tutoring platform. The 2009 and 2012 datasets represent different school years, whereas the 2017 subset originates from the ASSISTments Data Mining Competition.
- **Algebra2005.** From the KDD Cup 2010 educational data mining challenge, this dataset contains detailed algebra exercise logs with knowledge concept annotations. It captures responses from middle school students practicing multi-step algebra problems.
- **Bridge2006.** Also a cornerstone of the KDD Cup 2010 challenge, the Bridge2006 dataset captures an extensive array of student interactions with algebra exercises. Its massive interaction volume offers the statistical robustness required to benchmark models on long-term performance prediction and complex temporal dynamics.
- **NIPS34.** Originating from the NeurIPS 2020 Education Challenge, this dataset comprises diagnostic math questions annotated with leaf-node skill labels for each item. Its unique structure is particularly effective for evaluating the robustness and generalization of knowledge tracing models when confronted with high question sparsity.
- **Junyi.** Collected from Junyi Academy, the Junyi dataset contains student problem-solving logs, exercise-level metadata, and annotated relationships among exercises.
- **Ednet.** A large-scale hierarchical educational dataset collected by the Santa AI tutoring platform. It records a rich variety of student activities, ranging from exercise solving to content browsing, over more than two years, making it one of the largest interactive educational system datasets publicly available.
- **XES3G5M.** This benchmark, recently released at NeurIPS 2023, is derived from a real-world K-12 online math learning environment. Beyond standardized correctness sequences, it encompasses an enriched set of auxiliary metadata, including hierarchical knowledge concepts and the textual attributes of questions."

Table 3: Statistics of the datasets used in our benchmark.

Dataset	#Interactions	#Students	#Exercises	#KC
ASSISTments2009	346,860	4,217	26,688	123
ASSISTments2012	6,123,270	46,674	179,999	265
ASSISTments2017	942,816	686	3,162	102
Algebra2005	813,661	575	1,084	112
Bridge2006	3,686,871	1,146	19,258	493
NIPS34	1,382,727	4,918	948	57
Junyi	25,925,992	247,606	722	41
Ednet	95,293,926	784,309	13,169	188
XES3G5M	5,549,635	18,066	7,652	865

B Detailed Description of Algorithms in UniER

The description of benchmarking algorithms in UniER is demonstrated as follows.

Type I: End-to-end.

- **DRER** Huang et al. (2019). The paper proposes DRER, a deep reinforcement learning framework for adaptive exercise recommendation. It introduces Exercise Q-Networks to capture students' exercising states and histories, and designs multi-objective rewards to jointly optimize Review & Explore, Smoothness, and Engagement.

- **AKTRec** Ghosh et al. (2020). The AKT model introduces a context-aware attentive knowledge tracing framework that uses two self-attentive encoders and a monotonic attention mechanism to capture a student's past exercise history. Building on AKT, AKTRec identifies knowledge concepts that a student has not yet mastered and recommends exercises corresponding to these gaps, enabling personalized learning.
- **SimpleKTRec** Liu et al. (2023b) SimpleKT introduces a lightweight attention-based knowledge tracing framework. It represents each exercise embedding as a combination of its associated knowledge concepts and leverages a simple dot-product attention mechanism to extract temporal knowledge states efficiently. Building on SimpleKT, SimpleKTRec identifies unmastered knowledge concepts for each student and recommends corresponding exercises, supporting personalized learning.
- **MMER** Liu et al. (2023a) The paper proposes MMER, a meta multi-agent reinforcement learning framework for exercise recommendation, where each knowledge concept is treated as an agent with competitive and cooperative interactions. A meta-training stage enables the model to quickly adapt to new student groups with few-shot data, while the multi-agent design allows efficient long-term learning of KCs.

Type II: Two-stage.

- **KCPER** Wu et al. (2020). KCP-ER adopts a two-stage architecture for exercise recommendation, comprising a Knowledge Concept Prediction Layer (KCPL) and an Exercise Set Filtering Layer (ESFL). KCPL predicts the coverage and mastery of knowledge concepts using LSTM for coverage and DKT for mastery, generating vectors that represent both student knowledge state and the likelihood of each concept appearing in future exercises. Based on these predictions, ESFL first filters exercises to ensure suitable difficulty and novelty, and then uses a simulated annealing-based list generator to maximize diversity, producing a recommended exercise list.
- **KG4Ex** Guan et al. (2023). KG4Ex employs a two-stage architecture consisting of a feature extraction module and a knowledge graph-based recommendation module. The feature extraction module uses LSTM and exponential functions to compute students' mastery levels, the probability of knowledge concepts appearing in the next exercise, and forgetting rates. Based on these features, the knowledge graph module constructs relationships among students, exercises, and knowledge concepts, learns embeddings with TransE, TransE-adv, or RotatE, and recommends suitable exercises while providing explainable recommendation reasons.
- **MulOER** Ren et al. (2023). MulOER-SAN adopts a two-layer multi-objective framework for exercise recommendation, where the bottom layer uses self-attention networks to predict knowledge concept coverage and dynamically trace students' mastery levels. The top layer filters exercises from the candidate subset, applying a chaotic sparrow search algorithm to maximize diversity while a smoothness factor ensures gradual difficulty progression.
- **ER-TGA** Liu et al. (2025). ER-TGA adopts a two-stage framework combining cognitive diagnosis and a tribal-alliance genetic algorithm for exercise filtering. The cognitive diagnosis layer identifies students' weak knowledge components and constructs a candidate exercise set, addressing cold-start scenarios. Building on this, TGA selects the optimal subset of exercises minimizing a multi-objective recommendation cost (accuracy, novelty, diversity, proximity, coverage, quantity, and volatility), producing personalized exercise recommendations aligned with learners' knowledge gaps.
- **NR4DER** Cheng et al. (2025). NR4DER employs a two-stage framework consisting of an exercise filter module and a neural re-ranking module. The exercise filter predicts students' knowledge concept mastery using an mLSTM-based predictor and constructs candidate exercises with appropriate difficulty, enhanced for inactive students via a student representation enhancer. Building on this, the neural re-ranking module integrates exercise relevance and learning pattern diversity, generating personalized and diversified exercise recommendation lists that align with individual students' learning pace.

Type III: Full-path.

- **SRC** Chen et al. (2023). SRC formulates learning path recommendation as a set-to-sequence task, where a concept-aware encoder captures correlations among candidate learning concepts and

produces global concept representations. An attention-based decoder sequentially generates the learning path, guided by predicted student mastery and inter-concept dependencies. A knowledge-tracing auxiliary module predicts mastery of each concept at every step, stabilizing training and improving the quality of generated learning paths.

Type IV: Step-by-step.

- **AC** Konda & Tsitsiklis (1999). The classical Actor-Critic Algorithms framework is applied to step-by-step learning path recommendation, where the critic estimates the expected value of the current student state and the actor selects the next exercise to optimize learning outcomes. Temporal-difference learning stabilizes the critic's estimation, while the actor updates policy parameters along the gradient of the expected reward, allowing dynamic adaptation to students' evolving knowledge. This framework generates personalized learning paths sequentially, making decisions at each step based on current student knowledge and prior interactions.
- **DQN** Tang et al. (2019). DQN applies a deep convolutional neural network to reinforcement learning, where the agent observes high-dimensional pixel input and sequentially selects actions to maximize future rewards. The network is trained using a variant of Q-learning with experience replay, which stabilizes learning by breaking correlations between sequential data and enables off-policy updates. At each step, the agent updates its estimate of the Q-values for possible actions based on observed rewards, effectively learning step-by-step control policies that adapt to evolving states and sequential dependencies.
- **CSEAL** Liu et al. (2019). CSEAL models the learning path recommendation as a Markov Decision Process, combining knowledge tracing, cognitive navigation, and an actor-critic recommender to generate exercises sequentially. Knowledge Tracing estimates evolving student knowledge mastery levels, while the Cognitive Navigation module selects candidate items according to the prerequisite graph to ensure logical consistency and reduce search space.
- **RLTutor** Kubotani et al. (2021). RLTutor models adaptive tutoring as a step-by-step sequential decision problem, where an Inner Model estimates the student's current knowledge state using the DAS3H knowledge tracing method, incorporating temporal memory dynamics and prior learning history. A reinforcement learning agent (PPO-based) optimizes the teaching strategy for the Inner Model, generating the next exercise recommendation while minimizing interactions with actual students.
- **GEHRL** Li et al. (2023). GEHRL employs a hierarchical reinforcement learning framework for goal-oriented learning path recommendation, with a high-level agent selecting sub-goals and a low-level agent recommending learning items sequentially to achieve each sub-goal. A graph-based candidate selector constrains the low-level agent's action space to ensure the path includes only goal-related exercises, improving learning efficiency.
- **DLPR** Zhang et al. (2024a). DLPR addresses learning path recommendation as a step-by-step sequential decision problem, generating paths dynamically based on real-time student interactions and mastery feedback. It employs a Difficulty-driven Hierarchical Reinforcement Learning (DHRL) framework with a high-level L-Agent selecting learning items and a low-level P-Agent choosing associated practice items, incorporating a knowledge state estimation module (DIMKT) to track evolving student knowledge. A communication mechanism between the agents ensures smooth coordination, adapts to item difficulty, and produces efficient, smooth learning paths that maximize learning effectiveness.
- **PKSD** Li et al. (2024). PKSD introduces a privileged knowledge state distillation framework for reinforcement learning-based educational path recommendation, where the student's simulator-derived knowledge state is used as privileged information during training. A two-stage process first trains a knowledge state encoder with the privileged state and then trains a GRU-based knowledge state adapter to estimate the latent encoding from regular exercise logs during inference. GEPKSD further incorporates a knowledge graph encoder, combining concept dependencies with learner states, enabling RL agents to generate personalized, structured learning paths that adapt effectively to diverse student knowledge levels.
- **KnowLP** Cheng et al. (2026). KnowLP constructs personalized learning paths using a step-by-step reinforcement learning framework guided by dual knowledge structure graphs capturing both prerequisite and similarity relations among knowledge concepts. Three specialized

agents—prerequisite, similarity, and difficulty—sequentially select knowledge concepts and corresponding exercises while simulating the discrimination learning process, ensuring that blocked or confusing concepts are addressed. A knowledge tracing module (DIMKT) dynamically tracks student mastery and exercise difficulty, allowing the system to adaptively generate coherent and effective learning paths that improve student progression.

C Detailed Information of UniER

C.1 Other Metrics

In UniER, we consider 10 metrics for evaluation. Their definitions are given as follows.

- **NDCG.** NDCG evaluates position-aware ranking quality by assigning logarithmically decaying rewards to successful hits, defined as accurately recommending unmastered exercises, ensuring that critical pedagogical content is prioritized at the top.
- **F1-score.** The F1-score provides a holistic evaluation of predictive robustness by striking a harmonic balance between the exactitude of the recommendations and the exhaustive capture of the student's genuine cognitive deficits through successful hits.
- **Recall.** A metric that measures the coverage capability of the top-k recommendation list for students' unmastered KCs. It refers to the ratio of the number of students' really unmastered KCs hit in the recommendation list to the total number of all real unmastered KCs of the student, and a higher value indicates more comprehensive coverage of students' unmastered KCs.
- **MRR.** A core ranking metric that quantifies the first occurrence position of a student's weak KC in the recommendation list. A higher value indicates an earlier occurrence of the weak KC, reflecting the top-ranking hit performance of the recommendation results.
- **Novelty.** Novelty assesses the system's ability to recommend unpopular exercises containing KCs that have rarely appeared or were poorly mastered in the student's historical interactions.
- **Proximity.** Proximity measures the absolute difference between the average difficulty of the recommended exercises and the specific cognitive difficulty level of the student.
- **Volatility.** Volatility measures the degree of fluctuation in difficulty levels between consecutive exercises within the recommended list.
- **Accuracy.** Accuracy evaluates the algorithm's capability to recommend exercises with appropriate difficulty levels by assessing the distance between the recommended exercises' difficulty and the expected difficulty.
- **MAP.** A comprehensive evaluation metric that assesses both the ranking quality and recall capability of the recommendation system. It measures the overall ranking accuracy of all a student's weak KCs in the recommendation list, taking both the precision and coverage of recommendation results into account.
- **Diversity.** Diversity is formally defined as the average degree of dissimilarity across all pairs of exercises included in the final recommendation list.

C.2 Additional Experimental Details

Implementation Details. To ensure a comprehensive evaluation and maintain fairness across a broad spectrum of models, we introduce UniER, an open-spource toolkit. This toolkit is built on top of Pytorch 2.8.0 Paszke et al. (2019). Due to the high computational cost of training DRER, if the training process exceeds 72 hours, we terminate the training and use the latest saved model checkpoint for evaluation. For AKTRec, SimpleKTRec, MulOER, and NR4DER, we use 50% of the data on ASSISTments2012, Bridge2006, and Junyi to avoid out-of-memory issues during training.

Hardware Specifications. All our experiments were carried out on a Linux server with an Intel(R) Xeon(R) Silver 4410Y CPU, 256GB of RAM, and four NVIDIA GeForce RTX 4090 GPUs (24GB VRAM each).

Table 4: Hyper-parameter search space of ILER methods.

Algorithm	Hyper-parameter	Search Space
General Settings	DKT embedding size	32, 64, 128, 256
	DKT hidden size	64, 128, 256, 512
	DKT layers	1, 2, 3, 4
	DKT dropout rate	1e-5, 1e-4, 1e-3, e-2
	Episodes	2000, 5000, 10000
AKTRec	Dropout rate	0.05, 0.1, 0.2, 0.3
	Model dimension	64, 128, 256
	Feed-forward dimension	128, 256, 512
	Number of attention heads	2,4,8
	Number of blocks	1, 2, 4
	Learning rate	1e-5, 1e-4, 1e-3
SimpleKTRec	Embedding size	32, 64, 128, 256
	hidden size	32, 64, 128, 256
	Learning rate	1e-5, 1e-4, 1e-3
	Dropout rate	0.05, 0.1, 0.3, 0.5
	Number of blocks	1, 2, 4
	Number of attention heads	4, 8
DRER	exercise embedding size	30, 50, 70, 100
	KC embedding size	5, 10, 15, 20
	content embedding size	50, 100, 150, 200
	hidden size	64, 128, 256
	Batch size	16, 32, 64
MMER	γ	0.7, 0.8, 0.85, 0.9
	Hidden size	200
	Learning rate	1e-4,1e-3,5e-3
	Fine-tuning ratio	0.1,0.2,0.3
KCPER	δ	0.5, 0.6, 0.7, 0.8
MULOER	Embedding size	64, 128, 256
	Number of attention heads	2, 4, 6, 8
	Learning rate	1e-5, 1e-4, e-3
	Dropout rate	0.1, 0.2, 0.3
	N	20, 50, 100
	T	100, 200, 300
	d	10, 20, 30
KG4Ex	δ_1	0.5, 0.6, 0.7, 0.8
	δ_2	0.5, 0.6, 0.7, 0.8
	LSTM hidden size	64, 128, 256
	Learning rate	1e-4, 1e-3, 1e-2
	Batch size	32, 64, 128, 256
ER-TGA	Weak KC threshold	0.35, 0.4, 0.45, 0.5, 0.55, 0.6
	Crossover probability	0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9
	Mutation probability	1e-6, 1e-5, 1e-4, 1e-3, 1e-2, 1e-1
	Population size	30, 40, 50, 60, 70, 80
	Tribe size	6, 12, 24, 36, 48, 60
	Max generation	10, 30, 50, 100, 150
NR4DER	β	0.4, 0.6, 0.8, 1.0
	γ	0.1, 0.3, 0.5, 0.7, 1.0
	δ	0.5, 0.6, 0.7, 0.8
	Number of attention heads	2, 4, 6, 8
	Batch size	16, 32, 64, 128

Table 5: Hyper-parameter search space of PLER methods.

Algorithm	Hyper-parameter	Search Space
AC	Learning rate	1e-5, 1e-4, 1e-3
	Reward discount factor	0.85, 0.9, 0.95, 0.98
	Batch size	16, 32, 64
DQN	Learning rate	1e-5, 1e-4, 1e-3
	Reward discount factor	0.85, 0.9, 0.95, 0.98
	Batch size	16, 32, 64
CSEAL	LSTM dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
	Learning rate	1e-5, 1e-4, 1e-3
	Reward discount factor	0.85, 0.9, 0.95, 0.98
	Batch size	16, 32, 64
RLtutor	Learning rate	1e-5, 1e-4, 1e-3
	Reward discount factor	0.85, 0.9, 0.95, 0.98
	Batch size	16, 32, 64
	Clipping parameter	0.1, 0.2, 0.3
	Value function coefficient	0.2, 0.4, 0.5, 0.6, 0.8
SRC	Embedding size	32, 64, 128, 256
	Batch size	32, 64, 128, 256
	Dropout rate	0, 0.1, 0.2, 0.3, 0.4, 0.5
	Learning rate	1e-5, 1e-4, 1e-3
GEHRL	Learning rate	1e-5, 1e-4, 1e-3
	α	[0, 1.0]
	β	[0, 1.0]
DLPR	DIMKT embedding size	64, 128, 256
	DIMKT difficulty level	10, 50, 100
	DIMKT learning rate	1e-4, 1e-3, 2e-3
	Answering correctly probability	[0.4, 0.6]
PKSD	Learning rate	1e-5, 1e-4, 1e-3
	Reward discount factor	0.85, 0.9, 0.95, 0.98
	Batch size	16, 32, 64
KnowLP	DIMKT embedding size	64, 128, 256
	DIMKT difficulty level	10, 50, 100
	DIMKT learning rate	1e-4, 1e-3, 2e-3
	Answering correctly probability	[0.4, 0.6]
	Improvement threshold	1e-4, 1e-3, 1e-2

Table 6: Comparison of ILER sequence construction strategies in terms of TGA@10 and GPP@10. The top, middle, and bottom blocks correspond to DKT-guided adaptive sequencing, descending-score ordering, and random ordering, respectively. The top three scores in each column across all method-strategy combinations are highlighted by 1st, 2nd, and 3rd. "Avg." denotes the mean of TGA@10 and GPP@10 across all nine datasets.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M		Avg.
	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	T&G
DRER	-0.208	-0.183	0.055	0.033	-0.003	0.002	-0.487	-0.400	0.083	0.122	0.365	0.286	0.222	0.084	0.023
AKTRec	-0.083	-0.051	0.023	0.025	0.000	-0.001	-0.183	-0.016	0.090	0.137	0.315	0.188	0.224	0.089	0.057
S-KTRec	-0.087	-0.052	0.030	0.020	-0.001	-0.000	-0.171	-0.020	0.091	0.133	0.314	0.185	0.222	0.091	0.059
MMER	-0.141	-0.126	0.085	0.042	-0.009	-0.005	-0.014	0.044	0.082	0.043	0.338	0.247	0.249	0.032	0.070
KCP-ER	-0.174	-0.121	0.154	0.023	-0.002	0.001	-0.224	0.047	0.104	0.128	0.310	0.187	0.224	0.087	0.070
KG4Ex	-0.182	-0.118	0.116	0.057	-0.002	-0.002	0.061	0.060	0.095	0.136	0.282	0.140	0.236	0.130	0.087
MulOER	-0.083	-0.065	0.065	0.034	-0.006	-0.003	0.030	0.087	0.094	0.131	0.271	0.162	0.228	0.086	0.086
ER-TGA	-0.114	-0.083	0.104	0.052	0.002	0.006	0.015	0.027	0.091	0.142	-0.063	0.003	0.248	0.130	0.009
NR4DER	-0.144	-0.090	0.104	0.047	0.002	0.002	-0.117	0.039	0.095	0.145	0.314	0.183	0.228	0.099	0.074
DRER	-0.104	-0.090	-0.027	0.024	-0.005	-0.001	-0.162	-0.065	0.014	-0.098	0.050	0.029	0.124	0.003	-0.019
AKTRec	-0.123	-0.072	0.049	0.023	0.000	0.000	-0.483	-0.155	-0.067	-0.145	0.018	-0.032	0.134	-0.017	-0.057
S-KTRec	-0.123	-0.067	0.023	0.036	0.000	0.000	-0.454	-0.168	-0.082	-0.133	0.024	-0.007	0.125	-0.018	-0.049
MMER	-0.128	-0.121	0.063	0.049	-0.008	-0.004	-0.229	-0.121	0.062	-0.005	0.032	0.045	0.190	0.019	-0.016
KCP-ER	-0.160	-0.104	0.046	0.032	-0.369	0.001	0.143	-0.103	0.064	0.054	0.146	0.041	0.168	0.020	0.013
KG4Ex	-0.208	-0.116	0.116	0.046	-0.001	-0.002	-0.119	-0.128	0.011	-0.061	-0.087	-0.087	0.085	-0.075	-0.054
MulOER	-0.114	-0.088	0.065	0.034	-0.005	-0.002	-0.086	-0.059	0.027	0.009	0.045	0.008	0.195	0.054	0.016
ER-TGA	-0.117	-0.097	0.107	0.051	-0.003	0.004	-0.101	-0.125	0.068	0.118	-0.190	-0.124	0.063	-0.021	-0.035
NR4DER	-0.139	-0.089	0.088	0.042	0.002	0.003	-0.352	-0.153	0.019	-0.179	0.084	-0.020	0.162	-0.018	-0.033
DRER	-0.177	-0.163	0.015	0.030	-0.003	0.002	-0.279	-0.145	0.081	0.002	0.000	-0.026	0.113	-0.025	-0.031
AKTRec	-0.105	-0.067	-0.006	0.012	0.002	0.000	-0.438	-0.164	0.043	0.012	0.018	-0.003	0.136	-0.043	-0.039
S-KTRec	-0.118	-0.073	0.005	0.017	-0.001	0.000	-0.429	-0.158	0.051	0.008	0.029	-0.008	0.140	-0.009	-0.032
MMER	-0.133	-0.114	-0.004	0.014	-0.007	-0.003	-0.127	-0.101	0.057	0.017	0.143	0.095	0.151	-0.005	0.007
KCP-ER	-0.141	-0.090	0.055	0.030	-0.001	0.001	-0.378	-0.113	0.061	0.044	0.123	0.020	0.174	0.021	0.000
KG4Ex	-0.196	-0.127	0.117	0.045	-0.002	-0.001	-0.194	-0.131	0.042	-0.018	-0.046	-0.081	0.069	-0.045	-0.040
MulOER	-0.125	-0.103	0.074	0.041	-0.006	-0.004	-0.063	-0.091	0.029	-0.003	0.046	0.001	0.144	0.018	0.006
ER-TGA	-0.118	-0.094	0.117	0.054	0.002	0.005	-0.094	-0.113	0.047	-0.020	-0.117	-0.104	0.002	-0.142	-0.062
NR4DER	-0.137	-0.093	0.095	0.038	0.003	0.002	-0.315	-0.132	0.021	-0.179	0.099	0.002	0.174	0.021	-0.022

Hyperparameter Settings. Tables 4 and 5 provide a comprehensive list of all hyperparameters used in our random search complete with their search spaces. For the design of the default hyperparameters please refer to our code base in https://anonymous.4open.science/r/UniER-4532/.

This experimental setup enables us to measure the efficiency of each method robustly, providing a clear reflection of their practical viability in real-world applications.

C.3 Reproducibility

Ensuring the reproducibility of experimental results is a core principle of UniER. Below, we outline the measures we have taken to achieve this:

Accessibility. All datasets, algorithm implementations, and experimental configurations are freely accessible through our open-source project at https://anonymous.4open.science/r/UniER-4532/. No special requests or permissions are needed to access the resources.

Datasets. Our datasets are publicly available. All these datasets are permitted by their authors for academic use and contain no personally identifiable information or offensive content.

Documentation and Usage. We provide comprehensive documentation to facilitate easy use of our library. To facilitate seamless reproducibility, we provide comprehensive examples that detail the execution of our codebase, including specific configurations for data, methods, and GPU resource arguments. The code includes thorough comments to enhance readability.

Code Maintenance. We are committed to the regular maintenance and iterative refinement of our codebase, actively addressing user feedback and integrating community contributions. Furthermore, we implement rigorous version control protocols to ensure sustained reproducibility throughout the lifecycle of code maintenance.

Table 7: Comparison of ILER and PLER methods in terms of TGA@10 and GPP@10 using SimpleKT as an alternative student simulator. The top three scores in each column are highlighted by 1st, 2nd, and 3rd. "Avg." denotes the mean of TGA@10 and GPP@10 across all nine datasets.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M		Avg.
	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	TGA	GPP	T&G
DRER	-0.025	-0.032	-0.188	-0.170	-0.042	-0.135	0.265	0.067	0.030	0.140	0.091	0.095	0.060	0.061	0.022
AKTRec	0.000	-0.011	-0.207	-0.187	0.062	-0.204	0.138	-0.019	0.042	0.057	0.000	0.008	0.070	0.077	0.007
S-KTRec	0.000	-0.009	-0.188	-0.191	0.059	-0.209	0.125	-0.025	0.044	0.075	0.000	0.005	0.062	0.064	0.005
MMER	-0.022	-0.025	-0.226	-0.176	-0.037	-0.127	0.267	0.194	0.035	0.110	0.140	0.082	0.080	0.004	0.040
KCP-ER	0.000	0.004	-0.109	-0.229	0.073	-0.082	0.282	0.185	0.048	0.330	0.042	0.043	0.082	0.118	0.074
KG4Ex	-0.078	-0.054	-0.188	-0.434	0.105	-0.148	0.169	0.212	0.039	0.097	-0.044	-0.067	0.038	-0.009	-0.025
MulOER	-0.006	0.027	-0.160	-0.212	-0.001	-0.103	0.258	0.144	0.038	0.218	0.100	0.049	0.121	0.039	0.060
ER-TGA	0.008	0.010	-0.183	-0.183	-0.039	-0.103	0.265	0.242	0.044	0.147	0.048	0.054	0.046	0.140	0.029
NR4DER	-0.005	-0.013	-0.133	-0.129	0.062	-0.114	0.186	0.024	0.015	-0.011	0.028	0.023	0.075	0.112	0.028
SRC	0.043	0.014	0.432	0.060	0.112	-0.602	0.446	0.593	0.002	0.201	0.128	0.092	0.190	0.469	0.155
AC	0.098	0.118	-0.064	-0.255	0.304	-0.051	0.729	-0.303	0.443	0.324	0.120	0.064	0.626	0.014	0.140
DQN	0.090	0.120	-0.045	-0.224	0.462	0.029	0.723	-0.340	0.440	0.342	0.091	0.082	0.609	-0.054	0.165
CSEAL	0.180	0.368	-0.116	-0.202	0.256	-0.061	0.538	0.075	0.413	0.287	0.111	0.092	0.560	0.069	0.188
RLTutor	0.095	0.114	-0.042	-0.262	0.451	-0.016	0.724	0.156	0.433	0.348	0.136	0.045	0.619	-0.543	0.136
GEHRL	0.094	0.090	-0.032	-0.039	0.256	0.260	0.791	0.797	0.232	0.217	-0.111	-0.091	0.553	0.501	0.266
DLPR	0.044	0.006	0.119	-0.082	0.202	-0.014	0.541	0.136	0.097	-0.037	0.091	0.169	0.549	0.240	0.142
PKSD	0.075	0.133	-0.032	-0.222	0.131	-0.045	0.766	0.150	0.394	0.319	0.217	0.104	0.670	0.031	0.181
KnowLP	0.059	0.127	0.125	-0.075	0.208	-0.007	0.647	0.143	0.241	0.153	0.132	0.180	0.622	0.237	0.190

With these measures, UniER aims to foster transparency, accessibility, and collaboration within the research community.

D Additional Experimental Results and Analysis

In this section, we present additional experimental results to provide an exhaustively aligned perspective on the benchmarked models, encompassing virtually all conventional metrics alongside our comprehensive WCG evaluations across varying recommendation lengths.

Sensitivity of ILER to sequence construction. Table 6 shows that DKT-guided adaptive sequencing achieves the highest average score among the three strategies for each of the nine ILER methods. For example, KG4Ex achieves an average score of 0.087 under adaptive sequencing, compared with -0.054 and -0.040 under descending-score and random ordering, respectively. However, this aggregate advantage does not hold uniformly across datasets: DRER's GPP@10 on Ednet is -0.400 under adaptive sequencing, compared with -0.065 under descending-score ordering. Moreover, random ordering achieves higher average scores than descending-score ordering for five of the nine methods, indicating that following recommendation-score order does not consistently yield better multi-step gains. Overall, combining mastery-guided exercise selection with iterative recommendation updates yields higher average multi-step gains than the two static ordering strategies across all evaluated ILER methods.

Sensitivity to the evaluation simulator. To examine whether the observed ten-step advantage of PLER persists under an alternative evaluation simulator, we retain the recommendation models from the main experiments and replace DKT with SimpleKT only during evaluation. All recommendation model parameters remain fixed, without retraining or adaptation to SimpleKT. Table 7 reports the resulting TGA@10 and GPP@10 scores.

The average ten-step advantage of PLER remains evident under SimpleKT. All nine PLER methods achieve higher average scores than every ILER method, with scores ranging from 0.136 to 0.266, compared with -0.025 to 0.074 for ILER. GEHRL achieves the highest average score (0.266), followed by KnowLP (0.190) and CSEAL (0.188), whereas KCP-ER is the strongest ILER method (0.074). Even the lowest-scoring PLER method, RLTutor (0.136), exceeds KCP-ER by 0.062. These results extend the aggregate advantage observed in the main experiments to an evaluator different from the simulator used for policy training.

Table 8: Comparative experiments across nine datasets.

Dataset	Model	N@3	N@5	N@10	F@3	F@5	F@10	R@3	R@5	R@10	M@3	M@5	M@10
Assist09	DRER	0.549	0.549	0.502	0.199	0.295	0.385	0.123	0.206	0.357	0.232	0.240	0.241
	MMER	0.456	0.506	0.532	0.506	0.507	0.493	0.339	0.340	0.328	0.448	0.471	0.480
	AKTRec	0.585	0.587	0.583	0.436	0.354	0.241	0.367	0.274	0.165	0.519	0.514	0.522
	SimpleKTRec	0.576	0.578	0.576	0.425	0.352	0.241	0.354	0.271	0.164	0.505	0.505	0.514
	NR4DER	0.763	0.750	0.741	0.850	0.816	0.583	0.727	0.708	0.697	0.581	0.582	0.582
	ERTGA	0.256	0.262	0.264	0.337	0.410	0.412	0.232	0.235	0.236	0.274	0.280	0.281
	KCPER	0.334	0.348	0.362	0.308	0.321	0.334	0.265	0.269	0.277	0.270	0.284	0.292
	MulSUN-OER	0.312	0.326	0.329	0.268	0.265	0.222	0.281	0.290	0.291	0.304	0.308	0.309
	KG4Ex	0.422	0.453	0.446	0.341	0.346	0.341	0.288	0.283	0.276	0.379	0.395	0.402
Assist12	DRER	0.322	0.321	0.289	0.166	0.241	0.417	0.114	0.214	0.324	0.203	0.220	0.231
	MMER	0.509	0.565	0.605	0.500	0.499	0.481	0.318	0.314	0.310	0.483	0.515	0.532
	AKTRec	0.559	0.571	0.573	0.444	0.393	0.302	0.369	0.307	0.217	0.441	0.445	0.462
	SimpleKTRec	0.561	0.572	0.575	0.443	0.391	0.301	0.369	0.305	0.216	0.442	0.444	0.461
	NR4DER	0.612	0.633	0.635	0.611	0.589	0.605	0.377	0.368	0.345	0.309	0.313	0.319
	ERTGA	0.342	0.359	0.377	0.260	0.262	0.265	0.219	0.217	0.219	0.279	0.284	0.286
	KCPER	0.298	0.352	0.405	0.256	0.272	0.285	0.206	0.211	0.216	0.260	0.283	0.296
	MulSUN-OER	0.284	0.309	0.283	0.283	0.285	0.286	0.245	0.244	0.242	0.306	0.313	0.316
	KG4Ex	0.423	0.473	0.547	0.331	0.332	0.328	0.266	0.257	0.247	0.358	0.369	0.371
Assist17	DRER	0.326	0.389	0.470	0.140	0.239	0.440	0.091	0.173	0.407	0.166	0.197	0.211
	MMER	0.476	0.529	0.557	0.457	0.479	0.481	0.271	0.287	0.293	0.746	0.758	0.761
	AKTRec	0.438	0.475	0.507	0.320	0.294	0.239	0.237	0.201	0.148	0.268	0.294	0.338
	SimpleKTRec	0.461	0.503	0.524	0.332	0.305	0.241	0.249	0.209	0.150	0.291	0.317	0.360
	NR4DER	0.539	0.475	0.437	0.707	0.686	0.579	0.441	0.372	0.339	0.692	0.695	0.697
	ERTGA	0.152	0.156	0.159	0.210	0.258	0.259	0.114	0.113	0.113	0.177	0.190	0.196
	KCPER	0.223	0.282	0.334	0.171	0.192	0.197	0.121	0.127	0.123	0.170	0.208	0.238
	MulSUN-OER	0.238	0.273	0.299	0.174	0.173	0.174	0.123	0.114	0.109	0.215	0.237	0.256
	KG4Ex	0.222	0.269	0.255	0.163	0.170	0.189	0.118	0.113	0.121	0.209	0.245	0.262
Alge05	DRER	0.476	0.488	0.492	0.181	0.294	0.365	0.127	0.213	0.355	0.307	0.332	0.343
	MMER	0.488	0.541	0.588	0.492	0.493	0.478	0.312	0.316	0.305	0.532	0.546	0.559
	AKTRec	0.942	0.942	0.934	0.885	0.831	0.674	0.855	0.772	0.561	0.886	0.860	0.847
	SimpleKTRec	0.949	0.948	0.939	0.894	0.832	0.676	0.864	0.772	0.563	0.895	0.865	0.856
	NR4DER	0.662	0.634	0.606	0.795	0.793	0.787	0.567	0.531	0.501	0.746	0.746	0.746
	ERTGA	0.401	0.411	0.427	0.498	0.560	0.657	0.335	0.333	0.343	0.430	0.446	0.453
	KCPER	0.358	0.412	0.527	0.294	0.319	0.337	0.229	0.238	0.242	0.319	0.351	0.368
	MulSUN-OER	0.416	0.516	0.592	0.359	0.388	0.389	0.283	0.291	0.280	0.366	0.400	0.410
	KG4Ex	0.599	0.607	0.666	0.550	0.542	0.535	0.459	0.445	0.413	0.521	0.548	0.562
Brid06	DRER	0.404	0.405	0.380	0.202	0.277	0.340	0.092	0.180	0.378	0.262	0.284	0.295
	MMER	0.461	0.509	0.548	0.483	0.499	0.505	0.281	0.289	0.282	0.364	0.386	0.402
	AKTRec	0.569	0.589	0.595	0.447	0.390	0.285	0.357	0.285	0.186	0.410	0.428	0.464
	SimpleKTRec	0.570	0.590	0.595	0.443	0.388	0.285	0.353	0.284	0.186	0.406	0.426	0.462
	NR4DER	0.512	0.501	0.522	0.421	0.461	0.404	0.401	0.385	0.386	0.357	0.370	0.378
	ERTGA	0.316	0.278	0.281	0.505	0.605	0.607	0.250	0.215	0.217	0.363	0.363	0.365
	KCPER	0.270	0.280	0.359	0.195	0.199	0.199	0.149	0.148	0.148	0.203	0.214	0.215
	MulSUN-OER	0.225	0.321	0.255	0.152	0.169	0.173	0.108	0.111	0.112	0.174	0.203	0.213
	KG4Ex	0.365	0.415	0.323	0.297	0.316	0.306	0.229	0.234	0.222	0.323	0.334	0.351
Ednet	DRER	0.350	0.508	0.563	0.155	0.366	0.510	0.100	0.280	0.510	0.592	0.610	0.618
	MMER	0.527	0.582	0.614	0.528	0.535	0.517	0.403	0.408	0.394	0.444	0.462	0.467
	AKTRec	0.959	0.958	0.947	0.915	0.874	0.769	0.883	0.822	0.677	0.889	0.851	0.778
	SimpleKTRec	0.963	0.959	0.949	0.914	0.874	0.770	0.882	0.823	0.677	0.890	0.854	0.780
	NR4DER	0.655	0.654	0.656	0.760	0.782	0.778	0.564	0.543	0.534	0.743	0.762	0.766
	ERTGA	0.459	0.476	0.491	0.552	0.587	0.619	0.371	0.372	0.375	0.512	0.538	0.551
	KCPER	0.434	0.493	0.530	0.373	0.417	0.460	0.297	0.319	0.342	0.387	0.424	0.448
	MulSUN-OER	0.643	0.677	0.690	0.555	0.569	0.567	0.484	0.488	0.475	0.595	0.611	0.617
	KG4Ex	0.488	0.531	0.584	0.422	0.316	0.306	0.356	0.356	0.360	0.467	0.503	0.522
Junyi	DRER	0.365	0.390	0.403	0.183	0.311	0.460	0.155	0.301	0.436	0.300	0.309	0.326
	MMER	0.319	0.367	0.424	0.327	0.319	0.304	0.196	0.190	0.179	0.293	0.320	0.342
	AKTRec	0.633	0.641	0.640	0.411	0.405	0.376	0.477	0.451	0.405	0.423	0.451	0.436
	SimpleKTRec	0.652	0.661	0.658	0.428	0.406	0.382	0.481	0.448	0.413	0.434	0.422	0.420
	NR4DER	0.762	0.807	0.785	0.469	0.432	0.410	0.409	0.362	0.337	0.582	0.583	0.584
	ERTGA	0.296	0.390	0.551	0.235	0.244	0.247	0.194	0.197	0.199	0.255	0.267	0.268
	KCPER	0.475	0.529	0.498	0.323	0.361	0.382	0.393	0.490	0.551	0.381	0.396	0.400
	MulSUN-OER	0.542	0.597	0.605	0.362	0.378	0.375	0.367	0.369	0.356	0.447	0.473	0.502
	KG4Ex	0.487	0.533	0.541	0.354	0.389	0.395	0.324	0.330	0.323	0.533	0.549	0.558
NIPS34	DRER	0.821	0.811	0.828	0.301	0.430	0.611	0.194	0.318	0.568	0.558	0.589	0.594
	MMER	0.644	0.676	0.693	0.645	0.640	0.626	0.476	0.471	0.456	0.591	0.610	0.619
	AKTRec	0.852	0.851	0.843	0.765	0.741	0.677	0.683	0.639	0.556	0.646	0.612	0.625
	SimpleKTRec	0.850	0.850	0.843	0.769	0.742	0.678	0.687	0.641	0.558	0.648	0.615	0.627
	NR4DER	0.958	0.954	0.954	0.960	0.959	0.958	0.949	0.943	0.942	0.957	0.957	0.957
	ERTGA	0.576	0.589	0.594	0.681	0.712	0.695	0.508	0.513	0.518	0.607	0.616	0.617
	KCPER	0.600	0.645	0.665	0.521	0.556	0.594	0.427	0.445	0.470	0.497	0.530	0.543
	MulSUN-OER	0.584	0.649	0.713	0.488	0.501	0.508	0.395	0.393	0.388	0.525	0.549	0.556
	KG4Ex	0.754	0.779	0.794	0.663	0.661	0.653	0.569	0.545	0.523	0.715	0.731	0.735
XES3G5M	DRER	0.355	0.387	0.393	0.198	0.285	0.422	0.083	0.175	0.293	0.244	0.266	0.277
	MMER	0.340	0.391	0.444	0.342	0.340	0.338	0.185	0.228	0.345	0.309	0.338	0.361
	AKTRec	0.793	0.798	0.789	0.684	0.639	0.569	0.613	0.543	0.451	0.783	0.792	0.798
	SimpleKTRec	0.876	0.874	0.860	0.709	0.740	0.651	0.737	0.661	0.543	0.870	0.875	0.878
	NR4DER	0.849	0.854	0.871	0.739	0.675	0.548	0.673	0.590	0.443	0.834	0.835	0.836
	ERTGA	0.386	0.442	0.592	0.306	0.304	0.312	0.256	0.241	0.240	0.331	0.347	0.355
	KCPER	0.346	0.410	0.468	0.280	0.305	0.319	0.211	0.218	0.218	0.303	0.337	0.357
	MulSUN-OER	0.267	0.337	0.419	0.190	0.210	0.224	0.134	0.137	0.139	0.234	0.270	0.294
	KG4Ex	0.401	0.469	0.532	0.187	0.229	0.265	0.158	0.199	0.281	0.359	0.396	0.418

Table 9: Comparative experiments across nine datasets. Gray cells marked with "-" indicate that the corresponding metric is not applicable to the method.

Dataset	Model	ACC@m	ACC@s	Nov@m	Nov@s	Div@m	Div@s	Vol@m	Vol@s	Pro@m	Pro@s	MAP@3	MAP@10
Assist09	DRER	0.792	0.112	0.951	0.093	0.732	0.002	0.961	0.014	0.954	0.036	0.149	0.126
	MMER	0.768	0.148	0.978	0.061	0.394	0.318	0.965	0.041	0.835	0.158	0.297	0.448
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.579	0.584
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.570	0.575
	NR4DER	0.742	0.086	0.977	0.049	0.925	0.051	0.818	0.028	0.881	0.097	0.513	0.498
	ERTGA	0.504	0.082	0.999	0.001	0.508	0.158	0.992	0.019	0.780	0.137	0.159	0.128
	KCPER	0.689	0.063	0.977	0.038	0.608	0.066	0.983	0.013	0.896	0.082	0.192	0.195
	MulSUN-OER	0.787	0.100	0.997	0.006	0.804	0.166	0.914	0.044	0.861	0.094	0.184	0.161
	KG4Ex	0.794	0.119	0.720	0.246	0.664	0.235	0.971	0.034	0.872	0.091	0.315	0.340
Assist12	DRER	0.785	0.062	0.947	0.021	0.711	0.036	0.965	0.022	0.930	0.069	0.233	0.249
	MMER	0.852	0.090	0.983	0.053	0.609	0.290	0.970	0.028	0.920	0.074	0.368	0.741
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.538	0.551
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.542	0.555
	NR4DER	0.779	0.031	0.944	0.024	0.821	0.054	0.886	0.022	0.899	0.072	0.600	0.620
	ERTGA	0.872	0.066	1.000	0.000	0.934	0.003	0.907	0.032	0.833	0.076	0.155	0.118
	KCPER	0.841	0.055	0.986	0.016	0.738	0.060	0.896	0.026	0.907	0.074	0.168	0.189
	MulSUN-OER	0.821	0.106	0.997	0.007	0.723	0.200	0.962	0.047	0.900	0.076	0.174	0.143
	KG4Ex	0.764	0.091	0.998	0.004	0.938	0.009	1.000	0.000	0.902	0.077	0.171	0.102
Assist17	DRER	0.556	0.125	0.970	0.013	0.865	0.000	0.921	0.024	0.913	0.039	0.084	0.102
	MMER	0.578	0.145	0.977	0.015	0.590	0.243	0.959	0.035	0.899	0.074	0.584	0.955
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.402	0.443
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.432	0.472
	NR4DER	0.603	0.121	0.976	0.019	0.868	0.176	0.869	0.023	0.924	0.065	0.544	0.571
	ERTGA	0.509	0.132	0.974	0.022	0.845	0.133	0.925	0.045	0.949	0.044	0.095	0.103
	KCPER	0.666	0.103	0.980	0.013	0.679	0.060	0.872	0.032	0.915	0.061	0.100	0.166
	MulSUN-OER	0.895	0.060	0.991	0.007	0.892	0.051	0.908	0.038	0.944	0.040	0.100	0.112
	KG4Ex	0.514	0.110	0.939	0.057	0.611	0.138	0.896	0.015	0.895	0.077	0.099	0.136
Alge05	DRER	0.603	0.043	0.963	0.019	0.654	0.099	0.912	0.038	0.924	0.036	0.322	0.360
	MMER	0.719	0.115	0.953	0.031	0.792	0.096	0.850	0.074	0.804	0.134	0.427	0.723
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.941	0.946
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.949	0.952
	NR4DER	0.812	0.063	0.959	0.032	0.694	0.203	0.843	0.039	0.930	0.049	0.482	0.359
	ERTGA	0.614	0.019	0.986	0.019	0.624	0.228	0.863	0.063	0.882	0.066	0.257	0.159
	KCPER	0.721	0.014	0.992	0.004	0.639	0.016	0.874	0.047	0.922	0.055	0.165	0.136
	MulSUN-OER	0.911	0.086	0.997	0.034	0.598	0.171	0.934	0.062	0.866	0.097	0.200	0.135
	KG4Ex	0.512	0.215	0.929	0.051	0.509	0.105	0.984	0.017	0.925	0.055	0.369	0.323
Brid06	DRER	0.812	0.071	0.955	0.013	0.725	0.027	0.942	0.035	0.935	0.037	0.078	0.104
	MMER	0.828	0.074	0.992	0.014	0.736	0.217	0.953	0.037	0.926	0.073	0.238	0.391
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.538	0.558
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.540	0.560
	NR4DER	0.807	0.081	0.936	0.018	0.869	0.060	0.905	0.038	0.911	0.095	0.422	0.447
	ERTGA	0.545	0.096	0.997	0.016	0.669	0.081	0.903	0.047	0.949	0.028	0.176	0.156
	KCPER	0.786	0.015	0.995	0.005	0.363	0.066	0.824	0.027	0.942	0.045	0.111	0.107
	MulSUN-OER	0.960	0.048	0.996	0.004	0.598	0.097	0.973	0.029	0.906	0.081	0.072	0.046
	KG4Ex	0.765	0.046	0.995	0.008	0.937	0.019	0.993	0.003	0.853	0.047	0.183	0.192
Ednet	DRER	0.632	0.199	0.944	0.053	0.727	0.000	0.908	0.050	0.881	0.093	0.386	0.253
	MMER	0.646	0.231	0.955	0.053	0.622	0.093	0.920	0.067	0.902	0.078	0.240	0.486
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.955	0.957
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.959	0.960
	NR4DER	0.692	0.171	0.966	0.025	0.859	0.073	0.806	0.049	0.905	0.077	0.505	0.421
	ERTGA	0.698	0.161	0.984	0.015	0.948	0.030	0.832	0.075	0.905	0.076	0.290	0.204
	KCPER	0.759	0.121	0.974	0.011	0.883	0.038	0.930	0.025	0.905	0.074	0.232	0.232
	MulSUN-OER	0.851	0.097	0.994	0.013	0.766	0.136	0.884	0.077	0.895	0.083	0.393	0.290
	KG4Ex	0.568	0.230	0.887	0.091	0.565	0.149	0.962	0.020	0.864	0.111	0.316	0.247
Junyi	DRER	0.820	0.050	0.928	0.104	0.848	0.016	0.958	0.017	0.832	0.074	0.246	0.258
	MMER	0.856	0.087	0.904	0.206	0.525	0.061	0.993	0.013	0.923	0.043	0.199	0.373
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.387	0.399
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.379	0.392
	NR4DER	0.842	0.048	0.963	0.016	0.965	0.003	0.951	0.017	0.762	0.130	0.539	0.553
	ERTGA	0.840	0.050	0.985	0.014	0.975	0.005	0.862	0.044	0.809	0.119	0.147	0.118
	KCPER	0.863	0.057	0.955	0.023	0.936	0.017	0.965	0.018	0.842	0.097	0.318	0.368
	MulSUN-OER	0.902	0.044	0.964	0.015	0.922	0.067	0.933	0.036	0.835	0.088	0.372	0.309
	KG4Ex	0.811	0.094	0.932	0.015	0.901	0.035	0.903	0.021	0.822	0.065	0.313	0.361
NIPS34	DRER	0.791	0.125	0.964	0.018	0.816	0.000	0.953	0.026	0.975	0.021	0.339	0.212
	MMER	0.794	0.118	0.968	0.027	0.666	0.206	0.970	0.027	0.937	0.047	0.422	0.843
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.829	0.834
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.826	0.831
	NR4DER	0.779	0.111	0.969	0.025	0.900	0.030	0.893	0.034	0.925	0.050	0.931	0.882
	ERTGA	0.788	0.102	0.990	0.010	0.931	0.044	0.897	0.039	0.918	0.032	0.343	0.182
	KCPER	0.759	0.094	0.973	0.009	0.950	0.015	0.989	0.008	0.939	0.046	0.318	0.336
	MulSUN-OER	0.871	0.079	0.964	0.021	0.919	0.029	0.909	0.050	0.934	0.041	0.302	0.222
	KG4Ex	0.749	0.127	0.945	0.042	0.617	0.138	0.988	0.009	0.907	0.044	0.495	0.430
XES3G5M	DRER	0.825	0.048	0.992	0.005	0.873	0.022	0.907	0.033	0.914	0.036	0.133	0.120
	MMER	0.835	0.074	0.995	0.004	0.825	0.170	0.900	0.060	0.950	0.044	0.185	0.345
	AKTRec	-	-	-	-	-	-	-	-	-	-	0.591	0.434
	SimpleKTRec	-	-	-	-	-	-	-	-	-	-	0.733	0.577
	NR4DER	0.843	0.045	0.996	0.001	0.995	0.001	0.845	0.031	0.934	0.050	0.659	0.455
	ERTGA	0.584	0.140	0.999	0.001	0.902	0.053	0.864	0.056	0.828	0.076	0.195	0.103
	KCPER	0.849	0.042	0.996	0.002	0.870	0.045	0.911	0.031	0.933	0.049	0.151	0.119
	MulSUN-OER	0.648	0.060	0.991	0.006	0.872	0.060	0.923	0.049	0.938	0.045	0.092	0.055
	KG4Ex	0.755	0.109	0.986	0.012	0.916	0.048	0.882	0.013	0.906	0.060	0.202	0.157

Table 10: Comparison in terms of TGA@5 and TGA@20. The best three results are highlighted by 1st, 2nd, and 3rd.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M	
	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20
DRER	-0.165	-0.212	0.038	-0.049	-0.005	-0.002	-0.155	-0.980	0.078	0.021	-0.026	0.306	0.215	0.145
AKTRec	-0.009	-0.169	0.055	-0.272	-0.000	-0.001	-0.101	-0.731	0.096	0.027	0.331	-0.046	0.217	0.155
S-KTRec	0.003	-0.172	0.066	-0.262	0.001	-0.001	-0.076	-0.703	0.099	0.038	0.343	-0.050	0.223	0.149
MMER	-0.123	-0.164	0.010	0.037	-0.009	-0.009	0.028	-0.406	0.066	0.069	0.342	0.330	0.250	0.147
KCP-ER	-0.061	-0.171	0.057	0.000	-0.002	0.000	-0.085	-0.781	0.100	0.080	0.303	0.104	0.225	0.193
KG4Ex	-0.141	-0.203	0.116	0.135	-0.002	-0.002	0.042	-0.106	0.104	0.049	0.293	-0.091	0.232	0.211
MulOER	-0.036	-0.187	0.062	0.062	-0.007	-0.003	-0.009	-0.146	0.091	0.042	0.237	0.096	0.213	0.120
ER-TGA	-0.030	-0.158	0.104	0.102	-0.004	0.004	-0.019	-0.139	0.144	-0.016	0.109	-0.364	0.235	0.251
NR4DER	-0.092	-0.169	0.105	-0.078	-0.001	-0.000	-0.070	-0.707	0.087	0.033	0.323	0.076	0.229	0.165
SRC	0.041	0.119	0.201	0.212	0.000	0.000	0.297	0.186	0.083	0.049	0.494	0.083	-0.028	-0.011
AC	0.241	-0.088	0.225	-0.029	-0.002	0.003	-0.416	-0.265	0.229	0.408	0.405	0.439	0.859	0.896
DQN	0.326	0.085	0.099	0.126	0.006	0.089	0.116	0.101	0.353	0.455	0.612	0.534	0.908	0.792
CSEAL	0.175	0.145	0.199	0.036	-0.001	0.003	0.324	0.230	0.143	0.159	0.420	0.399	0.917	0.922
RLTutor	0.273	0.357	0.225	0.271	-0.001	0.000	-0.121	0.093	0.395	0.171	0.544	0.642	0.902	0.895
GEHRL	0.077	-0.040	0.151	0.169	0.008	0.033	-0.086	-0.048	-0.320	-0.202	-0.168	-0.183	0.912	0.901
DLPR	0.103	0.126	0.174	0.195	0.005	0.013	-0.121	-0.102	0.201	0.236	0.307	0.385	0.816	0.798
PKSD	0.288	0.274	0.222	0.211	0.001	0.003	0.334	0.376	0.371	0.310	0.660	0.650	0.864	0.902
KnowLP	-0.046	-0.005	0.207	0.231	0.005	0.009	0.124	0.127	0.249	0.279	0.430	0.463	0.796	0.869

Table 11: Comparison in terms of GPP@5 and GPP@20. The best three results are highlighted by 1st, 2nd, and 3rd.

	ASSIST2017		Algebra2005		Bridge2006		Ednet		Junyi		NIPS34		XES3G5M	
	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20	@5	@20
DRER	-0.163	-0.193	0.036	-0.013	-0.001	-0.002	-0.001	-0.772	0.114	-0.214	0.089	0.281	0.087	-0.053
AKTRec	0.001	-0.100	0.036	-0.074	0.000	0.001	0.012	-0.259	0.143	-0.118	0.188	-0.027	0.085	-0.172
S-KTRec	0.013	-0.096	0.037	-0.073	-0.000	0.001	0.016	-0.270	0.122	-0.158	0.195	-0.012	0.098	-0.190
MMER	-0.112	-0.146	0.018	0.014	-0.004	-0.006	0.064	-0.174	0.060	-0.011	0.274	0.259	0.051	-0.046
KCP-ER	-0.035	-0.106	0.036	0.000	-0.000	0.006	0.078	-0.303	0.120	0.049	0.198	-0.062	0.078	-0.015
KG4Ex	-0.101	-0.140	0.057	0.052	-0.002	-0.005	0.067	-0.043	0.159	-0.121	0.127	-0.025	0.117	0.038
MulOER	-0.018	-0.148	0.036	0.036	-0.004	-0.003	0.057	-0.146	0.126	-0.116	0.172	0.002	0.093	0.016
ER-TGA	0.005	-0.139	0.061	0.054	0.005	0.006	0.002	-0.168	0.381	-0.297	0.111	-0.309	0.126	0.076
NR4DER	-0.045	-0.117	0.046	-0.023	0.000	0.000	0.078	-0.320	0.150	-0.256	0.182	-0.040	0.106	-0.173
SRC	-0.026	-0.138	0.082	0.207	0.002	0.008	0.043	0.225	0.273	0.280	0.041	-0.063	-0.007	-0.013
AC	0.195	0.175	0.118	0.095	0.005	0.006	0.280	-0.161	0.208	0.184	0.368	0.321	-0.215	-0.163
DQN	0.246	0.155	0.126	0.126	0.006	0.006	0.375	-0.253	0.386	0.384	0.489	0.502	-0.141	0.131
CSEAL	0.245	0.221	0.105	0.009	0.005	0.004	0.360	0.428	0.382	0.354	0.489	0.430	-0.134	-0.129
RLTutor	0.211	0.314	0.122	0.127	0.005	0.005	0.346	0.461	0.390	0.367	0.473	0.437	-0.162	-0.044
GEHRL	0.131	0.017	0.042	0.035	0.007	0.006	0.372	0.147	-0.230	-0.174	0.098	0.099	0.184	0.250
DLPR	0.135	0.156	0.104	0.109	0.005	0.006	0.315	0.241	0.269	0.274	0.342	0.411	-0.095	0.073
PKSD	0.203	0.141	0.127	0.174	0.004	0.004	0.414	0.486	0.404	0.391	0.502	0.453	0.145	0.155
KnowLP	0.210	0.193	0.113	0.137	0.004	0.007	0.345	0.401	0.293	0.368	0.362	0.406	0.017	0.147

The simulator choice nevertheless affects individual method rankings. PKSD, which achieves the highest average score under DKT, ranks fourth under SimpleKT (0.181), while GEHRL becomes the leading method. Among ILER methods, KCP-ER replaces KG4Ex as the strongest average performer. The distinction between targeted and global gains also remains visible under SimpleKT. For example, on Ednet, DQN achieves a higher TGA@10 than SRC (0.723 vs. 0.446) but a lower GPP@10 (-0.340 vs. 0.593), showing that the preferred method still depends on the instructional objective. Taken together, these results support the persistence of PLER's average ten-step advantage across the two evaluated simulators, while revealing simulator-dependent differences in the relative performance of individual methods.

Complementary evaluation with conventional ILER metrics. Tables 8 and 9 report ILER performance across nine datasets using conventional metrics, including NDCG, F1, Recall, MRR, Accuracy, Diversity, and Novelty. These metrics assess several aspects of recommendation quality, including ranking quality, coverage of unmastered concepts, difficulty matching, and recommendation-list composition. WCG complements these measures by quantifying simulator-

Table 12: Experimental results on ASSIST2009 and ASSIST2012.

	ASSIST2009						ASSIST2012					
	T@5	T@10	T@20	G@5	G@10	G@20	T@5	T@10	T@20	G@5	G@10	G@20
DRER	0.193	0.098	0.164	0.181	0.116	0.056	0.120	0.144	-0.002	0.083	0.084	-0.019
AKTRec	0.066	-0.009	-0.455	0.067	0.034	-0.119	0.172	0.169	-0.136	0.085	0.077	-0.051
S-KTRec	0.072	0.003	-0.376	0.070	0.046	-0.077	0.177	0.180	-0.117	0.087	0.077	-0.041
MMER	0.165	0.152	0.074	0.085	0.084	0.035	0.071	0.088	0.079	0.067	0.077	0.066
KCP-ER	0.161	0.159	0.058	0.093	0.103	-0.011	0.155	0.150	0.032	0.093	0.095	-0.014
KG4Ex	0.156	0.176	0.112	0.093	0.063	-0.029	0.230	0.176	-0.094	0.141	0.138	-0.174
MulOER	0.235	0.230	0.083	0.102	0.096	-0.108	0.116	0.101	0.040	0.083	0.090	0.053
ER-TGA	-0.089	-0.076	-0.102	-0.525	-0.524	-0.546	0.137	0.104	0.088	0.075	0.098	0.024
NR4DER	0.152	0.134	0.052	0.076	0.058	-0.096	0.166	0.148	0.006	0.084	0.088	-0.026
SRC	0.070	0.164	0.214	0.214	0.054	0.184	0.020	-0.161	-0.259	-0.228	-0.234	-0.037
AC	-0.080	-0.005	0.035	0.132	0.010	-0.020	0.307	0.173	0.350	0.172	0.040	-0.087
DQN	0.148	0.019	0.208	0.180	0.155	0.08	0.533	0.375	0.531	0.283	0.003	0.176
CSEAL	0.002	0.002	0.009	0.021	0.001	0.001	0.466	0.299	0.179	0.263	0.051	0.013
RLTutor	0.159	0.184	0.100	0.272	0.270	0.303	0.467	0.291	0.189	0.183	0.292	0.140
GEHRL	0.055	-0.676	-0.672	-0.152	-0.306	-0.320	0.409	0.423	0.149	-0.375	-0.376	0.148
DLPR	0.091	0.110	0.131	0.171	0.160	0.185	0.284	0.103	0.301	0.019	-0.091	-0.053
PKSD	0.199	0.183	0.205	0.171	0.202	0.264	0.486	0.563	0.612	0.186	0.188	0.184
KnowLP	0.128	0.135	0.151	0.181	0.203	0.213	0.285	0.293	0.295	0.129	0.086	0.124

estimated changes in concept mastery after exercise interactions under specified instructional objectives and interaction budgets. By incorporating both conventional metrics and WCG, UniER provides a unified framework for examining recommendation quality alongside estimated learning gains.

Comprehensive WCG evaluation across varying path lengths. Tables 10, 11, and 12 present the extended performance of all models in terms of TGA and GPP at different learning path lengths (*e.g.*, @5, @10, and @20). We highlight the following remarkable observation:

PLER's average advantage widens over longer learning paths. Tables 10, 11, and 12 show a widening aggregate performance gap between ILER and PLER. Averaging TGA and GPP across methods and all nine datasets, ILER's score decreases from 0.075 at five steps to -0.070 at twenty steps. PLER's corresponding score decreases from 0.200 to 0.184, widening the gap from 0.125 to 0.254. Individual trends nevertheless differ: PKSD's average score across datasets and metrics increases from 0.310 to 0.322, whereas DQN's decreases from 0.281 to 0.235. These results identify the ability to sustain estimated gains over longer paths as a relative strength of PLER in the evaluated settings.

Short-horizon gains diminish in KT-based item recommendation. Tables 10, 11, and 12 reveal pronounced deterioration in AKTRec and S-KTRec as the interaction budget increases. Both methods shift from positive GPP@5 to negative GPP@20 on eight of the nine datasets. Their targeted gains also reverse on NIPS34: from five to twenty steps, AKTRec's TGA decreases from 0.331 to -0.046, while S-KTRec's decreases from 0.343 to -0.050. These consistent reversals reveal a gap between achieving short-horizon gains and sustaining them over extended learning paths in the two KT-based recommenders.

Novelty and diversity reveal different method strengths. Table 9 shows that ILER methods exhibit different strengths across recommendation-list properties. On ASSIST2009, NR4DER achieves higher diversity than MMER (Div@m: 0.925 vs. 0.394), despite their similar novelty scores (Nov@m: 0.977 vs. 0.978). On the same dataset, ER-TGA achieves higher novelty than KG4Ex (Nov@m: 0.999 vs. 0.720) but lower diversity (Div@m: 0.508 vs. 0.664). These comparisons show that highly novel recommendations can differ substantially in diversity, and that the two metrics can favor different methods. Reporting these complementary properties alongside WCG enables UniER to characterize both recommendation-list composition and estimated knowledge gains.

PKSD maintains positive gains across datasets and objectives. Tables 10, 11, and 12 reveal differences in the consistency of positive estimated gains across datasets. PKSD is the only evaluated method with positive TGA and GPP on all nine datasets at both five and twenty steps. Other leading PLER methods exhibit exceptions: DQN's GPP@20 on Ednet is -0.253, while CSEAL records negative GPP on XES3G5M at both budgets (-0.134 and -0.129). PKSD thus

combines strong aggregate performance with consistently positive estimated gains across the tested datasets and instructional objectives.