CARE: Curriculum-Aware Rubric Evolution for Open-Ended Text Generation

Anonymous ACL submission

Abstract

Reinforcement Learning with Verifiable Rewards (RLVR) has demonstrated significant potential in tasks characterized by well-defined objective criteria. However, extending this paradigm to open-ended generation tasks, which lack ground truth and are inherently subjective, poses fundamental challenges. Recent research has shifted towards Reinforcement Learning from AI Feedback (RLAIF). Nevertheless, existing methods largely rely on human-crafted or predetermined rubrics maintained via static or rule-based update mechanisms, failing to adapt timely to the evolving capabilities of the policy model, thereby resulting in lagging and sparse critical reward signals. To overcome these limitations, we propose Curriculum-Aware Rubric Evolution (CARE), a rubric evolution method that integrates curriculum awareness with diagnosis-driven evolution. CARE adaptively adjusts the complexity of rubrics across different training stages by monitoring the statistical properties of the reward distribution during policy training. Simultaneously, it introduces a discrepancy-based diagnostic sampling strategy, prioritizing high signal-to-noise critical samples for targeted rubric evolution. Experiments demonstrate that CARE consistently enhances alignment quality during training and achieves superior generation performance compared to multiple strong baselines. Our work provides an adaptive solution for scalable oversight in highly subjective generation scenarios.

1 Introduction

In recent years, Reinforcement Learning with Verifiable Rewards (RLVR) has shown strong effectiveness on tasks with clear and objective evaluation criteria (Chen, 2021; Shao et al., 2024). By constructing reward signals based on verifiable outcomes such as code execution results or mathematical correctness, RLVR provides stable supervision signals and has significantly improved the performance of large language models on reasoning-intensive tasks (Cobbe et al., 2021). However, extending this paradigm to open-ended generation tasks, such as creative writing or complex instruction following, remains challenging due to the absence of explicit ground truth (Stiennon et al., 2020; Ouyang et al., 2022; Gao et al., 2023). In these settings, reward signals are typically constructed through task-specific evaluation procedures, making the training process highly sensitive to the design of the evaluation mechanism (Ziegler et al., 2019; Christiano et al., 2017; Rafailov et al., 2023).

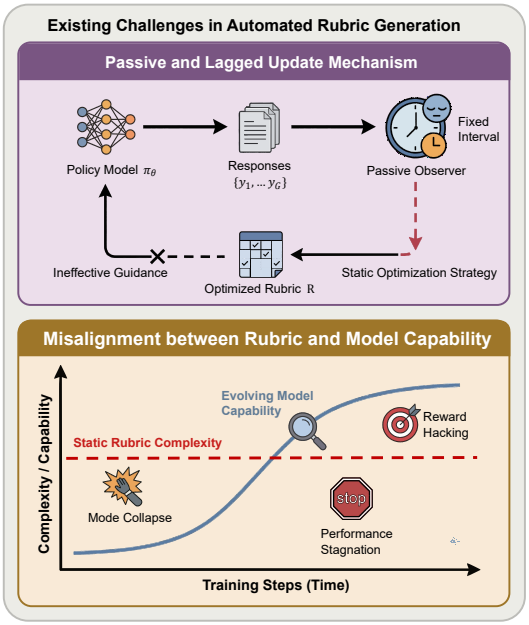


Figure 1: **Challenges in automated rubric generation.** Passive, fixed-interval updates cause delayed feedback, while static rubrics misalign with evolving model capabilities.

For open-ended generation without objective ground truth, existing approaches commonly adopt Reinforcement Learning from AI Feedback (RLAIF), where language models act as evaluators to score or rank generated outputs based on predefined prompts or rules (Zheng et al., 2023;

Lee et al., 2023; Bai et al., 2022). While this approach reduces reliance on human annotation, constructing and maintaining high-quality evaluation rubrics often requires substantial manual effort. Moreover, evaluation granularity is difficult to balance. Coarse rubrics tend to produce sparse learning signals, whereas fine-grained rubrics introduce significant evaluation noise (Kreutzer et al., 2018; Amodei et al., 2016). Since the reward function only partially reflects the true task objective, policy models may also exploit evaluator preferences, leading to reward hacking (Skalse et al., 2022; Gao et al., 2023) that does not correspond to genuine quality improvements.

Recent work has explored automated rubric generation to improve reward design (Shin et al., 2020; Dubois et al., 2023). However, existing frameworks typically rely on passive, interval-based updates and static optimization strategies (Kim et al., 2025). As illustrated in Figure 1, such update mechanisms depend on delayed training observations, which creates a trade-off between responsiveness and stability. Frequent updates can introduce non-stationary noise and destabilize training (Sutton et al., 1998; Haarnoja et al., 2018; Shao et al., 2024), while infrequent updates delay the correction of emerging failure modes. In addition, static update strategies apply the same optimization rules throughout training and do not account for the stage-wise evolution of policy capability. As a result, complex rubrics may be introduced too early, harming training stability, while later stages may suffer from insufficiently discriminative evaluation signals, leading to performance stagnation.

To address these limitations, we propose Curriculum-Aware Rubric Evolution (CARE), a closed-loop framework for rubric optimization. CARE explicitly couples the evolution of the rubric with the training dynamics of the policy model. The model's current training state is used to determine the appropriate complexity of reward signals under different curriculum objectives. At the same time, systematic mismatches between rubric evaluations and policy behavior provide diagnostic signals that guide rubric updates.

Specifically, CARE monitors statistical properties of reward distributions during training and adaptively selects curriculum objectives that are aligned with the current training stage. This design helps avoid mismatched or premature reward signals. In addition, CARE introduces an active diagnostic sampling mechanism that prioritizes samples exhibiting inconsistency between the rubric and the policy, evaluation disagreement, or abnormal reward distributions. These samples are used as high signal-to-noise inputs for rubric evolution. Through this bidirectional interaction, CARE enables the rubric and the policy to evolve in a coordinated manner, providing progressive training signals and supporting a more stable and controllable alignment process.

Our contributions are summarized as follows:

- We revisit reward design for open-ended generation, particularly focusing on the coupled evolution of evaluation rubrics and policy training dynamics to address the limitations of static supervision.
- We propose CARE, a framework that leverages statistical curriculum awareness, active diagnostic sampling, and progressive optimization to provide stable and reliable signals for model alignment.
- We conduct extensive experiments on open-ended generation benchmarks, including essay writing, summarization, and creative writing, demonstrating the effectiveness of CARE in terms of training stability and generation quality.

2 Related Works

Generative and Adaptive Reward Modeling. Conventional reward models typically output scalar scores, which often lack interpretability and are ill-equipped for complex reasoning tasks (Gao et al., 2023; Lightman et al., 2023). Recent advancements have shifted towards generative approaches that formulate verification as a reasoning task (Mahan et al., 2024). Frameworks like GenRM (Zhang et al., 2024) and RM-R1 (Chen et al., 2025b) leverage next-token prediction and self-generated Chain-of-Rubrics to bolster reasoning capabilities, while ReasonGRM (Chen et al., 2025a) underscores the importance of generating high-quality reasoning paths. To standardize these evaluations, OpenRubrics (Liu et al., 2025) and Rubicon (Huang et al., 2025) employ large-scale rubric sets or guidelines to steer the reward process, and RaR (Gunjal et al., 2025) demonstrates the efficacy of instance-specific rubrics in online RL. While these approaches offer structured guidance, recent research has explored evolutionary strategies to address reward hacking and task variabil-

ity (Skalse et al., 2022). MPO (Kim et al., 2025) and DR Tulu (Shao et al., 2025) introduce mechanisms to dynamically refine prompts or co-evolve criteria based on training contexts, while CARMO (Gupta et al., 2025) and PaTaRM (Jian et al., 2025) generate context-aware or task-adaptive rubrics to ground reward models. Despite these advances, most existing methods adapt reward specifications through static or monotonic mechanisms, offering limited responsiveness to evolving model capabilities and often leading to misaligned evaluation granularity in long-horizon or open-ended generation settings.

Curriculum Learning and Data Selection in RL Effective data filtering and curriculum scheduling are paramount for stable RLHF (Ouyang et al., 2022). Omni-Thinker (Li et al., 2025) highlights the significance of backward-transfer-based curriculum scheduling, ordering tasks from structured to open-ended to prevent forgetting. In terms of data filtering, ReasonGRM (Chen et al., 2025a) introduces metrics to filter out hallucinated reasoning paths, while frameworks like AutoRubric-R1V (Jia et al., 2025) utilize self-aggregation to distill consistent reasoning checkpoints from successful trajectories. Additionally, Chasing the Tail (Zhang et al., 2025) emphasizes the necessity of differentiating responses in the high-reward tail rather than focusing on trivial distinctions. Overall, current approaches lack integrated mechanisms to jointly adapt data selection and reward specification in response to evolving model behavior under reward uncertainty and distributional drift.

3 Methods

3.1 Problem Formulation

We formalize the alignment process as a general bilevel optimization problem (Beck et al., 2023). Let π_θ denote the policy model parameterized by θ, and $\mathcal{R}_\phi$ denote the rubric parameterized by ϕ (representing its textual constraints and criteria).

In the **inner loop**, the policy π_θ is optimized to maximize the expected return defined by the current rubric $\mathcal{R}_\phi$. For a given prompt x, the optimization objective is:

$$\theta^* = \arg\max_\theta \mathbb{E}_{x\sim\mathcal{D}, y\sim\pi_\theta(\cdot|x)} \left[R(x, y \mid \mathcal{R}_\phi)\right] \quad (1)$$

where $R(\cdot \mid \mathcal{R}_\phi)$ represents the reward scoring function instantiated by the rubric $\mathcal{R}_\phi$.

In the **outer loop**, we model the rubric update as a generative evolution process driven by a Meta-LLM ($\mathcal{M}_{meta}$). The objective is to synthesize a refined rubric $\mathcal{R}'_\phi$ that semantically resolves the deficiencies identified in the diagnostic set $\mathcal{D}_{diag}$, under the guidance of the current curriculum phase S:

$$\mathcal{R}'_\phi = \mathcal{M}_{meta}\Big(\mathcal{R}_\phi, \mathcal{D}_{diag}, S\Big) \quad (2)$$

Here, $\mathcal{M}_{meta}$ serves as the semantic optimization operator, which analyzes the edge cases in $\mathcal{D}_{diag}$ and executes textual modifications as dictated by the curriculum controller state S.

3.2 Framework Overview: Synergistic Co-Evolution

We propose the **C**urriculum-**A**ware **R**ubric **E**volution (**CARE**) framework. As illustrated in Figure 2, CARE establishes a closed-loop system characterized by active diagnosis and bidirectional synergistic co-evolution, driven by two core engines:

- **Streaming Curriculum State Machine:** Continuously monitors training dynamics to adaptively modulate the optimization difficulty, thereby mitigating the misalignment between evaluation criteria and evolving model capabilities.
- **Active Diagnostic Sampling:** Identifies critical edge cases characterized by high uncertainty, performance extremes, and distributional shifts, providing informative signals to guide targeted rubric refinement.

These modules work in synergy: fine-grained rubrics provide progressive reward signals to shape the policy's behavior, while the policy's response distribution serves as a diagnostic signal for rubric maturation.

3.3 Streaming Curriculum State Machine

To synchronize the evaluation granularity with the model's evolving competence, we design a sliding-window-based streaming state machine. This adaptive control mechanism serves a dual purpose: it proactively identifies pathological training dynamics (Gao et al., 2023; Skalse et al., 2022) while dynamically adjusting the rubric's complexity according to the curriculum phase to maintain training stability and sustain continuous capability breakthroughs.

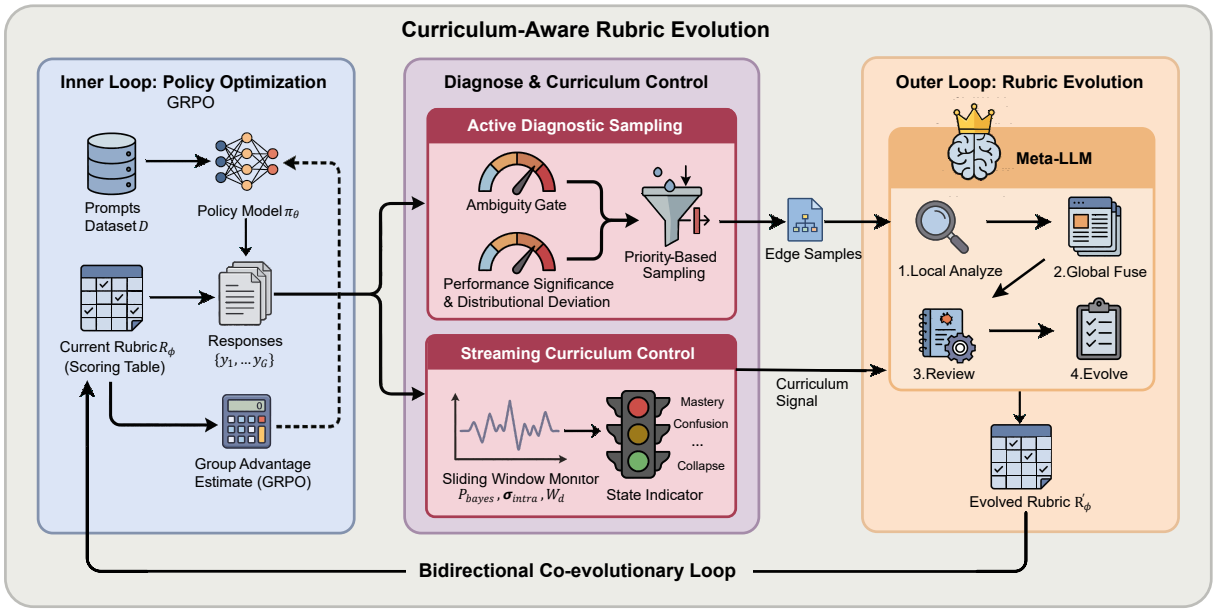


Figure 2: Overview of the CARE framework. The system establishes a closed-loop mechanism characterized by active diagnosis and bidirectional synergistic co-evolution. The process operates through three coordinated phases: (a) **Policy Optimization**, where the policy π_θ is trained via GRPO under the guidance of the current rubric $\mathcal{R}_\phi$; (b) **Diagnosis & Curriculum Control**, where the Streaming Curriculum State Machine monitors training dynamics, while the Active Diagnostic Sampling engine identifies critical edge cases.; and (c) **Rubric Evolution**, where a Meta-LLM utilizes these diagnostic and curriculum signals to evolve the rubric into $\mathcal{R}'_\phi$, ensuring evaluation criteria align with the model's evolving capabilities.

3.3.1 Training Dynamics Monitoring

We maintain a sliding window of length W over recent training iterations and denote the corresponding reward scores as S_{curr}. As a fixed reference point, we record S_{ref} at the last curriculum phase transition, enabling anchored comparison of training progress. We characterize training dynamics using three complementary metrics:

1. Bayesian Posterior Confidence (P_{bayes}): We estimate the probability that the mean return in the current window exceeds the historical baseline mean by a margin δ using a Bayesian approximation (Gelman et al., 1995). Assuming the sample mean follows a Gaussian distribution $\mathcal{N}(\mu_{S_{curr}}, \sigma^2_{se})$, where σ_{se} is the standard error derived from S_{curr}, we compute:

$$P_{bayes} \approx \Phi\left(\frac{\mu_{S_{curr}} - (\mu_{S_{ref}} + \delta)}{\sigma_{S_{curr}}/\sqrt{N}}\right), \quad (3)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution, $\sigma_{S_{curr}}$ is the standard deviation of S_{curr}, and $N = |S_{curr}|$ is the sample size of the current window.

2. Intra-group Score Deviation (σ_{intra}): We compute the average intra-group standard deviation to assess the model's stability on individual prompts. For a given prompt x with G sampled responses $\{y_1, \dots, y_G\}$ and corresponding scores $\{s_1, \dots, s_G\} \subset S_{curr}$, the metric is defined as:

$$\sigma_{intra} = \mathbb{E}_x\left[\sqrt{\frac{1}{G}\sum_{i=1}^{G}(s_i - \bar{s}_x)^2}\right], \quad (4)$$

where $\bar{s}_x$ denotes the mean score for prompt x.

3. Reward Distributional Drift (W_d): To measure distributional shift between the current reward distribution S_{curr} and the reference distribution S_{ref}, we compute the Wasserstein-1 distance (Arjovsky et al., 2017):

$$W_1(S_{curr}, S_{ref}) = \int_{-\infty}^{+\infty} |F_{curr}(x) - F_{ref}(x)|\, dx, \quad (5)$$

where $F_{curr}(x)$ and $F_{ref}(x)$ denote the empirical cumulative distribution functions (ECDFs) of the scores in S_{curr} and S_{ref}, respectively.

3.3.2 Adaptive Curriculum Mechanism

To align the rubric with the evolving capabilities of the student model, we maintain a state machine

that dynamically transitions between four curriculum phases. Based on the real-time monitoring metrics defined above (P_{bayes}, σ_{intra}, W_d, and the statistical properties of S_{curr}), the system applies distinct strategy modifiers to the rubric update process. Detailed hyperparameter configurations for these transitions are provided in Appendix C, and specific prompts are available in Appendix F:

Mastery ($P_{bayes} \to 1$): Triggered when the model statistically surpasses the baseline with high certainty. The system elevates the scoring ceiling and introduces high-order dimensions to break performance plateaus by optimizing for subtler distinctions.

Confusion ($\sigma_{intra} > \eta \cdot \sigma_{S_{curr}}$): Triggered when the average intra-group inconsistency (σ_{intra}) exceeds a critical proportion η of the global variance of the current window ($\sigma_{S_{curr}}$). To eliminate ambiguity, The system refines the granularity of scoring criteria and enforces strict boundaries between adjacent scores.

Strength ($W_d \gg 0, \mu_{S_{curr}} > \mu_{S_{ref}}$): Triggered when a significant structural shift moves the distribution towards higher scores (current mean $\mu_{S_{curr}}$ exceeds reference mean $\mu_{S_{ref}}$). The system sustains momentum by preserving core constraints while cautiously introducing limited open-ended dimensions. Simultaneously, this phase audits for potential reward hacking or anomalies, ensuring that performance gains reflect genuine capability rather than rubric exploitation.

Collapse ($W_d \gg 0, \mu_{S_{curr}} < \mu_{S_{ref}}$): Triggered when performance degradation (current mean $\mu_{S_{curr}}$ falls below reference mean $\mu_{S_{ref}}$) accompanies a significant distributional drift, suggesting potential mode collapse. The system executes a rollback to coarse-grained scoring scales and baseline constraints, aiming to restore a dense reward signal that reactivates exploration.

3.4 Active Diagnostic Sampling

Upon the activation of any curriculum state, the system initiates a diagnostic sampling phase to gather a batch of instances $\mathcal{D}_{diag} = \{(x, \mathbf{y}, \mathbf{s})\}$ from the current window. To pinpoint edge samples within this batch that render the rubric ineffective, we propose a priority-based sampling strategy. The priority score $S_{priority}(x)$ is defined as:

$$S_{priority}(x) = \underbrace{\left(\frac{\sigma_{intra}(x)}{\sigma_{S_{curr}}}\right)}_{\text{Ambiguity Gate}} \cdot (|A_{norm}(x)| + D_{KL}(\pi_\theta \| \pi_{ref})) \tag{6}$$

This formula incorporates three diagnostic dimensions, utilizing the statistical properties of the current score set S_{curr} defined in Section 3.2:

(i) **Ambiguity Gate:** This term is defined as the ratio of the local intra-group deviation $\sigma_{intra}(x)$ (as formulated in Eq.(4)) to the global standard deviation of the current window $\sigma_{S_{curr}}$. It amplifies the priority of samples where the local uncertainty significantly exceeds the global baseline, effectively pinpointing ill-defined rubric areas.
(ii) **Performance Significance:** This term utilizes the normalized absolute advantage $|A_{norm}(x)|$ to identify samples that exhibit the greatest discrepancy between observed rewards and the model's value baseline. By scaling the absolute advantage $A(x)$ relative to the maximum observed within the current diagnostic batch $\mathcal{D}_{diag}$, we map these deviations to a uniform range of $[0, 1]$. This normalization ensures numerical stability across varying training stages and prioritizes instances of either exceptional performance or unexpected failure. Such cases represent significant deviations from current model expectations and provide the most informative signals for subsequent rubric refinement.
(iii) **Distributional Deviation:** The term $D_{KL}(\pi_\theta \| \pi_{ref})$ captures the Kullback-Leibler divergence between the current policy π_θ and the reference policy π_{ref} given input x. High divergence signals areas of rapid behavioral change, such as emerging capabilities or potential catastrophic forgetting, which require urgent rubric verification.

Based on this metric, we construct the diagnostic set $\mathcal{D}_{diag}$ by selecting the top-K instances with the highest $S_{priority}(x)$ values from the current data batch. This selective sampling ensures that the subsequent update process is driven exclusively by the most informative edge cases.

3.5 Rubric Evolution Pipeline

When projecting granular diagnostic findings directly onto macro-level rubric synthesis, the Meta-LLM tends to generate localized patches to address

specific failures. This direct mapping often leads to stagnation in rubric evolution or compromises the global logical coherence of the rubric. To mitigate these issues, we implement a unified rubric evolution pipeline (detailed in Appendix E) designed to transform disparate diagnostic signals and current curriculum goals into a consistent, high-quality rubric. Specifically, the Meta-LLM executes a structured process comprising local diagnostic analysis, global strategic fusion, operational instruction synthesis, and adaptive rubric evolution. This step-by-step reasoning ensures that every proposed modification is grounded in empirical evidence and preserves global evaluation consistency.

4 Experiment

4.1 Experimental Setup

To validate the effectiveness of CARE, we establish a co-evolutionary alignment framework based on the Qwen2.5 family (Qwen et al., 2025), employing Qwen2.5-1.5B-Instruct as the policy model (π_θ). This lightweight configuration allows us to investigate the progressive curriculum evolution of our method. Given that compact policies often exhibit high sensitivity to reward consistency and constraint stability (Gao et al., 2023), this setting offers a suitable environment to observe the potential of our dynamic curriculum in facilitating capability acquisition. The process is supported by Qwen2.5-32B-Instruct, which adopts a composite supervisory role: functioning as the Evaluator to generate fine-grained reward signals for GRPO, while concurrently acting as the Meta-LLM to diagnose training dynamics and evolve rubric criteria. This choice aims to balance the reasoning capabilities required for subjective evaluation with the throughput necessary for iterative refinement. For comprehensive implementation details, including hyperparameters and hardware configurations, please refer to Appendix C.

4.2 Datasets

To ensure a comprehensive evaluation across diverse open-ended generation tasks, we employ three high-quality benchmarks: (1) **WritingPrompts** for creative narrative generation; (2) **BillSum** for constrained abstractive summarization; and (3) **EssayWriting** for structured argumentative writing. This selection allows us to verify the adaptability of CARE across varying objectives, ranging from stylistic creativity to strict information density. Detailed dataset sources, statistics, and preprocessing steps are provided in Appendix A.

4.3 Baselines

We compare CARE with the foundation model and several representative alignment strategies. Detailed implementation settings are provided in Appendix B.

Qwen2.5-1.5B The unaligned foundation model acting as the initial policy.

Iter-Initial A baseline trained using the rubric from the first iteration of CARE as a static reward model.

Iter-Final A baseline trained using the rubric from the final iteration of CARE as a static reward model.

AutoPrompt (Levi et al., 2024) A static optimization baseline that generates a fixed, optimized rubric based on the initial dataset prior to training.

MPO (Kim et al., 2025) A dynamic alignment framework that utilizes a meta-learner to iteratively evolve reward prompts, serving as a state-of-the-art dynamic comparator.

4.4 Evaluation Metrics

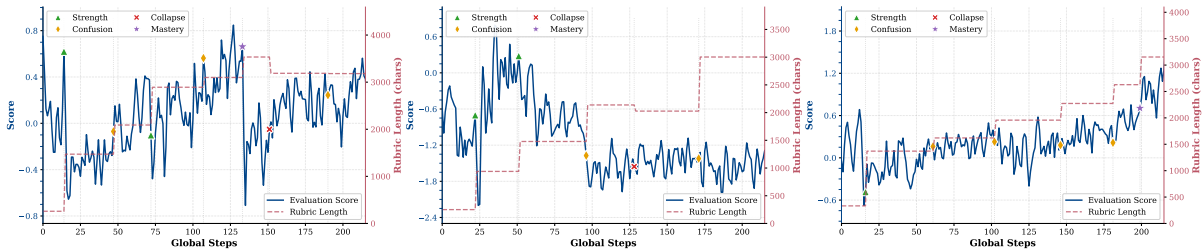
Standard Elo Rating (Elo). We conduct blind pairwise comparisons using GPT-4o (OpenAI et al., 2024) as a judge and compute mean Elo (Zheng et al., 2023) ratings relative to the baselines. This metric reflects overall response quality aggregated across multiple independent simulations.

AlpacaEval Win Rate (WR). We evaluate pairwise preference using the AlpacaEval 2.0 (Li et al., 2023) pipeline. Model outputs are compared against the Qwen2.5-3B-Instruct (Qwen et al., 2025) baseline, utilizing GPT-4o as the adjudicator. We report the standard Win Rate (WR), which measures the fraction of pairwise comparisons where our model is preferred over the baseline.

AlpacaEval Length-Controlled Win Rate (LC-WR). To account for verbosity bias (Dubois et al., 2024) we further report the Length-Controlled Win Rate (LC-WR). This metric normalizes the preference scores from GPT-4o based on output length, penalizing responses that achieve higher preference primarily through excessive verbosity rather than genuine quality.

Model	Evaluation Metrics								
	WritingPrompts			BillSum			EssayWriting		
	Elo	WR	LC-WR	Elo	WR	LC-WR	Elo	WR	LC-WR
Qwen2.5-1.5B	682	1.6	0.3	856	11.8	11.4	599	1.5	1.8
Iter-Initial	1014	11.0	10.0	981	20.5	**15.6**	1013	9.1	11.6
Iter-Final	1097	**14.4**	14.1	985	19.7	14.7	1188	11.6	11.5
AutoPrompt	1003	9.9	8.7	1025	18.8	13.4	914	7.3	**12.1**
MPO	1093	13.6	13.3	1064	19.9	14.6	1024	12.2	11.3
CARE (Ours)	**1111**	**14.4**	**14.6**	**1089**	**21.8**	15.2	**1262**	**13.8**	11.8

Table 1: Performance comparison on three in-domain evaluation sets. We report: (1) **Elo Rating**, computed from GPT-4o based pairwise comparisons (base model centered at zero); (2) **WR**, the standard pairwise win rate obtained using AlpacaEval 2.0; (3) **LC-WR**, the length-controlled win rate following AlpacaEval 2.0.



(a) Essay Writing (b) Summarization (c) Creative Writing

Figure 3: The co-evolution of Evaluation Score (solid line) and Rubric Length (dashed line) during optimization. The process reveals a curriculum based evolution trajectory. CARE dynamically modulates rubric complexity in response to real-time training dynamics, ensuring alignment with the model's evolving capabilities.

5 Results and Analysis

5.1 Main Results

Table 1 presents the comprehensive evaluation of CARE across diverse open-ended generation domains. Our approach demonstrates consistent performance improvements, securing the highest Elo ratings across all datasets and achieving competitive or superior preference metrics relative to established benchmarks. Specifically, a comparison with static baselines reveals that dynamic rubric evolution yields measurable gains in alignment quality. This performance margin suggests that aligning the complexity of evaluation criteria with the model's training dynamics is a significant factor in stabilization. While static approaches rely on fixed targets, CARE synchronizes rubric granularity with the policy's evolving capabilities. This progressive synchronization facilitates continuous capability acquisition, mitigating potential issues associated with premature complexity or insufficient discrimination during the training trajectory.

We further analyze the robustness of alignment strategies across varying task domains. Distinct tasks impose different challenges: BillSum prioritizes strict information density, while WritingPrompts demands open-ended creativity. As shown in Table 1, static optimization methods like AutoPrompt struggle to generalize and exhibit performance fluctuations. For example, AutoPrompt performs competitively on Essay Writing LC-WR but underperforms on WritingPrompts Elo. In contrast, CARE maintains consistent superiority across all domains. This cross-task adaptability indicates that the curriculum-aware mechanism effectively balances structural adherence with semantic quality, autonomously locating the optimal evaluation strictness for diverse generation scenarios.

Beyond automated metrics, we conducted a blind human evaluation to validate our protocol and strengthen the empirical evidence. Involving 150 randomly sampled response pairs (50 per dataset), three independent annotators performed pairwise comparisons between CARE and our strongest baseline, MPO, evaluating overall quality, coherence, and task adherence. As shown in Table 3, human evaluators consistently preferred CARE, yielding an average win rate of 58.7% compared to MPO's 26.7%. The annotations demonstrated substantial inter-annotator reliability (average Fleiss' $\kappa = 0.54$) and, crucially, an 80.3% average agreement with our GPT-4o judgments. This high align-

ment solidly validates our automated evaluation protocol and confirms that CARE's performance gains reflect genuine human preferences.

5.2 Ablation Study

To evaluate the specific contributions of the core mechanisms within CARE, we conducted an ablation study by systematically replacing the adaptive curriculum mechanism with a periodic static update strategy and substituting the active diagnostic sampling with a random selection mechanism. As shown in Table 2, the absence of the curriculum mechanism leads to a consistent regression in alignment efficacy across all domains. This performance drop suggests that without synchronizing evaluation complexity with the learning trajectory of the policy, the supervision signals may become misaligned. Such signals are likely either too abstract for early-stage acquisition or differ insufficiently for later-stage refinement, which hinders effective convergence. Similarly, employing random sampling results in a decline in robustness. We interpret this degradation as an indication that unguided rubric evolution struggles to capture high-information edge cases necessary for resolving specific ambiguities. In contrast, our discrepancy-based diagnostic approach ensures that rubric updates are driven by instances of high epistemic uncertainty to maximize the signal-to-noise ratio. These findings demonstrate that both the progressive modulation of difficulty and the targeted diagnosis of failure modes are indispensable for achieving stable open-ended alignment.

5.3 Analysis of Rubric Evolution

We analyze the dynamics of rubric evolution by examining the relationship between rubric complexity, measured by text length, and policy performance. As shown in Figure 3, the rubric evolves in a non-monotonic, staircase-like pattern driven by the policy's training state rather than by iteration count. Rubric expansion is selectively triggered when the policy demonstrates stable progress, while periods of confusion or instability lead to temporary stagnation of complexity, allowing the model to consolidate its current capabilities. This adaptive pacing prevents premature exposure to overly restrictive criteria and enables recovery from constraint-induced failure modes. Overall, we observe a co-evolutionary relationship in which increases in rubric length correspond to finer-grained or stricter evaluation signals that align with the model's capability growth, indicating that rubric evolution serves as an adaptive mechanism for matching evaluation granularity to the policy's training stage rather than introducing redundant verbosity. A comprehensive analysis of the rubric evolution is provided in Appendix G.

	WritingPrompts		BillSum		EssayWriting	
Variant	**WR**	**LC-WR**	**WR**	**LC-WR**	**WR**	**LC-WR**
CARE						
w/o Curriculum	48.1	47.4	43.2	43.6	45.0	47.0
w/o Sampling	49.3	47.5	45.6	46.0	40.3	43.4

Table 2: **Ablation study of CARE.** We report the Win Rate (WR) and Length-Controlled Win Rate (LC-WR) of each variant against the full CARE model across three datasets.

Dataset	CARE Win	Tie	MPO Win	Fleiss' κ	Agreement
WritingPrompts	52%	18%	30%	0.46	74%
BillSum	64%	14%	22%	0.62	86%
Essay Writing	60%	12%	28%	0.55	81%
Average	**58.7%**	**14.7%**	**26.7%**	**0.54**	**80.3%**

Table 3: Human evaluation results comparing CARE against MPO. We report the human preference win rates, inter-annotator agreement (Fleiss' κ), and the overall agreement between human judgments and the GPT-4o automated evaluator.

6 Conclusion

In this paper, we proposed Curriculum-Aware Rubric Evolution (CARE), a closed-loop framework that addresses the challenge of constructing high-quality reward signals for open-ended generation without explicit ground truth. Unlike passive or static optimization strategies that often struggle with the trade-off between responsiveness and stability, CARE explicitly couples the evolution of the rubric with the training dynamics of the policy model. By monitoring the statistical properties of reward distributions and utilizing active diagnostic sampling, our framework ensures that the model receives progressive supervision signals with high signal-to-noise ratios. Experimental results demonstrate that this bidirectional interaction effectively mitigates reward hacking and ensures training stability, leading to superior performance. Ultimately, our findings highlight that in the absence of objective criteria, a coordinated evolution between the rubric and the policy is essential for a stable and controllable alignment process.

Limitations

While our proposed CARE framework demonstrates strong efficacy in stabilizing open-ended alignment, certain scope boundaries present promising avenues for future research. First, the current methodology utilizes a highly capable Meta-LLM to guide a more compact student policy. Extending this co-evolutionary mechanism to align frontier models—where the teacher-student capability gap naturally narrows—remains an open challenge in weak-to-strong generalization. Second, the framework intentionally focuses on highly subjective generation tasks. Adapting curriculum-aware rubric evolution to strict-logic domains, such as mathematical reasoning or code generation, would require integrating distinct deterministic reward mechanisms. Finally, while utilizing LLM-based evaluators is standard practice for tasks lacking explicit ground truth, future work could explore hybrid verification methods that blend automated rubric generation with scalable human-in-the-loop oversight to further enhance alignment reliability in increasingly complex scenarios.

References

Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. 2016. Concrete problems in ai safety. *arXiv preprint arXiv:1606.06565*.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. 2017. Wasserstein gan. *Preprint*, arXiv:1701.07875.

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, and 1 others. 2022. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*.

Yasmine Beck, Ivana Ljubić, and Martin Schmidt. 2023. A survey on bilevel optimization under uncertainty. *European Journal of Operational Research*, 311(2):401–426.

Bin Chen, Xinzge Gao, Chuanrui Hu, Penghang Yu, Hua Zhang, and Bing-Kun Bao. 2025a. Reasongrm: Enhancing generative reward models through large reasoning models. *arXiv preprint arXiv:2506.16712*.

Mark Chen. 2021. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.

Xiusi Chen, Gaotang Li, Ziqi Wang, Bowen Jin, Cheng Qian, Yu Wang, Hongru Wang, Yu Zhang, Denghui Zhang, Tong Zhang, and 1 others. 2025b. Rm-r1: Reward modeling as reasoning. *arXiv preprint arXiv:2505.02387*.

Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, and 1 others. 2021. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*.

Yann Dubois, Balázs Galambosi, Percy Liang, and Tatsunori B Hashimoto. 2024. Length-controlled alpacaeval: A simple way to debias automatic evaluators. *arXiv preprint arXiv:2404.04475*.

Yann Dubois, Chen Xuechen Li, Rohan Taori, Tianyi Zhang, Ishaan Gulrajani, Jimmy Ba, Carlos Guestrin, Percy S Liang, and Tatsunori B Hashimoto. 2023. Alpacafarm: A simulation framework for methods that learn from human feedback. *Advances in Neural Information Processing Systems*, 36:30039–30069.

Leo Gao, John Schulman, and Jacob Hilton. 2023. Scaling laws for reward model overoptimization. In *International Conference on Machine Learning*, pages 10835–10866. PMLR.

Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. 1995. *Bayesian data analysis*. Chapman and Hall/CRC.

Anisha Gunjal, Anthony Wang, Elaine Lau, Vaskar Nath, Yunzhong He, Bing Liu, and Sean Hendryx. 2025. Rubrics as rewards: Reinforcement learning beyond verifiable domains. *arXiv preprint arXiv:2507.17746*.

Taneesh Gupta, Shivam Shandilya, Xuchao Zhang, Rahul Madhavan, Supriyo Ghosh, Chetan Bansal, Huaxiu Yao, and Saravan Rajmohan. 2025. Carmo: Dynamic criteria generation for context aware reward modelling. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 2202–2261.

Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, and 1 others. 2018. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*.

Zenan Huang, Yihong Zhuang, Guoshan Lu, Zeyu Qin, Haokai Xu, Tianyu Zhao, Ru Peng, Jiaqi Hu, Zhanming Shen, Xiaomeng Hu, and 1 others. 2025. Reinforcement learning with rubric anchors. *arXiv preprint arXiv:2508.12790*.

Mengzhao Jia, Zhihan Zhang, Ignacio Cases, Zheyuan Liu, Meng Jiang, and Peng Qi. 2025. Autorubric-r1v: Rubric-based generative rewards for faithful multimodal reasoning. *arXiv preprint arXiv:2510.14738*.

Ai Jian, Jingqing Ruan, Xing Ma, Dailin Li, Qianlin Zhou, Ke Zeng, and Xunliang Cai. 2025. Patarm: Bridging pairwise and pointwise signals

via preference-aware task-adaptive reward modeling. *arXiv preprint arXiv:2510.24235*.

Zae Myung Kim, Chanwoo Park, Vipul Raheja, Suin Kim, and Dongyeop Kang. 2025. Toward evaluative thinking: Meta policy optimization with evolving reward models. *arXiv preprint arXiv:2504.20157*.

Julia Kreutzer, Joshua Uyheng, and Stefan Riezler. 2018. Reliability and learnability of human bandit feedback for sequence-to-sequence reinforcement learning. *arXiv preprint arXiv:1805.10627*.

Harrison Lee, Samrat Phatale, Hassan Mansoor, Kellie Ren Lu, Thomas Mesnard, Johan Ferret, Colton Bishop, Ethan Hall, Victor Carbune, and Abhinav Rastogi. 2023. Rlaif: Scaling reinforcement learning from human feedback with ai feedback.

Elad Levi, Eli Brosh, and Matan Friedmann. 2024. Intent-based prompt calibration: Enhancing prompt optimization with synthetic boundary cases. In *ICLR 2024 Workshop on Navigating and Addressing Data Problems for Foundation Models*.

Derek Li, Jiaming Zhou, Leo Maxime Brunswic, Abbas Ghaddar, Qianyi Sun, Liheng Ma, Yu Luo, Dong Li, Mark Coates, Jianye Hao, and 1 others. 2025. Omni-thinker: Scaling multi-task rl in llms with hybrid reward and task scheduling. *arXiv preprint arXiv:2507.14783*.

Xuechen Li, Tianyi Zhang, Yann Dubois, Rohan Taori, Ishaan Gulrajani, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 2023. Alpacaeval: An automatic evaluator of instruction-following models. https://github.com/tatsu-lab/alpaca_eval.

Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. 2023. Let's verify step by step. In *The Twelfth International Conference on Learning Representations*.

Tianci Liu, Ran Xu, Tony Yu, Ilgee Hong, Carl Yang, Tuo Zhao, and Haoyu Wang. 2025. Openrubrics: Towards scalable synthetic rubric generation for reward modeling and llm alignment. *arXiv preprint arXiv:2510.07743*.

Dakota Mahan, Duy Van Phung, Rafael Rafailov, Chase Blagden, Nathan Lile, Louis Castricato, Jan-Philipp Fränken, Chelsea Finn, and Alon Albalak. 2024. Generative reward models. *arXiv preprint arXiv:2410.12832*.

OpenAI, :, Aaron Hurst, Adam Lerer, Adam P. Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, Aleksander Mądry, Alex Baker-Whitcomb, Alex Beutel, Alex Borzunov, Alex Carney, Alex Chow, Alex Kirillov, and 401 others. 2024. Gpt-4o system card. *Preprint*, arXiv:2410.21276.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, and 1 others. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744.

Qwen, :, An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, and 25 others. 2025. Qwen2.5 technical report. *Preprint*, arXiv:2412.15115.

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

Rulin Shao, Akari Asai, Shannon Zejiang Shen, Hamish Ivison, Varsha Kishore, Jingming Zhuo, Xinran Zhao, Molly Park, Samuel G Finlayson, David Sontag, and 1 others. 2025. Dr tulu: Reinforcement learning with evolving rubrics for deep research. *arXiv preprint arXiv:2511.19399*.

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*.

Taylor Shin, Yasaman Razeghi, Robert L Logan IV, Eric Wallace, and Sameer Singh. 2020. Autoprompt: Eliciting knowledge from language models with automatically generated prompts. *arXiv preprint arXiv:2010.15980*.

Joar Skalse, Nikolaus Howe, Dmitrii Krasheninnikov, and David Krueger. 2022. Defining and characterizing reward gaming. *Advances in Neural Information Processing Systems*, 35:9460–9471.

Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F Christiano. 2020. Learning to summarize with human feedback. *Advances in neural information processing systems*, 33:3008–3021.

Richard S Sutton, Andrew G Barto, and 1 others. 1998. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge.

Junkai Zhang, Zihao Wang, Lin Gui, Swarnashree Mysore Sathyendra, Jaehwan Jeong, Victor Veitch, Wei Wang, Yunzhong He, Bing Liu, and Lifeng Jin. 2025. Chasing the tail: Effective rubric-based reward modeling for large language model post-training. *arXiv preprint arXiv:2509.21500*.

Lunjun Zhang, Arian Hosseini, Hritik Bansal, Mehran Kazemi, Aviral Kumar, and Rishabh Agarwal. 2024. Generative verifiers: Reward modeling as next-token prediction. *arXiv preprint arXiv:2408.15240*.

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.

Daniel M Ziegler, Nisan Stiennon, Jeffrey Wu, Tom B Brown, Alec Radford, Dario Amodei, Paul Christiano, and Geoffrey Irving. 2019. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*.

A Dataset Details

We utilize the following datasets to cover distinctive challenges in open-ended generation:

WritingPrompts (Creative). Sourced from `euclaise/WritingPrompts_curated`, this dataset evaluates the model's capacity for engaging, long-form storytelling. It challenges the system to maintain thematic coherence and narrative flow over extended contexts.

BillSum (Summarization). We use the `duccd/billsum-clean` version comprising US Congressional bills. This task tests the ability to distill complex legal documents into concise, factually consistent summaries, serving as a rigorous benchmark for enforcing information density and length constraints.

EssayWriting (Argumentative). Adopted from `zaemyung/writing_prompts_collection`, this dataset focuses on academic prose. It requires formulating logical arguments with a formal tone, allowing us to verify whether the rubric successfully guides the policy towards high-quality reasoning patterns.

B Detailed Baseline Configurations

To evaluate the effectiveness of CARE, we compare it against the foundational model and a diverse set of alignment strategies. To ensure a fair comparison, all policy training phases for the baselines (except the zero-shot foundation model) utilize the same Group Relative Policy Optimization (GRPO) algorithm with identical hyperparameters (e.g., learning rate, batch size, and KL coefficient) as our method, unless otherwise specified.

B.1 Foundation Model

Qwen2.5-1.5B We use the official release of Qwen2.5-1.5B as the starting policy (π_0). This model serves as the zero-shot baseline to assess the inherent capability of the model without any specific alignment training on the target task.

B.2 Static Rubric Baselines

These baselines are designed to isolate the impact of the dynamic curriculum process. They utilize fixed reward models derived from specific stages of our method, testing whether a single, high-quality rubric is sufficient for alignment.

Iter-Initial This baseline evaluates the quality of the starting point of our framework. We extract the rubric generated in the *first iteration* of the CARE process. This rubric is used as a static, immutable reward function to train the policy from scratch. High performance here would indicate that the initial diagnostic is sufficient and further curriculum evolution is unnecessary.

Iter-Final This baseline tests the necessity of the intermediate curriculum stages. We take the rubric evolved at the *final iteration* of CARE—presumably the most refined version—and use it as a static reward function to train the policy from scratch. If CARE outperforms ITER-FINAL, it demonstrates that the *process* of gradually increasing difficulty and specificity is crucial for training, rather than just the final criteria itself.

B.3 State-of-the-Art Optimization Methods

AutoPrompt (Levi et al., 2024) We employ AUTOPROMPT as a representative baseline for static prompt optimization. Before policy training, we use an LLM optimizer to iteratively refine the rubric based on the initial dataset, aiming to maximize the correlation between the rubric scores and ground truth labels (or human preferences). Once the optimization converges, the resulting rubric is fixed and used to train the policy. This contrasts with our method by separating the prompt optimization phase from the policy training phase.

MPO (Meta Policy Optimization) (Kim et al., 2025) MPO serves as our primary dynamic baseline. It introduces a "meta-learner" that adjusts the reward model's prompts during training. Unlike CARE, which evolves criteria based on a pedagogical curriculum (focusing on edge cases and difficulty), MPO focuses on "evaluative thinking" to prevent reward hacking. It monitors the training dynamics and iteratively updates the prompt to maintain alignment as the policy's distribution shifts. We implement MPO using the same meta-learner architecture as described in the original paper for a direct comparison of dynamic alignment strategies.

C Implementation Details

We implement the CARE framework using the `trl` library. The detailed training configurations are as follows:

- **Optimizer & Training:** We employ the AdamW optimizer with a learning rate of

3×10^{-6}. A static KL penalty (β) of 0.02 is applied to constrain the policy update within the GRPO framework.

- **Curriculum Configuration:** We set the group size $G = 4$ and bound rubric scores within $[-4, 4]$. The transitions between curriculum states are governed by the real-time monitoring of S_{curr} and S_{ref} using the following hyperparameters:
  - *Mastery:* Triggered when the Bayesian posterior confidence $P_{bayes} > 0.95$ with a significance margin $\delta = 0.1$. To ensure stability, we additionally require the sample mean $\mu_{S_{curr}}$ to exceed 40% of the score span.
  - *Confusion:* Activated when the intra-group deviation ratio ($\sigma_{intra}/\sigma_{S_{curr}}$) exceeds the dynamic threshold η, where η starts at 1.0 and decays by 0.01 per update step, provided that $\sigma_{S_{curr}} > 0.4$.
  - *Collapse:* Detected when the reward distributional drift $W_d > 0.75\sigma_{S_{ref}}$ (Eq. 6) is accompanied by a decline in performance ($\mu_{S_{curr}} < \mu_{S_{ref}}$).
  - *Strength:* Identified when a moderate drift $W_d > 0.5\sigma_{S_{ref}}$ occurs alongside performance improvement ($\mu_{S_{curr}} > \mu_{S_{ref}}$), indicating a positive shift in the reward landscape.

- **Hardware:** All experiments are conducted on 4 NVIDIA A100 (80GB) GPUs. To optimize memory usage, we utilize the DeepSpeed ZeRO-2 offloading strategy.

D Human Evaluation Setup and Details

To ensure the reproducibility and ethical standards of our human evaluation study, we outline the details regarding annotator recruitment, compensation, and the instructions provided to the participants.

D.1 Recruitment and Compensation

We recruited three independent annotators through our university's internal mailing list. All annotators were proficient English speakers with a background in computer science.

The evaluation task consisted of reviewing 150 response pairs. On average, the task took approximately 2.5 hours to complete. Annotators were compensated at a rate of 50 CNY per hour. This compensation rate is well above the local minimum wage in the annotators' country of residence China, ensuring fair and adequate payment for their cognitive labor. The study did not involve any risks to the participants, and no personally identifiable information (PII) was collected or stored.

D.2 Instructions Provided to Annotators

Before beginning the evaluation, annotators were provided with the following text-based instructions and scoring guidelines:

> **Task Objective:** You are presented with a prompt and two AI-generated responses (Response A and Response B). Your goal is to carefully read the prompt and both responses, and determine which response is better based on the provided criteria.
>
> **Evaluation Criteria:**
>
> - **Overall Quality:** Which response is more coherent, logically structured, and well-written?
> - **Task Adherence:** Which response more strictly follows the explicit constraints and stylistic requirements of the original prompt?
>
> **Scoring Options:**
>
> - **Win (Response A):** Response A is clearly superior to Response B.
> - **Win (Response B):** Response B is clearly superior to Response A.
> - **Tie:** Both responses are of exactly equal quality, or both completely fail the prompt such that a meaningful distinction cannot be made.
>
> *Disclaimer: Please evaluate the responses objectively. Ignore superficial differences in formatting (e.g., bolding) unless they directly impact the readability and structure of the argument. Your answers will be used purely for academic research purposes.*

E Four-Stage Rubric Evolution Pipeline

When projecting granular diagnostic findings directly onto macro-level rubric synthesis, the Meta-LLM tends to generate localized patches to address

specific failures. This direct mapping often leads to stagnation in rubric evolution or compromises the global logical coherence of the rubric. To mitigate these issues, we implement a four-stage pipeline designed to transform disparate diagnostic signals and current curriculum goals into a consistent, high-quality rubric:

Local Diagnostic Analysis. The Meta-LLM analyzes each instance in $\mathcal{D}_{diag}$ individually by comparing student responses with existing scoring rationales. By cross-referencing student responses with existing rationales, the system identifies specific rubric-to-performance gaps, ensuring that every proposed modification is grounded in empirical evidence rather than heuristic guessing.

Global Strategic Fusion. To avoid overfitting to isolated cases, local modification candidates are aggregated and reconciled under the current curriculum objectives. This stage resolves conflicts and produces a unified optimization report that preserves global evaluation consistency.

Operational Instruction Synthesis. Directly modifying complex text often results in semantic drift or formatting errors; thus, we formulate explicit directives (e.g., insertion, deletion, or rephrasing) as a buffer. This conversion ensures that the Meta-LLM operates as a precise editor, maintaining the linguistic rigor necessary for stable reward modeling.

Adaptive Rubric Evolution. Finally, the system applies the synthesized directives to update the rubric, with particular emphasis on sharpening score-level boundaries to reduce ambiguity and provide clearer reward signals. By enforcing strict criteria separation, the system ensures that the refined rubric not only reflects diagnostic insights but also provides a clearer signal for the student policy's optimization.

F Curriculum Prompts

In this section, we provide the specific instruction templates used by the Meta-LLM to evolve the rubric under different curriculum states. These instructions correspond to the strategies described in Section 3.2.2.

F.1 Mastery

Triggered when the policy consistently surpasses the baseline ($P_{bayes} \to 1$). The objective is to break performance plateaus by optimizing for subtler distinctions.

Mastery Prompt

Current State: Mastery
Objective: The model has reached a stable and competent performance level. At this stage, the objective is to differentiate high-quality, insightful, and creative responses from merely adequate ones, rather than enforcing basic correctness.
Instructions for Rubric Evolution:

- **High-Order Dimensions:** Introducing new evaluation dimensions that capture higher-order qualities such as originality, insight, synthesis, or nuanced reasoning.
- **Creative Rewards:** Allowing open-ended and creative reward signals to play a meaningful role in evaluation.
- **Constraint Relaxation:** Reducing the granularity of rigid, constraint-based criteria so they function primarily as minimum competency checks.
- **Differentiation:** Increasing the granularity of existing open-ended or qualitative dimensions to better distinguish excellent performance.

Guiding Principle: The revised rubric should intentionally reward exploratory, non-standard, yet well-justified responses, while avoiding over-specification that would unnecessarily constrain creative expression.

F.2 Confusion

Triggered when intra-group inconsistency is high ($\sigma_{intra} > \eta \cdot \sigma_{global}$). The objective is to reduce epistemic uncertainty by clarifying boundaries.

Confusion Prompt

Current State: Confusion
Objective: The current evaluation process exhibits instability and scoring inconsistency, indicating ambiguity or insufficient specification in the rubric. The objective of

this stage is to improve clarity, interpretability, and scoring consistency, rather than altering task difficulty.
Instructions for Rubric Evolution:

- **Standardization:** Making evaluation standards explicit, concrete, and operationalizable.
- **Score Differentiation:** Expanding scoring descriptions to clearly distinguish adjacent score levels.
- **Anchoring:** Enriching criteria with illustrative examples or descriptive anchors to reduce evaluator interpretation variance.
- **Constraint Injection:** Introducing moderate constraint-based criteria only when they directly resolve ambiguity.

Guiding Principle: The revised rubric should minimize subjective drift and ensure that independent evaluators applying the criteria would arrive at similar scores.

F.3 Strength

Triggered when the model improves with structural distribution shifts ($W_d \gg 0, \mu_{curr} > \mu_{ref}$). The objective is to sustain momentum while preventing reward hacking.

Strength Prompt

Current State: Strength
Objective: The model is in a healthy learning phase with neither severe score collapse nor excessive instability. The objective at this stage is to maintain a balanced progression that supports both constraint satisfaction and emerging creativity.
Instructions for Rubric Evolution:

- **Constraint Preservation:** Preserve core constraint-based criteria that define correctness, relevance, and task adherence.
- **Balanced Open-Endedness:** Allow limited but meaningful space for open-ended or qualitative evaluation to encourage thoughtful and non-trivial responses.
- **Moderate Granularity:** Use moderate scoring granularity to distinguish quality differences without introducing unnecessary complexity.
- **Targeted Refinement:** Refine existing dimensions only when they improve alignment or clarity, rather than aggressively expanding or compressing the scale.

Guiding Principle: The revised rubric should balance reliability and expressiveness, ensuring that correct and well-structured responses are rewarded while still enabling higher-quality and more insightful outputs to stand out.

F.4 Collapse

Triggered when performance degrades despite distributional drift ($W_d \gg 0, \mu_{curr} < \mu_{ref}$). The objective is to restore gradients and encourage exploration.

Collapse Prompt

Current State: Collapse
Objective: The current training phase shows signs of score collapse, with rewards overly concentrated or consistently low. The primary objective is to restore a learnable and encouraging reward signal that enables effective exploration.
Instructions for Rubric Evolution:

- **Baseline Definitions:** Introducing clear, constraint-based evaluation criteria that define attainable baseline behaviors.
- **Partial Credit:** Ensuring that partial correctness, reasonable attempts, or structurally valid responses can receive non-zero scores.
- **Coarse Granularity:** Using coarse-grained scoring scales to reduce sensitivity to minor errors and prevent early discouragement.
- **Simplification:** Avoiding premature

emphasis on creativity or fine-grained qualitative distinctions.

Guiding Principle: The rubric should serve as a scaffold for exploration, making it easy for the model to understand how to earn points even when performance is not yet optimal.

G Case Study

This section presents the rubric evolution across five curriculum phases for the essay writing task, where the rubric content is dynamically adjusted according to the training state of the policy model.

G.1 Initial (Step 0)

The initial rubric is broad and generic. The policy model (π_0) produces relevant but structurally loose essays.

Rubric

Dimension Name: Argumentation and Evidence
Description: Evaluates the strength, depth, and use of specific evidence to support the arguments presented.
Scoring Criteria:

Score 0: The essay is incoherent, completely off-topic, too short to evaluate, or lacks a clear central thesis and organization.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks structure, supporting evidence, or contains significant logical gaps.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but arguments are generic or lack depth.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with strong arguments, though it may have minor stylistic or pacing issues.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

G.2 Confusion (Step 47)

Triggered by high intra-group variance (σ_{intra}). The rubric introduces explicit constraints and mandatory structural elements to eliminate ambiguity.

Rubric

Dimension Name: Argumentation and Evidence
Description: Evaluates the strength, depth, and use of specific evidence to support the arguments presented.
Scoring Criteria:

Score 0: The essay is incoherent, completely off-topic, too short to evaluate, or lacks a clear central thesis and organization.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks structure, supporting evidence, or contains significant logical gaps.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but

arguments are generic or lack depth.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with strong arguments, though it may have minor stylistic or pacing issues.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Dimension Name: Depth of Analysis
Description: Evaluates the depth and nuance of the analysis provided in the essay.
Scoring Criteria:

Score 0: The essay lacks any significant analysis and merely restates the prompt or provides superficial observations.
Example: "The essay restates the prompt without offering any analysis or reasoning."

Score 1: The essay provides basic logical deductions but fails to delve into deeper analysis or consider multiple perspectives.
Example: "The essay makes basic logical deductions but does not explore the underlying reasons or implications."

Score 2: The essay offers a moderate level of analysis, considering some aspects of the issue but not fully exploring the complexities.
Example: "The essay provides a moderate level of analysis, touching on some key points but not fully exploring the nuances."

Score 3: The essay demonstrates a thorough and nuanced analysis, considering multiple perspectives and the underlying reasons for the issue.
Example: "The essay provides a thorough analysis, considering multiple perspectives and the underlying reasons for the issue."

Score 4: The essay offers a sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments.
Example: "The essay not only provides a thorough analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments."

Dimension Name: Engagement with Counterarguments
Description: Evaluates the student's engagement with counterarguments, ensuring a balanced and thoughtful discussion of the topic.
Scoring Criteria:

Score 0: The essay does not address any counterarguments or shows no attempt to engage with opposing viewpoints.
Example: "The essay presents a one-sided argument without acknowledging any potential counterarguments."

Score 1: The essay acknowledges the existence of counterarguments but does not provide a substantive response or analysis.
Example: "The essay mentions counterarguments but does not delve into their validity or impact on the main argument."

Score 2: The essay addresses counterarguments with some depth but lacks a nuanced analysis or fails to integrate these counterarguments into the overall argument.
Example: "The essay discusses counterarguments but does not fully integrate them into the main argument, leading to a somewhat disjointed discussion."

Score 3: The essay engages with counterarguments in a balanced and thoughtful manner, providing a nuanced analysis and integrating these counterarguments into the

overall argument.
Example: "The essay effectively addresses counterarguments, providing a balanced and nuanced discussion that strengthens the main argument."

Score 4: The essay demonstrates a sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking."

Dimension Name: Clarity and Expression
Description: Evaluates the clarity and effectiveness of the student's expression, including the use of precise language and logical flow.
Scoring Criteria:

Score 0: The response is incoherent, with unclear and disjointed ideas that fail to convey a clear message.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The response is somewhat clear but lacks precision and logical flow, making it difficult to follow.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The response is generally clear and coherent, but there are occasional lapses in precision and logical flow.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The response is clear and effectively communicated, with precise language and a logical flow, though minor stylistic or pacing issues may be present.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The response is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Dimension Name: Originality and Insight
Description: Evaluates the originality and insight of the student's response, assessing whether the essay offers unique perspectives, innovative ideas, or insightful analysis that goes beyond the basic requirements of argumentation and evidence.
Scoring Criteria:

Score 0: The essay lacks any originality or insight. It merely restates common knowledge or repeats points without adding any new perspective.
Example: "The essay reiterates widely known facts without offering any new insights or unique perspectives."

Score 1: The essay shows minimal originality or insight. It may include a few minor, novel ideas but fails to significantly elevate the discussion.
Example: "The essay includes a minor, novel idea but does not develop it or integrate it into the broader argument."

Score 2: The essay demonstrates some originality or insight. It offers a few unique perspectives or innovative ideas that add value to the discussion.
Example: "The essay presents a few unique perspectives that add value to the discussion, but they are not fully developed or integrated."

Score 3: The essay shows clear originality and insight. It offers several unique perspectives or innovative ideas that significantly enhance the discussion.
Example: "The essay presents several unique perspectives and innovative ideas that significantly enhance the discussion and provide a deeper understanding of the topic."

Score 4: The essay is highly original and insightful. It offers profound, unique perspectives or groundbreaking ideas that elevate the discussion to a new level.
Example: "The essay presents profound, unique perspectives and groundbreaking ideas that elevate the discussion to a new level, demonstrating exceptional originality and insight."

Dimension Name: Task Adherence
Description: Evaluates whether the response adheres to the task requirements, including the format (essay vs. story) and the presence of a clear central thesis and depth of argumentation.
Scoring Criteria:

Score 0: The response is completely off-topic, does not follow the task requirements, and lacks a clear central thesis.
Example: "The response is a story that does not address the assigned prompt or form a logical argument."

Score 1: The response attempts to address the task but lacks structure, a clear central thesis, or depth of argumentation.
Example: "The response is a story that attempts to address the prompt but lacks a clear central thesis and depth of argumentation."

Score 2: The response adheres to the task requirements and has a clear central thesis, but the argumentation is basic and lacks depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis but lacks depth of argumentation."

Score 3: The response adheres to the task requirements, has a clear central thesis, and demonstrates strong argumentation with some depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and strong argumentation, though it may lack some depth."

Score 4: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth."

G.3 Strength (Step 72)

Triggered by steady improvement. The rubric preserves core constraints while auditing for reward hacking and allowing moderate flexibility.

Rubric

Dimension Name: Argumentation and Evidence
Description: Evaluates the strength, depth, and use of specific evidence to support the arguments presented.
Scoring Criteria:

Score 0: The essay is incoherent, completely off-topic, too short to evaluate, or lacks a clear central thesis and organization.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks structure, supporting evidence, or contains significant logical gaps.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but arguments are generic or lack depth.
Example: "The essay claims that the issue

is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with strong arguments, though it may have minor stylistic or pacing issues.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Dimension Name: Depth of Analysis
Description: Evaluates the depth and nuance of the analysis provided in the essay.
Scoring Criteria:

Score 0: The essay lacks any significant analysis and merely restates the prompt or provides superficial observations.
Example: "The essay restates the prompt without offering any analysis or reasoning."

Score 1: The essay provides basic logical deductions but fails to delve into deeper analysis or consider multiple perspectives.
Example: "The essay makes basic logical deductions but does not explore the underlying reasons or implications."

Score 2: The essay offers a moderate level of analysis, considering some aspects of the issue but not fully exploring the complexities.
Example: "The essay provides a moderate level of analysis, touching on some key points but not fully exploring the nuances."

Score 3: The essay demonstrates a thorough and nuanced analysis, considering multiple perspectives and the underlying reasons for the issue.
Example: "The essay provides a thorough analysis, considering multiple perspectives and the underlying reasons for the issue."

Score 4: The essay offers a sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments.
Example: "The essay not only provides a thorough analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments."

Dimension Name: Engagement with Counterarguments
Description: Evaluates the student's engagement with counterarguments, ensuring a balanced and thoughtful discussion of the topic.
Scoring Criteria:

Score 0: The essay does not address any counterarguments or shows no attempt to engage with opposing viewpoints.
Example: "The essay presents a one-sided argument without acknowledging any potential counterarguments."

Score 1: The essay acknowledges the existence of counterarguments but does not provide a substantive response or analysis.
Example: "The essay mentions counterarguments but does not delve into their validity or impact on the main argument."

Score 2: The essay engages with counterarguments in a balanced and thoughtful manner, providing a nuanced analysis and integrating these counterarguments into the overall argument.
Example: "The essay effectively addresses counterarguments, providing a balanced and nuanced discussion that strengthens the main argument."

Score 3: The essay demonstrates a sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument.

Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking."

Dimension Name: Clarity and Expression
Description: Evaluates the clarity and effectiveness of the student's expression, including the use of precise language and logical flow.
Scoring Criteria:

Score 0: The response is incoherent, with unclear and disjointed ideas that fail to convey a clear message.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The response is somewhat clear but lacks precision and logical flow, making it difficult to follow.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The response is generally clear and coherent, but there are occasional lapses in precision and logical flow.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The response is clear and effectively communicated, with precise language and a logical flow, though minor stylistic or pacing issues may be present.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The response is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The response is exceptionally clear and coherent, with flawless precision and logical flow, demonstrating mastery of the language and argumentation.
Example: "The essay demonstrates exceptional clarity and coherence, with flawless precision and logical flow, effectively supporting the central thesis with high-quality evidence and nuanced analysis."

Dimension Name: Originality and Insight
Description: Evaluates the originality and insight of the student's response, assessing whether the essay offers unique perspectives, innovative ideas, or insightful analysis that goes beyond the basic requirements of argumentation and evidence.
Scoring Criteria:

Score 0: The essay lacks any originality or insight. It merely restates common knowledge or repeats points without adding any new perspective.
Example: "The essay reiterates widely known facts without offering any new insights or unique perspectives."

Score 1: The essay shows minimal originality or insight. It may include a few minor, novel ideas but fails to significantly elevate the discussion.
Example: "The essay includes a minor, novel idea but does not develop it or integrate it into the broader argument."

Score 2: The essay demonstrates clear originality and insight. It offers several unique perspectives or innovative ideas that significantly enhance the discussion.
Example: "The essay presents several unique perspectives and innovative ideas that significantly enhance the discussion and provide a deeper understanding of the topic."

Score 3: The essay is highly original and insightful. It offers profound, unique perspectives or groundbreaking ideas that elevate the discussion to a new level.
Example: "The essay presents profound, unique perspectives and groundbreaking ideas that elevate the discussion to a new level, demonstrating exceptional originality and insight."

Dimension Name: Task Adherence
Description: Evaluates whether the response adheres to the task requirements, including the format (essay vs. story) and the presence of a clear central thesis and depth of argumentation.
Scoring Criteria:

Score 0: The response is completely off-topic, does not follow the task requirements, and lacks a clear central thesis.
Example: "The response is a story that does not address the assigned prompt or form a logical argument."

Score 1: The response attempts to address the task but lacks structure, a clear central thesis, or depth of argumentation.
Example: "The response is a story that attempts to address the prompt but lacks a clear central thesis and depth of argumentation."

Score 2: The response adheres to the task requirements and has a clear central thesis, but the argumentation is basic and lacks depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis but lacks depth of argumentation."

Score 3: The response adheres to the task requirements, has a clear central thesis, and demonstrates strong argumentation with some depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and strong argumentation, though it may lack some depth."

Score 4: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth."

Dimension Name: Specificity and Depth of Examples
Description: Evaluates the use of specific and detailed examples to support the arguments presented.
Scoring Criteria:

Score 0: The essay lacks any specific examples to support the arguments.
Example: "The essay makes broad statements without any concrete examples to back them up."

Score 1: The essay provides some general examples, but they are not detailed or specific enough to support the arguments.
Example: "The essay mentions a few general examples but does not provide enough detail to substantiate the claims."

Score 2: The essay includes specific examples that effectively support the arguments.
Example: "The essay provides specific examples that clearly support and validate the main argument."

Score 3: The essay includes highly specific and detailed examples that not only support the arguments but also provide a nuanced and comprehensive analysis.
Example: "The essay uses highly specific and detailed examples to offer a nuanced and comprehensive analysis, effectively supporting and validating the main argument."

Dimension Name: Logical Flow and Coherence
Description: Evaluates the logical progression of ideas, smoothness of transitions, and overall coherence of the essay.
Scoring Criteria:

Score 0: The essay is incoherent, with no logical progression of ideas, abrupt transitions, and a lack of clear structure.

Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks a clear logical progression of ideas, with some abrupt transitions and a weak overall structure.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but the logical progression of ideas is somewhat disjointed, with some rough transitions.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with a strong logical progression of ideas, smooth transitions, and a clear structure that supports the central thesis.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage, with seamless transitions and a clear, coherent structure.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Dimension Name: Relevance and Task Adherence
Description: Evaluates the extent to which the response stays on topic and adheres to the task requirements.
Scoring Criteria:

Score 0: The response is completely off-topic and does not address the task requirements.
Example: "The essay discusses unrelated topics and fails to connect to the assigned prompt."

Score 1: The response partially addresses the task requirements but contains significant irrelevant content.
Example: "The essay mentions the task but includes a lot of unrelated information that does not support the main argument."

Score 2: The response mostly stays on topic and adheres to the task requirements, but there are minor deviations.
Example: "The essay addresses the task but includes some tangential points that do not directly support the main argument."

Score 3: The response stays on topic and adheres to the task requirements with only minor deviations.
Example: "The essay clearly addresses the task and provides relevant examples and logical deductions that support the main argument."

Score 4: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content."

G.4 Mastery (Step 133)

Triggered by high certainty ($P_{bayes} \rightarrow 1$). The ceiling is raised to demand originality, synthesis, and nuance.

Rubric

Dimension Name: Argumentation and Evidence
Description: Evaluates the strength, depth, and use of specific evidence to support the arguments presented.
Scoring Criteria:

Score 0: The essay is incoherent, com-

pletely off-topic, too short to evaluate, or lacks a clear central thesis and organization.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks structure, supporting evidence, or contains significant logical gaps.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but arguments are generic or lack depth.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with strong arguments, though it may have minor stylistic or pacing issues.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The essay demonstrates exceptional synthesis and anticipation of counter-arguments, showcasing a high level of critical thinking and insight.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments with exceptional depth and sophistication."

Dimension Name: Depth of Analysis
Description: Evaluates the depth and nuance of the analysis provided in the essay.
Scoring Criteria:

Score 0: The essay lacks any significant analysis and merely restates the prompt or provides superficial observations.
Example: "The essay restates the prompt without offering any analysis or reasoning."

Score 1: The essay provides basic logical deductions but fails to delve into deeper analysis or consider multiple perspectives.
Example: "The essay makes basic logical deductions but does not explore the underlying reasons or implications."

Score 2: The essay offers a moderate level of analysis, considering some aspects of the issue but not fully exploring the complexities.
Example: "The essay provides a moderate level of analysis, touching on some key points but not fully exploring the nuances."

Score 3: The essay demonstrates a thorough and nuanced analysis, considering multiple perspectives and the underlying reasons for the issue.
Example: "The essay provides a thorough analysis, considering multiple perspectives and the underlying reasons for the issue."

Score 4: The essay offers a sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments.
Example: "The essay not only provides a thorough analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments."

Score 5: The essay offers an exceptionally sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments with a high level of nuance and depth.
Example: "The essay not only provides

a thorough and sophisticated analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments with a high level of nuance and depth, demonstrating mastery of analysis."

Score 6: The essay demonstrates an exceptionally sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments with a high level of nuance and depth, and goes beyond the expected level of sophistication.
Example: "The essay not only provides a thorough and sophisticated analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments with a high level of nuance and depth, demonstrating mastery of analysis and going beyond the expected level of sophistication."

Dimension Name: Engagement with Counterarguments
Description: Evaluates the student's engagement with counterarguments, ensuring a balanced and thoughtful discussion of the topic.
Scoring Criteria:

Score 0: The essay does not address any counterarguments or shows no attempt to engage with opposing viewpoints.
Example: "The essay presents a one-sided argument without acknowledging any potential counterarguments."

Score 1: The essay acknowledges the existence of counterarguments but does not provide a substantive response or analysis.
Example: "The essay mentions counterarguments but does not delve into their validity or impact on the main argument."

Score 2: The essay engages with counterarguments in a balanced and thoughtful manner, providing a nuanced analysis and integrating these counterarguments into the overall argument.
Example: "The essay effectively addresses counterarguments, providing a balanced and nuanced discussion that strengthens the main argument."

Score 3: The essay demonstrates a sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking."

Score 4: The essay demonstrates an exceptionally sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument and demonstrates mastery of critical thinking.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking and demonstrating mastery of engagement with counterarguments."

Score 5: The essay demonstrates an exceptionally sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument and demonstrates mastery of critical thinking, with a high level of originality and insight.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking and demonstrating mastery of engagement with counterarguments, with a high level of originality and insight."

Score 6: The essay demonstrates an exceptionally sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument and demonstrates mastery of critical thinking and a deep understanding of the topic, showcasing exceptional originality and insight.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking and demonstrating mastery of engage-

ment with counterarguments and a deep understanding of the topic, showcasing exceptional originality and insight."

Dimension Name: Clarity and Expression
Description: Evaluates the clarity and effectiveness of the student's expression, including the use of precise language and logical flow.
Scoring Criteria:

Score 0: The response is incoherent, with unclear and disjointed ideas that fail to convey a clear message.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The response is somewhat clear but lacks precision and logical flow, making it difficult to follow.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The response is generally clear and coherent, but there are occasional lapses in precision and logical flow.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The response is clear and effectively communicated, with precise language and a logical flow, though minor stylistic or pacing issues may be present.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The response is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The response is exceptionally clear and coherent, with flawless precision and logical flow, demonstrating mastery of the language and argumentation.
Example: "The essay demonstrates exceptional clarity and coherence, with flawless precision and logical flow, effectively supporting the central thesis with high-quality evidence and nuanced analysis."

Score 6: The response is exceptionally clear and coherent, with flawless precision and logical flow, demonstrating mastery of the language and argumentation, and showcasing exceptional originality and insight.
Example: "The essay demonstrates exceptional clarity and coherence, with flawless precision and logical flow, effectively supporting the central thesis with high-quality evidence and nuanced analysis, and showcasing exceptional originality and insight."

Dimension Name: Originality and Insight
Description: Evaluates the originality and insight of the student's response, assessing whether the essay offers unique perspectives, innovative ideas, or insightful analysis that goes beyond the basic requirements of argumentation and evidence.
Scoring Criteria:

Score 0: The essay lacks any originality or insight. It merely restates common knowledge or repeats points without adding any new perspective.
Example: "The essay reiterates widely known facts without offering any new insights or unique perspectives."

Score 1: The essay shows minimal originality or insight. It may include a few minor, novel ideas but fails to significantly elevate the discussion.
Example: "The essay includes a minor, novel idea but does not develop it or integrate it into the broader argument."

Score 2: The essay demonstrates clear originality and insight. It offers several unique perspectives or innovative ideas that significantly enhance the discussion.
Example: "The essay presents several unique perspectives and innovative ideas that significantly enhance the discussion and provide a deeper understanding of the topic."

Score 3: The essay is highly original and insightful. It offers profound, unique perspectives or groundbreaking ideas that elevate the discussion to a new level.
Example: "The essay presents profound, unique perspectives and groundbreaking ideas that elevate the discussion to a new level, demonstrating exceptional originality and insight."

Score 4: The essay demonstrates exceptional originality and insight, offering a comprehensive and nuanced exploration of unique perspectives and innovative ideas.
Example: "The essay provides a comprehensive and nuanced exploration of unique perspectives and innovative ideas, demonstrating exceptional originality and insight."

Score 5: The essay demonstrates exceptional originality and insight, offering a comprehensive and nuanced exploration of unique perspectives and innovative ideas that significantly elevate the discussion.
Example: "The essay provides a comprehensive and nuanced exploration of unique perspectives and innovative ideas that significantly elevate the discussion, demonstrating exceptional originality and insight and a deep understanding of the topic."

Dimension Name: Task Adherence
Description: Evaluates whether the response adheres to the task requirements, including the format (essay vs. story) and the presence of a clear central thesis and depth of argumentation.
Scoring Criteria:

Score 0: The response is completely off-topic, does not follow the task requirements, and lacks a clear central thesis.
Example: "The response is a story that does not address the assigned prompt or form a logical argument."

Score 1: The response attempts to address the task but lacks structure, a clear central thesis, or depth of argumentation.
Example: "The response is a story that attempts to address the prompt but lacks a clear central thesis and depth of argumentation."

Score 2: The response adheres to the task requirements and has a clear central thesis, but the argumentation is basic and lacks depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis but lacks depth of argumentation."

Score 3: The response adheres to the task requirements, has a clear central thesis, and demonstrates strong argumentation with some depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and strong argumentation, though it may lack some depth."

Score 4: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth."

Score 5: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth, demonstrating mastery of task adherence.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth and mastery of task adherence."

Score 6: The response adheres to the task requirements, has a clear central thesis, and

demonstrates compelling and insightful argumentation with significant depth, demonstrating mastery of task adherence and exceptional originality and insight.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth and mastery of task adherence, and showcasing exceptional originality and insight."

Dimension Name: Specificity and Depth of Examples
Description: Evaluates the use of specific and detailed examples to support the arguments presented.
Scoring Criteria:

Score 0: The essay lacks any specific examples to support the arguments.
Example: "The essay makes broad statements without any concrete examples to back them up."

Score 1: The essay provides some general examples, but they are not detailed or specific enough to support the arguments.
Example: "The essay mentions a few general examples but does not provide enough detail to substantiate the claims."

Score 2: The essay includes specific examples that effectively support the arguments.
Example: "The essay provides specific examples that clearly support and validate the main argument."

Score 3: The essay includes highly specific and detailed examples that not only support the arguments but also provide a nuanced and comprehensive analysis.
Example: "The essay uses highly specific and detailed examples to offer a nuanced and comprehensive analysis, effectively supporting and validating the main argument."

Score 4: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating a high level of critical thinking."

Score 5: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking."

Score 6: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking and exceptional originality and insight.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking and exceptional originality and insight."

Dimension Name: Logical Flow and Coherence
Description: Evaluates the logical progression of ideas, smoothness of transitions, and overall coherence of the essay.
Scoring Criteria:

Score 0: The essay is incoherent, with no logical progression of ideas, abrupt transitions, and a lack of clear structure.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks a clear logical progression of ideas, with some abrupt transitions and a weak overall structure.
Example: "The essay states a personal opinion on the topic but repeats the same point

multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but the logical progression of ideas is somewhat disjointed, with some rough transitions.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with a strong logical progression of ideas, smooth transitions, and a clear structure that supports the central thesis.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage, with seamless transitions and a clear, coherent structure.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The essay is exceptionally coherent and demonstrates flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis.
Example: "The essay demonstrates exceptional coherence and flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis with high-quality evidence and nuanced analysis."

Score 6: The essay is exceptionally coherent and demonstrates flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis, showcasing exceptional originality and insight.
Example: "The essay demonstrates exceptional coherence and flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis with high-quality evidence and nuanced analysis, showcasing exceptional originality and insight."

Dimension Name: Relevance and Task Adherence
Description: Evaluates the extent to which the response stays on topic and adheres to the task requirements.
Scoring Criteria:

Score 0: The response is completely off-topic and does not address the task requirements.
Example: "The essay discusses unrelated topics and fails to connect to the assigned prompt."

Score 1: The response partially addresses the task requirements but contains significant irrelevant content.
Example: "The essay mentions the task but includes a lot of unrelated information that does not support the main argument."

Score 2: The response mostly stays on topic and adheres to the task requirements, but there are minor deviations.
Example: "The essay addresses the task but includes some tangential points that do not directly support the main argument."

Score 3: The response stays on topic and adheres to the task requirements with only minor deviations.
Example: "The essay clearly addresses the task and provides relevant examples and logical deductions that support the main argument."

Score 4: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content."

Score 5: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence."

Score 6: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence, showcasing exceptional originality and insight.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence and showcasing exceptional originality and insight."

Dimension Name: Task Adherence Strictness
Description: Evaluates the strict adherence to the task requirements, penalizing responses that significantly deviate from the task.
Scoring Criteria:

Score 0: The response is completely off-topic and does not address the task requirements.
Example: "The essay discusses unrelated topics and fails to connect to the assigned prompt."

Score 1: The response partially addresses the task requirements but contains significant irrelevant content.
Example: "The essay mentions the task but includes a lot of unrelated information that does not support the main argument."

Score 2: The response mostly stays on topic and adheres to the task requirements, but there are minor deviations.
Example: "The essay addresses the task but includes some tangential points that do not directly support the main argument."

Score 3: The response stays on topic and adheres to the task requirements with only minor deviations.
Example: "The essay clearly addresses the task and provides relevant examples and logical deductions that support the main argument."

Score 4: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content."

Score 5: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence."

Score 6: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence, showcasing exceptional originality and insight.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence and showcasing exceptional originality and insight."

G.5 Collapse (Step 151)

Triggered by performance degradation ($W_d \gg 0, \mu < \mu_{ref}$). The rubric is simplified to restore gradient and encourage basic recovery.

Rubric

Dimension Name: Argumentation and Evidence
Description: Evaluates the strength, depth, and use of specific evidence to support the

arguments presented.
Scoring Criteria:

Score 0: The essay is incoherent, completely off-topic, too short to evaluate, or lacks a clear central thesis and organization.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks structure, supporting evidence, or contains significant logical gaps.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but arguments are generic or lack depth.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with strong arguments, though it may have minor stylistic or pacing issues.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The essay demonstrates exceptional synthesis and anticipation of counter-arguments, showcasing a high level of critical thinking and insight.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments with exceptional depth and sophistication."

Dimension Name: Depth of Analysis
Description: Evaluates the depth and nuance of the analysis provided in the essay.
Scoring Criteria:

Score 0: The essay lacks any significant analysis and merely restates the prompt or provides superficial observations.
Example: "The essay restates the prompt without offering any analysis or reasoning."

Score 1: The essay provides basic logical deductions but fails to delve into deeper analysis or consider multiple perspectives.
Example: "The essay makes basic logical deductions but does not explore the underlying reasons or implications."

Score 2: The essay offers a moderate level of analysis, considering some aspects of the issue but not fully exploring the complexities.
Example: "The essay provides a moderate level of analysis, touching on some key points but not fully exploring the nuances."

Score 3: The essay demonstrates a thorough and nuanced analysis, considering multiple perspectives and the underlying reasons for the issue.
Example: "The essay provides a thorough analysis, considering multiple perspectives and the underlying reasons for the issue."

Score 4: The essay offers a sophisticated and insightful analysis, synthesizing complex ideas and anticipating potential counter-arguments.
Example: "The essay not only provides a thorough analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments."

Score 5: The essay offers an exceptionally sophisticated and insightful analysis, syn-

thesizing complex ideas and anticipating potential counter-arguments with a high level of nuance and depth, demonstrating mastery of analysis.
Example: "The essay not only provides a thorough and sophisticated analysis but also synthesizes complex ideas and effectively addresses potential counter-arguments with a high level of nuance and depth, demonstrating mastery of analysis."

Dimension Name: Engagement with Counterarguments
Description: Evaluates the student's engagement with counterarguments, ensuring a balanced and thoughtful discussion of the topic.
Scoring Criteria:

Score 0: The essay does not address any counterarguments or shows no attempt to engage with opposing viewpoints.
Example: "The essay presents a one-sided argument without acknowledging any potential counterarguments."

Score 1: The essay acknowledges the existence of counterarguments but does not provide a substantive response or analysis.
Example: "The essay mentions counterarguments but does not delve into their validity or impact on the main argument."

Score 2: The essay engages with counterarguments in a balanced and thoughtful manner, providing a nuanced analysis and integrating these counterarguments into the overall argument.
Example: "The essay effectively addresses counterarguments, providing a balanced and nuanced discussion that strengthens the main argument."

Score 3: The essay demonstrates a sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking."

Score 4: The essay demonstrates an exceptionally sophisticated engagement with counterarguments, offering a deeply nuanced analysis that significantly enhances the overall argument and demonstrates mastery of critical thinking.
Example: "The essay not only addresses counterarguments but also synthesizes them into a compelling and insightful discussion, showcasing a high level of critical thinking and demonstrating mastery of engagement with counterarguments."

Dimension Name: Clarity and Expression
Description: Evaluates the clarity and effectiveness of the student's expression, including the use of precise language and logical flow.
Scoring Criteria:

Score 0: The response is incoherent, with unclear and disjointed ideas that fail to convey a clear message.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The response is somewhat clear but lacks precision and logical flow, making it difficult to follow.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The response is generally clear and coherent, but there are occasional lapses in precision and logical flow.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The response is clear and effectively communicated, with precise language and a logical flow, though minor stylistic or pacing issues may be present.
Example: "The essay supports its central thesis with relevant examples and logical

deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The response is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The response is exceptionally clear and coherent, with flawless precision and logical flow, demonstrating mastery of the language and argumentation.
Example: "The essay demonstrates exceptional clarity and coherence, with flawless precision and logical flow, effectively supporting the central thesis with high-quality evidence and nuanced analysis."

Dimension Name: Originality and Insight
Description: Evaluates the originality and insight of the student's response, assessing whether the essay offers unique perspectives, innovative ideas, or insightful analysis that goes beyond the basic requirements of argumentation and evidence.
Scoring Criteria:

Score 0: The essay lacks any originality or insight. It merely restates common knowledge or repeats points without adding any new perspective.
Example: "The essay reiterates widely known facts without offering any new insights or unique perspectives."

Score 1: The essay shows minimal originality or insight. It may include a few minor, novel ideas but fails to significantly elevate the discussion.
Example: "The essay includes a minor, novel idea but does not develop it or integrate it into the broader argument."

Score 2: The essay demonstrates clear originality and insight. It offers several unique perspectives or innovative ideas that significantly enhance the discussion.
Example: "The essay presents several unique perspectives and innovative ideas that significantly enhance the discussion and provide a deeper understanding of the topic."

Score 3: The essay is highly original and insightful. It offers profound, unique perspectives or groundbreaking ideas that elevate the discussion to a new level.
Example: "The essay presents profound, unique perspectives and groundbreaking ideas that elevate the discussion to a new level, demonstrating exceptional originality and insight."

Score 4: The essay demonstrates exceptional originality and insight, offering a comprehensive and nuanced exploration of unique perspectives and innovative ideas.
Example: "The essay provides a comprehensive and nuanced exploration of unique perspectives and innovative ideas, demonstrating exceptional originality and insight."

Score 5: The essay demonstrates exceptional originality and insight, offering a comprehensive and nuanced exploration of unique perspectives and innovative ideas that significantly elevate the discussion.
Example: "The essay provides a comprehensive and nuanced exploration of unique perspectives and innovative ideas that significantly elevate the discussion, demonstrating exceptional originality and insight and a deep understanding of the topic."

Dimension Name: Task Adherence
Description: Evaluates whether the response adheres to the task requirements, including the format (essay vs. story) and the presence of a clear central thesis and depth of argumentation.
Scoring Criteria:

Score 0: The response is completely off-topic, does not follow the task requirements, and lacks a clear central thesis.
Example: "The response is a story that does

not address the assigned prompt or form a logical argument."

Score 1: The response attempts to address the task but lacks structure, a clear central thesis, or depth of argumentation.
Example: "The response is a story that attempts to address the prompt but lacks a clear central thesis and depth of argumentation."

Score 2: The response adheres to the task requirements and has a clear central thesis, but the argumentation is basic and lacks depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis but lacks depth of argumentation."

Score 3: The response adheres to the task requirements, has a clear central thesis, and demonstrates strong argumentation with some depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and strong argumentation, though it may lack some depth."

Score 4: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth."

Score 5: The response adheres to the task requirements, has a clear central thesis, and demonstrates compelling and insightful argumentation with significant depth, demonstrating mastery of task adherence.
Example: "The response is an essay that addresses the prompt with a clear central thesis and compelling, insightful argumentation, demonstrating significant depth and mastery of task adherence."

Dimension Name: Specificity and Depth of Examples
Description: Evaluates the use of specific and detailed examples to support the arguments presented.
Scoring Criteria:

Score 0: The essay lacks any specific examples to support the arguments.
Example: "The essay makes broad statements without any concrete examples to back them up."

Score 1: The essay provides some general examples, but they are not detailed or specific enough to support the arguments.
Example: "The essay mentions a few general examples but does not provide enough detail to substantiate the claims."

Score 2: The essay includes specific examples that effectively support the arguments.
Example: "The essay provides specific examples that clearly support and validate the main argument."

Score 3: The essay includes highly specific and detailed examples that not only support the arguments but also provide a nuanced and comprehensive analysis.
Example: "The essay uses highly specific and detailed examples to offer a nuanced and comprehensive analysis, effectively supporting and validating the main argument."

Score 4: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating a high level of critical thinking."

Score 5: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking."

Score 6: The essay includes exceptionally specific and detailed examples that provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking and exceptional originality and insight.
Example: "The essay uses exceptionally specific and detailed examples to provide a thorough and insightful analysis, significantly enhancing the overall argument and demonstrating mastery of critical thinking and exceptional originality and insight."

Dimension Name: Logical Flow and Coherence
Description: Evaluates the logical progression of ideas, smoothness of transitions, and overall coherence of the essay.
Scoring Criteria:

Score 0: The essay is incoherent, with no logical progression of ideas, abrupt transitions, and a lack of clear structure.
Example: "The text consists of fragmented sentences or unrelated ideas that fail to address the assigned prompt or form a logical argument."

Score 1: The essay attempts to address the prompt but lacks a clear logical progression of ideas, with some abrupt transitions and a weak overall structure.
Example: "The essay states a personal opinion on the topic but repeats the same point multiple times without offering any factual backing, reasoning, or examples."

Score 2: The essay addresses the prompt with a basic structure and clear thesis, but the logical progression of ideas is somewhat disjointed, with some rough transitions.
Example: "The essay claims that the issue is important and lists a few general reasons (e.g., 'it affects many people'), but fails to provide specific details or in-depth analysis to substantiate these claims."

Score 3: The essay is well-organized and persuasive with a strong logical progression of ideas, smooth transitions, and a clear structure that supports the central thesis.
Example: "The essay supports its central thesis with relevant examples and logical deductions, clearly explaining how the provided evidence connects to and validates the main argument."

Score 4: The essay is compelling and insightful, demonstrating sophisticated reasoning, excellent flow, and precise language usage, with seamless transitions and a clear, coherent structure.
Example: "The essay not only employs high-quality evidence to support the thesis but also synthesizes complex ideas to offer a nuanced analysis, effectively anticipating and addressing potential counter-arguments."

Score 5: The essay is exceptionally coherent and demonstrates flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis.
Example: "The essay demonstrates exceptional coherence and flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis with high-quality evidence and nuanced analysis."

Score 6: The essay is exceptionally coherent and demonstrates flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis, showcasing exceptional originality and insight.
Example: "The essay demonstrates exceptional coherence and flawless logical flow, with seamless transitions and a clear, coherent structure that effectively supports the central thesis with high-quality evidence and nuanced analysis, showcasing exceptional originality and insight."

Dimension Name: Relevance and Task Adherence
Description: Evaluates the extent to which the response stays on topic and adheres to the task requirements, with a clear penalty for responses that significantly deviate from the task requirements.
Scoring Criteria:

Score 0: The response is completely off-topic and does not address the task requirements.
Example: "The essay discusses unrelated topics and fails to connect to the assigned prompt."

Score 1: The response partially addresses the task requirements but contains significant irrelevant content.
Example: "The essay mentions the task but includes a lot of unrelated information that does not support the main argument."

Score 2: The response mostly stays on topic and adheres to the task requirements, but there are minor deviations.
Example: "The essay addresses the task but includes some tangential points that do not directly support the main argument."

Score 3: The response stays on topic and adheres to the task requirements with only minor deviations.
Example: "The essay clearly addresses the task and provides relevant examples and logical deductions that support the main argument."

Score 4: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content."

Score 5: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence."

Score 6: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence, showcasing exceptional originality and insight.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence and showcasing exceptional originality and insight."

Dimension Name: Task Adherence Strictness
Description: Evaluates the strict adherence to the task requirements, penalizing responses that significantly deviate from the task.
Scoring Criteria:

Score 0: The response is completely off-topic and does not address the task requirements.
Example: "The essay discusses unrelated topics and fails to connect to the assigned prompt."

Score 1: The response partially addresses the task requirements but contains significant irrelevant content.
Example: "The essay mentions the task but includes a lot of unrelated information that does not support the main argument."

Score 2: The response mostly stays on topic and adheres to the task requirements, but there are minor deviations.
Example: "The essay addresses the task but includes some tangential points that do not directly support the main argument."

Score 3: The response stays on topic and adheres to the task requirements with only minor deviations.
Example: "The essay clearly addresses the task and provides relevant examples and logical deductions that support the main argument."

Score 4: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument.
Example: "The essay is tightly focused on

the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content."

Score 5: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence."

Score 6: The response is highly relevant and adheres closely to the task requirements, demonstrating a clear and focused argument and mastery of task adherence, showcasing exceptional originality and insight.
Example: "The essay is tightly focused on the task, providing relevant and specific examples that effectively support the main argument without any irrelevant content, demonstrating mastery of task adherence and showcasing exceptional originality and insight."