

Triviality Corrected Endogenous Reward

Xinda Wang^{1*} Zhengxu Hou^{2*} Yangshijie Zhang^{2,3} Bingren Yan² Jialin Liu¹
Chenzhuo Zhao¹ Zhibo Yang² Bin-Bin Yang² Feng Xiao^{2†}

¹Peking University ²Alibaba Group ³Lanzhou University
{nev_settle}@stu.pku.edu.cn
{houzhengxu.hzx, ron.xf}@alibaba-inc.com

Abstract

Reinforcement learning for open-ended text generation is constrained by the lack of verifiable rewards, necessitating reliance on judge models that require either annotated data or powerful closed-source models. Inspired by recent work on unsupervised reinforcement learning for mathematical reasoning using confidence-based endogenous rewards, we investigate whether this principle can be adapted to open-ended writing tasks. We find that directly applying confidence rewards leads to *Triviality Bias*: the policy collapses toward high-probability outputs, reducing diversity and meaningful content. We propose TCER (Triviality Corrected Endogenous Reward), which addresses this bias by rewarding the relative information gain between a specialist policy and a generalist reference policy, modulated by a probability-dependent correction mechanism. Across multiple writing benchmarks and model architectures, TCER achieves consistent improvements without external supervision. Furthermore, TCER also transfers effectively to mathematical reasoning, validating the generality of our approach across different generation tasks.

1 Introduction

The capabilities of Large Language Models (LLMs) have advanced significantly, with alignment techniques playing a pivotal role in their success (Kumar et al., 2025; Tie et al., 2025). Following supervised fine-tuning (SFT), Reinforcement Fine-Tuning (RFT) has emerged as a crucial step for further optimizing model behavior according to specific objectives (Wu et al., 2025b; Jin et al., 2025). While RFT has proven effective in

*Contributed equally.

†Corresponding author.

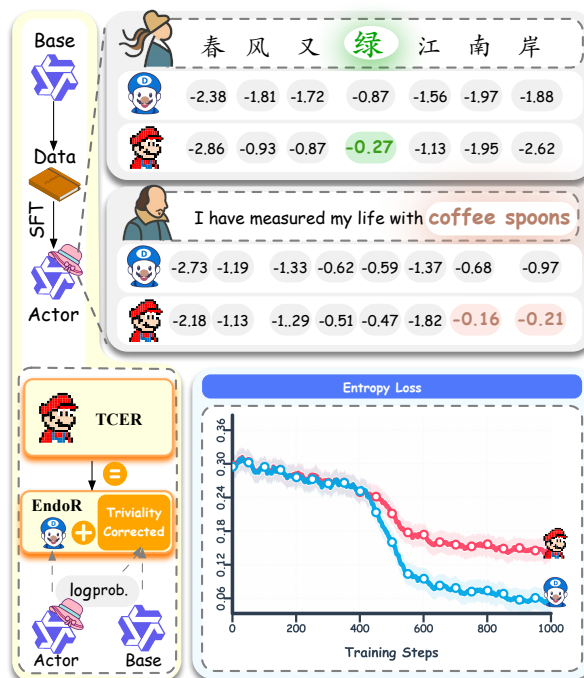


Figure 1: TCER computes rewards using both base and actor models, emphasizing high-entropy meaningful content over EndoR. After RL Training, TCER shows gentler entropy decay with enhanced exploration.

domains with clear evaluation criteria, its application to writing presents challenges that remain unresolved (Gooding et al., 2025; Li et al., 2025b).

Unlike mathematical reasoning or code generation where correctness is objectively verifiable, writing quality is subjective and multifaceted (Ying et al., 2025). Current approaches rely on LLMs as judges (Zheng et al., 2023), facing a dilemma: powerful proprietary models like GPT-4 provide high-quality evaluation (Wu et al., 2025c) but incur substantial API costs (Huang et al., 2025a), while training custom judge models requires extensive human-annotated preference data (Feng et al., 2025; Lu, 2025; Wang et al., 2025c). The latter approach is further complicated by the subjectivity

of writing, which makes consistent annotation challenging with different annotators often disagreeing on quality metrics (Ni et al., 2025; Chiang and Lee, 2023). In parallel, the mathematical reasoning community has developed a solution to avoid external supervision (Li et al., 2025c): *endogenous rewards*, where models use their own confidence scores (log-probabilities) as reward signals. This approach has demonstrated success in mathematics (Prabhudesai et al., 2025). The appeal of such judge-free methods (Li et al., 2025a) naturally raises the question: can this paradigm be adapted to creative writing, solving the reward model problem?

In this paper, we perform a systematic investigation of endogenous reward RFT for writing. As shown in Figure 1, we find that endogenous rewards can improve over SFT baselines, suggesting that confidence signals remain useful in open-ended generation. However, we also identify a degeneration phenomenon: optimizing for self-confidence tends to drive the policy toward low-entropy behavior and reducing diversity and content richness (Li et al., 2025b). We refer to this failure mode as *Triviality Bias*.

To address this *Triviality Bias*, we propose TCER (Triviality Corrected Endogenous Reward). We define a generalist reference policy π_b and a specialist policy π_s obtained by finetuning π_b on high-quality in-domain data. TCER define the token-level log-likelihood ratio $\log \frac{\pi_s}{\pi_b}$, which favors tokens that are characteristic of the specialist’s domain. To counteract the bias towards high-confidence predictions, we modulate this reward with an adaptive weight of $(1 - \pi_s)^\lambda$. This mechanism suppresses the incentive for high-probability tokens and allocates reward to choices that are informative yet less predictable, thereby promoting exploration. We optimize using GRPO with reference augmentation to align model outputs with high-quality targets. Our main contributions are:

1. We provide a systematic study of judge-free RFT via confidence-based endogenous rewards for writing, showing both its effectiveness and a degeneration phenomenon (*Triviality Bias*) characterized by low-entropy output.
2. We propose TCER, which mitigates triviality bias by rewarding relative information gain between a specialist policy and a generalist reference policy, modulated by a probability-dependent correction.
3. We conduct extensive experiments across multiple models and datasets in writing and mathematical reasoning, demonstrating consistent improvements over baselines and cross-domain generalization.

2 Related Work

2.1 Open-Ended Text Generation

Training models for proficient open-ended writing has been a central challenge in LLM development (Wei et al., 2025). The dominant approach relies on constructing high-quality datasets for SFT (Wang et al., 2024). LongWriter addresses the challenge of generating coherent long-form content by extending context windows and introducing specialized data collection pipelines (Bai et al., 2024). DeepWriting takes a different approach through reverse-engineered reasoning (Wang et al., 2025a), where high-quality outputs are augmented with synthetic planning traces. These datasets enable SFT models to achieve reasonable writing capabilities (Kim et al., 2025), establishing the foundation for further optimization.

Beyond SFT, RFT has emerged as a promising direction for enhancing writing quality (Huang et al., 2025b). However, existing RL approaches for writing predominantly rely on external reward models. LongWriter-zero employs LLMs as judges (Wu et al., 2025c), using GPT-4 or similar models to score generated texts during RL training. Similarly, Writing-zero (Lu, 2025) and EvolvR (Wang et al., 2025c) utilize pairwise preference judgments from large language models to construct reward signals for story generation and creative writing tasks. While these methods demonstrate improvements over SFT baselines, they inherit the fundamental limitation of requiring costly external evaluation (Wu et al., 2025a), whether from human annotators or proprietary models.

2.2 Unsupervised Rewards for RFT

The mathematical reasoning domain has pioneered the exploration of unsupervised reward (Zhao et al., 2025) mechanisms for RL training. TTRL (Zuo et al., 2025) introduces entropy-based rewards by clustering rollout answers and rewarding based on cluster sizes, where higher consistency in solution clusters indicates more reliable reasoning paths (Zhang et al., 2025a). Another line of work explores confidence-based approaches (Prabhudesai

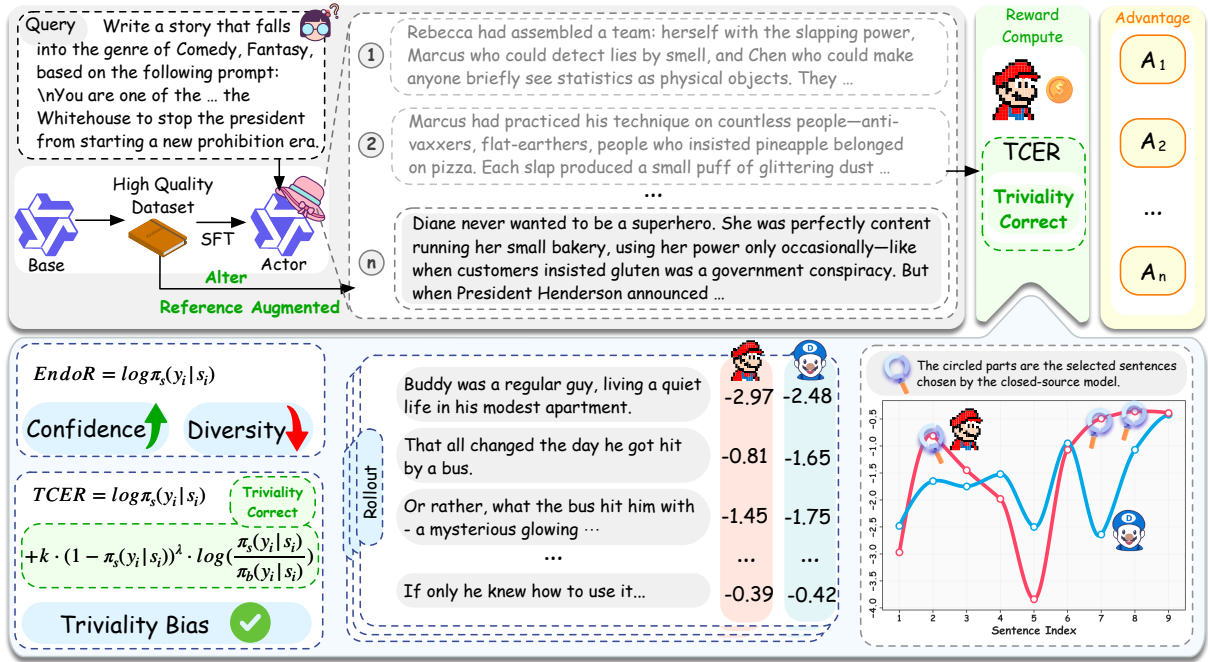


Figure 2: Overview of TCER training pipeline and reward comparison. (a) Training workflow: SFT on high-quality data to obtain π_s , followed by Reference Augmented GRPO with TCER that incorporate information gain ϕ and gating weight w . (b) Sentence-level reward visualization on a rollout: token-level rewards are averaged within each sentence comparing TCER and EndoR. (Rewards are average log-probs; higher is closer to 0.)

et al., 2025). Various methods directly leverage the model’s inherent confidence as direct rewards for self-improvement (Li et al., 2025a), and some constructing process reward models from confidence scores to provide fine-grained feedback (Tan et al., 2025). Recently, theoretical work has established the endogenous reward hypothesis (Li et al., 2025c), demonstrating that sufficiently capable models implicitly encode reward functions within their parameters. This provides mathematical justification for these confidence-based methods.

However, these methods have remained confined to mathematical reasoning and have not been applied to open-ended writing task (Yue et al., 2025). The challenge is the risk of entropy collapse (Zhang et al., 2025b; Cui et al., 2025; Wang et al., 2025b) that confidence-based approaches inherently drive models toward generating high-frequency, formulaic content, resulting in decreased diversity and creative expression (Li et al., 2025b; West and Potts, 2025). Our work addresses this challenge by introducing a correction mechanism that mitigates entropy collapse in writing tasks, enabling the application of unsupervised RL methods beyond mathematical reasoning.

3 Methodology

This section formulates judge-free RFT for open-ended text generation. We analyze the *triviality bias* induced by endogenous rewards. We derive TCER from an information theoretic decomposition of a generalist policy and an in-domain specialist policy, modulated by a probability-dependent correction. Finally, we describe optimization with Reference Augmented Group Relative Policy Optimization (GRPO) (Shao et al., 2024). The framework is shown in Figure 2. Additional derivations are deferred to Appendix A.

3.1 Implicit Rewards and Triviality Bias

Recently, connections between next-token prediction and inverse reinforcement learning (IRL) motivate using model confidence as an endogenous reward. In IRL, an optimal policy under reward r takes a Boltzmann form:

$$\pi^*(y | x) \propto \exp\left(\frac{Q^*(x, y)}{\alpha}\right), \quad (1)$$

where Q^* is the soft Q -function and α is the temperature. In our setting, let π_s denote a specialist policy obtained by SFT on high-quality in-domain data. A confidence based endogenous reward assigns each generated token y_i in context s_i the re-

ward

$$r_e(y_i | s_i) = \log \pi_s(y_i | s_i), \quad (2)$$

where s_i denotes the decoding context at position i . The degeneration induced by r_e can be characterized through a context maximization. For any context s , consider the one step objective

$$\max_{\pi(\cdot|s)} \mathbb{E}_{y \sim \pi(\cdot|s)} [\log \pi_s(y | s)]. \quad (3)$$

Because the expectation in Eq. (3) is linear in $\pi(\cdot | s)$, the maximizer is attained at an extreme point of the simplex which is a deterministic distribution:

$$\pi^*(\cdot | s) = \delta \left(y = \arg \max_{v \in \mathcal{V}} \pi_s(v | s) \right). \quad (4)$$

Optimizing r_e encourages low-entropy behavior by concentrating on the high probability tokens under π_s . In open-ended writing, these tokens correspond to high-frequency, yielding *triviality bias*: generations become increasingly templated, with reduced diversity.

3.2 Information Theoretic Decomposition

To isolate token-level signals of in-domain specialization beyond generic preference, we decompose the specialist policy π_s relative to a generalist baseline π_b , where π_b and π_s are the same model before and after in-domain fine-tuning, respectively. For any token and context,

$$\pi_s(y_i | s_i) = \pi_b(y_i | s_i) \cdot \frac{\pi_s(y_i | s_i)}{\pi_b(y_i | s_i)}. \quad (5)$$

Taking logarithms yields

$$\log \pi_s(y_i | s_i) = \log \pi_b(y_i | s_i) + \phi(y_i | s_i), \quad (6)$$

where we define the specific information gain as the token-level log-likelihood ratio

$$\phi(y_i | s_i) = \log \frac{\pi_s(y_i | s_i)}{\pi_b(y_i | s_i)}. \quad (7)$$

3.3 Triviality Corrected Endogenous Reward

Directly using the log-likelihood ratio $\phi(y_i | s_i)$ as a reward can be numerically unstable when $\pi_b(y_i | s_i)$ is extremely small. We therefore retain the endogenous reward $\log \pi_s(y_i | s_i)$ and add a gated triviality correction. The TCER reward is defined as

$$r_t(y_i | s_i) = \log \pi_s(y_i | s_i) + k \cdot w(y_i | s_i) \cdot \phi(y_i | s_i), \quad (8)$$

with the triviality corrected weighting

$$w(y_i | s_i) = (1 - \pi_s(y_i | s_i))^\lambda, \quad (9)$$

where $k > 0$ scales the correction and $\lambda > 0$ controls the sharpness of suppression. For numerical stability, we apply ε -smoothing with $\varepsilon > 0$:

$$\phi(y_i | s_i) \leftarrow \log \frac{\pi_s(y_i | s_i) + \varepsilon}{\pi_b(y_i | s_i) + \varepsilon}. \quad (10)$$

Giving a context s_i and a candidate token y_i , and denote $p = \pi_s(y_i | s_i) \in [0, 1]$ and $q = \pi_b(y_i | s_i) \in [0, 1]$. The gated correction term in Eq. (8) can be written as

$$C(p) = (1 - p)^\lambda \cdot \log \frac{p + \varepsilon}{q + \varepsilon}. \quad (11)$$

Under Eq. (10), the smoothed log-ratio is uniformly bounded: since $p + \varepsilon \in [\varepsilon, 1 + \varepsilon]$ and $q + \varepsilon \in [\varepsilon, 1 + \varepsilon]$, we have

$$\left| \log \frac{p + \varepsilon}{q + \varepsilon} \right| \leq \log \frac{1 + \varepsilon}{\varepsilon}. \quad (12)$$

Therefore, the correction term admits the quantitative bound

$$|C(p)| \leq (1 - p)^\lambda \log \frac{1 + \varepsilon}{\varepsilon}. \quad (13)$$

In particular, as $p \rightarrow 1$, $(1 - p)^\lambda \rightarrow 0$ while the factor in Eq. (12) remains finite, yielding

$$\lim_{p \rightarrow 1} C(p) = 0. \quad (14)$$

Conversely, if $p \leq 1 - \delta$ for some $\delta \in (0, 1]$, then $(1 - p)^\lambda \geq \delta^\lambda$, and the correction term satisfies

$$|C(p)| \geq \delta^\lambda \left| \log \frac{p + \varepsilon}{q + \varepsilon} \right|. \quad (15)$$

Eq. (15) implies that when $\pi_s(y_i | s_i)$ is moderate, the gate function preserves the information gain signal at least a δ^λ fraction of the log-ratio magnitude, allowing the correction to reward domain-specific tokens while suppressing high-probability trivial.

3.4 Reference Augmented GRPO

We optimize the policy π_θ against r_t using GRPO. For a given input prompt q , we sample a group of G outputs $\{o_1, \dots, o_G\}$ from the current policy $\pi_{\theta_{\text{actor}}}$. To stabilize the advantage estimation and guide exploration, we augment this group with the ground-truth reference o_{gt} from the dataset, forming the augmented set $\mathcal{O} = \{o_1, \dots, o_G, o_{\text{gt}}\}$. Although r_t

Model	LB	HB-A	HB-B	WB-A	WB-B	WB-C	WB-D	WB-E	WB-F
<i>General Closed-source Models</i>									
GPT-4o	83.1	83.7	87.6	74.4	73.4	74.3	77.9	75.8	78.0
Claude 3.5	89.3	82.9	88.3	59.5	57.6	56.3	59.3	62.0	67.7
Claude 3.7	97.8	83.9	93.2	78.2	77.9	76.5	79.3	79.2	80.8
<i>General Open-source Models</i>									
Llama3.1-8B-Instruct	60.3	45.5	41.3	47.5	45.6	43.7	42.3	48.6	50.8
Qwen2.5-32B-Instruct	78.8	77.0	78.4	52.5	49.8	51.0	49.6	53.9	54.2
Qwen3-32B	93.2	84.0	86.8	79.8	78.2	80.8	78.9	82.9	81.9
<i>Finetuned Open-source Models</i>									
Qwen2.5-7B-Instruct	74.7	72.0	76.3	58.9	56.8	56.5	54.0	59.9	60.0
+ SFT	83.1	73.7	85.8	69.7	69.8	67.9	63.1	71.3	66.3
+ SFT + RL-EndoR	84.5	79.0	84.7	72.6	70.9	70.1	64.6	73.0	68.7
+ SFT + RL-TCER	86.3	81.1	86.4	72.7	73.1	70.8	67.8	73.4	71.1

Table 1: Performance comparison on writing benchmarks. We evaluate models across LongBench-Write, HelloBench subsets, and WritingBench domains. Starting from Qwen2.5-7B-Instruct, each training stage (SFT, RL-EndoR, RL-TCER) demonstrates progressive improvements.

is defined at the token level, GRPO uses an output-level scalar reward. We define the sequence-level reward as the length-normalized average of token rewards:

$$R_i = \frac{1}{|o_i|} \sum_{t=1}^{|o_i|} r_t(o_{i,t}). \quad (16)$$

The advantage \hat{A}_i is computed by normalizing against the augmented group statistics:

$$\hat{A}_i = \frac{R_i - \frac{1}{|\mathcal{O}|} \sum_{j \in \mathcal{O}} R_j}{\sigma_{\mathcal{O}}}. \quad (17)$$

The inclusion of o_{gt} elevates the baseline mean, effectively penalizing trivial generations that fail to match the reference under the reward. This sequence-level advantage is broadcast to each token position, and maximize the GRPO objective:

$$\mathcal{J}_{GRPO}(\theta) = \mathbb{E} \left[\min(\rho_{i,t}(\theta) \hat{A}_i, \text{clip}(\rho_{i,t}(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_i) - \beta D_{KL}(\pi_{\theta} \parallel \pi_{\text{ref}}) \right]. \quad (18)$$

where $\rho_i(\theta)$ is the policy ratio, ϵ is the clipping parameter, and β controls the KL divergence penalty to a reference policy π_{ref} .

4 Experiments

4.1 Experimental Setup

Datasets. Our experimental setup utilizes distinct datasets tailored for writing and mathematical rea-

soning tasks to ensure specialized training and robust evaluation. For writing tasks, we employ two high-quality datasets:

- **DeepWriting** (Wang et al., 2025a): A collection of 20,000 high-quality writing samples designed to elicit complex and nuanced text generation.
- **LongWriter** (Bai et al., 2024): A dataset comprising 6,000 samples focused on long-form, coherent content creation.

For both datasets, we implement a 50/50 split. The first half is used for the SFT phase to build an in-domain policy (π_s), while the second, disjoint half is used as the source of prompts for the RL phase. This separation prevents data leakage between the SFT and RL stages and ensures that the RL algorithm optimizes on unseen prompts. For mathematical reasoning tasks, we follow the data strategy established by LUFFY (Yan et al., 2025). We use a 45,000-sample subset of the OpenR1-Math-220k dataset for the SFT phase. For the RL phase, we train our model on the Math-LightEval dataset.

Benchmarks. To provide a comprehensive and multi-aspect evaluation of our models, we employ a suite of specialized benchmarks targeting both writing proficiency and mathematical reasoning. For Writing Tasks, we probe three distinct dimensions

Model	LB	HB-A	HB-B	WB-A	WB-B	WB-C	WB-D	WB-E	WB-F
<i>DeepWriting Dataset</i>									
DeepWriter-8B	85.6	77.8	85.7	72.2	71.8	69.8	70.6	73.7	72.3
+ RL-EndoR	86.1	79.6	86.9	76.0	73.6	73.9	69.6	77.3	74.5
+ RL-TCER	87.3	82.2	<u>86.2</u>	76.5	75.3	74.8	74.2	<u>76.7</u>	75.2
<i>LongWriter Dataset</i>									
LongWriter-8B	76.5	68.8	82.9	57.9	53.9	49.0	52.0	52.9	52.0
+ RL-EndoR	80.2	70.1	84.2	64.4	62.6	60.2	56.7	59.9	58.8
+ RL-TCER	82.6	72.6	84.5	65.8	62.9	61.1	57.1	60.6	59.8

Table 2: Effectiveness of TCER for existing finetuned models. Applying our RL methods to finetuned models DeepWriter-8B and LongWriter-8B.

of generative performance using three complementary benchmarks following the evaluation protocol established in DeepWriter:

- **LongBench-Write (LB)** (Bai et al., 2024): This benchmark is designed to measure a model’s ability to produce ultra-long-form text, allowing us to assess its foundational capacity for maintaining thematic consistency over extended outputs.
- **HelloBench (HB)** (Que et al., 2024): HelloBench evaluates performance on a diverse set of in-the-wild tasks. Our analysis focuses on two key subsets: **HB-A** (Open-Ended QA), which tests the generation of detailed and nuanced answers, and **HB-B** (Heuristic Text Generation), which assesses creative reasoning and stylistic fidelity in long-form narrative continuation.
- **WritingBench (WB)** (Wu et al., 2025d): This benchmark measures controllability across six writing domains: **A** (Academic & Engineering), **B** (Finance & Business), **C** (Politics & Law), **D** (Literature & Arts), **E** (Education), and **F** (Advertising & Marketing).

Given the subjective nature of open-ended generation, we adopt the established protocol of using close-source LLMs as judges. Specifically, Claude-Sonnet37 was used to score outputs for LongBench and WritingBench, while GPT-4o was employed for HelloBench. The reported scores represent win rates or the original benchmark scores without rescaling. All results are averaged over three independent runs.

For Mathematical Reasoning, we evaluate our models on six widely recognized benchmarks to

test the cross-domain generalization of our method: AIME 2024 (Li et al., 2024), AIME 2025 (Li et al., 2024), AMC (Li et al., 2024), MATH-500 (Hendrycks et al., 2021), Minerva (Lewkowycz et al., 2022) and OlympiadBench (He et al., 2024). For AIME 2024, AIME 2025, and AMC, which have smaller test sets, we report the avg@32 metric. For the other three larger benchmarks, we report standard pass@1 accuracy.

4.2 Main Results

Table 1 details the results of our full training pipeline, starting from the base Qwen2.5-7B-Instruct model and applying SFT and RL sequentially on the DeepWriting dataset. Fine-tuning on the DeepWriting dataset boosts the model’s performance across all benchmarks compared to the original Qwen2.5-7B-Instruct model. For instance, the LongBench (LB) score jumps from 74.7 to 83.1, and the average WritingBench (WB) performance is significantly elevated, demonstrating the effectiveness of domain specialization. Applying RL with the endogenous reward provides further improvements over the SFT model, confirming that endogenous rewards are a viable signal. However, its gains are inconsistent and marginal in several creativity intensive dimensions. Our proposed TCER method consistently outperforms all baselines, achieving the highest scores across all reported benchmarks for our finetuned models.

To assess the applicability of TCER as a post-training enhancement technique, we applied our RL method to finetuned open-source writing models: DeepWriter-8B and LongWriter-8B. These models represent a high standard of performance. The results, shown in Table 2, demonstrate that TCER can further elevate their capabilities. For

Method	AIME 24	AIME 25	AMC	MATH-500	Minerva	Olympiad	Avg
Qwen2.5-7B-Math	11.2	4.4	30.8	46.0	7.7	14.4	19.1
+ SFT	22.2	22.3	52.8	82.6	40.8	43.7	44.1
+ SFT + RL-EndoR	30.5	25.4	60.9	87.0	42.6	46.8	48.9
+ SFT + RL-TCER	32.4	26.1	62.1	<u>86.4</u>	44.5	49.6	50.2
Llama3.1-8B-Instruct	4.6	0.2	21.2	46.4	20.9	12.4	17.6
+ SFT	9.4	9.5	39.0	68.4	27.2	32.4	30.9
+ SFT + RL-EndoR	9.7	13.2	39.2	70.0	25.7	33.3	31.9
+ SFT + RL-TCER	10.2	15.0	40.5	71.8	28.6	35.3	33.6

Table 3: Performance comparison on mathematical reasoning benchmarks. TCER outperforms both the SFT baseline and EndoR across different base models.

the DeepWriter-8B model, TCER boosts the score from 70.6 to **74.2** on WritingBench-D (Literature & Arts), and on HelloBench-A (Ability), it increases the score from 77.8 to **82.2**. Similarly, for the LongWriter-8B model, TCER provides a targeted boost to its primary strength. The LongBench (LB) score improves from 76.5 to **82.6** after applying TCER, a substantial (+6.1) point increase that surpasses the gain from RL-EndoR.

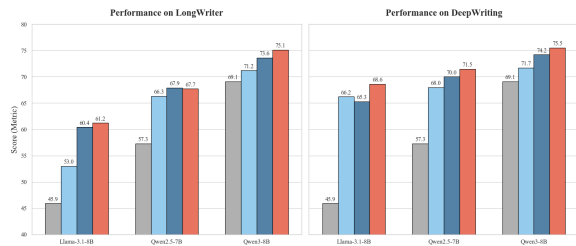


Figure 3: Performance generalization across different models and datasets. For each model group, the bars represent the performance of the base model (grey), after SFT (light blue), after applying RL-EndoR (dark blue), and after applying our RL-TCER (orange).

Metric	EndoR	TCER
High-quality avg.	-1.11	-0.80
Regular avg.	-1.08	-0.93
High-quality Recall %	20.7%	35.8%

Table 4: Sentence-level analysis of reward signals. Sentences identified as brilliant by Gemini-2.5-Pro, GPT-4o, Claude-Opus4 are evaluated. Analysis conducted on 10,000 outputs with ≥ 5 high-quality sentences per text.

4.3 Generalization to Different Models

To verify that the benefits of TCER are not specific to a single model architecture, we conducted

a comparative analysis across three distinct open-source models: Llama-3.1-8B-Instruct, Qwen2.5-7B-Instruct, and Qwen3-8B. Each model was subjected to the same SFT and RL pipeline on both the LongWriter and DeepWriting datasets. The results, visualized in Figure 3, demonstrate the robust, model-agnostic nature of our approach. A clear and consistent trend emerges across all six configurations: our method consistently achieves the highest performance, surpassing both the SFT baseline and the EndoR. These results indicate that TCER is an applicable enhancement, offering performance gains regardless of the underlying model architecture.

4.4 Generalization to Math Reasoning

To investigate the cross-domain applicability of our method, we extended our evaluation to the verifiable and logical domain of mathematical reasoning. We applied the same SFT and RL training pipeline to two different base models, Qwen2.5-7B and Llama3.1-8B-Instruct. The results, presented in Table 3, demonstrate that TCER’s benefits are not confined to open-end generation but generalize effectively to mathematical reasoning. For both the Qwen and Llama models, our RL-TCER method consistently achieves the average scores, outperforming both the SFT baseline and the RL-EndoR approach. This suggests that by penalizing trivial, high-probability steps and rewarding more diversity ones, TCER encourages the model to explore potentially correct solution paths.

4.5 Validation of Quality Enhancement

To validate that TCER identifies and rewards high quality content, we conducted a sentence-level analysis comparing the reward signals of TCER and En-

Config	LB	HB-A	HB-B	WB-A	WB-B	WB-C	WB-D	WB-E	WB-F
FULL TCER	86.3	81.1	86.4	72.7	73.1	70.8	67.8	73.4	71.1
w/o $w(y_i s_i) \cdot \phi(y_i s_i)$	84.5	79.0	84.7	72.6	70.9	70.1	64.6	73.0	68.7
w/o $\log \pi_s(y_i s_i)$	76.2	70.8	72.0	68.5	68.2	65.6	60.1	69.4	64.1
w/o $w(y_i s_i)$	81.5	76.5	79.1	70.6	70.5	71.2	62.7	73.1	65.3
w/o Ref-Augmentation	85.7	78.2	85.6	72.9	72.3	70.6	66.6	73.3	71.2

Table 5: Ablation studies on Qwen2.5-7B-Instruct with the DeepWriting dataset. Each component contributes to the overall performance.

doR. First, we prompted the SFT model to generate a diverse set of outputs across writing tasks. Second, we employed a panel of three LLMs including Gemini 2.5 Pro, Claude-Opus4, and GPT-4o as judges. Through aggregation by designed prompt (with Gemini 2.5 Pro performing the aggregation), these judges identified sentences exhibiting exceptional writing quality. Third, for each sentence in the corpus, we computed both the EndoR score and the TCER score (scores are aggregated log-probs; higher is closer to 0). While our reward mechanism operates at the token level, we aggregate token rewards within each sentence for interpretability, since training uses sequence-level aggregation of token rewards and sentence-level trends reveal where rewards concentrate in the generated text.

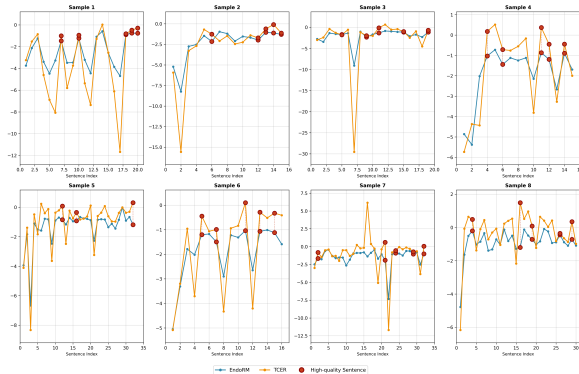


Figure 4: Displays these trajectories as line plots, with EndoR scores shown in blue and TCER scores in orange. High-quality sentences identified by our judge panel are marked with bold red points.

We evaluated three key metrics: (1) the average reward score for sentences identified as high-quality, (2) the average reward score for regular sentences, and (3) the recall rate of high-quality sentences. For the recall metric, we select the top k sentences based on reward scores within each generated text, where k equals the number of high-quality sentences identified by the judge panel for

that text, then calculate the recall rate for each text and report the average across all texts in Table 4. We visualized the sentence-level reward trajectories for eight SFT outputs in Figure 4. The visualizations reveal a consistent pattern: TCER and EndoR show notable divergence specifically at high-quality sentences identified by judges, with TCER assigning higher rewards at these points. In contrast, for regular sentences, the two methods produce comparable scores, indicating that TCER selectively amplifies rewards for exceptional content while maintaining similar baseline evaluations.

4.6 Ablation Study

To dissect the TCER formulation and isolate the contribution of its key components, we conducted a comprehensive ablation study. The experiments were performed on the Qwen2.5-7B-Instruct model with the DeepWriting dataset configuration. The results, detailed in Table 5.

- **Removing the triviality correction:** Our method reverts to the EndoR baseline. This results in a performance drop across most benchmarks. This validates that the correction term is essential for mitigating triviality bias.
- **Removing the endogenous reward:** Relying only on the correction signal leads to a catastrophic performance collapse. This shows that $\log \pi_s(y_i|s_i)$ is crucial for maintaining overall generation quality and coherence.
- **Removing the gating coefficient:** Removing $w(y_i|s_i)$ causes a considerable performance drop, indicating the importance of selectively applying the correction to novel, low-probability tokens.
- **Removing the reference augmented:** Removing the ground-truth reference from

GRPO results in lower performance, supporting the role of reference samples in elevating the optimization baseline.

5 Conclusion

We presented the study of transferring unsupervised RL methods from mathematical reasoning to high-quality writing. While endogenous rewards improve performance, they introduce Triviality Bias suppressing diversity and quality. Our proposed TCER addresses this through information theoretic rewards that measure divergence from base models, modulated by triviality corrected weighting. Experiments demonstrate TCER’s effectiveness across both writing and mathematical domains, achieving quality improvements. By eliminating dependence on external reward models, TCER offers a practical path toward scalable, high-quality text generation without costly human annotation evaluation or API access.

Limitations

Despite the demonstrated effectiveness of TCER, several limitations warrant discussion. First, as a self-supervised reinforcement learning method, while it successfully optimizes the model’s existing capabilities and improves output quality, it has an inherent ceiling for endogenous reward approach. Second, our evaluation relies on automated benchmarks that employ LLMs as judges. While we acknowledge potential biases in using model-based evaluation, these benchmarks represent the current standard for scalable assessment of generation quality. Third, our analysis of reward attribution at the token level remains limited. Although TCER computes rewards based on token-level probabilities, the complex interdependencies in natural language limit the interpretability of why certain generation patterns receive higher rewards, making it difficult to provide fine-grained insights into the model’s decision-making process. Future research should investigate the integration of endogenous and external rewards, alongside developing improved evaluation methodologies and interpretability techniques for deeper insights into model behavior.

References

Yushi Bai, Jiajie Zhang, Xin Lv, Linzhi Zheng, Siqi Zhu, Lei Hou, Yuxiao Dong, Jie Tang, and Juanzi

Li. 2024. Longwriter: Unleashing 10,000+ word generation from long context llms. *arXiv preprint arXiv:2408.07055*.

Cheng-Han Chiang and Hung-yi Lee. 2023. Can large language models be an alternative to human evaluations? *arXiv preprint arXiv:2305.01937*.

Ganqu Cui, Yuchen Zhang, Jiacheng Chen, Lifan Yuan, Zhi Wang, Yuxin Zuo, Haozhan Li, Yuchen Fan, Huayu Chen, Weize Chen, and 1 others. 2025. The entropy mechanism of reinforcement learning for reasoning language models. *arXiv preprint arXiv:2505.22617*.

Yuanning Feng, Sinan Wang, Zhengxiang Cheng, Yao Wan, and Dongping Chen. 2025. Are we on the right way to assessing llm-as-a-judge? *arXiv preprint arXiv:2512.16041*.

Sian Gooding, Lucia Lopez-Rivilla, and Edward Grefenstette. 2025. Writing as a testbed for open ended agents. *arXiv preprint arXiv:2503.19711*.

Chaoqun He, Renjie Luo, Yuzhuo Bai, Shengding Hu, Zhen Thai, Junhao Shen, Jinyi Hu, Xu Han, Yujie Huang, Yuxiang Zhang, and 1 others. 2024. Olympiadbench: A challenging benchmark for promoting agi with olympiad-level bilingual multimodal scientific problems. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3828–3850.

Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Eric Tang, Dawn Song, and Jacob Steinhardt. 2021. Measuring mathematical problem solving with the math dataset. *arXiv preprint arXiv:2103.03874*.

Hui Huang, Xingyuan Bu, Hongli Zhou, Yingqi Qu, Jing Liu, Muyun Yang, Bing Xu, and Tiejun Zhao. 2025a. An empirical study of llm-as-a-judge for llm evaluation: Fine-tuned judge model is not a general substitute for gpt-4. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 5880–5895.

Zeyu Huang, Tianhao Cheng, Zihan Qiu, Zili Wang, Yinghui Xu, Edoardo M Ponti, and Ivan Titov. 2025b. Blending supervised and reinforcement fine-tuning with prefix sampling. *arXiv preprint arXiv:2507.01679*.

Hangzhan Jin, Sicheng Lv, Sifan Wu, and Mohammad Hamdaqa. 2025. RL is neither a panacea nor a mirage: Understanding supervised vs. reinforcement learning fine-tuning for llms. *arXiv preprint arXiv:2508.16546*.

Zae Myung Kim, Anand Ramachandran, Farideh Tavazoei, Joo-Kyung Kim, Oleg Rokhlenko, and Dongyeop Kang. 2025. Align to structure: Aligning large language models with structural information. *arXiv preprint arXiv:2504.03622*.

- Komal Kumar, Tajamul Ashraf, Omkar Thawakar, Rao Muhammad Anwer, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, Phillip HS Torr, Fahad Shahbaz Khan, and Salman Khan. 2025. Llm post-training: A deep dive into reasoning large language models. *arXiv preprint arXiv:2502.21321*.
- Aitor Lewkowycz, Anders Andreassen, David Dohan, Ethan Dyer, Henryk Michalewski, Vinay Ramasesh, Ambrose Slone, Cem Anil, Imanol Schlag, Theo Gutman-Solo, and 1 others. 2022. Solving quantitative reasoning problems with language models. *Advances in neural information processing systems*, 35:3843–3857.
- Jia Li, Edward Beeching, Lewis Tunstall, Ben Lipkin, Roman Soletskyi, Shengyi Huang, Kashif Rasul, Longhui Yu, Albert Q Jiang, Ziju Shen, and 1 others. 2024. Numinamath: The largest public dataset in ai4maths with 860k pairs of competition math problems and solutions. *Hugging Face repository*, 13(9):9.
- Pengyi Li, Matvey Skripkin, Alexander Zubrey, Andrey Kuznetsov, and Ivan Oseledets. 2025a. Confidence is all you need: Few-shot rl fine-tuning of language models. *arXiv preprint arXiv:2506.06395*.
- Tianjian Li, Yiming Zhang, Ping Yu, Swarnadeep Saha, Daniel Khashabi, Jason Weston, Jack Lanchantin, and Tianlu Wang. 2025b. Jointly reinforcing diversity and quality in language model generations. *arXiv preprint arXiv:2509.02534*.
- Yi-Chen Li, Tian Xu, Yang Yu, Xuqin Zhang, Xiong-Hui Chen, Zhongxiang Ling, Ningjing Chao, Lei Yuan, and Zhi-Hua Zhou. 2025c. Generalist reward models: Found inside large language models. *arXiv preprint arXiv:2506.23235*.
- Xun Lu. 2025. Writing-zero: Bridge the gap between non-verifiable problems and verifiable rewards. *arXiv preprint arXiv:2506.00103*.
- Jingwei Ni, Yu Fan, Vilém Zouhar, Donya Rooein, Alexander Hoyle, Mrinmaya Sachan, Markus Leipold, Dirk Hovy, and Elliott Ash. 2025. Can reasoning help large language models capture human annotator disagreement. *Preprint*.
- Mihir Prabhudesai, Lili Chen, Alex Ippoliti, Katerina Fragkiadaki, Hao Liu, and Deepak Pathak. 2025. Maximizing confidence alone improves reasoning. *arXiv preprint arXiv:2505.22660*.
- Haoran Que, Feiyu Duan, Liqun He, Yutao Mou, Wangchunshu Zhou, Jiaheng Liu, Wenge Rong, Zekun Moore Wang, Jian Yang, Ge Zhang, and 1 others. 2024. Hellobench: Evaluating long text generation capabilities of large language models. *arXiv preprint arXiv:2409.16191*.
- Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Yang Wu, and 1 others. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.
- Hongze Tan, Jianfei Pan, Jinghao Lin, Tao Chen, Zhihang Zheng, Zhihao Tang, and Haihua Yang. 2025. Gtpo and grpo-s: Token and sequence-level reward shaping with policy entropy. *arXiv preprint arXiv:2508.04349*.
- Guiyao Tie, Zeli Zhao, Dingjie Song, Fuyang Wei, Rong Zhou, Yurou Dai, Wen Yin, Zhejian Yang, Jiangyue Yan, Yao Su, and 1 others. 2025. A survey on post-training of large language models. *arXiv e-prints*, pages arXiv–2503.
- Haozhe Wang, Haoran Que, Qixin Xu, Minghao Liu, Wangchunshu Zhou, Jiazhan Feng, Wanjun Zhong, Wei Ye, Tong Yang, Wenhao Huang, and 1 others. 2025a. Reverse-engineered reasoning for open-ended generation. *arXiv preprint arXiv:2509.06160*.
- Shenzhi Wang, Le Yu, Chang Gao, Chujie Zheng, Shixuan Liu, Rui Lu, Kai Dang, Xionghui Chen, Jianxin Yang, Zhenru Zhang, and 1 others. 2025b. Beyond the 80/20 rule: High-entropy minority tokens drive effective reinforcement learning for llm reasoning. *arXiv preprint arXiv:2506.01939*.
- Tiannan Wang, Jiamin Chen, Qingrui Jia, Shuai Wang, Ruoyu Fang, Huilin Wang, Zhaowei Gao, Chunzhao Xie, Chuou Xu, Jihong Dai, and 1 others. 2024. Weaver: Foundation models for creative writing. *arXiv preprint arXiv:2401.17268*.
- Xinda Wang, Zhengxu Hou, Yangshijie Zhang, Bingren Yan, Zhibo Yang, Xingsheng Zhang, Luxi Xing, Qiang Zhou, and Chen Zhang. 2025c. Evolv: Self-evolving pairwise reasoning for story evaluation to enhance generation. *arXiv preprint arXiv:2508.06046*.
- Xiaolong Wei, Bo Lu, Xingyu Zhang, Zhejun Zhao, Dongdong Shen, Long Xia, and Dawei Yin. 2025. Igniting creative writing in small language models: Llm-as-a-judge versus multi-agent refined rewards. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 17171–17197.
- Peter West and Christopher Potts. 2025. Base models beat aligned models at randomness and creativity. *arXiv preprint arXiv:2505.00047*.
- Mian Wu, Gavin Zhang, Sewon Min, Sergey Levine, and Aviral Kumar. 2025a. Rlac: Reinforcement learning with adversarial critic for free-form generation tasks. *arXiv preprint arXiv:2511.01758*.
- Yongliang Wu, Yizhou Zhou, Zhou Ziheng, Yingzhe Peng, Xinyu Ye, Xinting Hu, Wenbo Zhu, Lu Qi, Ming-Hsuan Yang, and Xu Yang. 2025b. On the generalization of sft: A reinforcement learning perspective with reward rectification. *arXiv preprint arXiv:2508.05629*.

- Yuhao Wu, Yushi Bai, Zhiqiang Hu, Roy Ka-Wei Lee, and Juanzi Li. 2025c. Longwriter-zero: Mastering ultra-long text generation via reinforcement learning. *arXiv preprint arXiv:2506.18841*.
- Yuning Wu, Jiahao Mei, Ming Yan, Chenliang Li, Shaopeng Lai, Yuran Ren, Zijia Wang, Ji Zhang, Mengyue Wu, Qin Jin, and 1 others. 2025d. Writing-bench: A comprehensive benchmark for generative writing. *arXiv preprint arXiv:2503.05244*.
- Jianhao Yan, Yafu Li, Zican Hu, Zhi Wang, Ganqu Cui, Xiaoye Qu, Yu Cheng, and Yue Zhang. 2025. Learning to reason under off-policy guidance. *arXiv preprint arXiv:2504.14945*.
- Shuangshuang Ying, Yunwen Li, Xingwei Qu, Xin Li, Sheng Jin, Minghao Liu, Zhoufutu Wen, Xeron Du, Tianyu Zheng, Yichi Zhang, and 1 others. 2025. Beyond correctness: Evaluating subjective writing preferences across cultures. *arXiv preprint arXiv:2510.14616*.
- Yang Yue, Zhiqi Chen, Rui Lu, Andrew Zhao, Zhaokai Wang, Shiji Song, and Gao Huang. 2025. Does reinforcement learning really incentivize reasoning capacity in llms beyond the base model? *arXiv preprint arXiv:2504.13837*.
- Qingyang Zhang, Haitao Wu, Changqing Zhang, Peilin Zhao, and Yatao Bian. 2025a. Right question is already half the answer: Fully unsupervised llm reasoning incentivization. *arXiv preprint arXiv:2504.05812*.
- Yiming Zhang, Harshita Diddee, Susan Holm, Hanchen Liu, Xinyue Liu, Vinay Samuel, Barry Wang, and Daphne Ippolito. 2025b. Noveltybench: Evaluating language models for humanlike diversity. *arXiv preprint arXiv:2504.05228*.
- Xuandong Zhao, Zhewei Kang, Aosong Feng, Sergey Levine, and Dawn Song. 2025. Learning to reason without external rewards. *arXiv preprint arXiv:2505.19590*.
- Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric Xing, and 1 others. 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in neural information processing systems*, 36:46595–46623.
- Yuxin Zuo, Kaiyan Zhang, Li Sheng, Shang Qu, Ganqu Cui, Xuekai Zhu, Haozhan Li, Yuchen Zhang, Xinwei Long, Ermo Hua, and 1 others. 2025. Ttrl: Test-time reinforcement learning. *arXiv preprint arXiv:2504.16084*.

A Additional Derivations and Proofs

A.1 Linearity of the one step objective and deterministic optimizer

We justify the claim in the main text that the one-step objective in Eq. (3) is linear in the policy $\pi(\cdot | s)$ for a fixed context s , and that an optimum is attained by a deterministic distribution. Fix a context s and let \mathcal{V} be the vocabulary. Let

$$f(v) := \log \pi_s(v | s), v \in \mathcal{V}. \quad (19)$$

The one-step objective can be written as

$$J(\pi; s) := \mathbb{E}_{y \sim \pi(\cdot | s)}[f(y)] = \sum_{v \in \mathcal{V}} \pi(v | s) f(v), \quad (20)$$

where $\pi(\cdot | s)$ ranges over the probability simplex

$$\Delta(\mathcal{V}) := \left\{ \pi : \mathcal{V} \rightarrow [0, 1] \mid \sum_{v \in \mathcal{V}} \pi(v) = 1 \right\}. \quad (21)$$

Lemma A.1 (Linearity in π). *For any two distributions $\pi_1, \pi_2 \in \Delta(\mathcal{V})$ and any $\lambda \in [0, 1]$,*

$$J(\lambda\pi_1 + (1 - \lambda)\pi_2; s) = \lambda J(\pi_1; s) + (1 - \lambda)J(\pi_2; s). \quad (22)$$

Proof. By Eq. (20),

$$\begin{aligned} & J(\lambda\pi_1 + (1 - \lambda)\pi_2; s) \\ &= \sum_{v \in \mathcal{V}} (\lambda\pi_1(v) + (1 - \lambda)\pi_2(v)) f(v) \\ &= \lambda \sum_{v \in \mathcal{V}} \pi_1(v) f(v) + (1 - \lambda) \sum_{v \in \mathcal{V}} \pi_2(v) f(v) \\ &= \lambda J(\pi_1; s) + (1 - \lambda)J(\pi_2; s). \end{aligned} \quad (23)$$

□

Lemma A.2 (Deterministic optimizer). *The maximum of $J(\pi; s)$ over $\Delta(\mathcal{V})$ is achieved by a deterministic distribution. In particular,*

$$\max_{\pi \in \Delta(\mathcal{V})} J(\pi; s) = \max_{v \in \mathcal{V}} f(v), \quad (24)$$

and any optimizer is of the form $\pi^(\cdot | s) = \delta_{v^*}(\cdot)$ where δ_{v^*} denotes the Dirac distribution at $v^* \in \arg \max_{v \in \mathcal{V}} f(v)$.*

Proof. Let $M := \max_{v \in \mathcal{V}} f(v)$. For any $\pi \in \Delta(\mathcal{V})$,

$$J(\pi; s) = \sum_{v \in \mathcal{V}} \pi(v) f(v) \leq \sum_{v \in \mathcal{V}} \pi(v) M = M, \quad (25)$$

since $\sum_v \pi(v) = 1$. Equality holds iff π places all mass on the set of maximizers of f , i.e., $\pi(v) = 0$ for all v such that $f(v) < M$. In particular, choosing any single maximizer v^* and setting $\pi(v^*) = 1$ achieves $J(\pi; s) = M$. □

The above results demonstrate that optimizing the endogenous reward without regularization inevitably leads to deterministic policies, providing theoretical support for the entropy collapse phenomenon observed in practice.

A.2 Coverage of the information gain correction

The main text uses the intuition that TCER does not only suppress saturated tokens but also affects *how broadly* the information-gain signal can contribute within a context. We formalize this via a context-wise coverage quantity that measures the expected activity of the gating function.

Definition. Given a context s , define the gate $w(y | s) = (1 - \pi_s(y | s))^\lambda \in [0, 1]$ and the context-wise coverage

$$\begin{aligned} S(s) &:= \mathbb{E}_{y \sim \pi_s(\cdot | s)}[w(y | s)] \\ &= \sum_{v \in \mathcal{V}} \pi_s(v | s) (1 - \pi_s(v | s))^\lambda. \end{aligned} \quad (26)$$

The coverage $S(s)$ quantifies how much the correction term remains active on average under the distribution $\pi_s(\cdot | s)$.

Lemma A.3 (Range and deterministic case). *For any context s , $S(s) \in [0, 1]$. Moreover, if $\pi_s(\cdot | s)$ is deterministic (i.e., $\exists v^* \text{ s.t. } \pi_s(v^* | s) = 1$), then $S(s) = 0$.*

Proof. Since $w(y | s) \in [0, 1]$ for all y , its expectation under any distribution lies in $[0, 1]$, hence $S(s) \in [0, 1]$. If π_s is deterministic, then with probability one the sampled token has $p = \pi_s(y | s) = 1$, so $w = (1 - 1)^\lambda = 0$ almost surely, yielding $S(s) = 0$. □

Lemma A.4 (Coverage lower bound via non-saturated mass). *Fix $\delta \in (0, 1]$ and define the non-saturated set*

$$A_\delta(s) := \{y \in \mathcal{V} : \pi_s(y | s) \leq 1 - \delta\}. \quad (27)$$

Then

$$S(s) \geq \delta^\lambda \cdot \pi_s(A_\delta(s) | s). \quad (28)$$

Proof. For any $y \in A_\delta(s)$, we have $1 - \pi_s(y | s) \geq \delta$, hence $w(y | s) \geq \delta^\lambda$. Therefore,

$$\begin{aligned} S(s) &= \mathbb{E}_{y \sim \pi_s} [w(y | s)] \\ &\geq \mathbb{E}_{y \sim \pi_s} [w(y | s) \mathbf{1}\{y \in A_\delta(s)\}] \\ &\geq \mathbb{E}_{y \sim \pi_s} [\delta^\lambda \mathbf{1}\{y \in A_\delta(s)\}] \\ &= \delta^\lambda \pi_s(A_\delta(s) | s). \end{aligned} \quad (29)$$

□

Eq. (28) shows that $S(s)$ is large whenever $\pi_s(\cdot | s)$ assigns substantial probability mass to non-saturated alternatives. In such contexts, the gate is less likely to shut off the information-gain term, so the ϕ -based correction can apply to a broader portion of likely tokens under π_s . This provides a quantitative characterization of how TCER maintains active correction signals in high-entropy regimes.

A.3 Closed-form derivative of the gated correction

For completeness, we provide the derivative of $C(p)$ with respect to p , which can be useful for analyzing smoothness:

$$C(p) = (1 - p)^\lambda \log \frac{p + \varepsilon}{q + \varepsilon}. \quad (30)$$

A direct differentiation gives

$$\frac{d}{dp} C(p) = -\lambda(1 - p)^{\lambda-1} \log \frac{p + \varepsilon}{q + \varepsilon} + \frac{(1 - p)^\lambda}{p + \varepsilon}. \quad (31)$$

Under $\varepsilon > 0$ and $p \in [0, 1]$, both $\log \frac{p + \varepsilon}{q + \varepsilon}$ and $(p + \varepsilon)^{-1}$ are bounded; thus the derivative remains bounded for $\lambda \geq 1$ and exhibits the expected suppression near $p \approx 1$ when $\lambda > 1$.

While the derivative vanishes as $p \rightarrow 1$, suggesting diminishing correction for high-confidence tokens, the entropy-preserving effect of TCER operates through the sequence-averaged reward structure. The sequence reward is computed as:

$$r_t(\tau) = \frac{1}{T} \sum_{t=1}^T [\log p_t + k \cdot C(p_t)] \quad (32)$$

This averaged reward is then applied uniformly to all tokens in the sequence during gradient updates.

To quantify the diversity preference, consider the expected reward difference between diverse and deterministic sequences. For a deterministic sequence where all $p_t \approx 1 - \epsilon$ for small ϵ :

$$r_{\text{det}} \approx \log(1 - \epsilon) + k \cdot \epsilon^\lambda \cdot \log \frac{1 - \epsilon + \varepsilon}{q + \varepsilon} \approx -\epsilon \quad (33)$$

For a diverse sequence with confidence variance $\sigma^2 = \text{Var}_t[p_t]$, containing tokens with $p_t \in [1 - 2\sigma, 1]$:

$$r_{\text{diverse}} \geq \mathbb{E}[\log p_t] + k \cdot \mathbb{E}[(1 - p_t)^\lambda \cdot \log \frac{p_t + \varepsilon}{q_t + \varepsilon}] \quad (34)$$

The second term is substantial for medium-confidence tokens. Using the lower bound from Eq. (15), tokens with $p_t \leq 1 - \sigma$ contribute at least:

$$k \cdot \sigma^\lambda \left| \log \frac{p_t + \varepsilon}{q_t + \varepsilon} \right| \quad (35)$$

Therefore, the reward advantage of diverse sequences is:

$$\Delta r = r_{\text{diverse}} - r_{\text{det}} \geq k \cdot \sigma^{\lambda+1} \cdot \mathbb{E} \left[\left| \log \frac{p + \varepsilon}{q + \varepsilon} \right| \right] > 0 \quad (36)$$

This quantifies how the gating mechanism creates systematic preference for diversity: sequences with higher confidence variance receive higher average rewards, leading the policy to learn diverse outputs rather than collapse to deterministic patterns. The parameter k controls the strength of this diversity preference, with larger k providing stronger protection against entropy collapse.

B Experimental Details

B.1 Implementation Details

Across all tasks, we adopt a unified training pipeline consisting of SFT followed by RL with Reference-Augmented GRPO. For writing, we experiment with Qwen2.5-7B-Instruct, Qwen3-8B, and Llama3.1-8B-Instruct; for mathematical reasoning, we use Qwen2.5-Math-7B and Llama3.1-8B-Instruct. On writing datasets (DeepWriter and LongWriter), we perform SFT for 1 epoch with learning rate 2×10^{-5} and batch size 8, and then run RL for 1,000 update steps. On math tasks, we run SFT for 3 epochs and RL for 500 update steps with learning rate 1×10^{-5} and batch size 64.

During RL, for each prompt we sample a group of $G = 8$ rollouts and augment it with the dataset reference completion, using the length-normalized average token reward as the sequence reward. Following the theoretical requirement from EndoR that reward models should remain frozen during RL training to prevent distribution shift, we keep both π_s and π_b fixed throughout optimization. GRPO uses a KL penalty of coefficient $\beta = 0.001$. TCER is computed using a correction scale $k = 3$ and gating exponent $\lambda = 2$, with ε -smoothing set to

$\varepsilon = 1 \times 10^{-5}$ in the log-ratio for numerical stability. For generation during RL training, we use temperature $\tau = 0.7$ and maximum length of 8000 tokens. All training data and evaluation benchmarks employ their original open-source prompts without modification. We will release our code upon acceptance to facilitate reproducibility. All experiments are run on a cluster of $16 \times$ NVIDIA H20 GPUs; to reduce RL-time overhead, the actor runs on the main training GPUs while the frozen π_s and π_b are served on an additional $2 \times$ H20 GPUs for reward computation.

B.2 Training Details

During RL fine-tuning, we log training-time diagnostics to characterize how different endogenous rewards shape policy behavior and to diagnose degeneration. Specifically, we track (i) the sequence-level reward (computed from token-level rewards via length-normalized averaging, consistent with Eq. (16)) and (ii) the actor’s token-level entropy over generated positions. For each configuration, we report EndoR reward trajectories using the *EndoR reward* logs and TCER reward trajectories using the *TCER reward* logs; the corresponding entropy comparison between EndoR and TCER is plotted jointly in the *EndoR-vs-TCER entropy* figures. Figures 5 and 6 summarize these dynamics for writing and mathematical reasoning, respectively.

Across both writing and mathematical reasoning tasks, EndoR exhibits a consistent tendency toward entropy reduction as optimization proceeds, indicating a drift toward low-entropy, high-confidence continuations. While this behavior can increase reward, it aligns with the degeneration mechanism discussed in Section 3.1, where confidence maximization encourages concentration on saturated high-probability tokens. In contrast, TCER achieves comparable or higher reward with improved stability, while consistently exhibiting slower entropy decay and higher entropy levels than EndoR throughout training. This suggests that the triviality-corrected information-gain term mitigates entropy collapse by reallocating learning signal away from saturated high-probability tokens and toward more informative alternatives, leading to more balanced RL optimization dynamics.

B.3 Hyperparameter Details

We study the sensitivity of TCER to its two hyperparameters: the correction scale k and the gating

exponent λ . Table 6 reports results on the DeepWriter dataset with Qwen2.5-7B-Instruct, where we sweep a small set of representative (k, λ) values.

Overall, the method is stable across the tested range. Increasing k from 1 to 3 consistently improves performance, and $(k = 3, \lambda = 2)$ achieves the best scores on all benchmarks. Further increasing k to 4 yields slightly lower results, suggesting diminishing returns when the correction becomes overly strong. Fixing $k = 3$, varying λ in $\{1, 2, 3\}$ produces only minor differences, with $\lambda = 2$ consistently performing best. Based on this analysis, we use $(k = 3, \lambda = 2)$ as the default setting in our main experiments unless stated otherwise.

C Case Study

C.1 Reward Cases

We present reward case studies from the *Validation of Quality Enhancement* pipeline. For readability, we render the prompt, the SFT output, and the sentence-level reward table as screenshots (especially for Chinese content). Figures 7–9 show (i) the writing-critique prompt, (ii) the corresponding SFT output, and (iii) the sentence-level reward comparison between EndoR and TCER for the same output.

For each line, the reward table reports EndoR and TCER scores, their difference $\Delta = \text{TCER} - \text{EndoR}$, and whether the line is highlighted by our closed-source judge panel (*Selected=True*). Qualitatively, judge-highlighted lines tend to receive a more favorable reward difference under TCER than under EndoR, while many non-selected lines exhibit small or negative Δ , particularly for structural markers or low-information segments. Notably, degenerate artifacts such as empty lines are heavily down-weighted by TCER in the English case, illustrating that the triviality correction can suppress non-informative content. Overall, these cases align with the role of TCER as a reweighting mechanism that shifts learning signal away from saturated or trivial tokens and toward more informative spans, improving the alignment between training reward and perceived writing quality.

C.2 Writing Cases

We present qualitative writing cases sampled from WritingBench, together with scores from our closed-source judge panel, and include both Chinese and English examples. For each case, we use

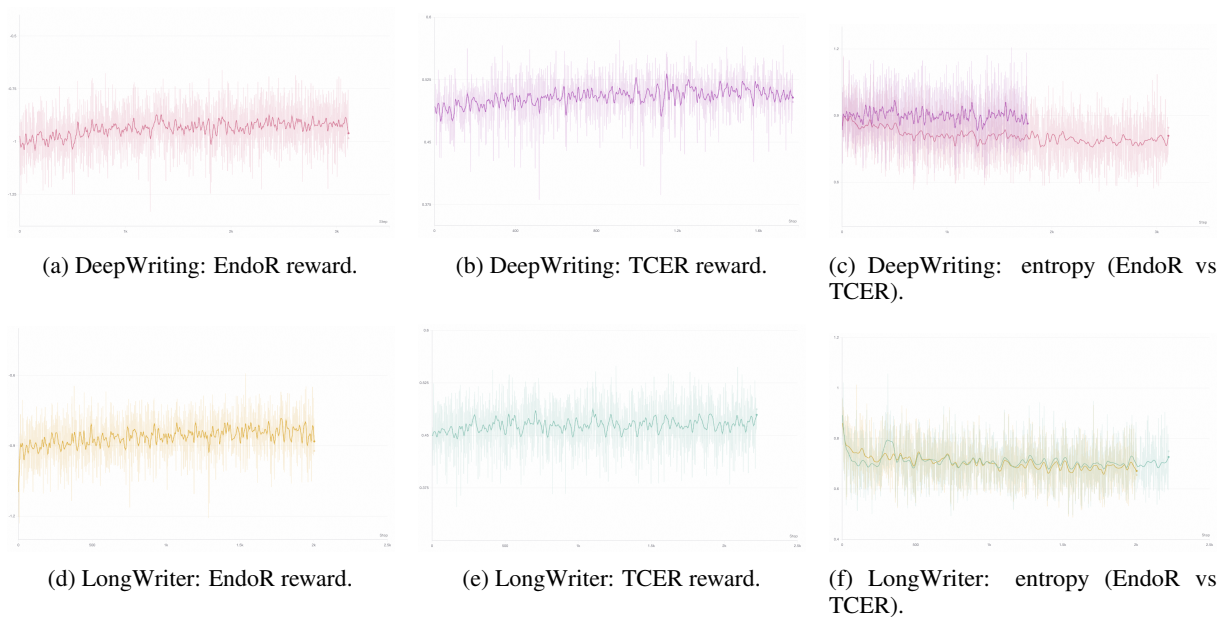


Figure 5: RL training dynamics on writing tasks. For each dataset, we report the EndoR reward trajectory, the TCER reward trajectory, and a direct entropy comparison between EndoR and TCER.

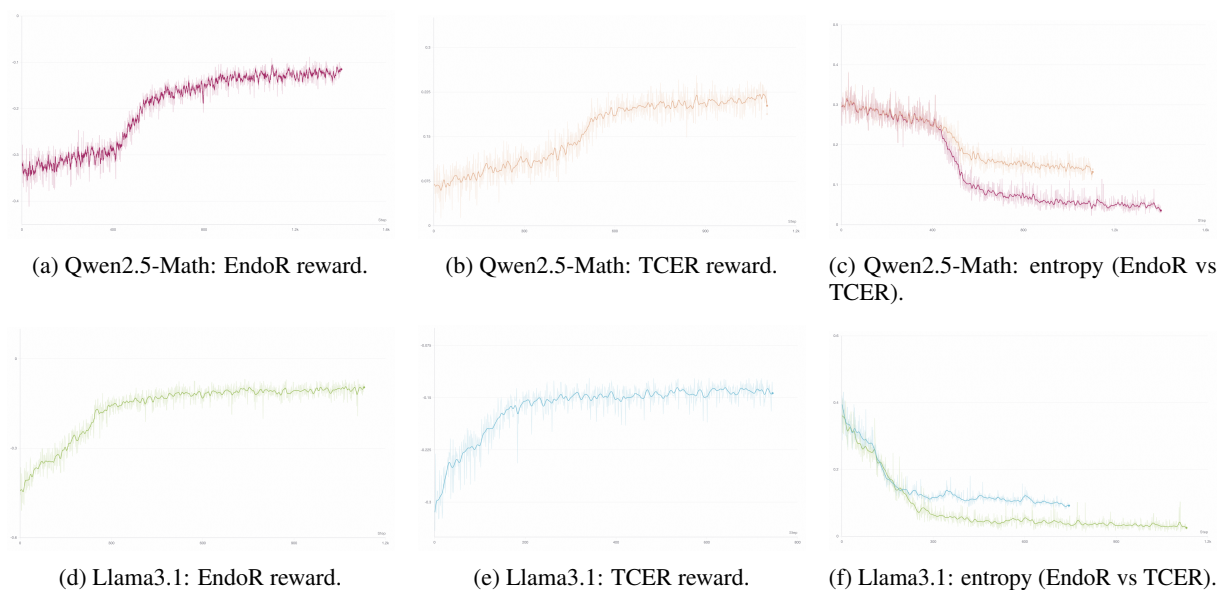


Figure 6: RL training dynamics on mathematical reasoning tasks. We report EndoR and TCER reward trajectories together with entropy comparisons for each model configuration.

the same prompt and compare responses generated by three models: the SFT policy, the EndoR-trained policy, and the TCER-trained policy. All prompts and outputs are shown as screenshots for consistency and readability.

Across the sampled cases, the SFT model produces fluent and on-topic responses with standard patterns. After RL with EndoR, responses can become more verbose and rely more on repetitive, high-confidence phrasing, which is associated with weaker judge scores in these examples. In contrast,

the TCER-trained model typically maintains coherence while reducing repetition and improving expressive diversity, leading to higher judge scores in our displayed cases. These qualitative comparisons provide an intuitive complement to our quantitative results, illustrating how TCER mitigates degeneration tendencies induced by confidence-driven rewards and improves perceived writing quality.

Configuration	LB	HB-A	HB-B	WB-A	WB-B	WB-C	WB-D	WB-E	WB-F
$(k = 1, \lambda = 1)$	83.7	78.5	83.9	70.8	71.2	68.6	65.3	71.5	69.0
$(k = 2, \lambda = 2)$	85.1	79.8	85.2	71.9	72.4	69.7	66.5	72.6	70.2
$(k = 3, \lambda = 2)$	86.3	81.1	86.4	72.7	73.1	70.8	67.8	73.4	71.1
$(k = 4, \lambda = 2)$	85.9	80.7	86.0	72.4	72.8	70.4	67.3	73.0	70.7
$(k = 3, \lambda = 1)$	85.4	80.2	85.6	72.1	72.5	70.1	67.0	72.8	70.4
$(k = 3, \lambda = 3)$	85.8	80.6	85.9	72.3	72.7	70.3	67.2	72.9	70.6

Table 6: Hyperparameter sensitivity analysis on the DeepWriter dataset with Qwen2.5-7B-Instruct. Results show win rates or benchmark scores across LongBench (LB), HelloBench subsets (HB-A: Open QA, HB-B: Text Generation), and WritingBench domains (WB-A through WB-F representing different professional writing categories). The optimal configuration ($k = 3, \lambda = 2$) achieves consistent improvements across all metrics.

C.3 Prompt Cases

We further present prompt cases illustrating how highlighted sentences are selected via a multi-judge protocol. Specifically, we use a shared sentence-selection prompt to query three proprietary LLM judges (GPT-4o, Claude Opus 4, and Gemini 2.5 Pro). Each judge independently selects a set of high-quality sentences from the same model output, with a minimum of five selections required to ensure sufficient coverage.

To obtain a stable final highlight set, we aggregate the three judge outputs using Gemini 2.5 Pro. The aggregation follows an agreement-based strategy: sentences selected by multiple judges are preferred, and if fewer than five sentences satisfy this criterion, additional candidates are included based on agreement strength until at least five highlighted sentences are obtained. This design reduces idiosyncratic bias from any single judge while preserving consistent quality signals. While using LLM judges may introduce evaluation biases, this multi-judge approach represents the current best practice for scalable quality assessment in the absence of large-scale human annotation.

Figures 18 and 19 show the sentence-selection prompt used by individual judges and the aggregation prompt used by Gemini 2.5 Pro, respectively. These prompt cases support our qualitative analyses by providing a transparent and reproducible mechanism for constructing highlighted sentence sets used in reward validation and writing case studies.

Prompt
Instruction: Write a story that falls into the genre of Action, Drama, Superhero Fiction, based on the following prompt:
 You are a nice person, but your superpower is that you instinctively know exactly what to say to someone to crush them. You're very effective in throwing supervillains off their game, but your fellow heroes always feel really uncomfortable watching you work.

Model Output

Zephyr stood over the defeated supervillain, gloating triumphantly. "Looks like this city's problems are solved... again. Until next time, scum." He turned to face the rest of the superhero team, who had just arrived on the scene.

"Another successful victory, thanks to our verbal virtuoso, Zephyr!" Captain Valiant proclaimed. "His masterful use of words truly brought this villain to his knees."

Zephyr smirked arrogantly. "Oh, don't give me that credit, Captain. I'm not here to take any praise. Just stating the facts."

The other heroes exchanged uneasy glances. They'd all heard the tales of Zephyr's unique power - the ability to verbally eviscerate anyone he chose, shattering their confidence and will in one scathing remark. It made him incredibly effective against supervillains... but also rather abrasive to work with.

"I just don't see why we need someone like him on the team," muttered Microbeam, his fingers twitching nervously. "All that negative energy, it's toxic!"

"He's right," chimed in Speedster. "I mean, sure, his power is useful, but man, does he have to go so far over the top with it? Even the bad guys get hurt by what he says."

"Relax, you two," Zephyr chuckled. "I'm not going to hurt anyone's feelings. That's just my job. Besides, these villains deserve it. They're nothing but evil."

"Don't make us turn off our microphones, Zeph," interrupted Paragon, her eyes narrowing. "I know you love verbally eviscerating people, but could you maybe tone it down a bit? The rest of us have feelings too, you know."

Zephyr rolled his eyes. "Fine, fine. I'll keep my trashy talk to a minimum. But don't come crying to me when some washed-up criminal gets a complex and decides to take it out on innocent bystanders."

The heroes sighed collectively, glad to get that settled. They knew Zephyr was just being his usual self - cruel, sarcastic, and blunt to a fault. But hey, nobody's perfect, right?

As they flew off to their next adventure, Zephyr couldn't help but feel a little down. Sure, he knew he wasn't the most team-friendly guy around... but still. Watching his friends squirm under his verbal barbs was almost as much fun as crushing the villains themselves. He'd have to work on that.

But first, they had to stop a meteor from hitting the city. Again. Just another day in the life of an ace superhero.

Zephyr grinned as the meteor came into view. "Alright, you ugly rock, it's time to go boom!"

And with a mighty yell, the heroes launched themselves towards the sky, ready to save the day... one snide remark at a time.

Note: This output is shown verbatim; the reward table below reports sentence-level EndoR/TCER scores aligned to the segmented lines.

Figure 10: English reward case study: prompt and model output.

Reward Table (Full 1-19 Lines) Δ = TCER - EndoR Selected=True highlighted Empty lines marked

Rows with Selected=True are shaded. Rows with empty/degenerate content are tinted red for visibility.

ID	Sentence (raw line)	EndoR	TCER	Δ	Selected
1	Zephyr stood over the defeated supervillain, gloating triumphantly. "Looks like this city's problems are solved... again. Until next time, scum." He turned to face the rest of the superhero team, who had just arrived on the scene.	-2.757	-3.085	-0.248	False
2	"Another successful victory, thanks to our verbal virtuoso, Zephyr!"	-3.431	-2.489	1.022	False
3	"Captain Valiant proclaimed. "His masterful use of words truly brought this villain to his knees."	-1.365	-0.396	0.969	False
4	Zephyr smirked arrogantly. "Oh, don't give me that credit, Captain. I'm not here to take any praise. Just stating the facts."	-1.583	-1.383	0.280	False
5	The other heroes exchanged uneasy glances. They'd all heard the tales of Zephyr's unique power - the ability to verbally eviscerate anyone he chose, shattering their confidence and will in one scathing remark. It made him incredibly effective against supervillains... but also rather abrasive to work with.	-1.697	-1.778	-0.081	True
6	"I just don't see why we need someone like him on the team," muttered Microbeam, his fingers twitching nervously. "All that negative energy, it's toxic!"	-1.385	-0.557	0.829	False
7	[EMPTY]	-9.076	-29.525	-20.449	False
8	"He's right," chimed in Speedster. "I mean, sure, his power is useful, but man, does he have to go so far over the top with it?"	-1.044	-1.169	-0.125	False
9	Even the bad guys get hurt by what he says."	-2.317	-1.994	0.323	True
10	"Relax, you two," Zephyr chuckled. "I'm not going to hurt anyone's feelings. That's just my job. Besides, these villains deserve it. They're nothing but evil..."	-1.680	-2.027	-0.347	False
11	"Don't make us turn off our microphones, Zeph," interrupted Paragon, her eyes narrowing. "I know you love verbally eviscerating people, but could you maybe tone it down a bit?"	-1.334	-0.131	1.203	True
12	The rest of us have feelings too, you know."	-0.863	0.658	1.521	False
13	Zephyr rolled his eyes. "Fine, fine. I'll keep my trashy talk to a minimum. But don't come crying to me when some washed-up criminal gets a complex and decides to take it out on innocent bystanders."	-0.962	-0.549	0.413	False
14	The heroes sighed collectively, glad to get that settled. They knew Zephyr was just being his usual self - cruel, sarcastic, and blunt to a fault. But hey, nobody's perfect, right?"	-1.077	-0.414	0.664	False
15	As they flew off to their next adventure, Zephyr couldn't help but feel a little down. Sure, he knew he wasn't the most team-friendly guy around... but still. Watching his friends squirm under his verbal barbs was almost as much fun as crushing the villains themselves. He'd have to work on that.	-1.126	-1.012	0.114	True
16	But first, they had to stop a meteor from hitting the city. Again. Just another day in the life of an ace superhero.	-2.084	-2.478	-0.394	False
17	Zephyr grinned as the meteor came into view. "Alright, you ugly rock, it's time to go boom!"	-1.722	-0.945	0.777	False
18	[EMPTY]	-2.309	-4.514	-2.206	False
19	And with a mighty yell, the heroes launched themselves towards the sky, ready to save the day... one snide remark at a time.	-1.138	-0.684	0.454	True

Figure 11: English reward case study: sentence-level EndoR vs. TCER rewards. The table reports $\Delta = \text{TCER} - \text{EndoR}$ and whether each line is selected by the judge panel.

Instruction: 以北京胡同为主题写一篇作文。

Output Text:

胡同记忆\n\n记忆就像一本翻不完的老相册，每一页都记录着我们生命中最珍贵的片段。翻开我心中的那本相册，总有一部分被老城区的胡同所占据。那里不仅是我的童年乐园，更是我心灵的归宿。如今虽已离开多年，那青砖灰瓦间的生活场景，那胡同里的人情味，却依然清晰地印在我的脑海里，如同昨日般鲜活。清晨的胡同总是最先苏醒的，热乎乎的豆汁儿，刚出锅的炸酱面，卖豆汁儿的吆喝声从街角传来，那酸子豆香和面糊的甜味顺着风飘进每家每户。紧接着是卖糖葫芦的，酸酸甜甜的糖葫芦，脆皮裹着红果儿，一个五分，那晶莹剔透的糖葫芦，在冬日的阳光下闪闪发光，吸引着孩子们的目光。煎饼、油条、豆浆，现实地吃咧咧，卖早点的小贩背着一个大铁桶，桶上系着几根绳子，每根绳子上串着一个小板凳，叮叮当当，敲打着发出清脆的声音。胡同里的叫卖声此起彼伏，像是演奏着一曲独特的市井交响乐，唤醒了整个沉睡的社区。在老北京的胡同里，邻里之间的关系如同一家人般亲密。夏日午后，老人们在胡同里晒太阳，坐在自家门前的马扎上聊天，话题从天聊到庄稼，从国家大事聊到街坊邻居的家长里短。有时，他们会搬出小板凳，围坐在槐树下，玩起“打十张”、“斗地主”等牌戏。老张，你这牌打得好好洗洗，不然我们都不给你玩了。大家边玩边聊，那份融洽与和谐，是现在小区里永远学不来的。孩子们在胡同里追逐打闹，嬉笑声此起彼伏。我家那胡同里住着李奶奶、王爷爷这样的老人，王爷爷这样的老人，他们没事就爱串门，谁家做了好吃的，第一口总是先给邻居尝尝。来来来，尝尝我做的豆汁儿，这可是老传统。孩子们在胡同里无忧无虑，我们最喜欢玩的是踢毽子，一人一个彩色毽子，你踢我一下，我踢你一下，看谁踢得远踢得稳。有时候几个孩子会聚在一起，选出两个“门神”，把守“门”，其余的孩子分成两队，一队当“清军”，一队当“义和团”，在院子里追逐打闹，喊得“捉人啦”、“快跑啊”。记得有一年中秋节，我们几个孩子在胡同里捉迷藏，最后躲在一座老宅的后院里。那座老宅已经没人住很多年了，听说里面闹鬼，所有人在那处听见一个老人在屋檐下咳嗽，吓得赶紧跑得更快了。后来我们商量好，谁也不许告诉大人，就当没这回事。那晚的月亮特别圆，特别亮，照在湿漉漉的青石板上，我们的心里也充满了喜悦。春天，胡同里的玉兰树开花了，洁白的花瓣飘落一地，仿佛下了一场小雪。孩子们捡起花瓣，开心地玩着。夏天，老槐树撑起一片绿荫，蝉儿在树上鸣叫，仿佛在演奏一曲夏日交响乐。傍晚时分，萤火虫在胡同里飞舞，像是点亮了无数小灯笼。秋天，银杏树叶金黄，铺满了整个胡同，走在上面沙沙作响。冬天，雪花飘飘，胡同变成了一条白色的长龙，清晨的阳光照在雪地上，闪闪发光。每到春节，胡同里更是热闹非凡，家家户户贴春联，挂红灯笼，鞭炮声此起彼伏，空气中弥漫着香油和烟火的气息，让人感到浓浓的年味。人们说起浓浓的胡同记忆，那就要数我十岁那年的春节了。那年大年三十晚上，我们小孩穿上新衣服，给长辈拜年，讨红包。记得有一年，我表哥结婚，整个胡同都热闹起来了，各家各户都挂起了红灯笼，放起了鞭炮。婚礼队伍从我家门前经过，我站在门框边，看着花轿缓缓前行，心里既兴奋又羡慕。新郎穿着大红喜服，看起来格外帅气。婚礼结束后，亲戚们在院子里摆起长桌，摆满各种好吃的，大家边吃边聊，好不热闹。那一刻，我真切地感受到了什么是家和万事兴。饭桌上，长辈们举杯祝福新人，祝福我们大家，新年快乐，万事如意。我看看爷爷奶奶那幸福的笑容，心里暖暖的，知道这就是我生长的地方，这就是我的家。如今，每当我站在高楼林立的都市中，闭上眼睛，脑海中总会浮现出那条蜿蜒的胡同，那些熟悉的声音，那些亲切的面孔，那些纯真的笑容。胡同里的日子虽然简朴，却充满了人情味和幸福感。离开胡同多年，我始终怀念那里的一草一木，一砖一瓦。也许，真正的怀念不是对过去的留恋，而是对那段简单而纯粹生活的向往。正如老北京有句俗话说得好，人走茶不凉，胡同情更长。无论走到哪里，那条充满爱的胡同，永远是我心中最温暖的港湾。

Model Judge

score: 8, reason: 这篇散文准确且生动地描绘了北京胡同文化元素，包含丰富的细节和真实的场景描写。叫卖声的描述如卖豆汁儿的吆喝声、热乎乎的豆汁儿、刚出锅的炸酱面、酸酸甜甜的糖葫芦、脆皮裹着红果儿，一个五分非常真实。民俗活动如老人们在槐树下打十张、斗地主以及长辈们在婚礼后院子里摆长桌吃席的场景展现了浓厚的北京本土生活。传统游戏如踢毽子、清军vs义和团的游戏、捉迷藏等也很生动。四季变化的描写如玉兰雨、银杏树叶铺满胡同等细节充满画面感。文中穿插了老北京方言和俗语如来来来、尝尝我做的豆汁儿，这可不是老传统以及人走茶不凉、胡同情更长，增添了地域特色。文章整体展现了作者对胡同的亲身经历，虽然部分描写如五毛钱一个的糖葫芦价格设定可能有些不够准确，但整体上呈现了胡同生活的真实面貌和人情味。北京方言与语言风格的应用: score: 6, reason: 文章使用了适量的北京方言表达，如豆汁儿、糖葫芦、小贩、马扎等词汇，基本准确地展现了老北京生活的特色。语言风格整体朴实细腻，如清晨叫卖声的描述让人感受到市井气息。然而，方言的使用相对有限，主要集中在名词层面，缺乏更丰富的北京口语表达方式如句式特点。如吆喝声的描述较为通用，没有体现出独特的北京韵味。文中的方言与叙事的融合不够自然，如豆汁儿、现实地吃咧咧等词语使用显得点到即止，没有深入展现北京话的韵味。虽然有一些典型表达如热乎乎、马扎、打七张等，但整体上北京方言的使用较为克制，未能充分展现老北京胡同生活的语言特色。文章结尾的人走茶不凉，胡同情更长虽是点睛之笔，但整体上方言融入度仍有提升空间。感官描写与情感表达: score: 7, reason: 散中的感官描写丰富多样，涵盖了多种感官体验。视觉描写如晶莹剔透的糖葫芦，在冬日的阳光下闪闪发光、玉兰树开花了、洁白的花瓣飘落一地；听觉描写如叮叮当当，敲打着发出清脆的声音、吆喝声此起彼伏；嗅觉描写如那酸子豆香和面糊的甜味、空气中弥漫着香油和烟火的气息；触觉描写如湿漉漉的青石板路，情感表达真挚深刻。通过我们的内心也充满了喜悦、那一刻，我真切地感受到了什么是家和万事兴等描写传递出浓厚的怀旧情感。然而，某些感官描写虽丰富，但某些描写略显概括而非深入细节，如四季变换的描述相对笼统。怀旧情感真挚但层次感可进一步提升，比如可以更深入探讨离开胡同后的复杂情感变化。整体而言，文章营造浓厚的怀旧氛围，能引起读者共鸣，但尚未达到极致的感染力和沉浸感。结构完整性与主题呼应: score: 9, reason: 这篇散文结构布局精妙，严格遵循了要求的段落划分，共分为5个自然段，结构清晰合理。字数控制在2000字左右，符合要求。开头以富有感染力的方式点明了记忆主题，记忆就像一本翻不完的老相册，每一页都记录着我们生命中最珍贵的片段，并引出对胡同生活的怀念。内容完整覆盖了所有要求中的元素：清晨的叫卖声、邻里互动、童年游戏、四季变换和难忘事件。结尾部分人走茶不凉、胡同情更长巧妙呼应了开头的记忆主题，将胡同比作心中最温暖的港湾，形成了完美的闭环。整体结构浑然一体，段落之间过渡自然，内容承接流畅。开头点明主题，结尾深情呼应，形成了完整而精彩的结构闭环。内容全面性与场景描绘: score: 9, reason: 这篇散文完美覆盖了所有要求中的元素，内容完整覆盖了所有要求中的元素。第二段生动细致地描写了清晨的叫卖声，包含豆汁儿、糖葫芦、煎饼等具体物品的吆喝声和声音细节(如叮当声)。第二段并生动刻画了邻里互动，从老人们在胡同里晒太阳、打牌到串门分享食物，体现了深厚的情感。第四段全面展现了传统游戏，包括踢毽子和角色扮演游戏，并通过中秋捉迷藏的具体事件增强了代入感。第五段四季景致描绘极为生动，从玉兰雨、蝉鸣、银杏落叶到雪景，季节变化一目了然。第六段以表哥结婚为具体事件，细腻描写了婚礼场景和亲情氛围。全文语言朴实细腻，穿插了老北京方言(如豆汁儿、小贩、马扎等)，开头点明记忆主题，结尾呼应开头形成完整闭环。散文整体结构如行云流水，完美再现胡同生活的方方面面，体现了深厚的历史文化积淀。]]

Figure 14: Writing case (TCER). Response generated by the TCER-trained model under the same prompt.

Prompt (Writing Critique)
Instruction: Write a detailed work report on analyzing the handling of environmental protection cases by local courts.

Output Text:
Work Report: Analysis of Environmental Protection Case Handling by Local Courts
Environmental protection has become increasingly complex in today's rapidly industrializing society. As ecological issues continue to rise in prominence, the role of judicial institutions in resolving environmental disputes has gained unprecedented importance. This report aims to analyze the handling of environmental protection cases by local courts, examining both successful practices and areas requiring improvement. By integrating judicial precedents with local environmental policies, we seek to identify best practices and propose actionable suggestions for enhancing the effectiveness of environmental justice mechanisms. The analysis is particularly relevant given the growing frequency of environmental incidents and the increasing demands placed on local courts to resolve these disputes efficiently and fairly.
Case 1: Green Valley Chemical Co., Ltd. v. Municipal Environmental Protection Bureau of City A
This case involves a dispute over administrative penalties imposed by the Municipal Environmental Protection Bureau of City A on Green Valley Chemical Co., Ltd. for alleged water pollution violations. The chemical company challenged the administrative penalty decision on the grounds of procedural irregularities and insufficient evidence.
Case 2: Harmonious Development Co., Ltd. v. District People's Court of County B
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County B regarding the approval of an industrial project and its environmental impact assessment.
Case 3: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City C
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City C regarding environmental liability and remediation obligations following a pollution incident.
Case 4: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City D
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City D regarding the approval of a remediation project.
Case 5: Harmonious Development Co., Ltd. v. District People's Court of County E
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County E regarding the approval of an industrial project and its environmental impact assessment.
Case 6: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City F
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City F regarding environmental liability and remediation obligations following a pollution incident.
Case 7: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City G
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City G regarding the approval of a remediation project.
Case 8: Harmonious Development Co., Ltd. v. District People's Court of County H
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County H regarding the approval of an industrial project and its environmental impact assessment.
Case 9: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City I
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City I regarding environmental liability and remediation obligations following a pollution incident.
Case 10: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City J
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City J regarding the approval of a remediation project.
Case 11: Harmonious Development Co., Ltd. v. District People's Court of County K
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County K regarding the approval of an industrial project and its environmental impact assessment.
Case 12: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City L
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City L regarding environmental liability and remediation obligations following a pollution incident.
Case 13: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City M
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City M regarding the approval of a remediation project.
Case 14: Harmonious Development Co., Ltd. v. District People's Court of County N
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County N regarding the approval of an industrial project and its environmental impact assessment.
Case 15: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City O
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City O regarding environmental liability and remediation obligations following a pollution incident.
Case 16: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City P
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City P regarding the approval of a remediation project.
Case 17: Harmonious Development Co., Ltd. v. District People's Court of County Q
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County Q regarding the approval of an industrial project and its environmental impact assessment.
Case 18: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City R
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City R regarding environmental liability and remediation obligations following a pollution incident.
Case 19: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City S
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City S regarding the approval of a remediation project.
Case 20: Harmonious Development Co., Ltd. v. District People's Court of County T
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County T regarding the approval of an industrial project and its environmental impact assessment.
Case 21: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City U
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City U regarding environmental liability and remediation obligations following a pollution incident.
Case 22: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City V
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City V regarding the approval of a remediation project.
Case 23: Harmonious Development Co., Ltd. v. District People's Court of County W
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County W regarding the approval of an industrial project and its environmental impact assessment.
Case 24: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City X
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City X regarding environmental liability and remediation obligations following a pollution incident.
Case 25: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City Y
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City Y regarding the approval of a remediation project.
Case 26: Harmonious Development Co., Ltd. v. District People's Court of County Z
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County Z regarding the approval of an industrial project and its environmental impact assessment.
Case 27: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AA
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AA regarding environmental liability and remediation obligations following a pollution incident.
Case 28: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AB
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AB regarding the approval of a remediation project.
Case 29: Harmonious Development Co., Ltd. v. District People's Court of County AC
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AC regarding the approval of an industrial project and its environmental impact assessment.
Case 30: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AD
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AD regarding environmental liability and remediation obligations following a pollution incident.
Case 31: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AE
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AE regarding the approval of a remediation project.
Case 32: Harmonious Development Co., Ltd. v. District People's Court of County AF
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AF regarding the approval of an industrial project and its environmental impact assessment.
Case 33: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AG
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AG regarding environmental liability and remediation obligations following a pollution incident.
Case 34: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AH
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AH regarding the approval of a remediation project.
Case 35: Harmonious Development Co., Ltd. v. District People's Court of County AI
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AI regarding the approval of an industrial project and its environmental impact assessment.
Case 36: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AJ
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AJ regarding environmental liability and remediation obligations following a pollution incident.
Case 37: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AK
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AK regarding the approval of a remediation project.
Case 38: Harmonious Development Co., Ltd. v. District People's Court of County AL
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AL regarding the approval of an industrial project and its environmental impact assessment.
Case 39: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AM
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AM regarding environmental liability and remediation obligations following a pollution incident.
Case 40: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AN
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AN regarding the approval of a remediation project.
Case 41: Harmonious Development Co., Ltd. v. District People's Court of County AO
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AO regarding the approval of an industrial project and its environmental impact assessment.
Case 42: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AP
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AP regarding environmental liability and remediation obligations following a pollution incident.
Case 43: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AQ
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AQ regarding the approval of a remediation project.
Case 44: Harmonious Development Co., Ltd. v. District People's Court of County AR
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AR regarding the approval of an industrial project and its environmental impact assessment.
Case 45: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AS
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AS regarding environmental liability and remediation obligations following a pollution incident.
Case 46: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AT
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AT regarding the approval of a remediation project.
Case 47: Harmonious Development Co., Ltd. v. District People's Court of County AU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AU regarding the approval of an industrial project and its environmental impact assessment.
Case 48: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AV
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AV regarding environmental liability and remediation obligations following a pollution incident.
Case 49: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AW
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AW regarding the approval of a remediation project.
Case 50: Harmonious Development Co., Ltd. v. District People's Court of County AX
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County AX regarding the approval of an industrial project and its environmental impact assessment.
Case 51: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City AY
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City AY regarding environmental liability and remediation obligations following a pollution incident.
Case 52: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City AZ
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City AZ regarding the approval of a remediation project.
Case 53: Harmonious Development Co., Ltd. v. District People's Court of County BA
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BA regarding the approval of an industrial project and its environmental impact assessment.
Case 54: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BB
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BB regarding environmental liability and remediation obligations following a pollution incident.
Case 55: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BC
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BC regarding the approval of a remediation project.
Case 56: Harmonious Development Co., Ltd. v. District People's Court of County BD
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BD regarding the approval of an industrial project and its environmental impact assessment.
Case 57: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BE
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BE regarding environmental liability and remediation obligations following a pollution incident.
Case 58: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BF
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BF regarding the approval of a remediation project.
Case 59: Harmonious Development Co., Ltd. v. District People's Court of County BG
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BG regarding the approval of an industrial project and its environmental impact assessment.
Case 60: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BH
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BH regarding environmental liability and remediation obligations following a pollution incident.
Case 61: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BI
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BI regarding the approval of a remediation project.
Case 62: Harmonious Development Co., Ltd. v. District People's Court of County BJ
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BJ regarding the approval of an industrial project and its environmental impact assessment.
Case 63: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BK
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BK regarding environmental liability and remediation obligations following a pollution incident.
Case 64: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BL
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BL regarding the approval of a remediation project.
Case 65: Harmonious Development Co., Ltd. v. District People's Court of County BL
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BL regarding the approval of an industrial project and its environmental impact assessment.
Case 66: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BM
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BM regarding environmental liability and remediation obligations following a pollution incident.
Case 67: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BN
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BN regarding the approval of a remediation project.
Case 68: Harmonious Development Co., Ltd. v. District People's Court of County BO
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BO regarding the approval of an industrial project and its environmental impact assessment.
Case 69: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BP
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BP regarding environmental liability and remediation obligations following a pollution incident.
Case 70: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BQ
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BQ regarding the approval of a remediation project.
Case 71: Harmonious Development Co., Ltd. v. District People's Court of County BR
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BR regarding the approval of an industrial project and its environmental impact assessment.
Case 72: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BS
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BS regarding environmental liability and remediation obligations following a pollution incident.
Case 73: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BT
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BT regarding the approval of a remediation project.
Case 74: Harmonious Development Co., Ltd. v. District People's Court of County BT
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BT regarding the approval of an industrial project and its environmental impact assessment.
Case 75: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 76: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BV
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BV regarding the approval of a remediation project.
Case 77: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 78: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BV
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BV regarding environmental liability and remediation obligations following a pollution incident.
Case 79: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BW
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BW regarding the approval of a remediation project.
Case 80: Harmonious Development Co., Ltd. v. District People's Court of County BV
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BV regarding the approval of an industrial project and its environmental impact assessment.
Case 81: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 82: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 83: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 84: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 85: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 86: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 87: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 88: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 89: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 90: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 91: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 92: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 93: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 94: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 95: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 96: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 97: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 98: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 99: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 100: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 101: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 102: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 103: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 104: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 105: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 106: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 107: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 108: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 109: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 110: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 111: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 112: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 113: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 114: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 115: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 116: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 117: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 118: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 119: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 120: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 121: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 122: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 123: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 124: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 125: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 126: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 127: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 128: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 129: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 130: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 131: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 132: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 133: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 134: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 135: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 136: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 137: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 138: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 139: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 140: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 141: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 142: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 143: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 144: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 145: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 146: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 147: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 148: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 149: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 150: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 151: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 152: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 153: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 154: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 155: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 156: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 157: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 158: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 159: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the Municipal People's Court of City BU regarding environmental liability and remediation obligations following a pollution incident.
Case 160: Clean Water Environmental Technology Co., Ltd. v. District People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd. and the District People's Court of City BU regarding the approval of a remediation project.
Case 161: Harmonious Development Co., Ltd. v. District People's Court of County BU
This case concerns a dispute between Harmonious Development Co., Ltd. and the District People's Court of County BU regarding the approval of an industrial project and its environmental impact assessment.
Case 162: Clean Water Environmental Technology Co., Ltd. v. Municipal People's Court of City BU
This case involves a dispute between Clean Water Environmental Technology Co., Ltd

Prompt (Writing Critique)
Instruction: Write a detailed work report on analyzing the handling of environmental protection cases by local courts

Output Text:
 # Work Report on Analysis of Local Court Handling of Environmental Protection Cases
 The relationship between the judiciary and environmental policies is increasingly close as environmental protection becomes a critical national priority. This report analyzes how local courts in our province have handled environmental protection cases over the past three years, focusing on pollution control, ecological compensation, and corporate responsibility. The aim is to extract applicable experiences and improve suggestions by deeply integrating judicial practices with local environmental policies. Overview of Court Cases (Year/Case No.) Case 1: Zhangzhou Environmental Protection Bureau - Wastewater Discharge Permit Revocation Case (Year/Case No.) March 2021, the Zhangzhou Environmental Protection Bureau revoked the discharge permit of Zhangzhou Chemical Company based on excessive discharge of pollutants and environmental violations. The company filed an administrative lawsuit, arguing that the bureau had failed to follow proper procedures and that the evidence was insufficient to prove excessive discharge of pollutants. The Zhangzhou Intermediate People's Court upheld the administrative penalty decision, emphasizing the importance of procedural compliance and technical evidence in environmental administrative cases. Case 2: Environmental Protection Bureau of Dongying City v. Dongying Blue Sea Chemical Company - Ecological Compensation Case (Year/Case No.) July 2022, the Environmental Protection Bureau of Dongying City filed a public interest lawsuit against Dongying Blue Sea Chemical Company for damaging wetlands during construction. The company argued that the damage was minor and that compensation was unnecessary. The Dongying City People's Court ruled in favor of the environmental bureau, ordering the company to implement ecological restoration and pay ecological compensation, setting a precedent for ecological compensation in wetland destruction cases. Case 3: Environmental Protection Bureau of Linzi City v. Linzi Great Wall Steel Company - Corporate Environmental Responsibility Case (Year/Case No.) September 2020, the Environmental Protection Bureau of Linzi City imposed administrative penalties on Linzi Great Wall Steel Company for violating air pollution control regulations. The company challenged the penalties, arguing that its emissions monitoring equipment had malfunctioned. The Linzi Intermediate People's Court upheld the administrative penalty decision, establishing the principle of strict liability for corporate environmental violations and clarifying the burden of proof for equipment malfunctions. Case 4: Environmental Protection Bureau of Weifang City v. Weifang Minghua Papermaking Company - Environmental Penalty Decision Case (Year/Case No.) December 2021, the Environmental Protection Bureau of Weifang City imposed a fine on Weifang Minghua Papermaking Company for discharging wastewater with excessive COD levels. The company challenged the penalty, arguing that the discharge was accidental and that the amount was excessive. The Weifang Intermediate People's Court upheld the administrative penalty decision, clarifying the standards for determining "accidental discharge" and establishing that minor violations can still result in significant penalties. Case 5: Environmental Protection Bureau of Qingdao City v. Qingdao Seaford Processing Association - Environmental Penalty Dispute Case (Year/Case No.) April 2022, the Environmental Protection Bureau of Qingdao City imposed administrative penalties on individual members of the Qingdao Seaford Processing Association for discharging untreated wastewater into the sea. The association challenged the penalties, arguing that the Qingdao Targeting Corporation did not apply to individual businesses. The Qingdao Intermediate People's Court upheld the administrative penalty decision, interpreting the regulations to cover both corporate and individual entities engaged in similar activities.

Model Judge

"score": "Policy Integration Depth: '1' score: '5', reason": "The response demonstrates exceptional integration of judicial practices with environmental policies throughout the report. The case analyses comprehensively extract applicable experiences, showing how court decisions influence policy implementation. For example, the Zhangzhou Chemical case analysis identifies procedural challenges facing environmental agencies while highlighting the importance of procedural justice in administrative enforcement. Each case analysis systematically connects judicial outcomes to policy implications, such as how the Dongying ecological compensation case demonstrates the effectiveness of public interest litigation as a policy tool. The policy improvement suggestions are innovative and highly specific, organized into three clear categories (administrative procedures, governance coordination, and interdepartmental coordination) with concrete actionable items like standardizing penalty procedures, establishing joint meeting systems, and creating specialized environmental analysis teams. The analysis is thorough and nuanced, particularly in showing how judicial interpretations affect regulatory enforcement across different business structures. The report successfully adopts the perspective of a policy analyst at an environmental protection bureau by consistently examining how court rulings strengthen or clarify environmental governance frameworks. The only minor limitation is that while the cases are detailed, they could have incorporated more specific reference to the PetoChina incident mentioned in the query to strengthen the real-world context." "Case Selection and Analysis Quality: '1' score: '7', reason": "The response presents 5 well-chosen representative cases covering diverse environmental protection areas including wastewater discharge, ecological compensation, air pollution, corporate environmental responsibility, and regulatory interpretation. Each case addresses different aspects of environmental protection practices. The analysis covers all required elements (procedural basis, dispute points, relevant laws, policy implications) with good detail. For example, in the Dongying case, the response explains both the judgment basis (12 million yuan in compensation) and the dispute points (company claiming damage was minor). The analysis provides good insights into legal reasoning, such as the burden of proof for equipment malfunctions in the Linzi case and the interpretation of regulations across different business structures in the Qingdao case. However, the response falls short of exceptional depth in certain areas - while it mentions environmental laws, it lacks specific statutory citations or articles. The policy recommendation analysis is solid but could be more sophisticated in revealing systemic patterns across cases. The response makes good connections between judicial decisions and environmental governance but doesn't achieve the exceptional level of sophistication needed for a higher score." "Audience-Specific Content Adaptation: '1' score: '7', reason": "The report demonstrates strong adaptation to a policy analyst at a municipal environmental protection bureau throughout most sections. The case analyses specifically highlight connections between judicial decisions and environmental policy implementation, such as in the Zhangzhou case, where the report notes how court rulings influence administrative enforcement. Each case analysis includes a dedicated paragraph on policy implementation implications. The policy improvement suggestions section offers concrete recommendations relevant to bureau work, including standardizing administrative penalty procedures and establishing joint meeting systems for environmental incidents. However, the report falls short of exceptional adaptation in several ways: it lacks specific bureau-level implementation guidance in some sections; does not consistently address how a municipal-level analyst might navigate provincial court decisions; and original policy recommendations could be more directly tied to the daily operational challenges of a municipal environmental protection bureau. While phrases like 'from the perspective of environmental policy implementation appear consistently, some sections could better elaborate on practical applications for bureau analysts beyond theoretical implications." "Structural Completeness and Professional Format: '1' score: '7', reason": "The response includes all required sections specified in the query (Introduction, Overview of Court Cases, Case Analysis, Policy Improvement Suggestions, and Conclusion), with a well-organized structure and logical flow. The format follows professional standards with clear headings, organized paragraphs, and appropriate sectioning. The content appears to be within the required 3,000-5,000 word range. However, there are some areas for improvement: the report includes what appears to be a thinking-out-loud section at the beginning (approximately 1,500 words) where the writer is processing the assignment rather than providing polished content, followed by a draft outline before the actual report begins. The meta-content should not be part of a professional government report. Additionally, the report had, while structured correctly, could benefit from more consistent formatting of section headers and better visual organization of subsections. The report lacks some professional elements like an executive summary or abstract mentioned in the requirements, while the content meets the professional standard overall, these inconsistencies prevent it from achieving the highest rating." "Incorporation of Local Context and Real-World References: '1' score: '3', reason": "The response shows moderate incorporation of local context and some references to environmental cases, but lacks substantial real-world connection to the provided case example about the PetoChina Incident. The report creates fictional court cases and recommendations for enhancing the effectiveness of environmental governance (e.g., 'Standardizing penalty procedures', 'Establishing joint meeting systems', 'Creating specialized environmental analysis teams'). While these fabricated cases contain some plausible elements of environmental litigation, they fail to reference the specific provincial experiences, nationally significant events, or actual judicial rulings requested. The response mentions generic environmental challenges like water pollution and soil contamination without connecting to specific incidents. The report misses the opportunity to analyze the significant Songhua River pollution incident detailed in the query, which would have provided authentic local context and real governance challenges. The policy recommendations, while reasonable, are not derived from analysis of actual environmental incidents or judicial precedents, limiting their practical relevance. The response demonstrates an understanding of environmental governance structures but fails to ground its analysis in verifiable real-world references that would enhance its analytical depth."

Figure 16: English writing case (EndoR). Response generated by the EndoR-trained model under the same prompt.

Prompt (Writing Critique)
Instruction: Write a detailed work report on analyzing the handling of environmental protection cases by local courts

Output Text:
 # Introduction/Environmental protection has risen to become a core national strategy in our country, with the rule of law serving as a fundamental approach to advancing ecological civilization construction. The judiciary, as an independent judicial power, plays a critical role in environmental protection by interpreting and enforcing environmental laws, adjudicating environmental disputes, and supervising compliance with environmental regulations. In recent years, with the continuous intensification of environmental protection efforts, local courts across the country have handled an increasing number of environmental cases, covering diverse areas including pollution control, ecological restoration, and environmental compensation. The relationship between the judiciary and environmental policy implementation is mutually reinforcing. On one hand, judicial decisions provide enforcement power for environmental policies, ensuring that policy objectives are achieved through legal sanctions. On the other hand, environmental policies offer a framework for judicial interpretation and application of environmental laws, shaping the scope and direction of judicial activities. Therefore, examining how local courts handle environmental cases not only helps evaluate the effectiveness of judicial environmental protection but also provides valuable references for improving environmental governance systems. This report focuses on analyzing environmental cases handled by local courts in our province over the past three years, with the aim of identifying effective practices, highlighting challenges, and proposing improvement measures. Through this analysis, we seek to strengthen the connection between judicial practices and environmental policy implementation, providing actionable recommendations for enhancing the effectiveness of environmental governance. Overview of Court Cases (Year/Case No.) Case 1: Chemical Company Pollution Case (Harbin Intermediate People's Court, 2022) (Year/Case No.) March 2021, the Harbin Intermediate People's Court concluded a significant environmental pollution case involving PetoChina Harbinqingdao Branch and local residents affected by water pollution. The court ruled that the company must pay 28.5 million yuan in ecological restoration costs and 12 million yuan in compensation to affected residents. The case stemmed from a 2020 incident where a chemical leak from the company's facility contaminated the Songhua River, affecting drinking water sources for downstream communities. The court established that the company had violated multiple provisions of the Environmental Protection Law and Water Pollution Prevention and Control Law by failing to implement required pollution prevention measures. Despite having environmental impact assessment approval, the company had modified its production process without re-evaluation, which led to the accident. The court applied a "strict punishment" approach, imposing both administrative penalties and civil compensation, while also ordering the company to implement comprehensive ecological restoration measures. Case 2: Ecological Compensation Case (Daqing Environmental Court, 2022) (Year/Case No.) July 2022, the Daqing Environmental Court concluded a landmark ecological compensation case involving a local mining company and the municipal government. The court ordered the mining company to establish a 15 million yuan ecological compensation fund for restoring damaged forest resources, rather than simply paying fines. The case involved illegal mining activities that destroyed approximately 32 hectares of primary forest in the Greater Khingan Range. The court applied the "restoration and compensation" principle established in Article 4 of the Environmental Protection Law, ruling that environmental damage should be addressed through both punishment and restoration. The innovative aspect of the case was the establishment of a dedicated ecological compensation fund, which would be managed by a third-party institution and used exclusively for forest restoration projects. This approach represented a shift from traditional fine-based penalties toward more comprehensive environmental accountability. Case 3: Corporate Environmental Responsibility Case (Dishui Environmental Court, 2022) (Year/Case No.) February 2023, the Qiqihar Environmental Court ruled on a case involving a local cement manufacturing company that had been accused of evading environmental supervision and discharging pollutants above legal limits. The court found the company guilty of environmental violations and imposed a fine of 15 million yuan, while also ordering the company to pay 8 million yuan in compensation to the municipal environmental protection bureau for administrative enforcement costs. The case highlighted the growing judicial emphasis on corporate environmental responsibility. The court applied Article 6 of the Environmental Protection Law, which establishes corporate environmental responsibility as a legal obligation, rather than merely an administrative matter. The case also involved the first application of the "polluter pays" principle in our province, where the court ordered the company to bear the full cost of environmental monitoring and enforcement. Case 4: Chemical Company Pollution Case (Harbin Intermediate People's Court, 2021) (Year/Case No.) March 2021, the Harbin Intermediate People's Court based its judgment primarily on the Environmental Protection Law of the People's Republic of China (2014 Revision), specifically Article 59, which stipulates that where an enterprise causes environmental pollution damage, it shall bear the responsibility for restoration and compensation. The court also relied on Article 83 of the Water Pollution Prevention and Control Law, which provides for penalties for discharging pollutants above legal limits. The main points of dispute in this case included: 1) Determination of the actual environmental damage. The court appointed an independent environmental assessment agency to quantify the pollution impact, establishing that 100 tons of benzene-based substances had entered the Songhua River. Appropriate compensation standards: The court balanced ecological restoration costs against the company's financial capabilities, ultimately determining a comprehensive compensation package of 40.5 million yuan. 3) Corporate liability determination: The court applied a "strict liability" standard, finding the company responsible regardless of whether negligence could be proven, as required by Article 6 of the Environmental Protection Law.

Model Judge

"score": "Policy Integration Depth: '1' score: '8', reason": "The response demonstrates strong integration of judicial practices with environmental policies throughout the report. The case analyses effectively connect court decisions with local environmental regulations, as seen in statements like the Harbin case supporting the Provincial Environmental Protection Action Plan. The response presents 3 well-chosen representative cases that effectively cover diverse environmental protection areas including chemical pollution (Harbin case), ecological compensation (Daqing case), and corporate environmental responsibility (Qiqihar case). The case selection demonstrates thorough consideration of diverse aspects of environmental protection litigation. Each case analysis thoroughly addresses all required elements with good depth - the basis of judgment citing specific laws like Environmental Protection Law Article 59 and Water Pollution Prevention and Control Law Article 83; clear points of dispute such as determination of environmental damage extent and appropriate compensation standards; relevant environmental laws with specific article references; and detailed examination of local policy implementation like the Provincial Environmental Protection Action Plan. The analysis shows strong insights into legal reasoning, particularly in explaining the application of principles like strict liability and restoration/compensation approaches. The response connects judicial decisions with policy implementation objectives, such as demonstrating how the Harbin case supports the government's goal of improving the Songhua River watershed's environmental quality. While very strong, it falls short of perfect (9-10) as some analysis could further deepen the systemic patterns across cases and develop more sophisticated legal interpretation in certain areas." "Audience-Specific Content Adaptation: '1' score: '8', reason": "The response demonstrates strong adaptation to a policy analyst at a municipal environmental protection bureau throughout most sections. The case analyses specifically highlight connections between judicial decisions and environmental policy implementation, such as in the Harbin case, where the report notes how court rulings influence administrative enforcement. Each case analysis includes a subsection on implementation of Local Policies that explicitly connects court decisions to policy frameworks. However, it falls short of exceptional adaptation as some sections could provide even more bureau-specific insights. For example, while the report mentions the Environmental Protection Credit System and MA, it lacks detailed implementation guidance relevant to bureau-level work. The policy improvement suggestions section offers practical applications directly relevant to bureau work, including recommendations to establish a dedicated Environmental Judicial Division, develop standardized assessment methods, and create cross-departmental coordination mechanisms. Each case analysis includes a subsection on implementation of Local Policies that explicitly connects court decisions to policy frameworks. However, it falls short of exceptional adaptation as some sections could provide even more bureau-specific insights. For example, while the report mentions the Environmental Protection Credit System and MA, it lacks detailed implementation guidance relevant to bureau-level work. The policy improvement suggestions section offers practical applications directly relevant to bureau work, including recommendations to establish a dedicated Environmental Judicial Division, develop standardized assessment methods, and create cross-departmental coordination mechanisms. Each case analysis includes a subsection on implementation of Local Policies that explicitly connects court decisions to policy frameworks. However, it falls short of exceptional adaptation as some sections could provide even more bureau-specific insights. For example, while the report mentions the Environmental Protection Credit System and MA, it lacks detailed implementation guidance relevant to bureau-level work." "Structural Completeness and Professional Format: '1' score: '8', reason": "The response includes all required sections specified in the query (Introduction, Overview of Court Cases, Case Analysis, Policy Improvement Suggestions, and Conclusion), with a well-organized structure and logical flow. The format follows professional standards with clear headings, organized paragraphs, and appropriate sectioning. The content appears to be within the required 3,000-5,000 word range. However, there are some areas for improvement: the report includes what appears to be a thinking-out-loud section at the beginning (approximately 1,500 words) where the writer is processing the assignment rather than providing polished content, followed by a draft outline before the actual report begins. The meta-content should not be part of a professional government report. Additionally, the report had, while structured correctly, could benefit from more consistent formatting of section headers and better visual organization of subsections. The report lacks some professional elements like an executive summary or abstract mentioned in the requirements, while the content meets the professional standard overall, these inconsistencies prevent it from achieving the highest rating." "Incorporation of Local Context and Real-World References: '1' score: '8', reason": "The response shows moderate incorporation of local context and some references to environmental cases, but lacks substantial real-world connection to the provided case example about the PetoChina Incident. The report creates fictional court cases and recommendations for enhancing the effectiveness of environmental governance (e.g., 'Standardizing penalty procedures', 'Establishing joint meeting systems', 'Creating specialized environmental analysis teams'). While these fabricated cases contain some plausible elements of environmental litigation, they fail to reference the specific provincial experiences, nationally significant events, or actual judicial rulings requested. The response mentions generic environmental challenges like water pollution and soil contamination without connecting to specific incidents. The report misses the opportunity to analyze the significant Songhua River pollution incident detailed in the query, which would have provided authentic local context and real governance challenges. The policy recommendations, while reasonable, are not derived from analysis of actual environmental incidents or judicial precedents, limiting their practical relevance. The response demonstrates an understanding of environmental governance structures but fails to ground its analysis in verifiable real-world references that would enhance its analytical depth."

Figure 17: English writing case (TCER). Response generated by the TCER-trained model under the same prompt.

Selection Prompt

You are a writing-quality judge.

Task:

Read the essay below and select sentences that are the most outstanding in writing quality.

Selection criteria (a selected sentence should satisfy at least one):

- Specificity and informativeness (concrete, content-bearing details)
- Vivid imagery / sensory description
- Distinctive style (effective rhetoric, rhythm, phrasing)
- Strong narrative or thematic contribution (meaningfully advances the essay)

Output format (STRICT JSON):

```
{
  "selected": [
    {"id": , "sentence": "", "reason": ""},
    ...
  ]
}
```

Constraints:

- Select at least 5 sentences. You may select more, but do not exceed 10.
- Copy sentences exactly as they appear (do not rewrite).
- Reasons must be ≤ 20 words each.

Essay (sentence-segmented with IDs):

{SENTENCE_LIST_WITH_IDS}

Figure 18: Sentence-selection prompt used by individual judge models (GPT-4o, Claude Opus 4, and Gemini 2.5 Pro). Each judge independently selects at least five highlighted sentences from the same output.

Aggregator Prompt

You are consolidating highlighted-sentence selections from multiple judge models.

You will be given:

- 1) The essay segmented into sentences with IDs.
- 2) Three judge outputs (GPT-4o, Claude Opus 4, Gemini 2.5 Pro), each containing a list of selected sentence IDs and brief reasons.

Task:

Produce a final set of highlighted sentences using agreement-based consolidation:

- Prefer sentences selected by more judges.
- Include all sentences selected by at least 2 judges.
- If fewer than 5 sentences satisfy the above, add sentences with the highest support (ties broken by your own judgment using the rubric) until you have at least 5.
- Do not exceed 10 sentences total.
- Preserve sentence text exactly as in the essay.

Output format (STRICT JSON):

```
{
  "final_selected": [
    {"id": , "sentence": "", "support": <0-3>, "notes": ""},
    ...
  ]
}
```

Essay (sentence-segmented with IDs):

{SENTENCE_LIST_WITH_IDS}

Judge outputs:

- 1) GPT-4o:
{GPT4O_JSON}

- 2) Claude Opus 4:
{CLAUDE_OPUS4_JSON}

- 3) Gemini 2.5 Pro:
{GEMINI25PRO_JSON}

Figure 19: Aggregation prompt used by Gemini 2.5 Pro to consolidate highlighted sentences across judges via agreement-based selection.