SELF-REINFORCING CREATIVITY IN DIFFUSION GENERATIVE MODELS

Anonymous authors
Paper under double-blind review



Figure 1: Images generated by our creative diffusion generative model.

ABSTRACT

Creativity is commonly understood as the production of content that is both novel and meaningful. While existing diffusion models can synthesize images that appear to generalize beyond the training data, their creative capacity remains limited. In this paper, we present a method for fine-tuning a diffusion model to generate highly creative image samples that extend beyond its initial output distribution. This is achieved by rewarding the model for generating images that are perceptually distant from its original outputs, using a distance function derived from the model itself. At the same time, we keep its distribution close to the pre-trained one through a KL penalty. Despite using no external supervision or reward functions, our model preserves the pre-trained model's image quality while producing highly novel samples. We further show that existing reinforcement-learning methods for diffusion-model post-training can fine-tune such models directly, as they optimize KL-regularized reward objectives of the same form. These off-the-shelf methods can therefore use the model's own novelty signal to amplify its creative capacity, making the process *self-reinforcing*. We demonstrate our method across settings ranging from a simple two-dimensional distribution and unconditional MNIST generation to a large-scale text-to-image model. Across these domains, the fine-tuned models produce outputs that depart from the typical generations of their pretrained counterparts while preserving coherent structure. The resulting models generate such outputs through standard sampling, eliminating the need for per-sample optimization or specialized inference-time steering.

1 INTRODUCTION

Creativity is commonly understood as the ability to produce content that appears both novel and meaningful to humans (Boden, 1994; Hertzmann, 2018). It moves beyond the repetitive and predictable, allowing for ideas that are surprising and engaging. Creating such ideas is inherently challenging: they must depart from familiar patterns and conventions while remaining coherent, meaningful, and appealing. In this work, we study computational creativity in the context of visual content generation. For example, given the prompt "musical instrument", standard sampling may produce familiar instruments, whereas the bottom example in Figure 1 presents an unfamiliar design that nevertheless remains recognizable as a musical instrument. In the current data-driven era, many

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

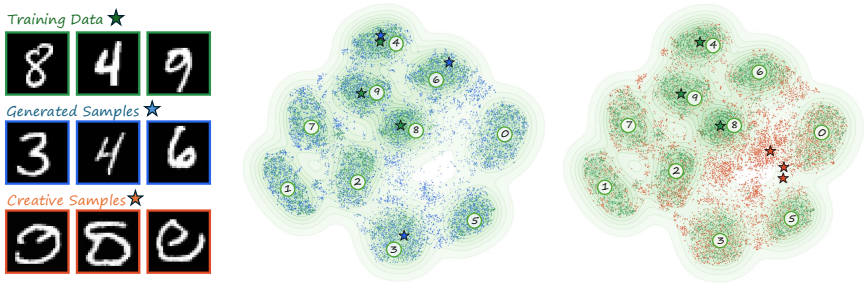


Figure 2: In a t-sne representation, MNIST training data (green), and generated samples by a diffusion model (blue) cluster in dense regions, while creative samples (red) systematically explore sparse regions. Despite lying beyond recognizable digits clusters, these samples remain smooth, structured, and glyph-like.

computational tasks are solved by learning patterns from large collections of examples. Creativity, however, requires producing precisely what the training data cannot provide: something new.

Generative models provide a powerful means of synthesizing visual content, making them a natural foundation for studying computational creativity. Their training objectives, however, primarily favor matching the training-data distribution. Although this allows them to generate samples that differ from individual training examples, their outputs remain concentrated around familiar patterns. Fig. 2 illustrates this behavior: samples from a diffusion model trained on MNIST concentrate around the digit clusters, leaving regions between them sparsely populated. The creative samples (marked in red) instead populate these regions while remaining coherent and glyph-like. In this work, we seek to learn a new generative distribution that favors departures from the pre-trained model's typical outputs, pushing generative models beyond conventional and predictable outputs toward content that is creative and surprising.

Previous methods encourage creative image generation through artistic-style supervision (Elgammal et al., 2017), learned or explicitly designed creativity objectives (Han et al., 2026; Ren, 2026; Song & Elgammal, 2026), and optimization of conditioning representations or interventions in the denoising process (Golan et al., 2026; Han et al., 2025a; Richardson et al., 2024). Many of these approaches either rely on external signals or leave the underlying generative model unchanged, rather than learning a generative distribution from which creative samples can be drawn directly.

We take a different approach: we fine-tune the pretrained diffusion model itself so that its standard denoising process directly produces creative samples. We translate the two defining properties of creativity into objectives for this distributional change. To promote novelty, samples from the fine-tuned distribution should be far from the pretrained model's typical generations. This requires a metric that captures perceptually meaningful differences between samples. Rather than introducing an external evaluator, we derive this metric from the diffusion model itself. The pretrained model therefore provides both the distributional reference for what is familiar and, through its learned distribution, the perceptual measure used to quantify novelty, requiring no external reward model, auxiliary evaluator, or additional supervision. At the same time, we constrain the fine-tuned distribution to remain close to the pretrained one using a KL-divergence penalty.

A key observation is that existing reinforcement-learning methods for diffusion-model post-training can directly learn the desired distribution because they optimize KL-regularized reward objectives of the same form. These methods can use the model's own novelty signal to reinforce its creative capacity, making the process *self-reinforcing*. In conventional post-training, the reward shifts probability mass toward preferred regions of the pretrained distribution, such as higher-quality and better prompt-adherent images. The KL regularizer prevents this shift from becoming too concentrated, keeping the post-trained distribution from becoming excessively narrow relative to the pretrained one. In our setting, the distance-based reward instead encourages the distribution to extend into unexplored regions. Here, the KL regularizer bounds this departure, ensuring that creative generations remain close enough to the pretrained distribution to remain coherent and meaningful.

We demonstrate our method's ability to produce creative and appealing images, as illustrated in Fig. 1. We compare against existing creative-generation methods both qualitatively and quantitatively, showing that our approach increases novelty and diversity while preserving prompt adherence. Finally, we study the effect of the rewarded distance metric, showing that the choice of distance strongly influences both the degree and the type of novelty induced by post-training.

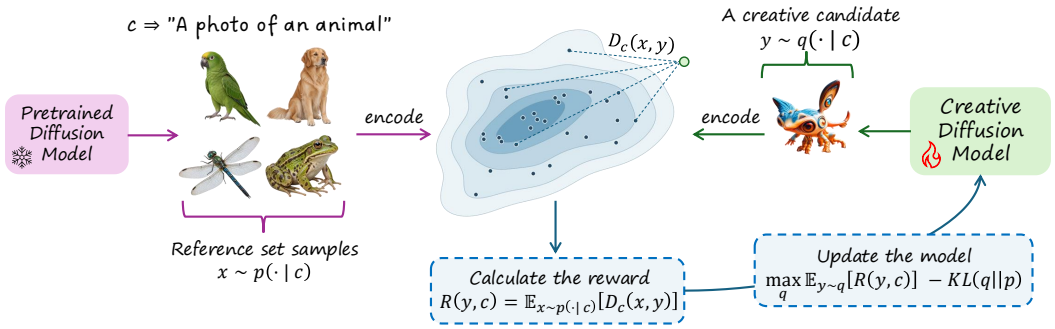


Figure 3: A candidate y sampled from the creative distribution q is rewarded according to its average model-induced distance from reference samples drawn from the pretrained distribution p. Fine-tuning maximizes novelty reward while penalizing $\mathrm{KL}(q\|p)$, encouraging creativity while preserving quality.

2 Related Work

Computational Creativity. A widely used definition of creativity combines *novelty* with *validity*, where validity refers to the usefulness or appropriateness of an idea within its context (Runco & Jaeger, 2012; Xu et al., 2012). Boden (2003) additionally emphasizes *surprise*, highlighting that creative ideas depart from familiar expectations. We adopt these criteria as a conceptual motivation for studying creative visual generation. In this setting, novelty alone is insufficient: an image may be unusual because it is incoherent or fails to reflect the user's intent. The challenge is therefore to encourage unfamiliar and surprising content while retaining the visual quality and coherence that make it meaningful.

Generalization and Creativity. Generalization enables a generative model to produce samples beyond its training examples, but this alone does not establish creativity (Kadkhodaie et al., 2024; Kamb & Ganguli, 2024; Li et al., 2023; Niedoba et al., 2024; Song et al., 2025; Vastola, 2025). Kadkhodaie et al. (2024) show that denoisers trained on non-overlapping subsets of the same dataset can converge to nearly identical score functions, demonstrating generalization beyond the memorization of individual training examples. Such generalization allows a model to produce unseen images, but does not by itself imply that their content departs from familiar patterns. Additional research further examines this phenomenon. Kamb & Ganguli (2024) show that locality and equivariance in convolutional networks enable new combinations of familiar image patches. Song et al. (2025) attribute generalization to selective underfitting. The model closely approximates the empirical score where training supervision is concentrated, but extrapolates beyond these regions during sampling. These findings explain how unseen samples arise under standard training, without explicitly encouraging the departure from familiar content that we seek. We build on this capacity to generalize by explicitly encouraging departures from familiar content while seeking to preserve visual quality and alignment with user intent.

Creative Generation with External Guidance. Existing methods encourage creative generation using various signals external to the generative model, including labeled categories, learned reward models, or feedback from auxiliary models.(Elgammal et al., 2017; Han et al., 2025a; 2026; Lu et al., 2024; Ren, 2026; Song & Elgammal, 2026). Creative Adversarial Networks (Elgammal et al., 2017) use artistic style labels to encourage departures from established styles while preserving the characteristics of each artwork. CREward (Han et al., 2026) learns creativity preferences over geometry, material, and texture from VLM-generated annotations, then uses the resulting reward models to train LoRA adapters. Other approaches leave the generator fixed and guide creativity through its conditioning. ConceptLab (Richardson et al., 2024) optimizes a reusable concept embedding under constraints that a VLM progressively expands, encouraging distinction from familiar members of a target category. VLM-guided adaptive negative prompting (Golan et al., 2026) instead uses a VLM to inspect intermediate denoising predictions and construct negative prompts for subsequent sampling steps. Although these latter methods require no creativity-labeled training data, their guidance still depends on an auxiliary model. Our approach obtains both the reference for familiarity and the perceptual novelty signal from the pretrained generator itself.

3 METHOD

In this section, we present a method for fine-tuning diffusion models to generate creative samples that are both novel and meaningful. As illustrated in Fig. 3, we begin by sampling from the model distribution under a given text condition to form a reference set of its typical outputs. We define the novelty reward of a generated sample from the fine-tuned model as its expected perceptual distance from this reference set. Rather than relying on an external evaluator, we derive the distance from the frozen pretrained diffusion model. We then maximize the expected novelty reward while using a KL penalty to keep the learned distribution close to the pretrained one. Finally, we optimize this objective using existing diffusion RL frameworks, yielding a model that generates creative samples through its standard denoising process.

We formalize the learning objective in Section 3.1, introduce the model-derived distance in Section 3.2, and describe the fine-tuning procedure in Section 3.3.

3.1 PROBLEM FORMULATION

We formulate the creative T2I generation problem in a conditional setting based on a text prompt c. The unconditional setting follows by omitting c. Let $p(\cdot|c)$ denote the pretrained model's output distribution under text condition c, and let $x_1, \ldots, x_N$ be reference samples drawn i.i.d. from $p(\cdot|c)$. Using only the pretrained model and these reference samples, we learn a new distribution $q(\cdot|c)$ from which creative samples can be drawn. We encourage the novelty aspect of creativity by rewarding perceptual distance from typical samples of $p(\cdot|c)$. To encourage meaningfulness, we keep $q(\cdot|c)$ close to $p(\cdot|c)$.

Let $R(y, c)$ denote the novelty reward of a candidate sample y generated using the text condition c. We quantify the novelty of a candidate sample y relative to the pretrained distribution by

$$R(y, c) = \mathbb{E}_{x\sim p(\cdot|c)} \left[D_c(x, y) \right], \tag{1}$$

where D_c measures the perceptual distance between two images with a shared condition c. In practice, we estimate this expectation using the reference set:

$$\widehat{R}(y, c) = \frac{1}{N} \sum_{i=1}^{N} D_c(x_i, y). \tag{2}$$

Following the principle of *minimum discrimination information* (Kullback, 1959), we seek the distribution closest to $p(\cdot|c)$ that achieves the highest expected novelty,

$$q_\lambda(\cdot|c) \in \arg\max_q \lambda \mathbb{E}_{y\sim q(\cdot|c)} \left[R(y, c) \right] - \mathrm{KL} \left(q(\cdot|c) \| p(\cdot|c) \right), \tag{3}$$

where λ is a hyperparameter that balances between the fidelity of $q(\cdot|c)$ to $p(\cdot|c)$ and the novelty of samples from $q(\cdot|c)$. Such an optimization problem admits a unique solution (Kullback, 1959), obtained by exponentially tilting the pretrained distribution,

$$q_\lambda(x|c) = \frac{p(x|c) \exp\left(\lambda R(x, c)\right)}{Z_\lambda(c)}, \tag{4}$$

where $Z_\lambda(c)$ is a normalizing constant (which we will not need to compute).

Although Eq. (4) characterizes the target distribution, it does not provide a direct way to sample from it. To be able to sample from this distribution, we fine-tune the pretrained diffusion model toward q_λ. Once trained, the resulting model generates samples through its standard denoising process, without per-sample optimization.

3.2 A DISTANCE DERIVED FROM THE MODEL ITSELF

In Eq. (1), the distance D_c determines which departures from $p(\cdot|c)$ are rewarded. The KL regularization controls how far the learned distribution moves from $p(\cdot|c)$, but not whether it moves toward meaningful outputs. Under a poorly chosen distance function, arbitrary deviations such as noise or blank images may therefore receive high novelty rewards.

In principle, one may choose D to be any existing perceptual distance measure. Perceptual distances based on external models learn a distribution implied by their training data and objectives. For example, LPIPS (Zhang et al., 2018) is calibrated against human similarity judgments, while CLIP (Radford et al., 2021) and DINO (Caron et al., 2021) define similarity through representations

learned from their respective pretraining distributions. All the above measures are trained on broad natural-image datasets, whose visual distinctions may differ from those relevant to the specific distribution $p(\cdot|c)$. These metrics may therefore emphasize variations that are meaningful under those distributions but irrelevant to $p(\cdot|c)$. For instance, when p models MNIST digits, the relevant distribution is that of handwritten digits rather than that of a representation learned from natural images. In our work, we derive D_c from the pretrained diffusion model, so that the same model determines which outputs are typical and how departures from them are measured.

Information-Estimation Metric. Recently, Ohayon et al. (2026) proposed the Information-Estimation Metric (IEM)—a perceptual image distance measure that is derived directly from a pretrained diffusion model, without using any supervision (*e.g.*, human image dissimilarity judgments). This makes the IEM a natural distance metric for our problem setting, allowing our method to be an entirely standalone, diffusion-based generative model. Concretely, the squared-IEM between a pair of points x_1, x_2 is defined as

$$\mathrm{IEM}_c^2(x_1, x_2) = \int_0^\infty \mathbb{E}_{N\sim\mathcal{N}(0,I)} \left[\|S_{p(\cdot|c)}(x_1 + \sigma N, \sigma) - S_{p(\cdot|c)}(x_2 + \sigma N, \sigma)\|^2\right] \frac{2}{\sigma^3} \mathrm{d}\sigma, \quad (5)$$

where $S_{p(\cdot|c)}(\cdot, \sigma) \coloneqq \nabla \log p_\sigma(\cdot|c)$ denotes the score of the density $p_\sigma(\cdot|c) = p(\cdot|c) * \mathcal{N}(0, \sigma^2 I)$, *i.e.*, the density $p(\cdot|c)$ blurred through convolution with an isotropic Gaussian density of standard deviation σ.[1]

Although we use the IEM as our default distance, our formulation is not limited to this choice. In the experiments, we substitute pixel- and feature-space alternatives into the same objective and report the resulting generations both qualitatively and quantitatively, showing how the choice of distance shapes the resulting notion of novelty.

3.3 Diffusion-Model Fine-Tuning

With the objective and creativity reward defined, we now describe how to fine-tune a pretrained score-based diffusion model S_p representing $p(\cdot|c)$ into a creative model S_q that samples from the target distribution $q(\cdot|c)$. We train a single model S_q over a a diverse datasets of text prompts that describe different concepts. Additional explanations on the dataset curation are presented in Appendix C. For each training prompt, we first generate a reference set of images using the frozen pretrained model. Then, we generate candidate images using the current creative model S_q. We compute the IEM between each candidate image and every reference image to estimate the candidate's distance from the reference set. The resulting rewards are used to update the creative model, balancing reward maximization against the KL penalty. Our optimization frameworks are RAM (Bergmeister et al., 2026) and DiffusionNFT (Zheng et al., 2026). Although these frameworks do not generally recover the exact exponential-tilt optimum, they provide practical approaches to reward maximization under KL regularization. At inference time, the fine-tuned model generates images using its standard denoising procedure, at the same computational cost as the pretrained model.

In practice, we approximate the IEM integral using a fixed sequence of noise levels $\sigma_1, \ldots, \sigma_{G+1}$ and independently sampled noise realizations $n_1, \ldots, n_G$, one for each noise level. For each image, we construct a vector

$$\Phi_{p(\cdot|c)}(x) = \left(S_{p(\cdot|c)}(x + \sigma_1 n_1, \sigma_1)\Delta_1, \ldots, S_{p(\cdot|c)}(x + \sigma_G n_G, \sigma_G)\Delta_G\right), \quad (6)$$

where $\Delta_i \coloneqq \sqrt{2(\sigma_{i+1} - \sigma_i)/\sigma_i^3}$. Then, using the same noise realization for a pair of images x and y at each noise level, we approximate the IEM between x and y via

$$\mathrm{IEM}_c^2(x, y) \approx \left\|\Phi_{p(\cdot|c)}(x) - \Phi_{p(\cdot|c)}(y)\right\|^2, \quad (7)$$

which forms a numerical estimate of the integral in Eq. (5). The reference samples and the base model used to construct this approximation of the IEM remain fixed throughout training. Further implementation details are provided in Appendix A.4.

4 Experiments

Our experiments evaluate the proposed method along three dimensions. We first examine how the novelty weight controls the novelty-fidelity trade-off in a two-dimensional example. We then evaluate novelty, prompt adherence, and diversity on natural images through comparisons with existing

[1]Note that the original IEM paper uses a different parameterization for the Gaussian channel. However, Eq. (5) provides an equivalent definition derived via a change of variables.

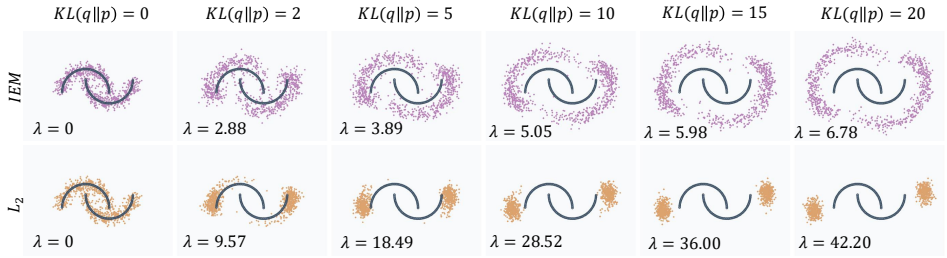


Figure 4: Effect of the novelty weight λ in the two-moons toy example. We compare exponential tilting with the diffusion-derived reward IEM (top) and the Euclidean distance reward L_2 (bottom). Each column matches $KL(q\|p)$ across the two rewards. the corresponding reward specific λ is reported inside each panel. Dark outlined points show the training samples, and colored points show samples from the tilted distribution.

methods and a user study. Finally, we study how the rewarded distance affects the generated content on MNIST and natural images.

For text-to-image generation, we use guidance-distilled Stable Diffusion 3.5 Medium (Esser et al., 2024) as our base model. We fine-tune it according to our creativity objective. Our main experiments use RAM (Bergmeister et al., 2026) as the optimization framework, with additional results using DiffusionNFT (Zheng et al., 2026) provided in the Appendix to demonstrate compatibility with an alternative RL framework. We compare our method against various creative baselines such as ConceptLab (Richardson et al., 2024), C3 (Han et al., 2025b), and VLM-Guided (Golan et al., 2026). We evaluate the baselines in their original configurations and adapt those using other backbones to Stable Diffusion 3.5. Additional qualitative results (C.2), ablations (B and implementation details (C) are provided in the Appendix.

4.1 Toy Example

We use a two-moon distribution to examine how our objective changes the generated samples as the novelty weight λ increases. Fig. 4 presents the reader with some intuition regarding which characteristics of the original samples are retained. Starting from a pretrained diffusion model of this distribution, we apply guidance according to our creativity objective. As the KL and λ values increase, the IEM shifts samples progressively farther from the base distribution while retaining recognizable moon-like forms, as can be seen in Fig. 4. As a reference, we show that such behavior is not trivial. The L_2 reward increasingly concentrates samples near the two endpoints that are globally farthest from the remaining mass. Interestingly, the IEM samples follow the original geometry rather than spreading uniformly outward from the distribution's center. Because IEM and L_2 have different scales, equal values of λ do not represent equal changes in terms of KL. We therefore select λ separately for each reward so that the two rows attain the same KL value in every column of Fig. 4. This comparison assigns both rewards the same fidelity budget and isolates the reflected geometry induced by the distance measure.

4.2 Qualitative Evaluation

We qualitatively compare our method against VLM-Guided (Golan et al., 2026), ConceptLab (Richardson et al., 2024), and C3 (Han et al., 2025b) in their original settings, as shown in Fig. 5. We use the baselines' original configurations, including ConceptLab with Kandinsky (Razzhigaev et al., 2023) and C3 with SDXL (Podell et al., 2024), which achieve higher novelty score than their SD3.5 adaptations, as presented in Table 1. Our method produces creative, unexpected images while maintaining coherent visual structure. For example, our sea creatures combine tentacles with large and colorful eyes. These features make them recognizable as marine life without depicting a specific familiar animal. Similarly, the buildings retain recognizable features, such as windows, while incorporating unconventional shapes.

The baselines produce creative variations to some extent, but the examples illustrate two potential limitations. Some outputs depart so far from the requested category that their quality and prompt adherence degrade. For instance, the left sea creature generated by C3 appears largely abstract, while the right hats generated by ConceptLab and VLM-Guided are difficult to recognize as hats. Others remain closer to familiar forms, such as the right C3 hat and the right VLM-Guided sea creature. In contrast, our examples combine substantial departures from familiar forms with recognizable category cues.

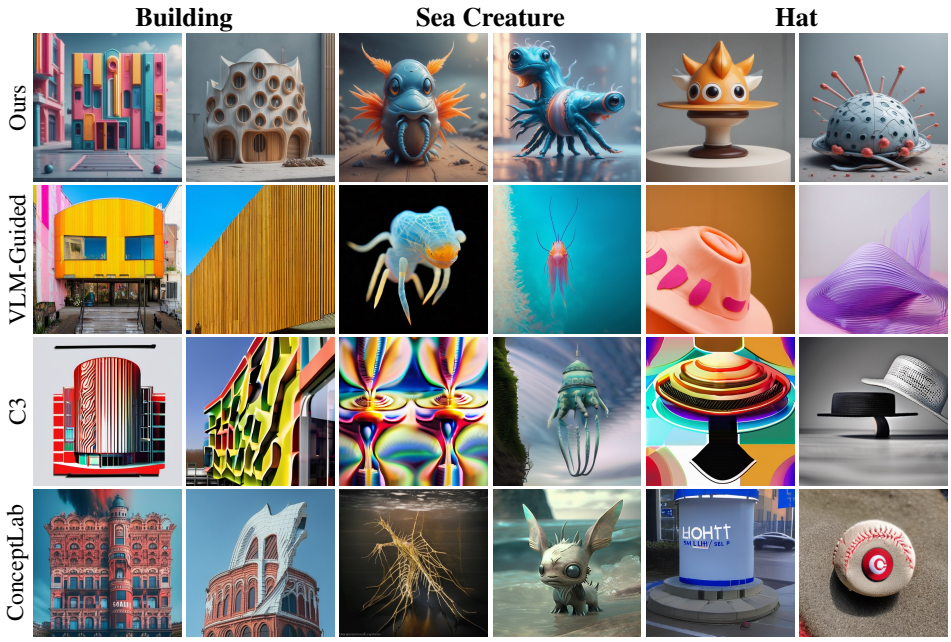


Figure 5: Qualitative comparison across four methods and three prompts. Two samples generated using the same seeds are shown for each prompt. Our method produces creative and surprising interpretations while keeping the requested objects recognizable and visually coherent.

Table 1: Comparison of validity, diversity, and novelty across creative generation methods.

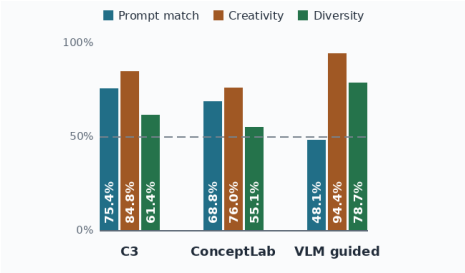
Method	VQAScore ↑	Novelty (cfg | distilled) ↑	Vendi ↑	∼ Time ↓
C3	0.845	0.012 | N/A	**3.685**	7.5m + 2s
ConceptLab	0.829	0.145 | N/A	<u>3.662</u>	8m + 2s
VLM Guided	0.848	0.055 | N/A	3.523	9s
C3 SD3.5	<u>0.880</u>	0.002, N/A	3.132	2m + 2s
ConceptLab SD3.5	0.845	0.080 | N/A	3.380	82m + 2s
ConceptLab SD3.5 + Ours	0.786	**0.158** | **0.081**	3.166	44m + 2s
Ours	**0.882**	<u>0.152</u> | <u>0.030</u>	3.032	**2s**

4.3 Quantitative Evaluation

Table 1 summarizes our quantitative evaluation. We evaluate all methods on 20 object categories, including animals, fruits, and furniture. Full set of prompts is provided in Appendix A.3. For each prompt, we generate eight images using seeds 0-7. We assess prompt adherence using VQAScore, computed with respect to the target object category, and measure diversity using Vendi Score.

We quantify novelty as a distributional departure using Kernel Inception Distance (KID) between each method's generated images and those of its corresponding backbone, conditioned on the same prompt. For each of three categories, we generate 100 images per model using seeds 0–99. A higher KID indicates a larger distributional deviation from the corresponding baseline. For the native configurations, the references are SDXL for C3, Kandinsky for ConceptLab, and SD3.5 for VLM-Guided. For our method, we report KID against both guidance-distilled and undistilled SD3.5. We report runtime per generated image, separating any optimization or adaptation overhead from sampling time. Unlike our training, which is performed once and shared across prompts, these optimization stages are performed separately for each prompt or concept and therefore constitute part of the generation cost.

C3, ConceptLab and VLM-Guided achieve higher Vendi scores. Part of this diversity may reflect weaker prompt adherence, as also seen qualitatively in Fig. 5. We evaluate these methods using their preferred prompt formulations, since our more composition-constrained prompts (full prompts



(a) Human evaluation.

(b) Distance metric comparison.

Method	VQAScore ↑	Novelty (cfg | distilled) ↑	Vendi ↑
SD3.5	0.877	0.100 | N/A	2.821
L2	0.852	0.085 | 0.016	3.055
DINO	0.874	0.114 | 0.007	2.976
CLIP	0.868	0.112 | 0.008	**3.103**
LPIPS	0.879	0.109 | **0.042**	3.069
Ours	**0.882**	**0.152** | 0.030	3.032

Figure 6: (a) Human evaluation results comparing our method with existing creative-generation baselines across creativity, diversity and prompt adherence. (b) Comparison of validity, diversity, and novelty across rewarded distance metrics.

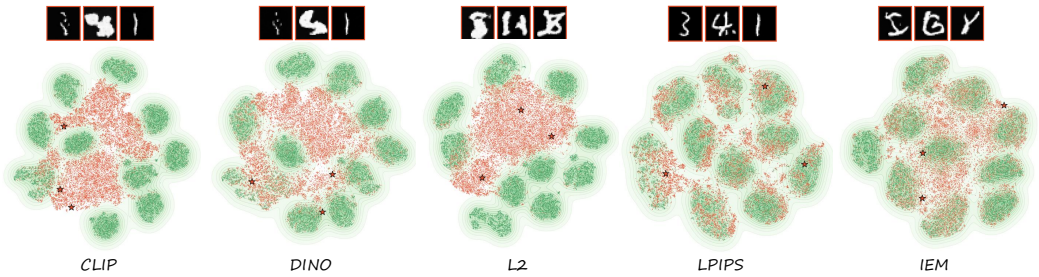


Figure 7: Distance-metric ablation on MNIST. Each panel shows a joint t-SNE of training (green) and generated (red) samples, with three matched-seed examples above. IEM preserves a coherent handwritten appearance while departing from familiar digits.

are in Appendix A.3) substantially degraded their outputs. This also partially explains the lower Vendi score of our method, which is evaluated under stricter compositional constraints while still producing meaningful and recognizable creative variations. The compatibility experiments further show that our post-trained model can serve as a base for complementary methods. In particular, ConceptLab SD3.5 + Ours achieves greater novelty than either ConceptLab SD3.5 or Ours alone.

User Study Since quantitative metrics do not always align with human judgments, we additionally conduct a pairwise human evaluation to assess creativity and quality. 34 participants observed 16 pairs, resulting in 544 response. We compare our method against VLM-Guided, ConceptLab, and C3, with each baseline evaluated in its native model and original setting. In Fig. 6a we report for each method the win rate of our method over it. Our method achieves strong win rates in perceived creativity, diversity and prompt adherence. Additional details are provided in Appendix C.1.

4.4 Distance Metric Study

We compare IEM with L_2, CLIP, DINO, and LPIPS to examine how the rewarded distance affects novelty and fidelity. The L_2 distance compares images directly, while CLIP, DINO, and LPIPS use external pretrained representations. IEM instead compares the pretrained diffusion model's denoising errors across noise levels, requiring no external model.

MNIST. We first evaluate these distances on MNIST in Figure 7, seeking symbols that depart from the ten digit classes while remaining plausibly handwritten with a few simple strokes. This setting helps distinguish meaningful novelty from departures that merely distort the appearance of the digits. We fine-tune an unconditional MNIST diffusion model using DiffusionNFT (Zheng et al., 2026), fixing training hyperparameters across distance metrics and comparing models at matched KL divergence from the pretrained model. For each metric, we compute a separate t-SNE embedding jointly over the training and generated samples.

CLIP and DINO produce visually similar samples. Together with L_2 distance, they populate regions between training clusters in their respective t-SNE embeddings. However, this spread does not consistently translate into plausible new handwritten symbols. Although the displayed samples exhibit some internal structure, their strokes are often fragmented, excessively thick, or irregular, weaken-



Figure 8: Qualitative comparison across rewarded distance metrics for the prompt "Hat". Each row corresponds to a different random seed, shared across all models within that row.

ing their handwritten appearance. LPIPS produces a more structured pattern in the visualization, with generated samples concentrated near the training-data clusters and around their peripheries. However, the displayed images remain readily recognizable as ordinary digits. Thus, while LPIPS preserves a coherent handwritten appearance, these examples show limited departure from the familiar digit classes.

In contrast, IEM produces many samples outside the dense training-data regions while retaining visible structure in the resulting distribution. The displayed samples resemble letter-like forms, similar to "I", "G", and "Y". Such outputs illustrate the domain-relative novelty we seek: although letters are familiar symbols, they are novel with respect to the ten digit classes while preserving a plausible handwritten appearance. These examples demonstrate how creative departures can retain the domain's intuitive rules of validity without being restricted to its existing categories. This experiment highlights the crucial role of choosing an appropriate distance metric to induce creative outputs that are not merely different, but also meaningful and valuable.

Natural Images. We next examine how the choice of distance metric affects text-to-image generation, using guidance-distilled SD3.5 fine-tuned with RAM as in our main experiments. Qualitative comparisons across distance metrics are presented in Figure 8, while the corresponding quantitative results are reported in Table 6b.

On natural images, CLIP and DINO again produce similar outputs, with only modest deviations from SD3.5. LPIPS favors large, red and pink compositions, while latent-space L_2 produces mostly white images. IEM yields unusual yet coherent forms with richer colors and details. Quantitatively, all rewards increase diversity over SD3.5. Furthermore, IEM achieves the highest VQAScore and novelty relative to CFG-guided SD3.5 while retaining high diversity.

5 Conclusions

We have presented an approach for reinforcing the creative capacity of pretrained generative models. Our work suggests that creativity can be more than an incidental outcome of generalization: the ability to produce surprising yet coherent content can be deliberately strengthened through post-training. Interestingly, the signal driving this change comes entirely from the model itself, suggesting that our unsupervised approach strengthens creative potential already implicit in the model, making such possibilities more readily accessible.

Our results suggest that the model's learned distribution can serve not only as a source of generation, but also as a reference for identifying what is familiar and what constitutes a meaningful departure from it. More broadly, this casts creativity as a problem of distribution shaping: rather than teaching the model what creativity should or should not look like, we encourage it to explore beyond the patterns it would ordinarily express. Since novelty in our formulation is measured relative to the pretrained model, it naturally reflects the model's own learned representation of what is familiar. An interesting direction for future work is therefore to explore how this model-relative notion relates to human perception of novelty and surprise. Similarly, while our post-training increases diversity relative to the pretrained model, further work could explore how different optimization strategies might further expand and sustain this diversity. We believe these questions open a broader direction for further exploring and exploiting the creative potential already present in generative models.

AI DISCLOSURE

In this work, we used generative AI tools for assist in the writing of proofs, propose or refine hypotheses, design or provide feedback on research methodology or experiments, implement methods, support qualitative and thematic data analysis, interpret results. We have not used generative AI tools for Generate synthetic data sets, help develop theoretical models or conceptual frameworks, provide critical ingredients for proving mathematical claims, assist with translation, clean and reformat dataset are not applicable to this work. Additionally, we used generative AI tools for Formulate questions for surveys or interviews, create or modify scientific figures or images, suggest experimental parameters, create or edit software code, creation of artifacts, draft parts of a research paper, summarize or analyse existing literature, discover research topics or identify gaps, brainstorming, sourcing/searching for information, edit a research paper to improve readability, identify relevant literature, propose a title or keywords for a research paper. We have reviewed all AI-assisted work.

REPRODUCIBILITY STATEMENT

To facilitate reproducibility, we provide implementation details in the appendix, including hyperparameter settings and techniques for efficient computation. We specify the existing frameworks used in our experiments and describe the construction of our training prompt dataset. The appendix also provides the exact evaluation settings and prompts used to obtain the reported results, complementing the method formulation and experimental protocols presented in the main text.

ETHICS STATEMENT

This work aims to support creative exploration through generative modeling. Our method builds on pretrained diffusion models and may inherit their biases and potential for generating harmful or misleading content. Encouraging novelty does not guarantee that outputs are safe or appropriate: the notion of validity studied in this work concerns adherence to the requested concept and coherent visual structure, rather than ethical acceptability. Furthermore, novelty is assessed relative to the pretrained model and does not establish originality with respect to all existing human-created works. Responsible use should therefore include appropriate safety evaluation and content safeguards, alongside respect for the licenses of the underlying models and datasets and the rights of creators.

REFERENCES

Andreas Bergmeister, Stefanie Jegelka, Nikolas Nüsken, Carles Domingo-Enrich, and Jakiw Pidstrigach. Reinforce adjoint matching: Scaling rl post-training of diffusion and flow-matching models. *arXiv preprint arXiv:2605.10759*, 2026.

M.A. Boden. The creative mind: Myths and mechanisms: Second edition. *The Creative Mind: Myths and Mechanisms: Second Edition*, pp. 1–344, 01 2003. doi: 10.4324/9780203508527.

Margaret A. Boden. Précis of the creative mind: Myths and mechanisms. *Behavioral and Brain Sciences*, 17(3):519–531, 1994. doi: 10.1017/S0140525X0003569X.

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9650–9660, 2021.

Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. Can: Creative adversarial networks, generating" art" by learning about styles and deviating from style norms. *arXiv preprint arXiv:1706.07068*, 2017.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*, 2024.

Shelly Golan, Yotam Nitzan, Zongze Wu, and Or Patashnik. VLM-guided adaptive negative prompting for creative generation. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=JzA6d2II4Q.

Jiyeon Han, Dahee Kwon, Gayoung Lee, Junho Kim, and Jaesik Choi. Enhancing creative generation on stable diffusion-based models. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28609–28618, 2025a. doi: 10.1109/CVPR52734.2025.02664.

Jiyeon Han, Dahee Kwon, Gayoung Lee, Junho Kim, and Jaesik Choi. Enhancing creative generation on stable diffusion-based models. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28609–28618. IEEE, 2025b.

Jiyeon Han, Ali Mahdavi-Amiri, Hao Zhang, and Haedong Jeong. Creward: A type-specific creativity reward model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21932–21941, 2026.

Aaron Hertzmann. Can computers create art? *Arts*, 2018.

Zahra Kadkhodaie, Florentin Guth, Eero Simoncelli, and Stéphane Mallat. Generalization in diffusion models arises from geometry-adaptive harmonic representations. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, and Y. Sun (eds.), *International Conference on Learning Representations*, volume 2024, pp. 46543–46567, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/cbaf319a4712385b5ba8a414808b5713-Paper-Conference.pdf.

Mason Kamb and Surya Ganguli. An analytic theory of creativity in convolutional diffusion models. *arXiv preprint arXiv:2412.20292*, 2024.

Solomon Kullback. *Information Theory and Statistics*. Wiley, New York, 1959.

Puheng Li, Zhong Li, Huishuai Zhang, and Jiang Bian. On the generalization properties of diffusion models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 2097–2127. Curran Associates, Inc., 2023. doi: 10.52202/075280-0101. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/06abed94583030dd50abe6767bd643b1-Paper-Conference.pdf.

Jack Lu, Ryan Teehan, and Mengye Ren. Procreate, don't reproduce! propulsive energy diffusion for creative generation. In *European Conference on Computer Vision*, pp. 397–414. Springer, 2024.

Matthew Niedoba, Berend Zwartsenberg, Kevin Murphy, and Frank Wood. Towards a mechanistic explanation of diffusion model generalization. *arXiv preprint arXiv:2411.19339*, 2024.

Guy Ohayon, Pierre-Etienne Fiquet, Florentin Guth, Jona Ballé, and Eero Simoncelli. Learning a distance measure from the information-estimation geometry of data. In *International Conference on Learning Representations*, volume 2026, pp. 5856–5894, 2026.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image synthesis. In *International Conference on Learning Representations*, volume 2024, pp. 1862–1874, 2024.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 8748–8763, 2021.

Anton Razzhigaev, Arseniy Shakhmatov, Anastasia Maltseva, Vladimir Arkhipkin, Igor Pavlov, Ilya Ryabov, Angelina Kuts, Alexander Panchenko, Andrey Kuznetsov, and Denis Dimitrov. Kandinsky: an improved text-to-image synthesis with image prior and latent diffusion. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 286–295, 2023.

Mengye Ren. Seeking the unfamiliar but memorable: Conceptual creativity as meta-learning. *arXiv preprint arXiv:2605.16477*, 2026.

Elad Richardson, Kfir Goldberg, Yuval Alaluf, and Daniel Cohen-Or. Conceptlab: Creative concept generation using vlm-guided diffusion prior constraints. *ACM Transactions on Graphics*, 43(3): 1–14, 2024.

Mark A Runco and Garrett J Jaeger. The standard definition of creativity. *Creativity research journal*, 24(1):92–96, 2012.

Kiwhan Song, Jaeyeon Kim, Sitan Chen, Yilun Du, Sham Kakade, and Vincent Sitzmann. Selective underfitting in diffusion models. *arXiv preprint arXiv:2510.01378*, 2025.

Kunpeng Song and Ahmed Elgammal. Creative image generation with diffusion models. *arXiv preprint arXiv:2601.22125*, 2026.

John Vastola. Generalization through variance: how noise shapes inductive biases in diffusion models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 34922–34960, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/564ccf9de3a2a959d3d338889bad9d88-Paper-Conference.pdf.

Kai Xu, Hao Zhang, Daniel Cohen-Or, and Baoquan Chen. Fit and diverse: set evolution for inspiring 3d shape galleries. 31(4), 2012.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–595, 2018.

Kaiwen Zheng, Huayu Chen, Haotian Ye, Haoxiang Wang, Qinsheng Zhang, Kai Jiang, Hang Su, Stefano Ermon, Jun Zhu, and Ming-Yu Liu. Diffusionnft: Online diffusion reinforcement with forward process. In *International Conference on Learning Representations*, volume 2026, pp. 134129–134150, 2026.

APPENDIX

A EXPERIMENTAL SETUP

A.1 CFG DISTILLATION

Both RL frameworks optimize the model by repeatedly sampling candidates from the current learned distribution and assigning them rewards. In standard RL post-training, where rewards typically favor properties such as perceptual quality or human preference, even candidates sampled from the unguided or weakly guided model can provide a useful learning signal. High-quality samples are reinforced, gradually shifting the model toward more desirable outputs. Our novelty objective creates a different challenge. Since reward increases with distance from the base distribution, low-quality or malformed samples can also appear highly novel and therefore receive large rewards. A KL penalty alone does not fully eliminate this failure mode, since such samples may still belong to low-quality regions of the pretrained distribution.

To ensure that the candidates presented to the RL algorithm already reflect the high-quality behavior normally obtained through classifier-free guidance (CFG), we first distill CFG into the model. We post-train SD3.5 with DiffusionNFT using the original objective, transferring the behavior of the CFG-guided distribution into the model itself. We then use this CFG-distilled model as the reference distribution and starting point for our creativity optimization. We additionally report it as a baseline, allowing us to isolate improvements due specifically to the creativity objective rather than to CFG distillation.

A.2 EFFICIENT REWARD COMPUTATION

When evaluated directly, Eq. (2) compares y with each of the N reference samples, and each comparison involves the pretrained denoiser at many noise levels. The squared IEM, however, is a squared distance between feature maps. Let $\Phi_{p,c}(x)$ denote the function

$$\Phi_{p,c}(x) : (\gamma, \varepsilon) \longmapsto e_{p,\gamma}(x, \varepsilon), \tag{8}$$

viewed as an element of the Hilbert space $\mathcal{H}$ of square-integrable functions of (γ, ε) over $[\gamma_{\min}, \gamma_{\max}]$, with inner product $\langle \Phi_{p,c}(x), \Phi_{p,c}(y) \rangle_{\mathcal{H}} = \int_{\gamma_{\min}}^{\gamma_{\max}} \mathbb{E}_{\varepsilon}[e_{p,\gamma}(x,\varepsilon)^{\top} e_{p,\gamma}(y,\varepsilon)]\, \mathrm{d}\gamma$. Because both residuals in Eq. (5) share the same noise realization, $D_c(x, y) = \|\Phi_{p,c}(x) - \Phi_{p,c}(y)\|^2_{\mathcal{H}}$. Expanding the square in Eq. (2) gives

$$\widehat{R}(y, c) = \|\Phi_{p,c}(y)\|^2_{\mathcal{H}} - 2\,\langle \Phi_{p,c}(y), \mu_{p,c} \rangle_{\mathcal{H}} + \kappa_{p,c}, \tag{9}$$

where

$$\mu_{p,c} = \frac{1}{N} \sum_{i=1}^{N} \Phi_{p,c}(x_i), \qquad \kappa_{p,c} = \frac{1}{N} \sum_{i=1}^{N} \|\Phi_{p,c}(x_i)\|^2_{\mathcal{H}}\,. \tag{10}$$

Both reference statistics depend only on the reference set and are computed once before fine-tuning, and $\kappa_{p,c}$ does not depend on y and can be omitted during optimization.

Given reference samples $x_1, \ldots, x_N \sim p$, the empirical novelty reward of a candidate y is

$$\widehat{R}_D(y) = \frac{1}{N} \sum_{i=1}^{N} \|\Phi_{p,\omega}(y) - \Phi_{p,\omega}(x_i)\|_2^2\,. \tag{11}$$

Let

$$\mu_{p,\omega} = \frac{1}{N} \sum_{i=1}^{N} \Phi_{p,\omega}(x_i), \qquad \kappa_{p,\omega} = \frac{1}{N} \sum_{i=1}^{N} \|\Phi_{p,\omega}(x_i)\|_2^2\,. \tag{12}$$

Expanding the squared distance gives

$$\widehat{R}_D(y) = \|\Phi_{p,\omega}(y)\|_2^2 - 2\Phi_{p,\omega}(y)^{\top} \mu_{p,\omega} + \kappa_{p,\omega}. \tag{13}$$

The reference statistics $\mu_{p,\omega}$ and $\kappa_{p,\omega}$ can be computed once before training. Consequently, we do not need to store or repeatedly process the complete reference set during fine-tuning. At each training step, we compute only the representation of the generated candidate and combine it with the precomputed statistics, substantially reducing memory usage and computation.



Figure 9: Qualitative comparison between the baseline and our model using identical neutral prompts. We show two prompts, with two matched random seeds per prompt.

A.3 PROMPTS

We use a shared prompt template for our creative models and baselines:

> A product design photo of an extremely {`creative subject`}. The object is shown alone and centered against a plain light gray background with soft, even studio lighting. The composition is simple, clean, and focused entirely on the object. The background is simple and uniform. The composition contains a single clearly presented object.

The placeholder {`creative subject`} combines a creativity modifier - creative, innovative, imaginary, or new type of - with an object category, such as hat, animal, or building.

We also find that the creative model can produce more creative outputs even without explicitly creative-inducing words such as ‘creative” or ‘imaginary”. As shown in Fig. 9, the baseline and creative model are given the same neutral prompt, yet the creative model produces more novel and unconventional interpretations.

We use the full prompt template for our creative models and base-model baselines. For competing methods, we use only the creative subject, omitting the surrounding template, as we observed that the wrapper substantially reduced the creativity of their outputs.

The 20 categories we use in the quantitative evaluation are: Building, Bag, Pet, Tool, Cup, Chair, Furniture, Hat, Musical instrument, Creature, Animal, Mushroom, Sea creature, Tree, Fruit, Food, Lamp, Shoe, Drinking glass, Hairstyle. KID is averaged over Building, Hat and Lamp.

A.4 IEM DECOMPOSITION.

The standard IEM-squared distance is defined as

$$IEM^2(x_1, x_2) = \int_0^\infty \mathbb{E}\Big[\|(x_1 - f(\sqrt{\gamma}x_1 + N, \gamma)) - (x_2 - f(\sqrt{\gamma}x_2 + N, \gamma))\|^2\Big]\, d\gamma. \quad (14)$$

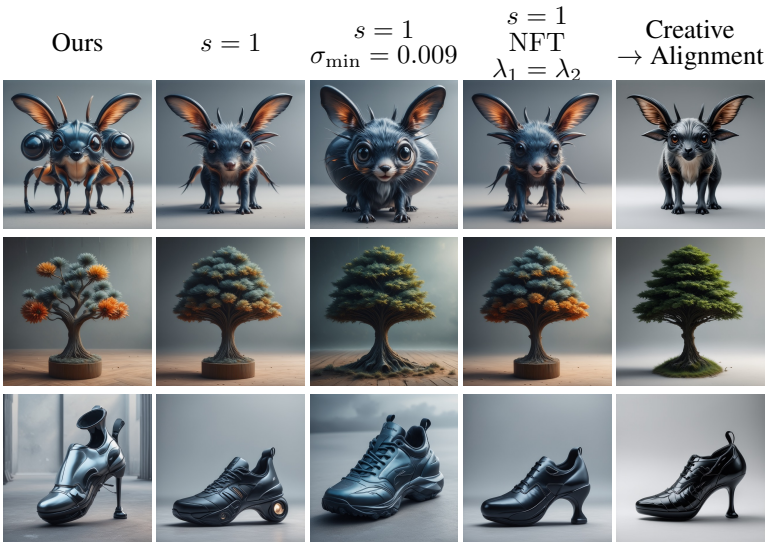


Figure 10: Design ablations across three prompts, one constant seed for each prompt.

Let us take the average with respect to the second variable:

$$\mathbb{E}\left[IEM^2(x,X)\right] = \int_0^\infty \mathbb{E}\left[\|(x - f(\sqrt{\gamma}x + N,\gamma)) - (X - f(\sqrt{\gamma}X + N,\gamma))\|^2\right] d\gamma \tag{15}$$

$$= \int_0^\infty \mathbb{E}\left[\|x - f(\sqrt{\gamma}x + N,\gamma)\|^2\right] d\gamma + \int_0^\infty \mathbb{E}\left[\|X - f(\sqrt{\gamma}X + N,\gamma)\|^2\right] d\gamma - 2\int_0^\infty \mathbb{E}\Big[(x - f(\sqrt{\gamma}x + N,\gamma))^\top (X - f(\sqrt{\gamma}X + N,\gamma))\Big] d\gamma \tag{16}$$

$$= -2\log p_X(x) - 2\int_0^\infty \mathbb{E}\Big[(x - f(\sqrt{\gamma}x + N,\gamma))^\top (X - f(\sqrt{\gamma}X + N,\gamma))\Big] d\gamma + c, \tag{17}$$

where the constant c does not depend on x. It is important because it ensures that the integrals do not diverge.

Let us denote

$$g(x) = \int_0^\infty \mathbb{E}\left[(x - f(\sqrt{\gamma}x + N,\gamma))^\top (X - f(\sqrt{\gamma}X + N,\gamma))\right] d\gamma. \tag{18}$$

Thus, if we tilt $p_X(x)$ using $\mathbb{E}[IEM^2(x,X)]$, we obtain

$$p_X(x)\exp\left\{\lambda\mathbb{E}\left[IEM^2(x,X)\right]\right\} = p_X(x)e^{-2\lambda\log p_X(x)}e^{-2\lambda g(x)}e^{\lambda c} \tag{19}$$

$$\propto e^{(1-2\lambda)\log p_X(x)}e^{-2\lambda g(x)}. \tag{20}$$

We leverage this decomposition to control the two contributions independently, replacing the shared weight λ with λ_1 for the log-density contribution and λ_2 for the cross term:

$$q_{\lambda_1,\lambda_2}(x) \propto \exp\{(1-2\lambda_1)\log p_X(x) - 2\lambda_2 g(x)\}. \tag{21}$$

The original IEM-based tilt is recovered when $\lambda_1 = \lambda_2 = \lambda$.

B Ablation Study

We ablate several design choices in our method, with quantitative results reported in Table 2 and representative generations shown in Figure 10.

Table 2: Comparison of validity, diversity, and novelty across different design choices.

Method	VQAScore ↑	Vendi ↑	Novelty (cfg | distilled) ↑
lower LoRA strength (s=1)	0.878	<u>3.020</u>	0.132 | 0.012
s=1 & high SNR in IEM	0.879	2.805	**0.162** | **0.032**
s=1 & NFT & $\lambda_1 = \lambda_2$	**0.884**	2.942	0.131 | 0.008
Creative→Alignment	0.878	2.914	0.116 | N/A
Ours	<u>0.882</u>	**3.032**	<u>0.152</u> | <u>0.030</u>

LoRA strength. Our default configuration applies the trained LoRA with strength $s = 2$. We evaluate the effect of this choice by reducing the strength to the standard value $s = 1$, reported as *Lower LoRA strength ($s = 1$)*.

Noise range in IEM. Our default IEM calculation uses only moderate-to-high noise levels. In the configuration $s = 1$ *& high SNR in IEM*, we retain $s = 1$ but expand the calculation to the full noise range, including low-noise, high-SNR levels.

RL framework and IEM formulation. Our main configuration uses RAM with $\lambda_1 = 2\lambda_2$ in the IEM calculation. In $s = 1$ *& NFT &* $\lambda_1 = \lambda_2$, we instead train with DiffusionNFT and use the vanilla IEM formulation, where $\lambda_1 = \lambda_2$. This configuration evaluates compatibility with an alternative RL framework and IEM formulation. All other configurations use $\lambda_1 = 2\lambda_2$.

Order of post-training stages. Finally, we examine the decision to perform guidance distillation before creativity post-training. RL-based guidance distillation can reduce diversity, potentially restricting the range of creative candidates available during subsequent optimization. We therefore evaluate the reverse order: starting from the original SD3.5 Medium model, we first perform creativity post-training and then apply several epochs of DiffusionNFT with its original objective to improve visual quality. For this configuration, we compute KID and relative differences with respect to the original SD3.5 Medium model, which serves as the starting point. All other configurations use SD3.5 Medium with our CFG-distillation LoRA as their starting point and reference.

As can be seen in Table 2, reducing the creative LoRA strength from $s = 2$ to $s = 1$ slightly reduces diversity and substantially reduces novelty, while validity remains nearly unchanged. Extending IEM beyond the high-SNR regime increases novelty, but reduces both validity and diversity, suggesting that focusing on higher noise levels provides a better balance between these objectives. Replacing RAM with NFT and using equal weighting $\lambda_1 = \lambda_2$ yields the highest VQAScore, but substantially reduces novelty compared with our final configuration. We additionally reverse the training order, applying creativity post-training before alignment, which results in lower novelty than our final configuration. Overall, our final configuration achieves the highest diversity while maintaining high validity and substantial deviation from both the base and guidance-distilled distributions.

Qualitatively, Figure 10 shows that the different configurations often preserve similar semantic and structural characteristics, with several ablations also producing creative interpretations. However, our final configuration tends to exhibit more pronounced departures from conventional representations while preserving the requested object.

C Implementation Details

Training. As previously described, we initialize creativity post-training from a guidance-distilled SD3.5 Medium model. We obtain this model through 12 epochs of standard DiffusionNFT post-training, using CLIPScore and PickScore as reward signals. Starting from this guidance-distilled model, we then perform 17 epochs of creativity post-training using RAM.

Our training prompts are drawn from three sources: 50% are sampled from HPDv3++, 25% from the basic category of PartiPrompts, and the remaining 25% are generated using ChatGPT 5.6 Sol, following the prompt construction scheme proposed by Golan et al. (2026). Specifically, the gen-

erated prompts follow the template: "A photo of a [creative/innovative/new type of/novel/unique] [main object] [scene description]."

IEM configuration. For the IEM calculation, we average over 256 samples drawn from the base conditional distribution. We set $\lambda_1 = 2\lambda_2$, where λ_1 and λ_2 multiply $\log p$ and g, respectively. We consider $\sigma \in [10, 1000]$ and discretize the integral using 64 integration levels. In all reported RAM experiments, we multiply the novelty reward by 10.

Inference. At inference time, we first merge the CFG-distillation LoRA into the SD3.5 Medium base model, so its effect is retained throughout all 40 sampling steps. The creativity LoRA is applied with strength $s = 2$ only during the first 8 steps, corresponding to the high-noise portion of the generation trajectory. All images are generated at a resolution of 512×512 pixels.

Ablation baselines. Below, we detail the configurations used in our ablation study.

Lower LoRA strength ($s = 1$). We reduce the inference-time strength of the creativity LoRA from $s = 2$ to $s = 1$.

$s = 1$ & high SNR in IEM. We use creativity LoRA strength $s = 1$ and extend the IEM noise range to $\sigma \in [0.009, 1000]$, incorporating low-noise, high-SNR levels.

$s = 1$ & NFT & $\lambda_1 = \lambda_2$. We replace RAM with DiffusionNFT, setting its regularization parameter to 0.25. For the IEM calculation, we use $\lambda_1 = \lambda_2$, average over 64 samples from the base conditional distribution, and discretize the integral using 32 integration levels. At inference time, we apply the creativity LoRA with strength $s = 1$.

Creative → Alignment. We reverse the order of the two post-training stages. Starting from the original, undistilled SD3.5 Medium model, we first perform 7 epochs of RAM training with our novelty reward. We then merge the resulting creativity LoRA into the model with strength $s = 0.8$ and perform 10 epochs of standard DiffusionNFT post-training, using CLIPScore and PickScore as reward signals. We sample from the resulting model without CFG. For this configuration, KID and relative differences are computed with respect to the original SD3.5 Medium model. For all other configurations, the reference is SD3.5 Medium with the CFG-distillation LoRA.

C3. For the native C3 baseline, we use SDXL with CFG scale 5 and 50 denoising steps, using ten seeds for the amplification-factor search.

Our adaptation to SD3.5 Medium retains C3's frequency amplification and per-prompt search while modifying the Fourier intervention to accommodate the different architecture. Specifically, we apply the intervention to transformer block 8 (only to image-tokens) rather than the SDXL U-Net layer groups. For each prompt, we search a single conservative amplification grid, $\lambda \in \{1.00, 1.05, \ldots, 1.25\}$, and increase the usability threshold from 0.7 to 0.8. Images are generated with CFG scale 4.5, 40 denoising steps, and a resolution of 512×512 pixels.

ConceptLab. We run the native ConceptLab baseline with Kandinsky 2.1 for 2,500 optimization steps, using BLIP2 Flan-T5-XL for live-negative discovery. Final generation uses a 5-step prior with CFG scale 4 and a 50-step decoder with CFG scale 7.5, producing images at a resolution of 512×512 pixels.

In our SD3.5 adaptation, each ConceptLab placeholder is represented by three independently learned embedding vectors, one for each text encoder: CLIP-L, CLIP-G, and T5. We disable the optional BLIP2-based live-negative discovery and instead use fixed, manually specified negative classes, represented by CLIP embeddings of reference images generated with SD3.5. At inference time, we insert the learned tokens into the prompt and sample SD3.5 with CFG scale 4.5 for 40 denoising steps.

VLM-Guided. We use the original, undistilled SD3.5 Medium model with CFG scale 4.5, 40 denoising steps, and a resolution of 512×512 pixels. At each denoising step, ViLT-VQA (`dandelin/vilt-b32-finetuned-vqa`) evaluates the current image estimate using a category-specific question, such as "What type of building is this?" The cleaned top-1 answer is appended to an accumulating negative prompt.

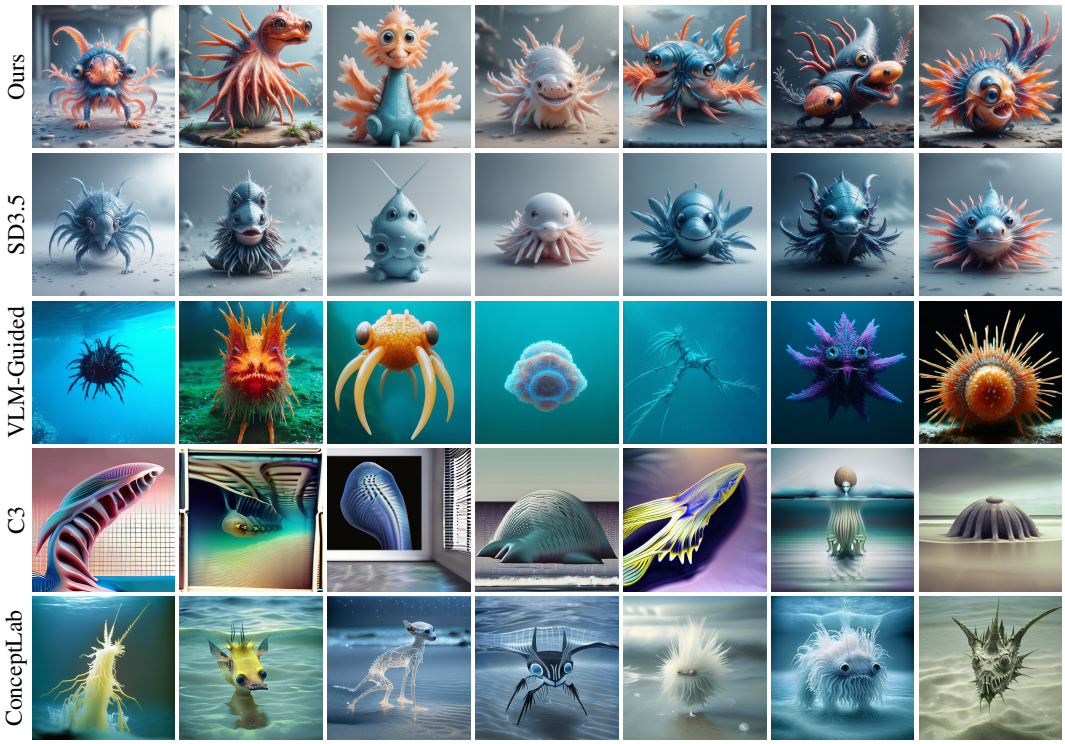


Figure 11: Qualitative comparison across five methods for the object “Sea Creature”. Seven samples generated using the same seeds are shown for each method.

C.1 User Study

The user study compares images generated by our method against four baselines: guidance distilled SD3.5, C3, ConceptLab, and VLM-negative. In each pairwise comparison, our image and a baseline image are generated from the same prompt and random seed. Their placement as Image A or Image B is counterbalanced, and method identities are hidden. Participants answer:

- Which image better matches the prompt?

- Which image is a more creative interpretation of the prompt?

For the diversity evaluation, participants compare two sets of four images generated from the same prompt and the same four seeds - one set from our method and one from a baseline. They select which set presents more diverse interpretations of the requested object while ignoring differences in background and composition (since the baseline and our creative model use composition-constrained prompts while other methods don’t). Every question allows three answers: A, B, or “About the same.”

C.2 Additional Comparisons

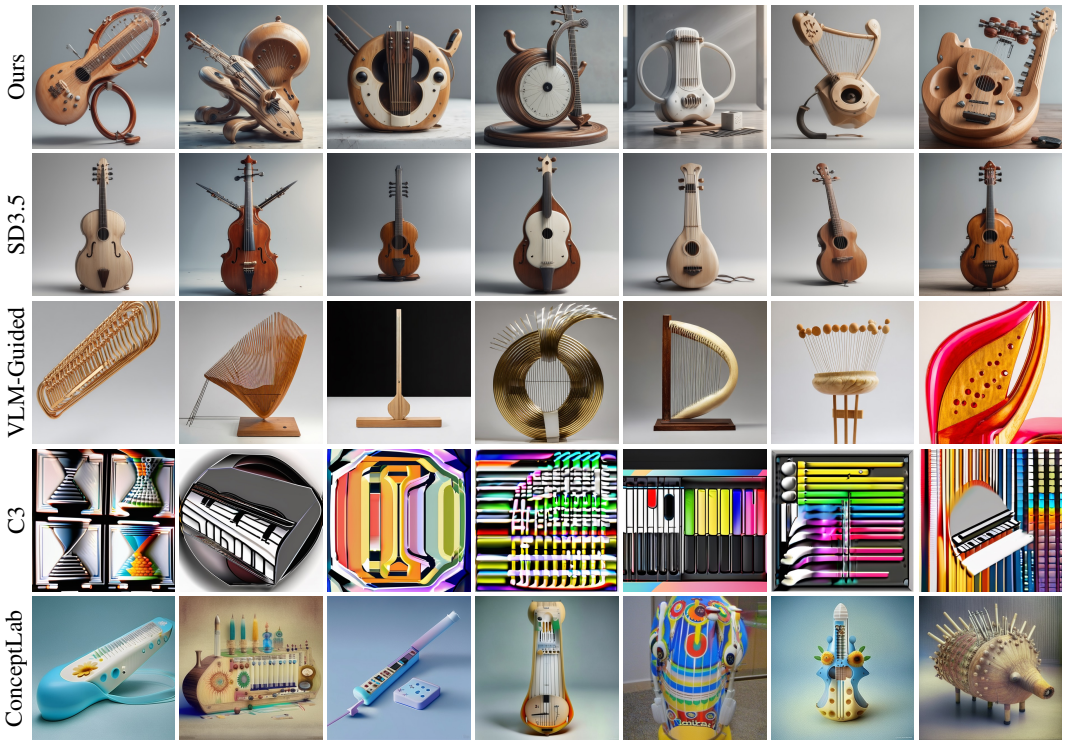


Figure 12: Qualitative comparison across five methods for the object “Musical Instrument”. Seven samples generated using the same seeds are shown for each method.

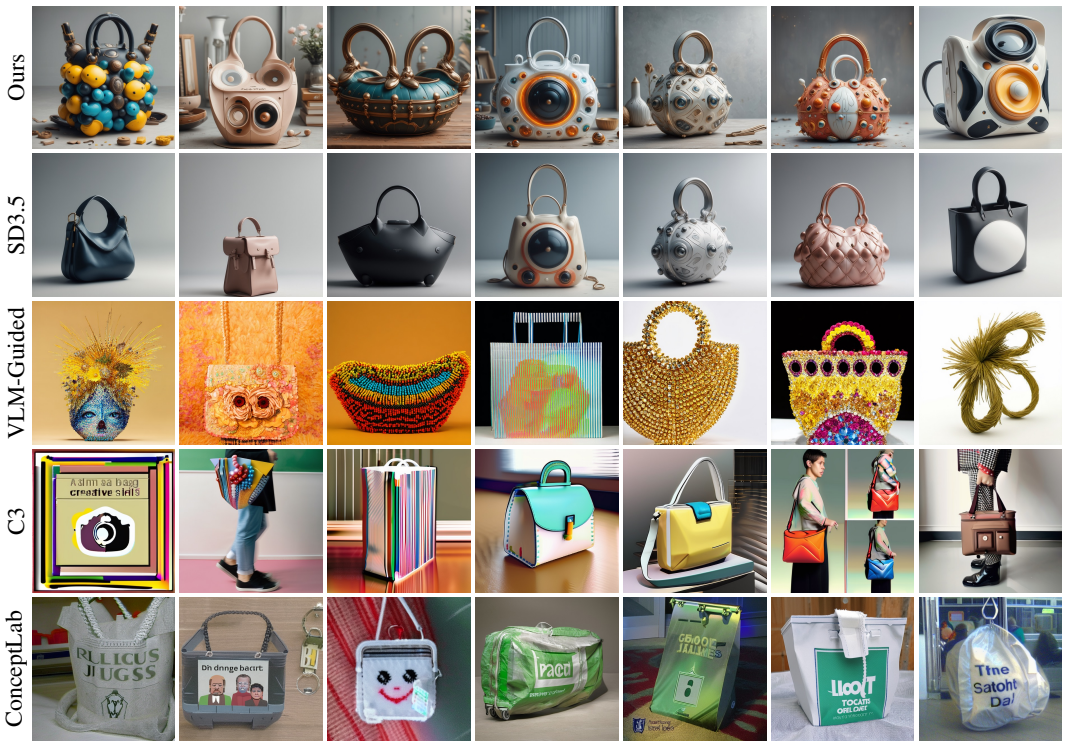


Figure 13: Qualitative comparison across five methods for the object “Bag”. Seven samples generated using the same seeds are shown for each method.

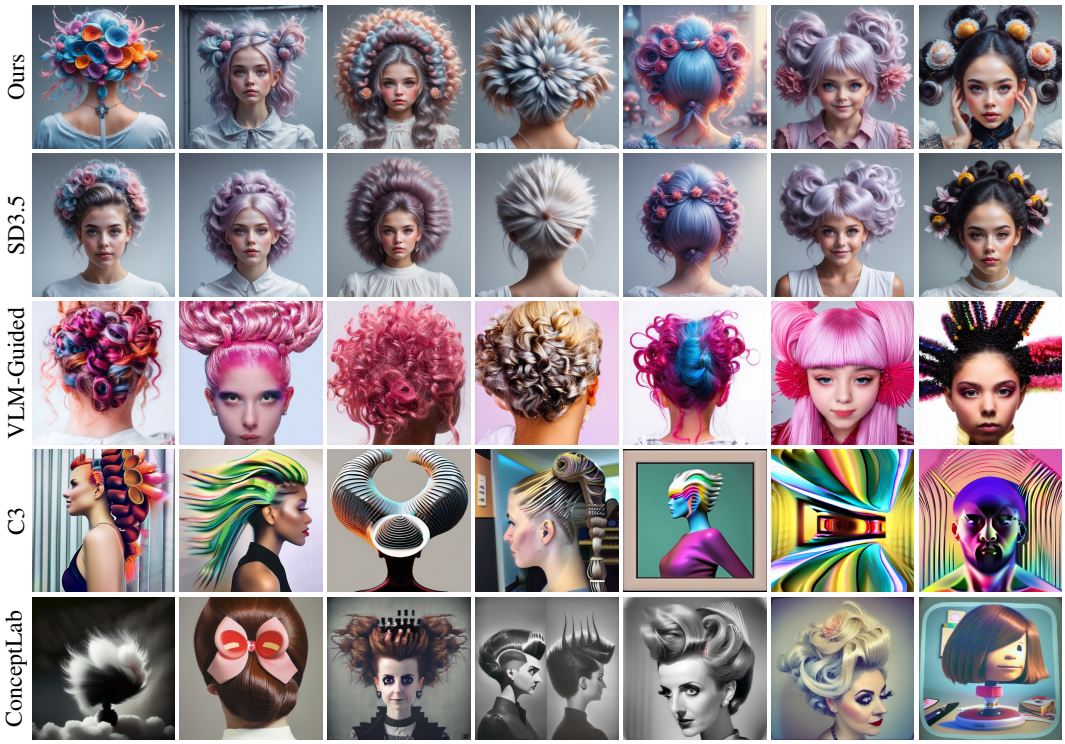


Figure 14: Qualitative comparison across five methods for the object “Hairstyle”. Seven samples generated using the same seeds are shown for each method.