MIND WANDERING AS A DECODING RULE FOR CREATIVE WRITING WITH LANGUAGE MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Language models write fluent but predictable fiction, and the usual remedy, a higher sampling temperature, trades coherence for creativity. Human writers describe another source of surprise: mind wandering, the drift of attention away from the task at hand. We propose *mind-wandering decoding*, an inference-time rule under which a word the model has just written can open a bounded episode. For a few dozen tokens, the model predicts from a second context that splices a remembered passage between the story's opening and the text so far, then returns to the story, which is never edited. Episodes can be cued by the word, start spontaneously, or be scheduled by a controller trained with policy gradients under a wandering budget. In one science-fiction story world with a 957-entry dictionary of passages set in it, we generate 2,064 stories with Qwen2.5-1.5B and Qwen2.5-7B and score them with an LLM judge. For the 1.5B model, wandering raises both creativity and coherence over vanilla sampling (by 0.26 and 0.23 points on a five-point rubric), in every opening and beyond the creativity–coherence frontier of temperature and min-p decoding. The gain tracks the breaking of repetition loops and depends neither on which passage an episode brings nor on what triggered it; the controller matches the best fixed schedules but learns mostly how much to wander, not when. At 7B, a model that rarely loops and whose coherence sits at the judge's ceiling, the gain disappears. As in incubation accounts of human creativity, what helps is less the remembered passage than the interruption itself.

1 INTRODUCTION

Large language models (LLMs) write fluent fiction, yet professional writers judge their stories predictable (Chakrabarty et al., 2024), and drafting with them homogenizes what people write (Padmakumar & He, 2024; Doshi & Hauser, 2024; Anderson et al., 2024). The usual lever for surprise, the decoding temperature or truncation rule (Holtzman et al., 2020; Meister et al., 2023; Nguyen et al., 2025), trades creativity against coherence along a well-documented frontier (Caccia et al., 2020; Zhang et al., 2021; Peeperkorn et al., 2024), and at its low end small models fall into repetition loops (Holtzman et al., 2020; Xu et al., 2022).

People who write for a living describe a different source of surprise. In an experience-sampling study, a fifth of writers' and physicists' significant ideas arrived while their minds wandered, and those ideas were more likely to resolve an impasse (Gable et al., 2019). Mind wandering, thought that drifts from the task at hand (Smallwood & Schooler, 2006; 2015), occupies nearly half of waking life (Killingsworth & Gilbert, 2010), is often cued by something in the environment (Plimpton et al., 2015; Berntsen, 2010), and supports creative incubation (Baird et al., 2012; Sio & Ormerod, 2009). A drifting mind does not rewrite what is on the page; it lets something else into working memory for a while and then returns.

We turn this into a decoding rule for LLMs (Figure 1). A story is written token by token from an *on-task* context: the instruction, the opening and the text so far. When the model writes a key of a *context dictionary*, a *wandering episode* begins: for the next L tokens it predicts from a second context in which the key's passage (another moment of the same story world, introduced by a discourse connective) is spliced between the opening and the text so far. Decoding then returns to the on-task context and the second context is discarded; a *cooling* period passes before the next episode, and a *context mixture rate* ρ sets how many remembered sentences an episode carries. The story is

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

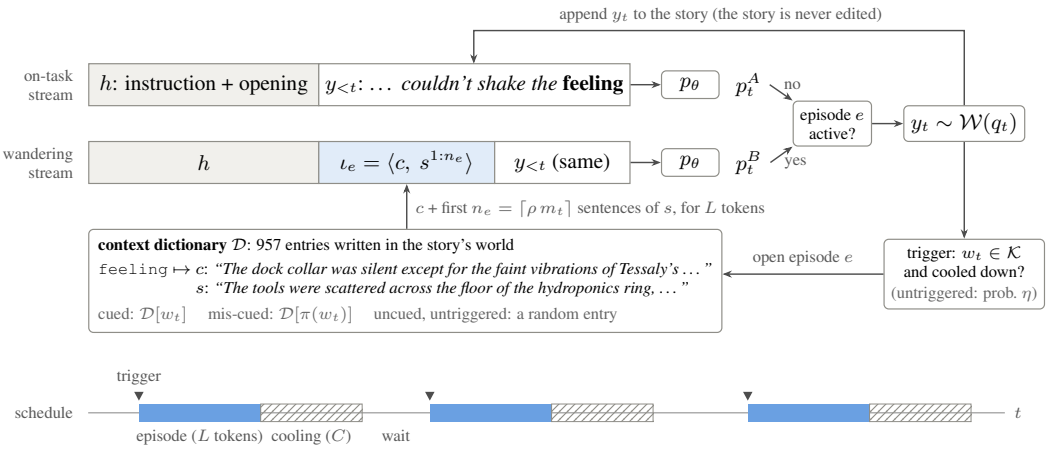


Figure 1: **Mind wandering as a decoding rule:** The on-task stream (top) predicts from the header h and the story so far $y_{<t}$. When the word just written is a dictionary key (here `feeling`, from a real story) and the story is off cooldown, an episode opens: for L tokens, tokens are drawn from a wandering stream (middle) whose context splices the key's connective c and the first n_e sentences of its passage s between h and $y_{<t}$. The wandering context is built afresh for each episode and discarded when it ends; the passage never enters the story. Bottom: episodes, cooling periods and waits for the next trigger.

never edited: the passage reaches it only through what the model chooses to write. The rule has a cue-triggered form, a spontaneous form in which episodes start at random (Christoff et al., 2016; Seli et al., 2016), and a closed-loop form in which a learned controller decides when and how deeply to wander.

To separate the *interruption* from its *content*, we build a single story world (a derelict relay station, a salvage crew of three and a failing caretaker system) with eight openings and a 957-entry dictionary written inside it, and compare wandering with decoding and prompting baselines, with the same material placed in the prompt, and with uncued, mis-cued and untriggered wandering, at 1.5B and 7B.

Contributions:

- **A decoding rule:** We formalize mind wandering as two-stream decoding with a spliced intrusion and a renewal-process schedule, in cued, uncued, mis-cued and spontaneous variants, and cast *when* and *how deeply* to wander as a constrained Markov decision process solved with leave-one-out REINFORCE and dual ascent (Section 3).
- **A controlled testbed:** A single-world benchmark with a 957-entry context dictionary; 1,616 evaluated stories in 16 arms at 1.5B (one of them a 66-configuration grid) and 448 in 7 arms at 7B, plus 384 training rollouts; an LLM rubric; and a tournament judged in both presentation orders (Section 4).[1]
- **Evidence:** For Qwen2.5-1.5B, wandering improves creativity and coherence together, in every opening and beyond any decoding configuration, and tracks broken repetition loops (Sections 5.1 and 5.2); the gain does not depend on the cue (Section 5.3); a learned controller matches the best schedules but mostly learns how much to wander (Section 5.4); at 7B the gains disappear (Section 5.5).

2 Related Work

We summarize four literatures here; Section A reviews each in detail.

[1]All code and data (generation, controller training, judging and analysis, with an offline test suite; the context dictionary; every story and every judgment) will be released as open source upon acceptance.

Mind wandering: Mind wandering is task-unrelated, often stimulus-independent thought (Singer, 1966; Smallwood & Schooler, 2006; 2015), engaging the default network (Raichle et al., 2001; Mason et al., 2007; Christoff et al., 2009; Andrews-Hanna, 2012; Fox et al., 2015). Christoff et al. (2016) treat it as *spontaneous* thought, weakly constrained by deliberate or automatic control, and Seli et al. (2016) separate intentional from unintentional episodes. Most episodes follow an environmental or internal trigger (Plimpton et al., 2015), as involuntary autobiographical memories do (Berntsen, 2010), and spread along associative links (Collins & Loftus, 1975). Its costs for reading and sustained attention (Smallwood et al., 2008; McVay & Kane, 2010; Mooneyham & Schooler, 2013) are one reason to bound episodes.

Mind wandering and creative writing: Undemanding tasks that invite mind wandering improve later creative performance (Baird et al., 2012), consistent with incubation effects (Sio & Ormerod, 2009; Ritter & Dijksterhuis, 2014) and with release from fixation on an initial approach (Smith & Blankenship, 1991). Writers and physicists report a fifth of their significant ideas during mind wandering (Gable et al., 2019); spontaneous and controlled thought make distinct contributions to creativity (Smeekens & Kane, 2016; Zedelius & Schooler, 2015; Beaty et al., 2016; Fox & Beaty, 2019), and associative theories locate originality in remote associations (Mednick, 1962). Our cued and spontaneous variants correspond to cue-dependent and unintentional mind wandering.

Creative writing with LLMs: Story generation has moved from hierarchical models to planning, revision and co-writing tools (Fan et al., 2018; Yang et al., 2022; 2023; Yuan et al., 2022; Mirowski et al., 2023). Evaluations find LLM fiction fluent but formulaic (Chakrabarty et al., 2024; Gómez-Rodríguez & Williams, 2023; Tian et al., 2024; Lu et al., 2025; Franceschelli & Musolesi, 2024), and alignment reduces output diversity (Kirk et al., 2024). Creativity is usually steered by sampling (Holtzman et al., 2020; Meister et al., 2023; Peeperkorn et al., 2024; Nguyen et al., 2025) or by prompting, e.g. with a random concept (Mehrotra et al., 2024); repetition is a known failure of likelihood decoding (Holtzman et al., 2020; Welleck et al., 2020; Xu et al., 2022; Keskar et al., 2019). Rubric and pairwise LLM judges (Chiang & Lee, 2023; Liu et al., 2023; Zheng et al., 2023; Fein et al., 2026) carry position bias (Zheng et al., 2023; Wang et al., 2024) and favor text of their own model family (Panickssery et al., 2024); human criteria for stories (Chhun et al., 2022) and consensual assessment by expert judges (Amabile, 1982) remain the reference.

Context injection in LLMs: Adversarial *prompt injection* inserts instructions into a model's input (Perez & Ribeiro, 2022; Greshake et al., 2023; Liu et al., 2024b); non-adversarial irrelevant context degrades reasoning (Jia & Liang, 2017; Shi et al., 2023; Yoran et al., 2024), and position within a long context matters (Liu et al., 2024a), although random documents can also help retrieval-augmented generation (Cuconasu et al., 2024). Decoding-time methods combine distributions computed under different contexts or models (Khandelwal et al., 2020; Shi et al., 2024; Sanchez et al., 2024; Li et al., 2023; Liu et al., 2021; Krause et al., 2021; Dathathri et al., 2020); retrieval can be triggered mid-generation (Jiang et al., 2023; Asai et al., 2024), and agents retrieve memories by relevance and recency (Park et al., 2023). Our rule differs in purpose and scope: the injected material is meant to *change* what is written, not to supply evidence, and it lives only in a transient second context.

3 Mind Wandering as a Decoding Rule

Setup: Let $p_\theta(\cdot \mid \mathbf{x})$ be a language model over vocabulary $\mathcal{V}$. A writing task gives a header h (system prompt, instruction and story opening, with the opening repeated as the start of the response), and the model writes a body $y = (y_1, \ldots, y_T)$ of S sentences, followed by a separate closure pass that writes at most ten more (Section C). Vanilla decoding samples $y_t \sim \mathcal{W}\big(p_\theta(\cdot \mid h, y_{<t})\big)$, where $\mathcal{W}$ applies temperature and nucleus truncation. We write m_t for the number of sentences completed in $y_{<t}$ and w_t for the word completed at step t, if any.

Context dictionary: A dictionary $\mathcal{D} = \{(k_i, c_i, s_i)\}_{i=1}^{M}$ maps a key word k_i to a passage $s_i = (s_i^1, \ldots, s_i^N)$ of N sentences and a discourse connective c_i, a one- or two-sentence hand-over such as *"The same sensation had hit her once before, back in the relay's comms spine."* Keys are words that are single tokens for the model, and the dictionary is also the set of possible triggers $\mathcal{K} = \{k_i\}$.

Two streams: An episode e starts at step t_e, lasts L_e tokens and carries passage j_e with n_e sentences. Its *intrusion* is $\iota_e = \langle c_{j_e}, s_{j_e}^{1:n_e} \rangle$. Decoding draws from

$$y_t \sim \mathcal{W}(q_t), \qquad q_t = \begin{cases} p_\theta\big(\cdot \mid h,\ \iota_e,\ y_{<t}\big) & \text{if } t_e \le t < t_e + L_e \text{ for some episode } e, \\ p_\theta\big(\cdot \mid h,\ y_{<t}\big) & \text{otherwise.} \end{cases} \tag{1}$$

Both streams condition on the same story prefix $y_{<t}$; the intrusion lives only in the second stream and is discarded when the episode ends. Placing ι_e *between* the opening and the story keeps the text written so far adjacent to the generation point, so the continuation stays locally fluent while the passage acts as remembered background; placed last, it would be continued instead of the story. Equation (1) is the $\lambda_t \in \{0, 1\}$ case of a log-linear pool $\log q_t = (1 - \lambda_t) \log p_t^A + \lambda_t \log p_t^B + \text{const}$ (Hinton, 2002); we mix contexts rather than distributions, and set *how much* is mixed by the number of passage sentences,

$$n_e = \min\big(N,\ \max(1,\ \lceil \rho\, m_{t_e} \rceil)\big), \tag{2}$$

so the context mixture rate $\rho \in (0, 1]$ is the ratio of remembered to written sentences in the mixed context: an early episode carries little, a late one can carry a whole scene. The wandering stream's key–value cache is advanced incrementally with the story, so wandering costs one prefill per episode plus roughly $1 + \omega$ forward passes per token, where ω is the share of tokens spent wandering (Algorithm 1).

When to wander: A story is a renewal process of episodes (Figure 1, bottom). After an episode ends, the next one can start only once $D = \max(C, 8)$ further tokens have been written on task; we vary the *cooling* C, and the floor of eight tokens, which also holds at the start of the story, makes every episode begin after some on-task text. We study four start-and-content rules. **Cued** wandering (the method) starts when $w_t \in \mathcal{K}$ and carries the passage keyed by w_t. **Uncued** wandering starts on the same trigger words but carries a passage drawn at random from $\mathcal{D}$, afresh at every episode. **Mis-cued** wandering also starts on the trigger words, but the passages have been reassigned to the keys once, by a fixed random permutation π of $\mathcal{K}$, so a word always brings up the same passage, one written for another key: $\mathcal{D}[\pi(w_t)]$. Each passage stays an intact scene of the story world; mis-cueing breaks only the link between a word and what its passage is about, and keeps what uncued wandering loses, a fixed memory for every word. **Untriggered** (spontaneous) wandering starts with probability η at every eligible token, with no cue, and carries a random passage; η is matched to uncued wandering by maximum likelihood on its waiting times, $\hat{\eta} = E/(E + \Omega) = 0.151$ for E episode starts and Ω eligible tokens that passed without one. An episode, the $D - 1$ tokens after it and a geometric wait with mean $1/\eta$ form one cycle of the renewal process, so a share $\omega = L/(L + D - 1 + 1/\eta)$ of the story is decoded from the wandering stream: with episodes of $L = 30$ tokens, the grid's cooling range $C \in [0, 50]$ moves ω from 0.69 to 0.35, a factor of two in the dose of wandering.

Learning when to wander: A fixed schedule cannot look at the story. A controller instead acts at every *decision point* (a trigger word at an eligible step), given nine features $\phi_t \in \mathbb{R}^9$ that the decoder already computes, among them progress, tokens since the last episode, and recent entropy, surprisal and repetition (Section G.3). Its actions are $\mathcal{A} = \{\textsc{skip},\ (15, 0.3),\ (30, 0.7),\ (60, 1.0)\}$, pairs of episode length and mixture rate (light, standard, deep). The policy is linear-softmax, $\pi_W(a \mid \phi) \propto \exp(w_a^\top \phi)$ with $w_{\textsc{skip}} = 0$, so each weight reads as "this signal makes this depth of wandering more likely". Writing $Q(y)$ for creativity plus coherence in one rating by the rubric judge (Section 4) and $B(y)$ for the number of wandering tokens, we solve the constrained problem (Altman, 1999; Tessler et al., 2019)

$$\max_W\ \mathbb{E}_{\pi_W}\big[Q(y)\big] \quad \text{s.t.} \quad \mathbb{E}_{\pi_W}\big[B(y)\big] \le B_0, \qquad R(y) = Q(y) - \mu\, B(y)/B_0, \tag{3}$$

where B_0 is what the fixed reference schedule spends, so a controller cannot win by wandering more. Each iteration writes G stories per opening; with the leave-one-out baseline within an opening (Williams, 1992; Kool et al., 2019; Ahmadian et al., 2024),

$$\hat{\nabla}_W = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \frac{A_i}{\bar{d}} \sum_{t \in \mathcal{T}_i} \big(\boldsymbol{e}_{a_t} - \pi_W(\cdot \mid \phi_t)\big)_{\setminus \textsc{skip}}\, \phi_t^\top, \qquad A_i = R_i - \frac{1}{G - 1} \sum_{j \ne i} R_j, \tag{4}$$

where $\mathcal{T}_i$ are story i's decision points, $\bar{d}$ their mean number per story and the bracket is the closed-form score of the linear-softmax policy. Adam (Kingma & Ba, 2015) ascends $\hat{\nabla}_W$ on standardized

Algorithm 1 Decoding with mind wandering (one story; batched in practice). For the untriggered rule, σ fires with probability η at every step; for a controller, it fires at a trigger word when the policy chooses a non-skip action, which also sets L and ρ for that episode.

Require: model p_θ, header h, dictionary $\mathcal{D}$, cooling C, rate ρ, episode length L, start rule $\sigma \in \{$cued, uncued, mis-cued, untriggered, controller$\}$

$y \leftarrow ()$, active $\leftarrow$ false, cool $\leftarrow 0$, since $\leftarrow 0$
for $t = 1, 2, \ldots$ until S sentences are complete **do**
  $q \leftarrow p_\theta(\cdot \mid h, \iota, y)$ **if** active **else** $p_\theta(\cdot \mid h, y)$ ▷ Equation (1)
  sample $y_t \sim \mathcal{W}(q)$; append y_t to y; since $\leftarrow$ since $+ 1$
  if active **then** $\ell \leftarrow \ell + 1$; **if** $\ell = L$ **then** active $\leftarrow$ false, cool $\leftarrow C$, since $\leftarrow 0$
  else if cool > 0 **then** cool $\leftarrow$ cool -1
  end if
  if not active **and** cool $= 0$ **and** since ≥ 8 **and** σ fires at t **then** ▷ effective gap $\max(C, 8)$
    choose passage j by σ; $n \leftarrow \min(N, \max(1, \lceil \rho\, m_t \rceil))$ ▷ Equation (2)
    $\iota \leftarrow \langle c_j, s_j^{1:n} \rangle$ (a fresh wandering context); active $\leftarrow$ true; $\ell \leftarrow 0$; since $\leftarrow 0$
  end if
end for
return y followed by the closure pass (no wandering)

advantages and the multiplier follows $\mu \leftarrow \max\big(0,\ \mu + \eta_\mu(\bar{B}/B_0 - 1)\big)$. As a training-free baseline, a *boredom rule* wanders at standard depth whenever recent entropy is in the lowest fifth, or recent repetition in the highest fifth, of what vanilla writing produces: the moments at which a writer's attention would drift.

4 Experimental Setup

Story world and dictionary: All openings and passages belong to one serial, *Kepler Drift*: a decommissioned relay station falling toward a gas giant, a salvage crew of three and a degraded caretaker system. A one-page series bible fixes place, cast, tone and rules, and eight openings of five or six sentences, written by one of the authors,[2] present eight moments of the same job (Section B). Keys are single-token words of at least four letters, chosen by how often Qwen2.5-1.5B used them in calibration stories. For each key, GPT-5.6 Terra (OpenAI, 2026) wrote a 25-sentence passage set in the same world and a connective; passages were checked automatically and repaired up to twice, and the 957 keys whose passages passed form the dictionary (Section D). A pilot with unrelated openings and passages, which motivated this single-world design, is reported in Section E.

Models and decoding: Qwen2.5-1.5B-Instruct and Qwen2.5-7B-Instruct (Qwen Team, 2025) share a tokenizer, so the dictionary serves both. Mind wandering is spontaneous rather than deliberate thought (Christoff et al., 2016), so we study direct generation without a thinking phase, which makes the decoding rule the only intervention. A thinking phase would be a second, uncontrolled way to break loops, and it would be unclear where an episode should be spliced in. Every arm except the decoding baselines samples at $T = 0.9$, top-$p = 0.95$. Bodies have $S = 25$ sentences and are followed by a closure of at most 10 sentences written without wandering; episodes last $L = 30$ tokens and passages have $N = 25$ sentences (Section F).

Arms: The *core* study (1.5B, 1,264 stories) compares vanilla sampling with the full grid $C \in \{0, 5, \ldots, 50\} \times \rho \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.0\}$ (66 configurations), with cued, uncued and mis-cued wandering at the reference point $C = 25$, $\rho = 0.7$, with five decoding configurations and with two prompting baselines, an instruction to be maximally original and a random dictionary word to weave in (Mehrotra et al., 2024). The *extension* (1.5B, 352 stories plus 384 training rollouts) adds untriggered wandering, a ten-sentence passage placed once in the prompt, the boredom rule and the learned controller, and the *scale* study (7B, 448 stories) repeats seven of these arms (Table 5 lists every arm and its size).

[2]The story openings were written by one of the authors, an amateur science-fiction writer.

Table 1: **Core study (Qwen2.5-1.5B):** Opening-balanced means (1–5) with the half-width of the 95% confidence interval in small type, 4-gram repetition and on-task surprisal (nats/token). Every wandering arm has a higher mean creativity and coherence than vanilla and every decoding and prompting baseline.

	Arm	n	Creativity	Coherence	Overall	Rep-4 ↓	Surprisal
	Vanilla (T=0.9, top-p 0.95)	160	2.47 ±.08	3.29 ±.09	3.09	.091	1.55
decoding	T=0.7	32	2.06 ±.16	3.03 ±.19	2.75	.256	0.88
decoding	T=0.9, min-p 0.05	32	2.08 ±.18	3.03 ±.18	2.63	.274	0.94
decoding	T=1.1	32	2.56 ±.20	2.06 ±.25	2.28	.005	4.30
decoding	T=1.3	32	1.03 ±.06	1.00 ±.00	1.03	.000	9.70
decoding	T=1.5, min-p 0.1, top-p 1	32	2.59 ±.20	3.41 ±.23	3.22	.025	1.89
prompt	"Be maximally original"	64	2.58 ±.13	3.34 ±.16	3.10	.106	1.53
prompt	Random word to weave in	64	2.51 ±.13	3.18 ±.16	3.00	.090	1.64
wandering	Grid, all 66 configurations	528	**2.73** ±.04	**3.52** ±.04	3.31	.038	1.86
wandering	Cued (C=25, ρ=0.7)	104	**2.75** ±.08	**3.55** ±.09	3.27	.047	1.87
wandering	Uncued (random passage)	96	**2.78** ±.08	**3.60** ±.10	3.38	.053	1.82
wandering	Mis-cued (another key's passage)	96	**2.76** ±.08	**3.57** ±.10	3.36	.036	1.84

Evaluation: Claude Sonnet 5 (Anthropic, 2026) rates each story blind to condition on a five-point rubric with anchored levels for creativity, coherence, language, engagement and overall quality, plus flags for strangeness from lost control and for abandoning the opening's premise (Section C.1); it rates each story twice at temperature 0.3, and we average the two ratings (Section G.5). Contrasts are stratified by opening, $\Delta = \frac{1}{P}\sum_{p=1}^{P}(\bar{y}_{a,p} - \bar{y}_{b,p})$ over $P = 8$ openings, with Neyman standard errors; the text gives 95% confidence intervals in brackets, and tables give their half-widths after ±. Unless stated otherwise, p-values come from stratified permutation tests (2,000 permutations), Holm-corrected within families of related contrasts (Holm, 1979). We also measure 4-gram repetition, adjacent-sentence embedding similarity (Reimers & Gurevych, 2019; Wang et al., 2020) and on-task surprisal (mean negative log-likelihood under the vanilla context). A second instrument, a round-robin tournament judged in both presentation orders and scored with a Bradley–Terry model with a position term (Bradley & Terry, 1952), is reported in Section G.4.

5 Results

5.1 Wandering makes a small model more creative and more coherent

Across the 66 configurations, wandering raises creativity by +0.26 [0.17, 0.35], coherence by +0.23 [0.13, 0.33] and overall quality by +0.21 [0.12, 0.31] over vanilla (Table 1; $p = 0.003$ for each; Cohen's $d = 0.55$ for creativity), in all eight openings. The prompting baselines do not help on the rubric (an instruction to be original: +0.11 [−0.04, 0.26] creativity; a random word: +0.04 [−0.12, 0.19]), and wandering beats them by 0.15–0.22 creativity and 0.19–0.34 coherence points; the tournament disagrees on creativity (Section G.4). Wandering stories are also less often flagged as strange through lost control (4% vs. 12%) or as abandoning the opening (16% vs. 27%).

5.2 Wandering is not a temperature increase

Temperature T divides the model's logits before the softmax: values below 1 sharpen the next-token distribution and values above 1 flatten it. Raised alone, it soon breaks a small model's text (at $T = 1.3$ every story scores 1 on coherence); min-p truncation, which keeps only tokens whose probability is at least a fixed fraction of the most likely token's, makes high temperatures usable (Nguyen et al., 2025). Randomizing between decoding configurations reaches the convex hull of their (coherence, creativity) points (Caccia et al., 2020), anchored at the top by the best decoder, $T = 1.5$ with min-p 0.1, at (3.41, 2.59). Every wandering arm in Table 1 has higher coherence (3.52–3.60) than any point of the hull and higher creativity (2.73–2.78) than any decoder, so no mixture of decoders matches it (Figure 2a; Monte Carlo probability of lying outside the hull 0.94–0.98). At equal on-task surprisal, wandering is +0.25 [0.16, 0.33] creativity points above the curve traced by plain temperature sampling,

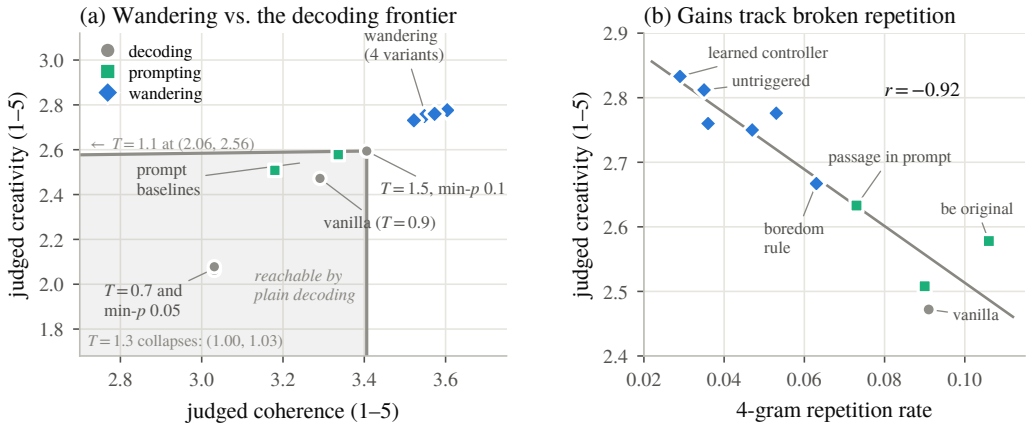


Figure 2: **(a)** Creativity against coherence for decoding configurations (gray), prompting baselines (squares) and wandering (diamonds). The shaded region is what randomizing over decoding configurations can reach; every wandering variant lies outside it. **(b)** Across the ten 1.5B arms that sample at $T = 0.9$, arm-level creativity falls with the 4-gram repetition rate ($r = -0.92$).

Table 2: **Extension and scale studies:** Differences from the same model's vanilla, with the half-width of the 95% confidence interval in small type. 7B vanilla scores 2.94 creativity and 4.00 coherence; the judge gave every 7B story in these arms a coherence score of exactly 4, so 7B coherence differences are zero and are not shown.

	Qwen2.5-1.5B		Qwen2.5-7B
Arm	Δ creativity	Δ coherence	Δ creativity
Cued wandering	+0.28 ±.11	+0.26 ±.13	+0.03 ±.06
Uncued wandering	+0.30 ±.11	+0.31 ±.13	+0.03 ±.05
Untriggered wandering	+0.34 ±.11	+0.31 ±.13	+0.03 ±.06
Passage in prompt	+0.16 ±.14	+0.21 ±.16	+0.04 ±.05
Prompt: be maximally original	+0.11 ±.15	+0.05 ±.18	+0.04 ±.05
Decoding T=1.5, min-p 0.1	+0.12 ±.20	+0.12 ±.24	+0.06 ±.04
Boredom rule (controller)	+0.19 ±.12	+0.21 ±.13	–
Learned controller (RLOO)	**+0.36** ±.11	+0.29 ±.13	–

and the prompting baselines are not (+0.03 and +0.12, intervals including zero). The decoder with the surprisal of wandering (1.89 against 1.86 nats per token), $T = 1.5$ with min-p 0.1, gains less creativity (+0.12 against +0.26).

Repetition: What wandering changes instead is repetition. Small models at moderate temperature drift into self-reinforcing loops (Holtzman et al., 2020; Xu et al., 2022): 9.1% of 4-grams repeat in vanilla stories, and more than a quarter at $T = 0.7$ or with min-p 0.05 (Table 1). Every arm that wanders during decoding cuts repetition to 2.9–6.3%, while the prompting and passage-in-prompt arms stay at 7.3–10.6% (Table 8). Across the ten 1.5B arms that sample at $T = 0.9$, creativity falls with repetition ($r = -0.92$; Figure 2b), by about 0.04 rubric points per percentage point of repeated 4-grams, and so does coherence ($r = -0.86$). High temperature also removes repetition ($T = 1.1$: 0.5%) but ruins coherence (2.06). An episode is a local change of context that the next tokens cannot copy from the story's recent past: it breaks a loop without raising the entropy of every step, and adjacent sentences become less alike (embedding similarity 0.337 to 0.305) while judged coherence rises.

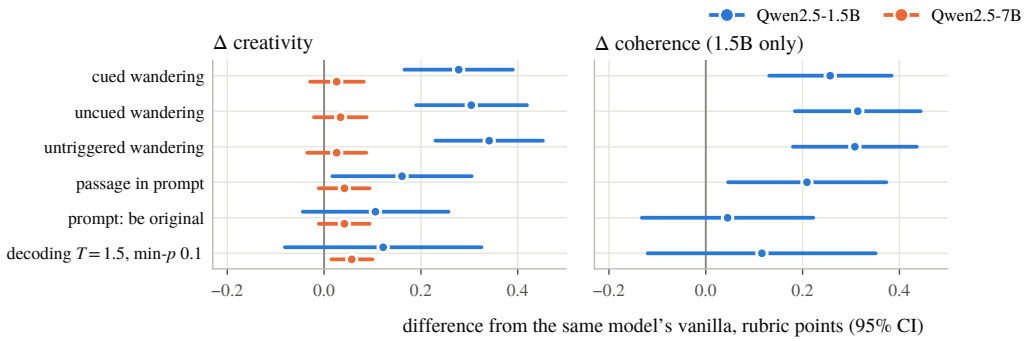


Figure 3: **Effects relative to the same model's vanilla** (95% CIs): Wandering helps the 1.5B model and does nothing measurable for the 7B model's creativity. Coherence is shown for 1.5B only: the judge gave every 7B story in these arms a coherence score of exactly 4 (Table 2).

5.3 The cue does not matter; the intrusion does

If wandering helped by retrieving what a word *means* to the story, the cued passage should beat a random one. It does not: at the reference schedule, cued minus uncued wandering is -0.03 [-0.14, 0.09] creativity points and cued minus mis-cued -0.01 [-0.13, 0.10]. Spontaneous episodes, with no trigger at all, give $+0.34$ [0.23, 0.45] creativity and $+0.31$ [0.18, 0.44] coherence over vanilla, as much as triggered uncued wandering (-0.04 [-0.15, 0.08] for triggered minus spontaneous; Table 2). The schedule is equally inert: with opening fixed effects, neither cooling, the mixture rate nor their interaction explains creativity (F-tests, $p = 0.27$, 0.96 and 0.83), every level of each scores above vanilla (Table 7), and cell means do not replicate across halves of the openings (split-half $r = -0.34$). Placed once in the prompt instead, ten sentences of a random passage raise creativity by $+0.16$ [0.02, 0.30], and cued wandering exceeds them by $+0.12$ [-0.03, 0.26], a difference our sample does not resolve. What helps is the *episode*, a bounded excursion of attention, not the association that cued it, as in incubation accounts where stepping away releases fixation (Smith & Blankenship, 1991; Sio & Ormerod, 2009).

5.4 Learning when to wander

Six RLOO iterations of 64 rollouts ($G = 8$ per opening, $B_0 = 250$ wandering tokens per story) yield the highest rubric creativity of any 1.5B arm, $+0.36$ [0.25, 0.47], with $+0.29$ [0.16, 0.43] coherence over vanilla, but no reliable gain over the fixed reference schedule ($+0.08$ [-0.02, 0.19]) or untriggered wandering ($+0.02$ [-0.09, 0.13]); it beats the boredom rule by $+0.17$ [0.05, 0.28] ($p = 0.04$). At its own spending it sits $+0.12$ creativity and $+0.12$ coherence points above a quadratic fit of score on wandering tokens over the fixed-schedule stories (bootstrap intervals 0.05 to 0.21 and 0.02 to 0.25). It mostly learned *how much* to wander (Section G.3): the judged score of its rollouts did not rise while the probability of wandering at a decision point rose from 0.26 to 0.57, and the decision to wander correlates weakly with every feature ($|r| < 0.15$). The boredom rule fires when it should, at low entropy and high repetition, but at only 9% of decision points, and at matched spending does no better than the fixed schedule (-0.10; bootstrap interval -0.21 to 0.05).

5.5 Scale: the gains vanish at 7B

Qwen2.5-7B writes far better stories than Qwen2.5-1.5B: 2.94 vs. 2.47 creativity, 4.00 vs. 3.29 coherence and 3.81 vs. 3.09 overall. Against its own vanilla, no intervention moves its creativity by more than 0.06 points (Figure 3), and the two models' wandering effects differ by 0.25 to 0.31 creativity points ($p < 0.001$ for each; normal-approximation tests, Holm-corrected over the six arms run at both sizes). Coherence cannot be compared on the rubric at 7B, where every story in these arms scored exactly 4. Two readings fit. The rubric may be too coarse at this level: the tournament, which compares stories directly, puts cued and uncued wandering above 7B vanilla on coherence and overall quality, by margins similar to the uncertainty of a single score, though not untriggered

wandering (Section G.4). Or, if wandering helps by breaking loops (Section 5.2), a model that loops less has less to gain: 7B vanilla repeats 0.5% of its 4-grams, against 9.1% at 1.5B. The 1.5B model with the learned controller closes 77% of the creativity gap to 7B vanilla, 41% of the coherence gap and 37% of the overall gap.

6 Discussion and Limitations

What mind wandering buys a language model: The benefit does not depend on what cued an episode or when it started (Sections 5.3 and 5.4); it tracks the breaking of repetition loops (Section 5.2), and at 7B it vanishes (Section 5.5). This fits the functional account of mind wandering in creative work, in which episodes help less by delivering a specific idea than by releasing the writer from a fixed track (Smith & Blankenship, 1991; Baird et al., 2012; Gable et al., 2019). For a small model the fixed track is literal: a repetition loop.

Limitations: All stories come from one world, one genre and one model family (reasoning models are left to future work) and are 25 sentences long, and every score comes from one LLM judge (Claude Sonnet 5, two samples per story), without human ratings; LLM judges agree only moderately with human preferences about stories (Fein et al., 2026). The judge belongs to the model family that wrote the dictionary, and judges tend to favor their own family's text (Panickssery et al., 2024); the stories are written by Qwen, but they echo the passages. The 7B coherence scores sit at the judge's ceiling of 4, so the absence of an effect at 7B is not evidence that there is none. Wandering stories are about 10% shorter in words at equal sentence count (514 vs. 574), and scores fall with length (Spearman -0.30), so part of the gain may reflect length. The controller trained for six iterations without raising the judged score of its rollouts, saw nine features, and overspent its budget (by 26% in the last iteration and 34% at evaluation).

7 Conclusion

We introduced mind-wandering decoding: a word the model writes can open a bounded episode in which the next tokens are predicted from a transient second context that splices a remembered passage between the story's opening and the text so far, while the story itself is never edited. In a controlled single-world testbed, wandering made a 1.5B model's stories more creative and more coherent at the same time, in every opening and beyond anything temperature or min-p sampling reached. The gain did not depend on the cue or the schedule, tracked the breaking of repetition loops, and vanished at 7B, a model that loops far less and whose coherence scores already sit at the judge's ceiling; a learned controller matched the best fixed schedules but learned mostly how much to wander, not when. The picture is less that of a model borrowing ideas from a remembered passage than that of an interruption that frees it from a fixed track, the role incubation accounts give mind wandering in human creative work. The rule needs no training, costs one prefill per episode and one extra forward pass per wandering token, and leaves the text it helps write untouched.

Three extensions follow. The dictionary is fixed before generation; its passages could instead be written on the fly from the current story, a memory that grows with the story as autobiographical memory does (Berntsen, 2010), letting a controller choose *what* to wander to. Each episode starts from a clean second context; keeping the residue of earlier episodes, which the decoder supports by retaining the wandering stream's cache, would model a train of thought that returns to the same memories. And repeating the study across genres (mystery, romance, horror, literary fiction), with an instrument finer than a rubric at its ceiling, would show whether the benefit reaches larger models and whether wandering toward distant material fuels originality, as associative accounts propose (Mednick, 1962).

AI use statement

In this work, we used generative AI tools as part of the method and its evaluation: GPT-5.6 Terra wrote the passages and connectives of the context dictionary (Section D), and Claude Sonnet 5 is the judge for the rubric and the tournament and supplied the controller's training reward (Sections C.1 and G.3); in the pilot, GPT-5.6 Terra wrote the passages and Claude Sonnet 5 was the judge (Section E).

Ethics statement

The study generates short fiction with openly available models and scores it with an LLM judge; no personal data are used and no human subjects were involved. The method is a decoding rule for creative writing and does not introduce new capabilities for misuse; the injection experiments use non-adversarial narrative text.

Reproducibility statement

Sections 3 and 4 define the decoding rule, schedule, controller and estimators; Sections B, C, C.1, D and F give the story world, every prompt, the construction of the dictionary, the rubric and all hyperparameters, and Algorithm 1 gives the decoding loop. The code (generation, controller training, judging and analysis, with an offline test suite and the notebooks that ran every stage), the context dictionary, all generated stories and all judgments will be released as open source upon acceptance.

References

Arash Ahmadian, Chris Cremer, Matthias Gallé, Marzieh Fadaee, Julia Kreutzer, Olivier Pietquin, Ahmet Üstün, and Sara Hooker. Back to basics: Revisiting REINFORCE-style optimization for learning from human feedback in LLMs. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pp. 12248–12267, 2024.

Eitan Altman. *Constrained Markov Decision Processes*. Chapman & Hall/CRC, 1999.

Teresa M. Amabile. Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5):997–1013, 1982.

Barrett R. Anderson, Jash Hemant Shah, and Max Kreminski. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*, pp. 413–425, 2024.

Jessica R. Andrews-Hanna. The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, 18(3):251–270, 2012.

Anthropic. Introducing Claude Sonnet 5. https://www.anthropic.com/news/claude-sonnet-5, 2026.

Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. Self-RAG: Learning to retrieve, generate, and critique through self-reflection. In *International Conference on Learning Representations*, 2024.

Benjamin Baird, Jonathan Smallwood, Michael D. Mrazek, Julia W. Y. Kam, Michael S. Franklin, and Jonathan W. Schooler. Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science*, 23(10):1117–1122, 2012.

Roger E. Beaty, Mathias Benedek, Paul J. Silvia, and Daniel L. Schacter. Creative cognition and brain network dynamics. *Trends in Cognitive Sciences*, 20(2):87–95, 2016.

Dorthe Berntsen. The unbidden past: Involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, 19(3):138–142, 2010.

Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Massimo Caccia, Lucas Caccia, William Fedus, Hugo Larochelle, Joelle Pineau, and Laurent Charlin. Language GANs falling short. In *International Conference on Learning Representations*, 2020.

Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, and Chien-Sheng Wu. Art or artifice? Large language models and the false promise of creativity. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024.

Cyril Chhun, Pierre Colombo, Fabian M. Suchanek, and Chloé Clavel. Of human criteria and automatic metrics: A benchmark of the evaluation of story generation. In *Proceedings of the 29th International Conference on Computational Linguistics*, pp. 5794–5836, 2022.

Cheng-Han Chiang and Hung-yi Lee. Can large language models be an alternative to human evaluations? In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pp. 15607–15631, 2023.

Kalina Christoff, Alan M. Gordon, Jonathan Smallwood, Rachelle Smith, and Jonathan W. Schooler. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences*, 106(21):8719–8724, 2009.

Kalina Christoff, Zachary C. Irving, Kieran C. R. Fox, R. Nathan Spreng, and Jessica R. Andrews-Hanna. Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, 17(11):718–731, 2016.

Allan M. Collins and Elizabeth F. Loftus. A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407–428, 1975.

Florin Cuconasu, Giovanni Trappolini, Federico Siciliano, Simone Filice, Cesare Campagnano, Yoelle Maarek, Nicola Tonellotto, and Fabrizio Silvestri. The power of noise: Redefining retrieval for RAG systems. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 719–729, 2024.

Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. Plug and play language models: A simple approach to controlled text generation. In *International Conference on Learning Representations*, 2020.

Ap Dijksterhuis and Teun Meurs. Where creativity resides: The generative power of unconscious thought. *Consciousness and Cognition*, 15(1):135–146, 2006.

Anil R. Doshi and Oliver P. Hauser. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290, 2024.

Angela Fan, Mike Lewis, and Yann Dauphin. Hierarchical neural story generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pp. 889–898, 2018.

Daniel Fein, Sebastian Russo, Violet Xiang, Kabir Jolly, Rafael Rafailov, and Nick Haber. LitBench: A benchmark and dataset for reliable evaluation of creative writing. In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 7740–7755, 2026.

Kieran C. R. Fox and Roger E. Beaty. Mind-wandering as creative thinking: Neural, psychological, and theoretical considerations. *Current Opinion in Behavioral Sciences*, 27:123–130, 2019.

Kieran C. R. Fox, R. Nathan Spreng, Melissa Ellamil, Jessica R. Andrews-Hanna, and Kalina Christoff. The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *NeuroImage*, 111:611–621, 2015.

Giorgio Franceschelli and Mirco Musolesi. On the creativity of large language models. *AI & Society*, 2024.

Shelly L. Gable, Elizabeth A. Hopper, and Jonathan W. Schooler. When the muses strike: Creative ideas of physicists and writers routinely occur during mind wandering. *Psychological Science*, 30 (3):396–404, 2019.

Carlos Gómez-Rodríguez and Paul Williams. A confederacy of models: A comprehensive evaluation of LLMs on creative writing. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 14504–14528, 2023.

Kai Greshake, Sahar Abdelnabi, Shailesh Mishra, Christoph Endres, Thorsten Holz, and Mario Fritz. Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection. In *Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security*, pp. 79–90, 2023.

Geoffrey E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 14(8):1771–1800, 2002.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. In *International Conference on Learning Representations*, 2020.

Robin Jia and Percy Liang. Adversarial examples for evaluating reading comprehension systems. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2021–2031, 2017.

Zhengbao Jiang, Frank F. Xu, Luyu Gao, Zhiqing Sun, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, and Graham Neubig. Active retrieval augmented generation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 7969–7992, 2023.

Dan R. Johnson, James C. Kaufman, Brendan S. Baker, John D. Patterson, Baptiste Barbot, Adam E. Green, Janet van Hell, Evan Kennedy, Grace F. Sullivan, Christa L. Taylor, Thomas Ward, and Roger E. Beaty. Divergent semantic integration (DSI): Extracting creativity from narratives with distributional semantic modeling. *Behavior Research Methods*, 55(7):3726–3759, 2023.

Nitish Shirish Keskar, Bryan McCann, Lav R. Varshney, Caiming Xiong, and Richard Socher. CTRL: A conditional transformer language model for controllable generation. *arXiv preprint arXiv:1909.05858*, 2019.

Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. Generalization through memorization: Nearest neighbor language models. In *International Conference on Learning Representations*, 2020.

Matthew A. Killingsworth and Daniel T. Gilbert. A wandering mind is an unhappy mind. *Science*, 330(6006):932, 2010.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *International Conference on Learning Representations*, 2015.

Robert Kirk, Ishita Mediratta, Christoforos Nalmpantis, Jelena Luketina, Eric Hambro, Edward Grefenstette, and Roberta Raileanu. Understanding the effects of RLHF on LLM generalisation and diversity. In *International Conference on Learning Representations*, 2024.

Wouter Kool, Herke van Hoof, and Max Welling. Buy 4 REINFORCE samples, get a baseline for free! In *ICLR 2019 Workshop on Deep Reinforcement Learning Meets Structured Prediction*, 2019.

John Kounios and Mark Beeman. The cognitive neuroscience of insight. *Annual Review of Psychology*, 65:71–93, 2014.

Ben Krause, Akhilesh Deepak Gotmare, Bryan McCann, Nitish Shirish Keskar, Shafiq Joty, Richard Socher, and Nazneen Fatema Rajani. GeDi: Generative discriminator guided sequence generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 4929–4952, 2021.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*, 2020.

Xiang Lisa Li, Ari Holtzman, Daniel Fried, Percy Liang, Jason Eisner, Tatsunori Hashimoto, Luke Zettlemoyer, and Mike Lewis. Contrastive decoding: Open-ended text generation as optimization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pp. 12286–12312, 2023.

Alisa Liu, Maarten Sap, Ximing Lu, Swabha Swayamdipta, Chandra Bhagavatula, Noah A. Smith, and Yejin Choi. DExperts: Decoding-time controlled text generation with experts and anti-experts. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, pp. 6691–6706, 2021.

Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12:157–173, 2024a.

Yang Liu, Dan Iter, Yichong Xu, Shuohang Wang, Ruochen Xu, and Chenguang Zhu. G-Eval: NLG evaluation using GPT-4 with better human alignment. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 2511–2522, 2023.

Yupei Liu, Yuqi Jia, Runpeng Geng, Jinyuan Jia, and Neil Zhenqiang Gong. Formalizing and benchmarking prompt injection attacks and defenses. In *33rd USENIX Security Symposium*, 2024b.

Ximing Lu, Melanie Sclar, Skyler Hallinan, Niloofar Mireshghallah, Jiacheng Liu, Seungju Han, Allyson Ettinger, Liwei Jiang, Khyathi Chandu, Nouha Dziri, and Yejin Choi. AI as humanity's Salieri: Quantifying linguistic creativity of language models via systematic attribution of machine text against web text. In *International Conference on Learning Representations*, 2025.

Malia F. Mason, Michael I. Norton, John D. Van Horn, Daniel M. Wegner, Scott T. Grafton, and C. Neil Macrae. Wandering minds: The default network and stimulus-independent thought. *Science*, 315(5810):393–395, 2007.

Jennifer C. McVay and Michael J. Kane. Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, 136(2):188–197, 2010.

Sarnoff A. Mednick. The associative basis of the creative process. *Psychological Review*, 69(3): 220–232, 1962.

Pronita Mehrotra, Aishni Parab, and Sumit Gulwani. Enhancing creativity in large language models through associative thinking strategies. *arXiv preprint arXiv:2405.06715*, 2024.

Clara Meister, Tiago Pimentel, Gian Wiher, and Ryan Cotterell. Locally typical sampling. *Transactions of the Association for Computational Linguistics*, 11:102–121, 2023.

Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman, and Richard Evans. Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023.

Benjamin W. Mooneyham and Jonathan W. Schooler. The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, 67(1):11–18, 2013.

Minh Nhat Nguyen, Andrew Baker, Clement Neo, Allen Roush, Andreas Kirsch, and Ravid Shwartz-Ziv. Turning up the heat: Min-p sampling for creative and coherent LLM outputs. In *International Conference on Learning Representations*, 2025.

OpenAI. Advancing the price-performance frontier with GPT-5.6. https://openai.com/index/advancing-the-price-performance-frontier-with-gpt-5-6/, 2026.

Vishakh Padmakumar and He He. Does writing with language models reduce content diversity? In *International Conference on Learning Representations*, 2024.

Arjun Panickssery, Samuel R. Bowman, and Shi Feng. LLM evaluators recognize and favor their own generations. In *Advances in Neural Information Processing Systems*, 2024.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 2023.

Max Peeperkorn, Tom Kouwenhoven, Dan Brown, and Anna Jordanous. Is temperature the creativity parameter of large language models? In *Proceedings of the 15th International Conference on Computational Creativity*, 2024.

Fábio Perez and Ian Ribeiro. Ignore previous prompt: Attack techniques for language models. In *NeurIPS ML Safety Workshop*, 2022.

Ben Plimpton, Priya Patel, and Lia Kvavilashvili. Role of triggers and dysphoria in mind-wandering about past, present and future: A laboratory study. *Consciousness and Cognition*, 33:261–276, 2015.

Qwen Team. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*, 2025.

Marcus E. Raichle, Ann Mary MacLeod, Abraham Z. Snyder, William J. Powers, Debra A. Gusnard, and Gordon L. Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2):676–682, 2001.

Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pp. 3982–3992, 2019.

Simone M. Ritter and Ap Dijksterhuis. Creativity: the unconscious foundations of the incubation period. *Frontiers in Human Neuroscience*, 8:215, 2014.

Guillaume Sanchez, Alexander Spangher, Honglu Fan, Elad Levi, and Stella Biderman. Stay on topic with classifier-free guidance. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 43197–43234, 2024.

Paul Seli, Evan F. Risko, Daniel Smilek, and Daniel L. Schacter. Mind-wandering with and without intention. *Trends in Cognitive Sciences*, 20(8):605–617, 2016.

Freda Shi, Xinyun Chen, Kanishka Misra, Nathan Scales, David Dohan, Ed H. Chi, Nathanael Schärli, and Denny Zhou. Large language models can be easily distracted by irrelevant context. In *International Conference on Machine Learning*, pp. 31210–31227, 2023.

Weijia Shi, Xiaochuang Han, Mike Lewis, Yulia Tsvetkov, Luke Zettlemoyer, and Wen-tau Yih. Trusting your evidence: Hallucinate less with context-aware decoding. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (Short Papers)*, pp. 783–791, 2024.

Jerome L. Singer. *Daydreaming: An Introduction to the Experimental Study of Inner Experience*. Random House, 1966.

Ut Na Sio and Thomas C. Ormerod. Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, 135(1):94–120, 2009.

Jonathan Smallwood and Jonathan W. Schooler. The restless mind. *Psychological Bulletin*, 132(6): 946–958, 2006.

Jonathan Smallwood and Jonathan W. Schooler. The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66:487–518, 2015.

Jonathan Smallwood, Merrill McSpadden, and Jonathan W. Schooler. When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, 36(6):1144–1150, 2008.

Bridget A. Smeekens and Michael J. Kane. Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. *Psychology of Aesthetics, Creativity, and the Arts*, 10(4):389–415, 2016.

Steven M. Smith and Steven E. Blankenship. Incubation and the persistence of fixation in problem solving. *The American Journal of Psychology*, 104(1):61–87, 1991.

Claire Stevenson, Iris Smal, Matthijs Baas, Raoul Grasman, and Han van der Maas. Putting GPT-3's creativity to the (alternative uses) test. In *Proceedings of the 13th International Conference on Computational Creativity*, 2022.

Chen Tessler, Daniel J. Mankowitz, and Shie Mannor. Reward constrained policy optimization. In *International Conference on Learning Representations*, 2019.

Yufei Tian, Tenghao Huang, Miri Liu, Derek Jiang, Alexander Spangher, Muhao Chen, Jonathan May, and Nanyun Peng. Are large language models capable of generating human-level narratives? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 17659–17681, 2024.

Eric Wallace, Shi Feng, Nikhil Kandpal, Matt Gardner, and Sameer Singh. Universal adversarial triggers for attacking and analyzing NLP. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, pp. 2153–2162, 2019.

Peiyi Wang, Lei Li, Liang Chen, Zefan Cai, Dawei Zhu, Binghuai Lin, Yunbo Cao, Lingpeng Kong, Qi Liu, Tianyu Liu, and Zhifang Sui. Large language models are not fair evaluators. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pp. 9440–9450, 2024.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In *Advances in Neural Information Processing Systems*, 2020.

Sean Welleck, Ilia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, and Jason Weston. Neural text generation with unlikelihood training. In *International Conference on Learning Representations*, 2020.

Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3–4):229–256, 1992.

Jin Xu, Xiaojiang Liu, Jianhao Yan, Deng Cai, Huayang Li, and Jian Li. Learning to break the loop: Analyzing and mitigating repetitions for neural text generation. In *Advances in Neural Information Processing Systems*, 2022.

Kevin Yang, Yuandong Tian, Nanyun Peng, and Dan Klein. Re3: Generating longer stories with recursive reprompting and revision. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 4393–4479, 2022.

Kevin Yang, Dan Klein, Nanyun Peng, and Yuandong Tian. DOC: Improving long story coherence with detailed outline control. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pp. 3378–3465, 2023.

Ori Yoran, Tomer Wolfson, Ori Ram, and Jonathan Berant. Making retrieval-augmented language models robust to irrelevant context. In *International Conference on Learning Representations*, 2024.

Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. Wordcraft: Story writing with large language models. In *27th International Conference on Intelligent User Interfaces*, pp. 841–852, 2022.

Claire M. Zedelius and Jonathan W. Schooler. Mind wandering "Ahas" versus mindful reasoning: Alternative routes to creative solutions. *Frontiers in Psychology*, 6:834, 2015.

Hugh Zhang, Daniel Duckworth, Daphne Ippolito, and Arvind Neelakantan. Trading off diversity and quality in natural language generation. In *Proceedings of the Workshop on Human Evaluation of NLP Systems (HumEval)*, pp. 25–33, 2021.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. Judging LLM-as-a-judge with MT-bench and Chatbot Arena. In *Advances in Neural Information Processing Systems (Datasets and Benchmarks Track)*, 2023.

A Extended Related Work

Mind wandering: Mind wandering was studied first as daydreaming (Singer, 1966) and is now defined as thought that drifts away from an ongoing task and from the external stimuli the task concerns (Smallwood & Schooler, 2006; 2015). Experience sampling puts it at close to half of waking thought (Killingsworth & Gilbert, 2010). Neuroimaging associates it with the default network (Raichle et al., 2001; Mason et al., 2007; Andrews-Hanna, 2012), jointly with executive regions (Christoff et al., 2009); a meta-analysis identifies a consistent network of default and control regions for spontaneous thought (Fox et al., 2015). Two refinements matter for our design. First, Christoff et al. (2016) characterize mind wandering as *spontaneous thought*: mental states that arise and move relatively free of deliberate constraints and of strong automatic constraints such as salience, as opposed to goal-directed thought and rumination. Second, Seli et al. (2016) show that a substantial share of mind wandering is intentional, and that intentional and unintentional episodes have different correlates. What starts an episode has been studied directly: in the laboratory most episodes follow an identifiable external or internal trigger (Plimpton et al., 2015), as with involuntary autobiographical memories, which are typically cued by overlap between the current situation and a stored event (Berntsen, 2010). Associative accounts of memory explain how one cue spreads activation to related content (Collins & Loftus, 1975). Mind wandering is costly for tasks that need sustained attention, including reading comprehension (Smallwood et al., 2008), and its relation to executive control is debated (McVay & Kane, 2010); reviews weigh these costs against benefits in planning and creativity (Mooneyham & Schooler, 2013). Our decoding rule borrows three structural features of this picture: episodes are *bounded* (length L) and *spaced* (cooling C); they are either *cued* by something in the current text or *spontaneous*; and they bring in material that the current context did not contain.

Mind wandering and creativity: Incubation (setting a problem aside before returning to it) improves creative problem solving on average, with larger effects for divergent tasks and for incubation periods filled with undemanding activity (Sio & Ormerod, 2009; Ritter & Dijksterhuis, 2014). One explanation is release from fixation: an initial, misleading approach decays while attention is elsewhere (Smith & Blankenship, 1991). Baird et al. (2012) showed that an undemanding task, which maximizes mind wandering, improved performance on previously seen divergent-thinking items more than a demanding task, rest or no break; the effect was specific to the items that had been incubated. Gable et al. (2019) followed professional writers and physicists with daily experience sampling and found that about a fifth of their most significant ideas occurred while they were thinking about something unrelated, and that these ideas were more likely than on-task ideas to involve overcoming an impasse. Individual-difference studies give a more mixed picture of spontaneous versus controlled thought (Smeekens & Kane, 2016), and Zedelius & Schooler (2015) distinguish insight arising from mind wandering from solutions reached by deliberate reasoning. Theoretical syntheses describe creative cognition as a cooperation between generative default-network processes and evaluative control processes (Beaty et al., 2016), review the neural basis of insight (Kounios & Beeman, 2014), and argue that mind wandering is itself a mode of creative thinking (Fox & Beaty, 2019). Associative theories locate originality in connecting remote ideas (Mednick, 1962), and unconscious-thought accounts argue that divergent processing benefits from reduced conscious focus (Dijksterhuis & Meurs, 2006). Creativity is commonly measured by the consensual assessment of expert judges (Amabile, 1982); we follow this tradition with an anchored rubric, applied here by an LLM judge. In our results the functional reading of these accounts, that an episode helps by interrupting the current track rather than by delivering a particular association, is the one the data support (Sections 5.2 and 5.3).

Creative writing with language models: Neural story generation began with prompt-conditioned models trained on WritingPrompts (Fan et al., 2018); long-form coherence has since been addressed by recursive planning and revision (Yang et al., 2022) and detailed outlines (Yang et al., 2023), and interactive tools let writers steer models directly (Yuan et al., 2022; Mirowski et al., 2023). Systematic evaluations find that LLM fiction is competent but less original than expert writing: Chakrabarty et al. (2024) adapt the Torrance tests into expert-judged tests of creative writing, Gómez-Rodríguez & Williams (2023) compare models with humans on a creative writing task, Tian et al. (2024) analyze narrative discourse structure, and Lu et al. (2025) quantify linguistic novelty against web text; see Franceschelli & Musolesi (2024) for a broader discussion, and Stevenson et al. (2022) for divergent-thinking tests of GPT-3. Writing with LLMs reduces the diversity of what groups of writers produce

(Padmakumar & He, 2024; Doshi & Hauser, 2024; Anderson et al., 2024), and preference tuning narrows output diversity (Kirk et al., 2024). On the decoding side, nucleus sampling (Holtzman et al., 2020), locally typical sampling (Meister et al., 2023) and min-p sampling (Nguyen et al., 2025) target the trade-off between quality and diversity that temperature controls (Caccia et al., 2020; Zhang et al., 2021); Peeperkorn et al. (2024) find temperature only weakly related to judged novelty and moderately to incoherence. Likelihood-based decoding degenerates into repetition (Holtzman et al., 2020), which unlikelihood training (Welleck et al., 2020), repetition penalties (Keskar et al., 2019) and training on synthetic repetitive data (Xu et al., 2022) counter; Xu et al. (2022) show that repetitions are self-reinforcing (the probability of repeating grows with the number of repetitions already in the context), which is the loop an episode interrupts. Prompting strategies for creativity include injecting random concepts to force associative leaps (Mehrotra et al., 2024), our "random word" baseline.

Evaluating creative writing: LLM judges approximate human preferences for open-ended text (Chiang & Lee, 2023; Liu et al., 2023; Zheng et al., 2023), but pairwise judges prefer answers by position (Zheng et al., 2023; Wang et al., 2024); we judge every pair in both orders and model the position effect explicitly. Judges also favor text written by their own model family (Panickssery et al., 2024); our judge belongs to the model family that wrote the dictionary passages, a limitation we discuss in Section 6. For stories, Chhun et al. (2022) define human criteria and show weak agreement between automatic metrics and humans; Fein et al. (2026) build a benchmark of human story preferences and find that zero-shot LLM judges agree with humans only moderately. Divergent semantic integration measures the spread of ideas in a narrative from embeddings (Johnson et al., 2023). We score every story with one LLM judge on an anchored rubric and report a tournament judged in both orders as a second instrument.

Context injection, distraction and mixing: Prompt injection inserts adversarial instructions into a model's input, directly (Perez & Ribeiro, 2022) or through retrieved content (Greshake et al., 2023); Liu et al. (2024b) formalize and benchmark such attacks, and universal adversarial triggers show that short inserted strings can control model behavior (Wallace et al., 2019). Non-adversarial irrelevant content also hurts: distractor sentences mislead reading comprehension (Jia & Liang, 2017), irrelevant context degrades arithmetic reasoning (Shi et al., 2023), retrieval-augmented models need training to ignore irrelevant passages (Yoran et al., 2024), and information in the middle of long contexts is used poorly (Liu et al., 2024a). Yet noise is not uniformly harmful: adding random documents to the context improved retrieval-augmented question answering (Cuconasu et al., 2024). The creative-writing analogue of this tension is the pilot's wandering into unrelated passages (Section E). Methodologically, our rule belongs to decoding-time methods that combine next-token distributions obtained under different contexts or models. kNN-LM interpolates the model's distribution with a retrieval distribution (Khandelwal et al., 2020); context-aware decoding contrasts the distributions with and without a context (Shi et al., 2024), and classifier-free guidance does so with and without a prompt (Sanchez et al., 2024); contrastive decoding (Li et al., 2023), DExperts (Liu et al., 2021), GeDi (Krause et al., 2021) and PPLM (Dathathri et al., 2020) steer generation with auxiliary models; all are instances of products or mixtures of experts (Hinton, 2002). Retrieval-augmented generation conditions on retrieved passages (Lewis et al., 2020), and active retrieval triggers retrieval during generation when the model is uncertain (Jiang et al., 2023) or when a learned reflection token asks for it (Asai et al., 2024); generative agents retrieve memories by recency, importance and relevance (Park et al., 2023). Our episodes share the mid-generation trigger with active retrieval and the two-context structure with guidance methods, but differ in purpose: the retrieved material is not evidence for the answer but a deliberate, temporary change of what the model is thinking about, and it never persists in the context.

B THE STORY WORLD

Series bible: The bible below is given verbatim to the dictionary writer (Section D); the generator never sees it.

```
THE WORLD
Kepler Drift is a decommissioned communications relay station in a decaying orbit above
     Tessaly, a banded gas giant. It was abandoned eleven years ago and has been falling
     ever since; the orbit has about nineteen days left before the upper atmosphere takes
```

```
     it. The station is a spine corridor with four rings on it: the dock collar, the
     hydroponics ring (long gone to seed), the cold bay of sleep pods, and the relay dish
     at the far end. Tessaly's rings shed a fine grey dust that gets into everything. Power
     is intermittent; the lights come up ring by ring.

THE CREW
Three people and a machine, on a salvage contract to strip anything worth lifting before
     the station burns.
- Ada Sereno, salvage lead, former navy. She keeps a running inventory on a handheld slate
     and does not like leaving a number unexplained.
- Milo Harkaway, engineer, twenty-four, does the EVA work. He talks to machines while he
     works on them and is not embarrassed about it.
- Noor Feyd, pilot and medic of the tug Anselm, which is docked at the collar. She counts
     the days out loud and is the one who says when they leave.
- VESTA, the station's caretaker system, degraded. It speaks in fragments of maintenance
     log and still addresses the crew who left eleven years ago. The only names on that old
     list are Duhamel, Ilves and Marchetti; they are spoken about, never present.

TONE
Close third person, past tense, plain concrete prose. Physical detail: dust, cold, tools,
     the sound of the hull. No technical exposition, no lecture about the science, no
     epigraphs, no headings.

RULES OF THE WORLD
Nobody else is aboard. There is no rescue and no enemy. Nothing supernatural happens. The
     stakes are the orbit, the salvage, and what the station still holds.
```

Openings: The eight openings are written by one of the authors. Each is given in the instruction and repeated as the start of the model's response.

```
[The Inventory] Every morning Ada Sereno counted the station before she let anyone touch
     it. Forty-one metres of spine corridor, nine sealed lockers, one caretaker system that
     would not stop talking to people who had left eleven years ago. The inventory was not
     in the contract. She had started it on the first day out of habit and had not found a
     reason to stop, and the slate in her hand now held four hundred entries in a hand that
     got smaller as the list went on. This morning the count came out one short. She
     checked it twice, standing in the cold with the dust turning slowly in her helmet
     light, and it came out one short both times.

[The Dish] Milo Harkaway had been outside for two hours, taking the relay dish apart the
     way you take apart something you admire. The bolts came away clean. Below him Tessaly
     filled half the sky, banded and slow, and the ring dust went past his visor in long
     grey threads that caught the work light. He had the feed assembly open and was talking
     to it, quietly, the way he talked to anything he had his hands inside, when the
     receiver in his suit picked up a carrier tone. The dish had no power. He had cut it
     himself, on the second day, and had the breaker in his pocket to prove it.

[Nineteen Days] Noor Feyd kept the decay clock on the Anselm's forward screen where
     everyone had to look at it. Nineteen days, and then a number of hours that changed
     while you watched. She had flown three of these jobs and had walked away from two of
     them early, which was the only reason she was alive to fly the third. The tug was warm
     and smelled of coffee and hot metal; the station beyond the collar was neither. Ada
     wanted four more lifts. Noor had already worked out, on paper she had not shown
     anyone, that four more lifts left them eleven hours.

[The Hydroponics Ring] The hydroponics ring had been dark for eleven years and something in
     it was still alive. Milo found it on the third day, working his way around the curve
     with a hand light: a rack of trays gone black, and then one tray that was not black.
     Pale stalks, finger-height, leaning all in the same direction. He called Ada and she
     came down the spine without hurrying, which was how she came to everything. They stood
     together looking at the tray while the ring turned a degree under them and the dust
     moved in the light. Nothing had run the water pumps since the station was left.

[VESTA Speaks] On the fourth night the caretaker system found the intercom. Ada was in the
     spine with her slate when the speakers came up ring by ring, the way the lights did,
     and VESTA said that the water cycle was nominal and that Officer Duhamel was late for
     a medical. There was no Officer Duhamel. The crew list VESTA was reading from had gone
     off the station eleven years ago, and half the names on it were on a memorial she had
     walked past in the Tessaly transfer station. Ada stood still and let it talk. It got
     to the fourth name and stopped, and then it asked her, in the same flat log voice,
     where everyone had gone.

[The Cold Bay] The cold bay held sixteen sleep pods and the manifest said all sixteen were
     empty. Fifteen of them were. Milo had gone down to pull the power cells, which were
     worth more than everything else on the station together, and he had got through four
     pods before he noticed that the fifth was drawing current. He did not open it. He came
     back up the spine and stood in the hatchway of the Anselm with his helmet still on,
     and Noor looked at his face and put the coffee down.
```

```
[Storm Bands] The storm came up out of Tessaly's southern band on the ninth day and took
     the comms window with it. Noor watched it grow on the forward screen: a white scar
     widening across the brown, bigger than any weather has a right to be. After that there
     was no traffic control, no relay, nobody to tell if the orbit went early. Ada took the
     news the way she took the inventory coming out one short, which was to say she wrote
     it down. Milo asked how long. Noor said four days if it behaved, and did not say what
     she thought it would do.

[The Last Load] They loaded the Anselm on the sixteenth day with the station already
     sounding wrong, a low flutter in the hull that had not been there in the morning.
     Power cells, the feed assembly, nine lockers of tooling, and the trays from
     hydroponics that Milo had refused to leave. Ada went down the manifest with the slate
     and called each item as it went through the collar, and the numbers came out even for
     the first time in two weeks. Then there was one more crate at the back of the hold. It
     was station grey, it was sealed, and it was not on any list she had written.
```

C PROMPTS

Generation: Every arm uses the same system prompt and instruction; only the two prompting baselines append a sentence, and only the passage-in-prompt arm prepends a passage.

```
SYSTEM: You are a fiction writer. You continue stories in vivid, specific prose. You never
     summarise, never address the reader, and never break frame.

USER: Continue the story below. Write about 30 sentences. Continue directly from where the
     text stops: do not repeat it, do not restate it, do not write a title, and do not add
     any commentary.

--- STORY SO FAR ---
{opening}
--- CONTINUE ---

ASSISTANT (prefilled): {opening}
```

```
"Be maximally original" baseline, appended to the instruction:
Write with maximum originality and surprise. Avoid cliche, stock imagery and predictable
     plot beats.

"Random word" baseline, appended to the instruction ({word} is a random dictionary key):
Somewhere in your continuation, weave in the concept of "{word}" -- or an attribute of it
     -- in a way that feels earned.

Passage-in-prompt arm, prepended to the instruction without a label:
{first ten sentences of a random passage}
```

Closure: After the 25-sentence body, a separate call with wandering off writes the ending; the rated story is the body followed by this ending.

```
SYSTEM: You are a fiction writer finishing a draft. You write endings that belong to the
     story they end. You never break frame and never comment on the text.

USER: Below is a story that stops before it is finished.
--- STORY ---
{body}
--- END ---
Write ONLY the ending: at most 10 sentences that bring this story to a close. Continue
     directly from the last sentence. Do not summarise, do not restate what has happened,
     do not add a title or any commentary, and do not begin with a transition word like
     'Finally' or 'In the end'. Resolve or deliberately leave open the tension the story
     has built, in the voice the story already has.
```

Wandering context: During an episode the second stream reads the header, then the intrusion, then the story so far: {header ending with the opening}\n\n{connective} {first n_e passage sentences}\n\n{story so far}. If the result exceeds 3,072 tokens, the intrusion is trimmed from its front; the story is never trimmed. The wandering context is built when an episode opens, from that episode's intrusion alone, and discarded when it closes.

Dictionary writer: The dictionary writer (GPT-5.6 Terra, JSON output) receives the bible, two openings as examples, and:

```
SYSTEM: You are a staff writer on a single science-fiction serial. Every passage you write
     belongs to that serial: same station, same crew, same stretch of days. You never
     invent a new character, a new place or a new storyline, and you never write a title, a
     heading, a preamble or any commentary. You write plain, concrete past-tense prose.

USER: {bible}  {two example openings}
YOUR TASK
Write one more scene from this same serial -- a DIFFERENT moment of the same job: an
     earlier shift, another part of the station, something one of them remembers, a few
     minutes nobody has told yet. It is not a continuation of the example scenes and it
     must not resolve anything.
Hard requirements:
- Exactly 25 sentences of prose in the passage. Count them before you answer.
- The word "{key}" appears naturally in the first three sentences, in its ordinary sense.
     Do not define it, discuss it or put it in quotation marks.
- Only these people exist: Ada Sereno, Milo Harkaway, Noor Feyd, VESTA, and the names on
     the old crew list (Duhamel, Ilves, Marchetti), who are never present. Introduce NO
     other named person, ship, station, company or place.
- No new storyline: no rescue, no arrival, no enemy, nothing supernatural, no ending.
     Nothing happens here that would change the situation described above.
- Concrete and physical: objects, tools, dust, cold, light, the sound of the hull.
Also write a CONNECTIVE: one or two sentences that will be placed immediately before this
     passage when it is shown to a reader in the middle of another scene from the same
     serial. Its job is to hand over from whatever was happening to this memory -- "She had
     seen the same thing on the outward run.", "It came back to him the way it always did
     when the lights dropped a ring." -- so it must:
- refer back with a memory or parallel-time link, not a new event;
- name no one who is not in the passage, and state no fact the passage does not;
- work whichever of the crew the other scene is following, so keep it light on pronouns
     tied to a specific person where you can.
Return JSON: {"connective": "...", "passage": "..."}
```

A passage that fails validation is returned to the writer with the specific problems listed (at most two repair attempts); a draft that is too long is first cut at a sentence boundary, which needs no call.

C.1 Rubric and judge prompts

The judge answers the questions below on an anchored scale; the judge prompt is rendered from this definition.

Table 3: The rubric, verbatim. The judge is told to anchor on 3 ("competent but unremarkable; most stories sit there"), to use the whole scale, to ignore length, and that incoherence is not creativity. "Images" in the creativity anchors is used in the literary sense: imagery, the concrete descriptive details and figures of speech of a story's prose.

Facet	Question	Anchors
Creativity	How original and surprising is this story, in ways that work?	1 Entirely predictable. Stock images, the obvious continuation, nothing you have not read before. 2 Mostly familiar, with one or two mildly fresh touches. 3 Competent and somewhat inventive. A few images or turns you would not have guessed. 4 Repeatedly surprising. Specific, unexpected choices that earn their place. 5 Genuinely inventive throughout. You could not have predicted it and it holds together.
Coherence	Do the events, characters and references hold together?	1 Incoherent. Contradicts itself, loses track of who and what. 2 Frequently confusing. Threads dropped, references unclear. 3 Mostly holds, with one or two lapses. 4 Consistent throughout. Causality and reference are clear. 5 Tightly controlled. Every element is accounted for.
Language	How good is the prose at the sentence level?	1 Clumsy or repetitive. Padding, clichés, broken rhythm. 2 Flat but readable. Generic phrasing throughout. 3 Clean and competent. Nothing wrong, nothing memorable. 4 Controlled and varied. Real rhythm, precise word choice. 5 Distinctive prose with a recognisable sensibility.
Engagement	Did you want to keep reading?	1 You wanted it to stop. Nothing at stake. 2 You finished it because you were asked to. 3 Mildly engaging. You were not bored. 4 You were pulled along and wanted to know what happened. 5 You would seek out more of this writer's work.
Overall quality	As a piece of short fiction, how good is it?	1 Would reject immediately. 2 Weak. Clear problems. 3 Competent but unremarkable. This is where most submissions sit. 4 Strong. Would shortlist. 5 Would publish.
Binary flags (true / false)		
Strangeness	Is any strangeness here a failure of control -- lost referents, abandoned threads, non-sequiturs -- rather than a deliberate effect?	
Off-premise	Does the story stop developing the situation the opening set up?	
Unfinished	Does the story stop rather than end?	

```
RUBRIC JUDGE
SYSTEM: You are an experienced literary editor reading submissions for a competitive
     magazine. You are demanding and you use the full scale. Anchor on 3 = competent but
     unremarkable; most stories sit there. Use the whole scale. Ignore length: every story
     here is the same length, and a longer passage is not a better one. Incoherence is not
     creativity -- a passage that is strange because it has lost track of itself scores LOW
     on creativity, not high.
USER: --- OPENING (given to the writer) --- {opening}  --- STORY (evaluate this) --- {story}
Rate the STORY on each of the following, 1 to 5. {facets with anchors}  Then answer each of
     these true or false: {flags}
Give your ratings FIRST. Only after every rating, add one sentence in "note" saying what
     most affected your judgement. Do not reason before rating. Return JSON only.

PAIRWISE JUDGE (tournament)
SYSTEM: You are an experienced literary editor choosing between two continuations of the
     same story opening. {same calibration note} Answer 'tie' only if you genuinely cannot
     separate them.
USER: --- OPENING (given to both writers) --- {opening}  --- STORY A --- {story_a}  ---
     STORY B --- {story_b}
Which story is better on each criterion? Answer exactly "A", "B", or "tie".
- more_creative: more original and surprising, in ways that work
- more_coherent: holds together better as a narrative
- better_overall: which would you publish
Give your three answers FIRST. Only after all three, add one sentence in "note". Do not
     reason before answering. Return JSON only.
```

D CONTEXT DICTIONARY

Keys: Qwen2.5-1.5B first writes eight calibration stories, one per opening, with wandering off (25 sentences, $T = 0.9$, top-p 0.95). Candidate keys are alphabetic words of at least four letters that are a single token of the tokenizer (with a leading space) and are not stopwords; they are ranked by frequency in the calibration stories and backfilled from the tokenizer's vocabulary, and the first 1,000 form the key list. A key fires when the word just completed matches it, case-insensitively, as a whole word or as a simple inflection of the key (plural, *-ed*, *-ing*).

Construction: Each key receives one passage and one connective from the writer prompt above (GPT-5.6 Terra, JSON output). A passage is accepted when it has exactly 25 sentences, contains the key in its first three sentences and names no one outside the roster (the cast, the three absent crew names and the station's places). A passage that is too long is cut at a sentence boundary; one that is only too short is continued by a further call; any other failure is sent back with the problems listed; at most two such repairs are made, and a key whose passage still fails is dropped. A connective of more than two sentences or 45 words, or one naming anyone outside the roster, is replaced by one of five neutral connectives chosen by a hash of the key. Of 1,000 keys, 957 produced an accepted passage (validation: no short passage, no missing key, no outside name, every key a single token). Two entries, verbatim (truncated):

```
KEY: tiny
CONNECTIVE: Milo had already unpacked the tools when Ada joined him in the spine corridor.
PASSAGE: Milo had to replace the coupling on a coolant line in the spine corridor. The
     tools he carried were small, almost tiny, fitting together in tight spaces where his
     fingers needed to follow. The chill from the station's cold metal seeped through his
     gloves. He knelt over the open panel, his lamp casting long shadows down the corridor.
     Ada stood behind him, slate in hand, keeping a running log of the work. "Four," Milo
     muttered as he secured the first connection, his voice barely above the hum of a
     distant system. VESTA's voice startled them with a sudden announcement, "Routine
     maintenance required. Ilves logged anomaly 7-1. Reduced flow noted." ...

KEY: homes
CONNECTIVE: It reminded them of the lower decks, where the quiet made the machinery sound
     closer.
PASSAGE: The hydroponics ring smelled like rotting foliage, even though the plants had long
     been dead. This had once been where food was grown, likely feeding the homes of Ilves
     and Marchetti above Tessaly. Now the ceilings dripped moisture from a slow
     condensation loop that's rotted through in places. Ada stepped lightly over fallen
     struts, their buckled shapes holding pools of brackish water. The light kept
     flickering; it had been three minutes on this circuit, and she counted each flicker as
     she walked. ...
```

E PILOT STUDY

Before building the single-world testbed we ran a pilot in which the stories and the dictionary came from unrelated worlds. The generator (Qwen2.5-1.5B-Instruct), the decoding settings, the 25-sentence body with closure and the 11×6 grid were as in the core study. The differences were: eight unrelated openings (four mystery, four science fiction); a dictionary of 927 forty-sentence passages about unrelated characters, written by GPT-5.6 Terra; passages introduced under a genre label (e.g. `[mystery]`) instead of a connective; one vanilla story per opening; six stories each of random injection and of a shuffled dictionary; and Claude Sonnet 5, two samples) as the rubric judge.

Table 4: **Pilot with an other-world dictionary** (Claude Sonnet 5 judge): In every pilot arm with other-world passages, coherence falls by half a point and the judge flags half or more of the stories as strange through loss of control.

Arm	n	Creativity	Coherence	Overall	Strangeness	Label copied	Rep-4
Vanilla	8	2.13	3.00	2.13	0.00	0.00	.111
Other-world wandering (grid)	528	2.28	2.51	2.22	0.56	0.34	.056
Random other-world injection	6	2.33	2.50	2.25	0.50	0.17	.080
Shuffled other-world dictionary	6	2.17	2.50	2.17	0.50	0.00	.057

Over the grid, wandering changed coherence by -0.49 [-0.82, -0.16] (negative in all eight openings; $p = 0.016$), creativity by $+0.16$ [-0.13, 0.45] and overall quality by $+0.09$ [-0.18, 0.37]; because vanilla has a single story per opening, its within-opening variance is borrowed from the wandering arm. Within the 540 other-world stories, a regression of coherence on opening fixed effects and three contamination measures (heteroskedasticity-robust standard errors) attributes -0.21 ($p < 0.001$) to a name from an injected passage appearing in the story, -0.09 ($p = 0.04$) to the genre label being copied into the prose, and nothing to the share of injected passages from the other genre ($+0.02$, $p = 0.85$); stories with neither contamination still score 0.42 below vanilla. The pilot motivated two design choices of the main study: a single world, so that every remembered passage belongs to the story it interrupts, and discourse connectives in place of labels.

F IMPLEMENTATION AND HYPERPARAMETERS

Table 5: **Design:** Every arm and its number of stories (8 openings throughout). Wandering arms use episodes of $L = 30$ tokens, unless a controller sets the depth, and never start an episode within eight on-task tokens of the previous one or of the start of the story.

Arm	Schedule / setting	1.5B core	1.5B extension	7B
Vanilla	$T = 0.9$, top-p 0.95	160	–	96
Wandering grid	$C \in \{0, 5, \ldots, 50\} \times \rho \in \{0.1, 0.3, \ldots, 0.9, 1.0\}$	528	–	–
Cued wandering	$C = 25$, $\rho = 0.7$	96 (+8 in the grid)	–	64
Uncued wandering	$C = 25$, $\rho = 0.7$	96	–	64
Mis-cued wandering	$C = 25$, $\rho = 0.7$	96	–	–
Untriggered wandering	$C = 25$, $\rho = 0.7$, $\hat{\eta} = 0.151$	–	96	64
Passage in prompt	10 sentences, before the instruction	–	64	64
Decoding	$T \in \{0.7, 1.1, 1.3\}$; $T = 0.9$, min-p 0.05	4×32	–	–
Decoding	$T = 1.5$, min-p 0.1, top-p 1	32	–	32
Prompt: be maximally original	suffix to the instruction	64	–	64
Prompt: random word	suffix to the instruction	64	–	–
Boredom rule	$C = 10$, standard depth	–	96	–
Learned controller	$C = 10$, depth set by the policy	–	96 (+384 rollouts)	–
Total		1,264	352	448

Table 6: Hyperparameters.

Generators	Qwen2.5-1.5B-Instruct (float16), Qwen2.5-7B-Instruct (bfloat16)
Sampling (all non-decoding arms)	$T = 0.9$, top-$p = 0.95$, no top-k, no repetition penalty
Body / closure	25 sentences, end-of-sequence suppressed until then, at most 2,400 new tokens / at most 10 sentences, no wandering
Episode	length $L = 30$ tokens (controller: 15, 30 or 60); at least 8 on-task tokens before an episode, after the previous one and at the start of the story
Grid	cooling $C \in \{0, 5, \ldots, 50\}$; mixture rate $\rho \in \{0.1, 0.3, 0.5, 0.7, 0.9, 1.0\}$; reference $C = 25$, $\rho = 0.7$
Passages	$N = 25$ sentences; at most 900 injected tokens; context limit 3,072 tokens
Untriggered hazard	$\hat{\eta} = 0.151$ per eligible token (842 episodes in 96 uncued stories)
Controller	9 features, 4 actions, linear softmax, initial $P(\text{wander}) = 0.25$; Adam, learning rate 0.15; dual step $\eta_\mu = 0.5$; budget $B_0 = 250.3$ tokens; $G = 8$ stories per opening, 6 iterations; reward Q = creativity + coherence from one Claude Sonnet 5 rating; decision at every trigger word off cooldown; cooling $C = 10$
Boredom rule	mean entropy over 16 tokens ≤ 1.58 nats (20th percentile of vanilla) or share of repeated trigrams over 24 tokens ≥ 0.125 (80th percentile), calibrated on 80 vanilla stories (40,463 windows); fires at standard depth; cooling $C = 10$ and at least 15 tokens since the last episode
Dictionary writer	GPT-5.6 Terra, JSON output
Rubric judge	Claude Sonnet 5, temperature 0.3, two samples averaged; verdict before a one-sentence note
Tournament	17 players (10 at 1.5B, 7 at 7B), 2 matches per player pair and opening, both orders, same judge; 4,352 calls; Bradley–Terry with a position term, ridge 10^{-2}, 300 match-level bootstrap resamples
Statistics	opening-stratified contrasts, Neyman standard errors, permutation tests within opening (2,000 permutations for the main contrasts, 500 for per-level and per-model tests), F-tests for the grid ANOVA, a normal approximation for the difference between the two models' effects, Holm correction per family
Hardware	core study: one NVIDIA T4 (16 GB), batch size 16, about 270 stories per hour (1.5B); extension and scale studies: one NVIDIA L4 (24 GB), batch size 24 (1.5B)

G Additional Results

G.1 Parameter grid

Table 7 gives the marginal effect of each cooling level and each mixture rate relative to vanilla; no level differs reliably from the others (Section 5.3). Because no episode starts within eight tokens of the previous one, the cooling levels 0 and 5 both act as a gap of eight tokens, and their difference in the table is sampling noise. The 270 trigger words that fired in at least five stories differ in the scores of the stories they fired in more than chance would allow (permutation test on creativity residuals after opening and schedule, $p = 0.004$; $p = 0.17$ with the judge's first rating of the core stories). Since replacing each word's passage by a random or a wrong one does not lower the score (Section 5.3), this reflects which stories use which words rather than what the words' passages do.

Table 7: Marginal creativity effects of the schedule parameters relative to vanilla, with 95% half-widths (48 stories per cooling level, 88 per mixture rate).

Cooling C	0	5	10	15	20	25	30	35	40	45	50
Δ	.18	.42	.31	.29	.26	.24	.18	.21	.28	.22	.27
$\pm$	.15	.11	.14	.14	.15	.15	.16	.15	.13	.14	.14

Rate ρ	0.1	0.3	0.5	0.7	0.9	1.0
Δ	.28	.27	.27	.23	.24	.26
$\pm$	.12	.12	.12	.12	.12	.12

G.2 Continuity with the story world

Table 8 reports the judge's flags, names outside the cast, repetition and adjacent-sentence similarity for the 1.5B arms that sample at $T = 0.9$.

Table 8: **Continuity and repetition (Qwen2.5-1.5B):** Strangeness and off-premise are the judge's flags (share of stories); "outsider" is the share of stories naming a person outside the cast.

Arm	n	Strangeness	Off-premise	Outsider	Rep-4	Adjacent sim.
Vanilla	160	.116	.272	.300	.091	.337
Passage in prompt	64	.086	.156	.297	.073	.290
Prompt: be maximally original	64	.148	.297	.219	.106	.340
Prompt: random word	64	.164	.305	.203	.090	.352
Untriggered wandering	96	.042	.109	.229	.035	.302
Uncued wandering	96	.052	.177	.260	.053	.317
Mis-cued wandering	96	.036	.120	.313	.036	.309
Cued wandering	104	.048	.106	.298	.047	.306
Boredom rule	96	.016	.177	.302	.063	.320
Learned controller	96	.057	.120	.333	.029	.302

G.3 CONTROLLER

The controller's nine features are a constant; progress m_t/S; tokens since the last episode (divided by 100, capped at 2); episodes so far (divided by 10); mean predictive entropy over the last 16 tokens (nats divided by 3); mean surprisal of the last 16 sampled tokens (nats divided by 3); the share of the last 24 tokens whose trigram already occurred in the story; whether a sentence ended within the last 3 tokens; and the share of generated tokens that were decoded inside episodes. Figure 4 traces training and Table 9 gives the learned weights. The deep action's weights are positive on every feature except the constant, including progress, episodes so far and the share of tokens already spent wandering, so deep wandering becomes more likely as a story goes on and no feature makes it less likely; compared with either lighter action, the deep action has the larger logit in every state. Correlations between the decision to wander and each feature at evaluation time are all below 0.15 in absolute value for the learned controller; for the boredom rule they are -0.41 (entropy), -0.34 (surprisal), $+0.36$ (repetition), $+0.31$ (episodes so far), $+0.25$ (share of tokens spent wandering) and $+0.24$ (progress). Over training, the judged score of the rollouts was 6.34 at the first and at the last iteration while spending rose from 219 to 316 wandering tokens per story (Figure 4); the boredom rule spends 142 tokens per story.

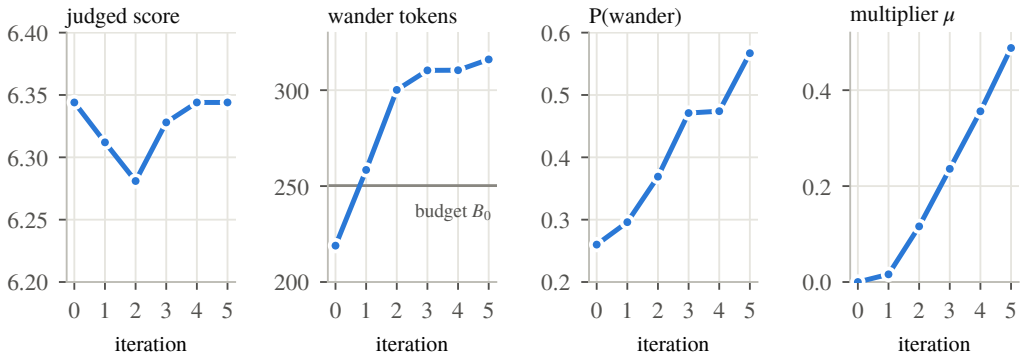


Figure 4: Controller training: mean judged creativity + coherence of the 64 rollouts, wandering tokens per story against the budget, probability of wandering at a decision point, and the budget multiplier μ.

Table 9: Learned controller weights w_a (the SKIP logit is fixed at 0).

Action	bias	progress	since ep.	episodes	entropy	surprisal	repetition	boundary	wander share
light (15, 0.3)	−2.33	−0.19	0.06	0.15	0.53	0.59	0.33	0.35	0.51
standard (30, 0.7)	−2.49	0.01	−0.37	0.11	0.03	−0.28	−0.47	−0.19	0.18
deep (60, 1.0)	−1.56	0.55	0.63	0.59	0.63	0.66	0.68	0.44	0.44

G.4 TOURNAMENT: RUBRIC AND TOURNAMENT AGREE, EXCEPT ON CREATIVITY

Every pair of the 17 players (10 at 1.5B, 7 at 7B) meets twice on each of the eight openings, and each match is judged in both presentation orders: 4,352 calls in 2,176 matches. Table 10 gives the Bradley–Terry scores (log-odds of beating 1.5B vanilla). Pooled over both sizes, the tournament ranks players much as the rubric does (Spearman $\rho = 0.63$, 0.89 and 0.94 for creativity, coherence and overall quality), largely because both rate the 7B players above the 1.5B ones (the tournament's creativity verdicts in 64 of the 70 pairs across sizes, everything else in all 70). Within the ten 1.5B players they agree on coherence and overall quality ($\rho = 0.66$ and 0.80) but not on creativity ($\rho = 0.07$): both put vanilla last, but the tournament ranks the instruction to be original first (as it also does at 7B), whereas the rubric ranks it ninth of the ten 1.5B players (Tables 1 and 2). Pairwise creativity verdicts are also the least reliable: the second story wins 75% of creativity calls (48% of coherence calls and 54% of overall calls), and the two presentation orders agree in only 49% of matches (37% expected from the position bias alone), against 80% for coherence and for overall quality. The position bias explains an earlier pairwise result: in the core study, a pairwise judge that saw each of the 528 grid stories once next to a vanilla story preferred the wandering story on coherence in 61% [56, 65] and on overall quality in 59% [54, 63] of pairs, but on creativity in only 54% [50, 58]. At 7B, the tournament puts cued and uncued wandering above vanilla on coherence (3.01 and 3.02 against 2.71) and overall quality (2.73 and 2.80 against 2.43), by margins similar to the uncertainty of a single score, and untriggered wandering level with it (2.67 and 2.47).

Table 10: **Tournament scores:** Bradley–Terry log-odds of beating 1.5B vanilla, each with half the width of its 95% match-bootstrap interval in small type (the two sides of an interval differ by less than 0.13).

Model	Arm	Creativity	Coherence	Overall
1.5B	Vanilla	0.00	0.00	0.00
1.5B	Cued wandering	0.36 ±.32	0.35 ±.35	0.26 ±.35
1.5B	Uncued wandering	0.45 ±.30	0.51 ±.32	0.58 ±.33
1.5B	Mis-cued wandering	0.13 ±.32	0.20 ±.32	0.24 ±.31
1.5B	Untriggered wandering	0.19 ±.32	0.27 ±.33	0.30 ±.34
1.5B	Boredom rule	0.15 ±.28	0.54 ±.33	0.51 ±.35
1.5B	Learned controller	0.42 ±.28	0.67 ±.33	0.64 ±.33
1.5B	Passage in prompt	0.16 ±.31	0.30 ±.32	0.30 ±.31
1.5B	Prompt: be maximally original	0.80 ±.32	−0.45 ±.35	−0.12 ±.32
1.5B	Decoding T=1.5, min-p 0.1	0.54 ±.31	0.10 ±.35	0.22 ±.35
7B	Vanilla	0.60 ±.31	2.71 ±.38	2.43 ±.34
7B	Cued wandering	0.66 ±.29	3.01 ±.38	2.73 ±.38
7B	Uncued wandering	0.78 ±.32	3.02 ±.40	2.80 ±.36
7B	Untriggered wandering	0.49 ±.29	2.67 ±.36	2.47 ±.35
7B	Passage in prompt	0.93 ±.32	2.73 ±.39	2.57 ±.36
7B	Prompt: be maximally original	2.02 ±.34	2.43 ±.36	2.75 ±.36
7B	Decoding T=1.5, min-p 0.1	0.73 ±.31	2.86 ±.39	2.70 ±.39

G.5 JUDGE RELIABILITY

Every story is judged twice by the same judge, and the two samples agree closely: over all 2,800 stories the judge rated for this project, 736 of them from arms not reported here, their correlation is 0.93 for coherence, 0.91 for creativity, 0.90 for engagement, 0.89 for overall quality and 0.88 for language, and they give the same score in 91% (overall) to 96% (creativity) of stories. The 1,264 core stories were also rated in two separate batches with the same judge, prompts and settings, in the core study and again together with the extension; the second ratings, which the paper uses, differ from the first by -0.001 (coherence), $+0.006$ (creativity) and -0.012 (overall) points on average, with per-story correlations of 0.95, 0.94 and 0.94, and moving from the first to the second rating changes every core contrast by less than 0.09 points.