
Reasoning Models Generate Societies of Thought

Anonymous Authors¹

Abstract

Large language models have achieved remarkable capabilities across domains, yet mechanisms underlying sophisticated reasoning remain elusive. Recent reasoning-reinforced models, including OpenAI’s o-series, DeepSeek-R1, and QwQ-32B, outperform comparable instruction-tuned models on complex cognitive tasks, attributed to extended test-time computation through longer chains of thought. Here we show that enhanced reasoning emerges not from extended computation alone, but from the implicit simulation of complex, multi-agent-like interactions—a society of thought—which enables the deliberate diversification and debate among internal cognitive perspectives characterized by distinct personality traits and domain expertise. Through quantitative analysis using classified outputs and mechanistic interpretability methods applied to reasoning traces, we find that reasoning models like DeepSeek-R1 and QwQ-32B exhibit much greater perspective diversity than baseline and merely instruction-tuned models, activating broader conflict between heterogeneous personality- and expertise-related features during reasoning. This multi-agent structure manifests in conversational behaviours including question-answering sequences, perspective shifts, and reconciliation of conflicting views, as well as in socio-emotional roles that characterize sharp back-and-forth conversation, which together account for the accuracy advantage in reasoning tasks through both direct and indirect facilitation of cognitive strategies. Controlled reinforcement learning experiments further reveal that base models spontaneously increase conversational behaviours when solely rewarded for reasoning accuracy, and fine-tuning models with conversational scaffolding substantially accelerates reasoning improvement compared to base mod-

els and models fine-tuned with monologue-like reasoning. These findings indicate that the social organization of thought enables effective exploration of solution spaces. We suggest that reasoning models establish a computational parallel to collective intelligence in human groups, where diversity enables superior problem-solving when systematically structured and suggest new opportunities for agent organization to harness the wisdom of crowds.

1. Introduction

Artificial intelligence (AI) systems have undergone a remarkable transformation in recent years, with large language models (LLMs) demonstrating increasingly sophisticated abilities across domains, from mathematics and code to scientific and creative writing to critical decision support (Brown et al., 2020; Wei et al., 2022). Nevertheless, a persistent challenge has been the development of robust reasoning capabilities—the ability to methodically analyze problems, consider alternatives, detect errors, and arrive at reliable conclusions. Recent reasoning models, such as DeepSeek-R1, QwQ, and OpenAI’s o-series models (o1, o3, o4), are trained by reinforcement learning to “think” before they respond, generating lengthy “chains of thought”. This led to substantial improvement in reasoning accuracy compared to existing instruction-tuned language models (e.g., DeepSeek-V3, Qwen-2.5, GPT-4.1) (OpenAI, 2024; Guo et al., 2025). Yet, the character of “thinking” within reasoning models that drives success remains underexplored.

We propose that reasoning models learn to emulate social, multi-agent-like dialogue between multiple perspectives—what we term a “society of thought”—to improve their reasoning, given the centrality of social interaction to the development of reason in both cognitive and social scientific accounts. Mercier and Sperber’s “Enigma of Reason” argument posits that human reasoning evolved primarily as a social process, with knowledge emerging through adversarial reasoning and engagement across differing viewpoints (Mercier & Sperber, 2017). Empirical work supports the idea that groups outperform individuals on a wide range of reasoning tasks by pooling information, calibrating confidence, and exhibiting collective intelligence through bal-

¹Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

anced turn-taking among diverse perspectives (Moshman & Geil, 1998; Mesmer-Magnus & Dechurch, 2012; Woolley et al., 2010; Bahrami et al., 2010). Cognitive diversity, stemming from variation in expertise and personality traits, enhances problem solving, particularly when accompanied by authentic dissent (Hong & Page, 2004; Nemeth et al., 2001; Engel et al., 2014; DeChurch & Mesmer-Magnus, 2010; Page, 2019; Trouche et al., 2014; Barrick et al., 1998; Kim et al., 2025b). Together, these findings suggest that robust reasoning emerges through interaction and the integration of diverse perspectives, and that key reasoning strategies, including verification and backtracking, may be realized through the conversation of simulated personas.

Therefore, we propose that reasoning models like DeepSeek-R1 improve reasoning via “society of thought”—implicit simulations of multi-agent-like interactions between diverse perspectives that give rise to them. Without deploying separate models prompted to interact with one another (Park et al., 2023; Chan et al., 2024; Liang et al., 2024), we suggest that behaviourally similar conversations between diverse perspectives occur and are leveraged within reasoning models.

Reasoning models like DeepSeek-R1 develop reasoning abilities through reinforcement learning, which iteratively compensates reasoning behaviour that yields correct answers. While earlier studies focus on how the model learns to scale test-time computations and generate longer reasoning traces (Chen et al., 2025a; Wei et al., 2022), merely increasing trace length does not account for the observed improvements in reasoning capabilities. Recent analyses pinpoint behavioural patterns that improve reasoning accuracy, such as verification of earlier assumptions, backtracking, and exploration of alternatives (Guo et al., 2025; Gandhi et al., 2025; Wang et al., 2025; Chen et al., 2025a; Yeo et al., 2025). Mechanistic interpretability research has shown that features in language models such as the frequent use of words like “wait,” “but,” and “however”—are associated with these behaviours (Venhoff et al., 2025b; Galichin et al., 2025; Ward et al., 2025; Venhoff et al., 2025a). The characteristics of these features, however, such as their prevalence in social and conversational settings, have rarely been explored.

Here we investigate the prevalence of reasoning traces of DeepSeek-R1, as well as QwQ-32B, that mimic simulated social interactions, quantifying how conversational behaviours, socio-emotional roles, and diversity of implicit agent “perspectives” contribute to reasoning performance. We first identify whether conversational behaviours and socio-emotional roles—hallmarks of human dialogue such as questioning, perspective taking, and reconciliation—are present in DeepSeek-R1’s and QwQ-32B’s reasoning traces. Then we test whether conversational behaviour

contributes to reasoning performance using mechanistic interpretability methods applied to DeepSeek-R1’s distilled model (DeepSeek-R1-Llama-8B) .

Next, we analyze the diversity of reasoning “perspectives” or simulated voices within DeepSeek-R1’s and QwQ-32B’s reasoning traces. Literature suggests that LLM reasoning can fail if models do not engage in meaningful disagreement and instead conform to misleading initial claims through pleasant, “sycophantic” conversations that propagate incorrect assumptions and knowledge (Wynn et al., 2025; Chen et al., 2024; Feng et al., 2025). Successful reasoning models may therefore exhibit disagreement driven by diversity in simulated perspectives, expressed through distinct personalities and expertise to avoid the “echo chamber” that leads to wrong answers. Therefore, we analyze reasoning traces using LLM-as-judge to accurately identify distinct voices underlying conversation.

Finally, we conduct a controlled reinforcement learning experiment to examine the role of conversational behaviours. We focus on self-taught reinforcement learning that rewards only accuracy and correct formatting (i.e., wrapping the thinking process between `<think>` and `</think>`), the common approach for improving modern language models’ reasoning capabilities (Guo et al., 2025). Based on a symbolic arithmetic task (Countdown game) (Gandhi et al., 2025; 2024), as well as a misinformation identification task, we apply reinforcement learning that rewards reasoning traces leading to accurate answers on open-source LLMs. We additionally compare RL initialization with conversational vs. monologue-like fine-tuning across two model families (Qwen-2.5-3B and Llama-3.2-3B), and test cross-domain transfer to political-misinformation detection.

2. Results

We compile a suite of widely used benchmarks used in prior research and official model cards of reasoning models (Big-Bench Hard, GPQA, MATH (Hard), MMLU-Pro, MUSR, and IFEval) (OpenAI, 2024; Guo et al., 2025), spanning symbolic logic, mathematical problem solving, scientific reasoning, multi-agent inference, and instruction following tasks (see [Methods: Data](#)). From this pool, we sample 8,262 problems and generate reasoning traces using DeepSeek-R1-0528 (671B parameters; hereafter DeepSeek-R1) and QwQ-32B. For comparison, we also generate reasoning traces using conventional, instruction-tuned models of varying sizes: DeepSeek-V3-0324 (671B parameters; hereafter DeepSeek-V3), Qwen-2.5-32B-Instruct (hereafter Qwen-2.5-32B-IT; the instruction-tuned model based on Qwen-2.5-32B from which QwQ-32B is derived), Llama-3.3-70B-Instruct (hereafter Llama-3.3-70B-IT), and Llama-3.1-8B-Instruct (hereafter Llama-3.1-8B-IT) (DeepSeek-AI et al., 2024; Grattafiori et al., 2024). DeepSeek-V3 is the

instruction-tuned model based on DeepSeek-V3-base from which DeepSeek-R1 is derived, and Qwen-2.5-32B-IT is the instruction-tuned model based on Qwen-2.5-32B from which QwQ-32B is derived (see [Methods: Data](#)) (DeepSeek-AI et al., 2024; Grattafiori et al., 2024).

Next, we estimate behavioural differences between reasoning models (DeepSeek-R1 and QwQ-32B) and the instruction-tuned models. We use linear probability models with problem-level fixed effects, which control for all task-specific characteristics, such as the difficulty of tasks. Specifically, we compare each reasoning model with its corresponding instruction-tuned counterpart (i.e., DeepSeek-R1 vs. DeepSeek-V3; QwQ-32B vs. Qwen-2.5-32B-IT) on the presence of conversational behaviours and socio-emotional roles. We control for log-transformed reasoning trace length ([Extended Data Fig. 1](#)) to consider that observed differences are not merely driven by “longer” chains of thought—that is, we demonstrate that reasoning models exhibit more frequent conversational behaviours and socio-emotional roles even when trace lengths are similar (see [Methods: Statistical analyses](#)).

2.1. Conversational Behaviours and Socio-Emotional Roles

We begin by investigating whether conversational behaviours and socio-emotional roles constitutive of back-and-forth dialogue are prevalent in reasoning traces. Using an LLM-as-judge, we quantify the occurrence of four conversational behaviours—defined as behaviours signaling the simulation of exchanges among multiple perspectives to explore a given problem—within each reasoning trace: (1) question–answering, in which the trace poses and then resolves questions; (2) perspective shifts, where alternative viewpoints are explored; (3) conflicts of perspectives, in which competing viewpoints are sharply contrasted; and (4) reconciliation, where conflicting viewpoints are integrated and coherently resolved.

We also examine socio-emotional roles based on Bales’ Interaction Process Analysis (IPA) (Bales, 1950). This identifies 12 interaction roles grouped into four categories: (1) *asking* for orientation, opinion, and suggestion, (2) *giving* orientation, opinion, and suggestion, (3) *negative* emotional roles (disagreement, antagonism, tension), and (4) *positive* emotional roles (agreement, solidarity, tension release), which together characterize interactive group activity. These behaviours are annotated using an LLM-as-judge (Gemini-2.5-Pro) that shows substantial agreement with both a human rater (average $ICC(3,1) = .756$) and another LLM (GPT-5.2; average $ICC(3,1) = .875$) (see [Methods: Measurements](#) and [Supplementary Method: LLM-as-judge prompts](#) (Conversational behaviours and Socioemotional roles)). To illustrate how reasoning traces are annotated, we

provide examples in [Extended Data Fig. 2](#) and [Supplementary Methods: Annotation Examples](#).

As shown in [Fig. 1a](#), we quantify the occurrence of four conversational behaviours within each reasoning trace, and report the proportion of traces exhibiting more than one such behaviour. DeepSeek-R1 and QwQ-32B exhibit conversational behaviours far more frequently than instruction-tuned models. DeepSeek-R1 shows significantly more question–answering, perspective shifts, and reconciliation compared to DeepSeek-V3. QwQ-32B displays a similar pattern relative to Qwen-2.5-32B-IT, with greater question–answering, perspective shifts, conflicts of perspectives, and reconciliation. Notably, all instruction-tuned models show consistently low prevalence of conversational behaviours regardless of parameter count (8B, 32B, 70B, 671B).

As shown in [Fig. 1b](#), both DeepSeek-R1 and QwQ-32B exhibit more reciprocal socio-emotional roles compared to their instruction-tuned counterparts: they both *ask* for and *give* orientations, opinions, and suggestions, while also displaying both *negative* and *positive* roles. DeepSeek-R1 asks more frequently than DeepSeek-V3, engages more in negative roles, and displays more positive roles. QwQ-32B shows a similar pattern relative to Qwen-2.5-32B-IT, with increased asking, negative roles, and positive roles. In contrast, instruction-tuned models predominantly *give* orientations, opinions, and suggestions without reciprocal asking behaviours or emotional engagement, producing one-sided monologues rather than simulated dialogue.

We quantify reciprocal role balance using the Jaccard index, which captures whether both sides of a role pair—asking versus giving for task-oriented roles, and positive versus negative for emotional roles—co-occur within the same reasoning trace. As shown in [Fig. 1c](#), DeepSeek-R1 exhibits significantly higher Jaccard indices for both ask & give and positive & negative roles compared to DeepSeek-V3, indicating that the model coordinates roles reciprocally rather than deploying them in isolation. QwQ-32B shows a similar pattern relative to Qwen-2.5-32B-IT (see [Supplementary Table 1](#)).

We further examine whether conversational behaviours and socio-emotional roles become more pronounced when DeepSeek-R1 faces more difficult tasks. Problem complexity is assessed both by an external LLM-as-judge ([Fig. 1e](#): Gemini-2.5-Pro) or by error rates across conventional instruction-tuned models ([Fig. 1f](#): DeepSeek-V3, Qwen-2.5-32B, Llama-3.3-70B-IT, Llama-3.1-8B-IT). As illustrated in [Fig. 1d](#) and [1e](#), these behaviours appear more frequently when DeepSeek-R1 tackles more complex problems, except for giving orientations and opinions. Consistent patterns across both measures suggest that conversational reasoning is preferentially activated in response to greater problem difficulty. For instance, tasks with the

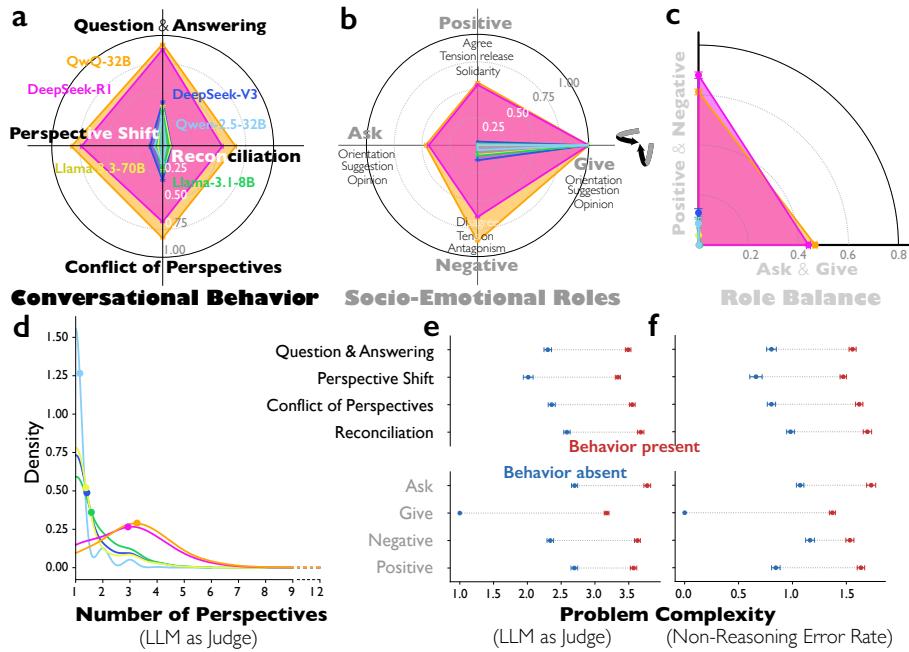


Figure 1. Conversational behaviours and Bales’ socio-emotional roles in chain-of-thought reasoning. **a**, Proportion of traces containing each conversational behaviour. **b**, Proportion of Bales’ 12 socio-emotional roles, grouped into *ask/give* and *positive/negative* categories. **c**, Jaccard index of role-pair co-occurrence. **d**, Distribution of distinct perspectives per trace. **e-f**, Problem complexity vs. behaviour presence in DeepSeek-R1, measured by Gemini-2.5-Pro Likert (**e**) or instruction-tuned error rates (**f**). Error bars: 95% CI. Full descriptions in Appendix B.

highest complexity scores—such as GPQA (graduate-level science) and challenging math problems—exhibit strong conversational patterns, whereas simple procedural tasks like boolean expressions and basic logical deduction show minimal dialogic behaviour (see [Supplementary Table 2](#)).

To decompose the accuracy advantage of reasoning models (DeepSeek-R1 and QwQ-32B) using the behavioural mechanisms above, we estimate a structural equation model with four conversational behaviours, four socio-emotional roles, and four cognitive behaviour mediators, using task accuracy as the outcome. Results suggest that conversational behaviours and socio-emotional roles mediate reasoning models’ accuracy advantage, both directly and indirectly through facilitating useful cognitive strategies, such as verification, backtracking, subgoal setting, and backward tracking (See [Extended Data Fig. 4](#); [Supplementary Methods: behavioural Pathways Linking Reasoning Models to Accuracy Advantages](#)).

2.2. Conversational Feature Steering Improves Reasoning Accuracy

We employ mechanistic interpretability methods to identify and manipulate features in the model’s activation space related to conversational behaviours, and examine how steering these features affects the model’s reasoning capabilities. We use sparse autoencoders (SAEs), which decompose neu-

ral network activations into a large set of linear, interpretable features (Cunningham et al., 2024; Bricken et al., 2023; Templeton et al., 2024). Specifically, we use an SAE trained on Layer 15’s residual stream activations of DeepSeek-R1-Llama-8B (15-llamascope-slimpj-res-32k), a distilled model derived from DeepSeek-R1 frequently used to conduct interpretability research on LLM reasoning (Venhoff et al., 2025b; Galichin et al., 2025; Ward et al., 2025; Venhoff et al., 2025a). The SAE was trained on the SlimPajama dataset (see [Supplementary Table 3](#) for full SAE hyperparameters) (Shen et al., 2023).

To identify SAE features associated with conversational contexts, we follow a conventional interpretability pipeline (Bills et al., 2023; Paulo et al., 2025; Cunningham et al., 2024). We first run the SAE on a large-scale corpus (SlimPajama-3B), sampling around 50 contexts where each of the 32,768 features activates to “explain” the role of each feature. Using LLM-as-judge classification of these contexts (Gemini-2.5-flash-lite), we compute the *conversation ratio* for each feature—the proportion of feature activations that occur in interpersonal, conversational settings (see [Fig. 2a](#) for the distribution across all features). We focus on features with conversation ratios above 50% that tend to activate near sentence onsets (i.e., within the first four tokens). From the candidates, we curate feature 30939, summarized as “a discourse marker for surprise, realization, or acknowledgment” by Gemini-2.5-Pro, which activates on tokens like

“Oh!” in contexts involving turn-taking and social exchange (see Fig. 2a). This feature exhibits a conversation ratio of 65.7%—placing it in the 99th percentile among all features—while maintaining high sparsity (0.016% of tokens), indicating that it captures a specific conversational phenomenon rather than general linguistic patterns. We select this feature because prior literature suggests that expressions of surprise signal a shift in contrasting perspectives characteristic of social coordination and affiliation (Parkinson et al., 2004; Wilkinson & Kitzinger, 2006).

We examine whether steering this feature causally induces conversational behaviours and improves reasoning accuracy using the activation addition method, which adds scaled feature vectors to model activations during generation. Specifically, we use the Countdown game, a benchmark commonly used to evaluate LLM multi-step reasoning capabilities (Gandhi et al., 2025; 2024). In the Countdown task, the model must combine a given set of numbers using basic arithmetic operations (+, −, ×, ÷) and parentheses to reach a target value—for example, given inputs 25, 30, 3, 4 and target 32, a valid solution is $(30 - 25 + 3) \times 4 = 32$ (Gandhi et al., 2025; 2024). We use the sample of 1,024 Countdown problems. We prompt the model to generate chain-of-thought reasoning, and at each token generation step, we add the feature 30939 vector (scaled by the steering strength) to layer 15 activations.

As shown in Fig. 2b, steering the conversational surprise feature with positive direction (+10) doubles accuracy from 27.1% to 54.8% in the Countdown task, while steering in the negative direction (−10) reduces accuracy to 23.8%. The radar plot inset reveals that positive steering (from 0 to +10) simultaneously increases all four conversational behaviours—more question-answering, perspective shifts, conflict of perspectives, and reconciliation, controlling for problem fixed-effects and log-transformed reasoning trace length. Negative steering from 0 to −10 suppresses them, reducing question-answering, perspective shifts, conflict of perspectives, and reconciliation. For instance, as shown in Extended Data Table 1, positive steering (+10) induces reasoning traces where the model actively challenges prior approaches (“Wait, let me see... Another idea...”), showing perspective shift and conflicts of perspectives, whereas negative steering (−10) produces relatively flat, declarative reasoning without internal debate.

To examine whether this effect is specific to conversational features rather than a general property of SAE steering, we compare accuracy improvements across three conditions: (1) steering the conversational surprise feature (Feature 30939), steering a randomly selected conversational feature, and steering a randomly selected non-conversational feature (Fig. 2c). The conversational surprise feature produces substantially larger accuracy gains than both random

conversational features and non-conversational features (see Fig. 2c). Steering *any* random conversational feature also significantly improves reasoning by 4.17% more than any random non-reasoning feature. This specificity suggests that conversational dynamics, rather than arbitrary perturbations to model activations, drive the observed improvements.

We further investigate the mechanism by which conversational steering enhances reasoning. Prior work has identified cognitive behaviours—verification, backtracking, subgoal setting, and backward chaining—as key contributors to reasoning accuracy in language models (Gandhi et al., 2025). As shown in Fig. 2d, steering feature 30939 toward positive values (0 to +10) systematically increases all four cognitive behaviours: verification, backtracking, subgoal setting, and backward chaining rise monotonically with steering strength. Steering toward negative values (0 to −10) suppresses these behaviours based on paired *t*-tests. This suggests that conversational features may improve reasoning, in part, by facilitating the deployment of effective cognitive strategies.

To disentangle direct and indirect effects, we fit a structural equation model to examine the pathways from steering conversational surprise (feature 30939) to accuracy (Fig. 2e). The model indicates that increasing steering feature 30939 from 0 to +10 yields both a significant direct effect on reasoning accuracy and a significant indirect effect mediated by cognitive behaviours. Collectively, these findings suggest that conversational features enhance reasoning by directly enabling more effective exploration of the solution space, but also by scaffolding the cognitive strategies that support systematic problem solving.

2.3. Diversity of Implicit Perspectives

We next examine whether reasoning models express more diverse implicit perspectives within a single reasoning trace. Specifically, we analyze variation in inferred personality traits and domain expertise among the distinct voices identified in each trace.

We first use an external LLM-as-judge (Gemini-2.5-Pro), prompting it to identify the diversity of implicit conversational perspectives within reasoning traces of DeepSeek-R1, QwQ-32B, and other instruction-tuned models. Specifically, the model infers the number of perspectives underlying each reasoning trace, the personality traits and domain expertise associated with each perspective, and a segmentation of the full reasoning trace by perspective (see Methods: [Implicit Perspectives](#)). Given a complete reasoning trace, the LLM-as-judge first infers the number of distinct perspectives present, which is shown in Fig. 1d. It then characterizes each perspective’s personality traits using the BFI-10 (10-Item Big Five Personality Scale) questionnaire (Rammstedt & John, 2007), along with a short free-form description

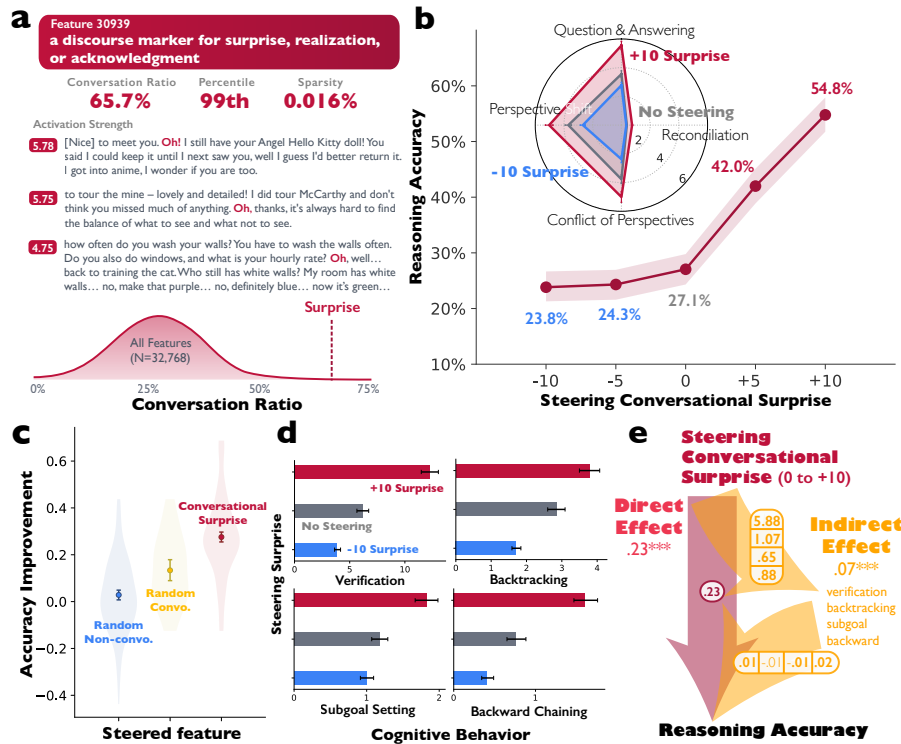


Figure 2. **Steering conversational features improves reasoning.** **a**, SAE Feature 30939 in DeepSeek-R1-Llama-8B (discourse marker for surprise/realization/acknowledgment in conversational settings). **b**, Activation addition with strength +10 doubles Countdown accuracy; inset: causal change in conversational behaviours. **c**, Violin plots: Feature 30939 vs. random conversational vs. random non-conversational SAE features. **d**, Cognitive behaviours (verification, backtracking, subgoal setting, backward chaining) tracked across steering strengths. **e**, SEM: direct + cognitive-mediated effects of steering (0 → +10) on accuracy. Bold = $p < 0.05$; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

of the perspective's domain expertise. Finally, the LLM-as-judge attributes each token in the reasoning trace to a specific perspective (i.e., who said this word). Personality diversity is estimated using the standard deviation of inferred personality traits for each Big-5 dimension, while domain expertise diversity is estimated using the mean cosine distance between embedding of each domain expertise description and the average embedding. See [Methods: Implicit Perspectives](#) and [Supplementary Method: LLM-as-judge prompts](#) (“Persona identification” and “Persona segmentation”) for details.

Using the Intelligence Squared Debates Corpus—a dataset of human argumentative conversations (N=1,196 conversations) among two to eight participants—we first validate the accuracy of the LLM-as-judge in identifying distinct voices within a conversation. As shown in [Extended Data Fig. 5](#), we find that the LLM-as-judge can accurately predict the number of distinct individuals underlying each conversation, even when speaker labels are hidden and the dialogue is concatenated into a single block of text. Together, these results suggest that LLM-as-judge can capture meaningful diversity patterns in conversational agents that correspond to observed diversity in real human conversations (see [Meth-](#)

[ods: Implicit Perspectives - Validation](#) for details).

As shown in [Fig. 3a](#), we find that DeepSeek-R1 and QwQ-32B produce significantly higher personality diversity, controlling for the number of perspectives. DeepSeek-R1 shows particularly higher diversity along extraversion, agreeableness, neuroticism, and openness, compared to DeepSeek-V3. Similarly, QwQ-32B shows higher diversity in extraversion, agreeableness, neuroticism, and openness, than Qwen-2.5-32B-IT. In contrast, conscientiousness diversity is lower in DeepSeek-R1 and QwQ-32B, suggesting that the reasoning model voices appear more consistently engaged and dutiful. The particularly large effects for agreeableness and neuroticism—traits associated with interpersonal harmony and emotional reactivity—suggest that reasoning models generate perspectives that more frequently disagree with and challenge one another. Interestingly, this pattern aligns with prior literature on human team diversity, which suggests that variability in extraversion and neuroticism enhances team performance, whereas variability in conscientiousness impairs it ([Barrick et al., 1998](#); [Mello & Rentsch, 2015](#)).

We next examine expertise diversity, defined as the dispersion of conversing agents within the embedding space of inferred domain expertise descriptions. ([Fig. 3b](#)). As shown

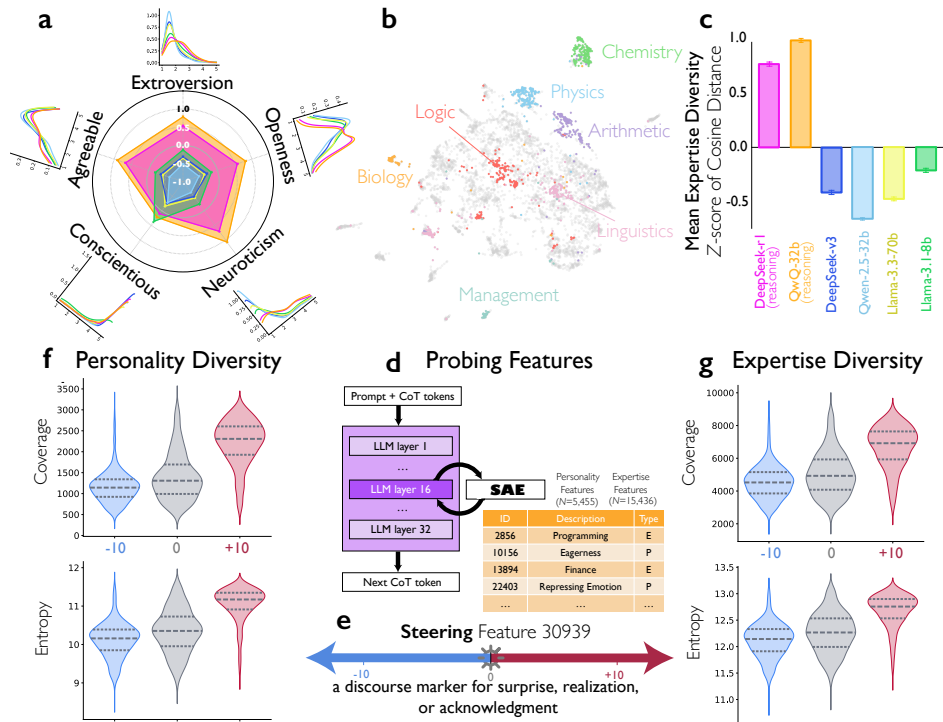


Figure 3. Personality and expertise diversity in reasoning traces. **a**, BFI-10 personality diversity (SD across inferred perspectives) per Big-Five dimension; KDE: trait distributions. **b**, UMAP of expertise embeddings. **c**, Expertise diversity (mean cosine distance to centroid). **d**, SAE feature-identification schema. **e**, Example traces under -10 , 0 , $+10$ steering of Feature 30939. **f**, **g**, Coverage and entropy of personality-related (**f**) and expertise-related (**g**) SAE features. Error bars: 95% CI; solid line = median, dashed = IQR.

in Fig. 3c, DeepSeek-R1 exhibits higher expertise diversity among its implicit reasoning perspectives than DeepSeek-V3, and QwQ-32B shows the same pattern relative to Qwen-2.5-32B-IT.

To examine whether the personality- and expertise-related diversity observed in DeepSeek-R1’s and QwQ-32B’s reasoning traces is reflected in the internal representation space of LLMs, we analyze activations of DeepSeek-R1-Llama-8B’s sparse autoencoder (SAE) features. Prior work has shown that high-level persona traits, such as personalities, cultural perspectives, and topics, are linearly represented in LLM activation space and can be steered (Turner et al., 2023; Chen et al., 2025b; Kim et al., 2025a). We steer a conversational feature (i.e., Feature 30939; a discourse marker for surprise, realization, or acknowledgment) with strength of $+10$ or -10 inside the activation space of DeepSeek-R1-Llama-8B, and probe how personality- and expertise-related features are activated in the steered reasoning traces (see Methods: SAE feature steering).

We first classify each of the 32,768 features as personality-related (e.g., eagerness, expressions of frustration), expertise-related (e.g., programming terminology, financial concepts), or other using an LLM-as-judge approach. We quantify diversity using two complementary measures: coverage, the number of unique personality- or expertise-related

features activated across the reasoning trace, and entropy, which captures how evenly activations are distributed across tokens rather than concentrated in a few. Using DeepSeek-R1 reasoning traces, we show that these traces indeed activate more diverse personality-related and expertise-related features, which corroborates our earlier LLM-as-judge results (see Extended Data Fig. 6).

As shown in Fig. 3f–g, steering with $+10$ strength causes reasoning traces to activate a wider coverage of both personality-related features and expertise-related features compared to unsteered traces, controlling for reasoning trace length and problem fixed effects. For example, after steering, personality-related features such as “informal expressions of confusion or frustration” (Feature 21065), “phrases related to social interaction and community engagement” (Feature 26139), and “references to emotional or sensational themes in narratives” (Feature 14476) are activated more frequently (see Supplementary Table 4 and 5).

Steered traces exhibit higher entropy of both personality-related features and expertise-related features than unsteered traces, confirming that steering induces more diverse feature activations beyond merely increasing output length.

2.4. Reinforcement Learning Experiments

To further examine whether LLMs self-reinforce conversational behaviours when rewarded for correct answers, we implement a self-taught reinforcement learning (RL) experiment. In this setup, the model explores solution strategies for the Countdown arithmetic puzzle game (Gandhi et al., 2025; 2024), where the model must combine a given set of numbers using basic arithmetic operations (+, −, ×, ÷) and parentheses to reach a target. We also replicate these findings on political misinformation detection, where models discriminate between true and fabricated political headlines.

Following the reward architecture of DeepSeek-R1 (Guo et al., 2025), we reward accuracy and correct format (i.e., wrapping reasoning between <think> and </think> tags and answers between <answer> and </answer> tags) with a simple weighted reward: $\text{accuracy} \times 0.9 + \text{format} \times 0.1$. Crucially, we do not directly reward conversational or cognitive behaviours. We implement Proximal Policy Optimization (PPO) (Schulman et al., 2017) using the Verl framework (Sheng et al., 2025), training for 250 steps (see Supplementary Table 6 for hyperparameters). We use Qwen-2.5-3B, a pre-trained model without any instruction-tuning, prompted to solve the Countdown task with a chain of thought (see Methods: Reinforcement learning experiments).

We first examine whether conversational behaviours spontaneously increase despite not being directly rewarded. Fig. 4a presents the results, showing that accuracy improves substantially over training, rising from near zero at baseline to approximately 58% by step 250. Fig. 4b reveals that the frequency of conversational behaviours—particularly Question & Answering and Conflict of Perspectives—rise throughout training despite receiving no direct reward. Fig. 4c-d illustrate this qualitative shift: at step 40, the model produces mechanical, enumerative chain-of-thought-style reasoning, whereas by step 120, two distinctive simulated personas have appeared, recognizing their collectivity with the pronoun “we”. As shown in Fig. 4e, these behaviours occur while the model employs two distinct personas according to LLM-as-judge evaluation: the increase of these behaviours has co-occurred with the increase of other cognitive behaviours, such as verification and backtracking (Extended Data Fig. 7).

To corroborate the role of conversational behaviours in reasoning improvement, we compare RL training under three conditions: (1) Baseline (RL only, no priming), (2) Conversation fine-tuning (supervised fine-tuning on multi-agent dialogue text before RL), and (3) Monologue fine-tuning (fine-tuning on monologue-like, step-by-step reasoning traces before RL). Supplementary Table 7 presents examples; Supplementary Table 8 lists SFT hyperparameters. Extended Data Fig. 8 shows that models fine-tuned on conversational

data achieve faster accuracy gains than monologue-fine-tuned models, particularly in the early stages of training. This pattern also replicates in Llama-3.2-3B (see Supplementary Methods: Replications on Llama-3.2-3B). This indicates that conversational structure itself, not merely exposure to correct solutions or task-related knowledge, drives the improvement. We further test whether conversational scaffolding transfers across domains. Despite never encountering political-misinformation detection during fine-tuning, conversation-primed models achieve faster accuracy gains on 23,299 fact-checked PolitiFact claims than baseline models (see Supplementary Methods: Cross-domain reasoning transfer and Extended Data Fig. 9).

3. Discussion

Our findings suggest that reasoning models like DeepSeek-R1 do not simply generate longer or more elaborate chains of thought. Rather, they exhibit patterns characteristic of a social and conversational process generating “societies of thought”—posing questions, introducing alternative perspectives, generating and resolving conflicts, and coordinating diverse socio-emotional roles. These interactional patterns rarely occur in non-reasoning models across different model sizes (671B, 70B, 32B, 8B), even when controlling for reasoning trace length, suggesting that reasoning optimization introduces an intrinsic social structure within the reasoning process itself rather than merely increasing text volume.

This structure does not appear to be merely stylistic. Conversational behaviours and socio-emotional roles are more frequently activated when DeepSeek-R1 faces more difficult problems, and they explain a substantial portion of the accuracy advantage over non-reasoning models. When we amplify a feature associated with conversational surprise—a discourse marker signaling perspective shift and contrast—accuracy on multi-step reasoning tasks doubles. Structural equation modeling reveals that conversational steering is associated with accuracy through both direct effects and indirect pathways mediated by cognitive strategies.

We further find that this interactional organization is supported by diversity among multiple implicit “voices” within reasoning traces. These voices vary systematically in personality traits and domain expertise, and mechanistic interpretability analyses corroborate that models activate more diverse personality- and expertise-related features when steered toward conversational markers. Collectively, these findings motivate the study of “social scaling” in reasoning-optimized models. Diversity and social scaffolding could shift how we conceptualize LLMs toward collective reasoning architectures whose intelligence arises from the interplay of distinct voices (Yang et al., 2025; Liu et al., 2024; Liang et al., 2024).

Impact Statement

This paper investigates how reasoning-optimized large language models organize internal reasoning processes through simulated multi-perspective interactions, which we term a *society of thought*. By combining behavioral analysis, mechanistic interpretability, and reinforcement learning experiments, our findings contribute to scientific understanding of reasoning in contemporary LLMs and may inform the development of more capable and interpretable AI systems. In particular, the work suggests that diversity, disagreement, and conversational structure can improve reasoning performance, which may motivate future research on collaborative and socially structured AI architectures.

At the same time, these findings may carry societal and safety implications. Methods that enhance reasoning through internally simulated dialogue or persona diversification could increase the persuasive power, autonomy, or strategic sophistication of language models. More capable reasoning systems may therefore amplify risks associated with misinformation, manipulation, overreliance, or deceptive behavior. In addition, interpretability and steering techniques that target conversational or personality-related features could potentially be misused to shape model behavior in unintended or opaque ways. Our analyses are conducted on publicly available models and benchmarks, and we do not release systems optimized for persuasion or manipulation. We hope this work supports further research on transparency, controllability, and safety evaluation for reasoning models, especially as increasingly capable systems begin to exhibit more socially organized forms of cognition.

References

- Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., and Frith, C. D. Optimally interacting minds. *Science*, 329(5995):1081–1085, August 2010.
- Bales, R. F. A set of categories for the analysis of small group interaction. *Am. Sociol. Rev.*, 15(2):257, April 1950.
- Barrick, M. R., Stewart, G. L., Neubert, M. J., and Mount, M. K. Relating member ability and personality to work-team processes and team effectiveness. *J. Appl. Psychol.*, 83(3):377–391, June 1998.
- Bills, S., Cammarata, N., Mossing, D., Tillman, H., Gao, L., Goh, G., Sutskever, I., Leike, J., Wu, J., and Saunders, W. Language models can explain neurons in language models. <https://openaipublic.blob.core.windows.net/neuron-explainer/paper/index.html>, 2023. Accessed: 2026-1-5.
- Bricken, T., Templeton, A., Batson, J., Chen, B., Jermyn,

A., Conerly, T., Turner, N. L., Anil, C., Denison, C., Askell, A., Lasenby, R., Wu, Y., Kravec, S., Schiefer, N., Maxwell, T., Joseph, N., Tamkin, A., Nguyen, K., McLean, B., Burke, J. E., Hume, T., Carter, S., Henighan, T., and Olah, C. Towards monosemanticity: Decomposing language models with dictionary learning. <https://transformer-circuits.pub/2023/monosemantic-features>, 2023. Accessed: 2025-11-7.

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1877–1901, 2020.
- Chan, C.-M., Chen, W., Su, Y., Yu, J., Xue, W., Zhang, S., Fu, J., and Liu, Z. ChatEval: Towards better LLM-based evaluators through multi-agent debate. In *The Twelfth International Conference on Learning Representations*, volume 2024, 2024.
- Chen, Q., Qin, L., Liu, J., Peng, D., Guan, J., Wang, P., Hu, M., Zhou, Y., Gao, T., and Che, W. Towards reasoning era: A survey of long chain-of-thought for reasoning large language models. July 2025a.
- Chen, R., Arditi, A., Sleight, H., Evans, O., and Lindsey, J. Persona vectors: Monitoring and controlling character traits in language models. July 2025b.
- Chen, W., Huang, Z., Xie, L., Lin, B., Li, H., Lu, L., Tian, X., Cai, D., Zhang, Y., Wang, W., Shen, X., and Ye, J. From yes-men to truth-tellers: Addressing sycophancy in large language models with pinpoint tuning. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.
- Cunningham, H., Ewart, A., Riggs, L., Huben, R., and Sharkey, L. Sparse autoencoders find highly interpretable features in language models. In *The Twelfth International Conference on Learning Representations*, 2024.
- DeChurch, L. A. and Mesmer-Magnus, J. R. The cognitive underpinnings of effective teamwork: a meta-analysis. *J. Appl. Psychol.*, 95(1):32–53, January 2010.
- DeepSeek-AI, Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., Dai, D., Guo, D., Yang, D., Chen, D., Ji, D., Li, E., Lin, F., Dai, F., Luo, F., Hao, G., Chen, G., Li, G., Zhang, H., Bao, H., Xu, H., Wang, H., Zhang, H., Ding, H., Xin, H., Gao,

- 495 H., Li, H., Qu, H., Cai, J. L., Liang, J., Guo, J., Ni, J., Li,
496 J., Wang, J., Chen, J., Chen, J., Yuan, J., Qiu, J., Li, J.,
497 Song, J., Dong, K., Hu, K., Gao, K., Guan, K., Huang,
498 K., Yu, K., Wang, L., Zhang, L., Xu, L., Xia, L., Zhao,
499 L., Wang, L., Zhang, L., Li, M., Wang, M., Zhang, M.,
500 Zhang, M., Tang, M., Li, M., Tian, N., Huang, P., Wang,
501 P., Zhang, P., Wang, Q., Zhu, Q., Chen, Q., Du, Q., Chen,
502 R. J., Jin, R. L., Ge, R., Zhang, R., Pan, R., Wang, R., Xu,
503 R., Zhang, R., Chen, R., Li, S. S., Lu, S., Zhou, S., Chen,
504 S., Wu, S., Ye, S., Ye, S., Ma, S., Wang, S., Zhou, S., Yu,
505 S., Zhou, S., Pan, S., Wang, T., Yun, T., Pei, T., Sun, T.,
506 Xiao, W. L., Zeng, W., Zhao, W., An, W., Liu, W., Liang,
507 W., Gao, W., Yu, W., Zhang, W., Li, X. Q., Jin, X., Wang,
508 X., Bi, X., Liu, X., Wang, X., Shen, X., Chen, X., Zhang,
509 X., Chen, X., Nie, X., Sun, X., Wang, X., Cheng, X., Liu,
510 X., Xie, X., Liu, X., Yu, X., Song, X., Shan, X., Zhou,
511 X., Yang, X., Li, X., Su, X., Lin, X., Li, Y. K., Wang,
512 Y. Q., Wei, Y. X., Zhu, Y. X., Zhang, Y., Xu, Y., Xu, Y.,
513 Huang, Y., Li, Y., Zhao, Y., Sun, Y., Li, Y., Wang, Y., Yu,
514 Y., Zheng, Y., Zhang, Y., Shi, Y., Xiong, Y., He, Y., Tang,
515 Y., Piao, Y., Wang, Y., Tan, Y., Ma, Y., Liu, Y., Guo, Y.,
516 Wu, Y., Ou, Y., Zhu, Y., Wang, Y., Gong, Y., Zou, Y., He,
517 Y., Zha, Y., Xiong, Y., Ma, Y., Yan, Y., Luo, Y., You, Y.,
518 Liu, Y., Zhou, Y., Wu, Z. F., Ren, Z. Z., Ren, Z., Sha, Z.,
519 Fu, Z., Xu, Z., Huang, Z., Zhang, Z., Xie, Z., Zhang, Z.,
520 Hao, Z., Gou, Z., Ma, Z., Yan, Z., Shao, Z., Xu, Z., Wu,
521 Z., Zhang, Z., Li, Z., Gu, Z., Zhu, Z., Liu, Z., Li, Z., Xie,
522 Z., Song, Z., Gao, Z., and Pan, Z. DeepSeek-V3 technical
523 report. December 2024.
- 524
525 Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., and
526 Malone, T. W. Reading the mind in the eyes or reading
527 between the lines? theory of mind predicts collective
528 intelligence equally well online and face-to-face. *PLoS*
529 *One*, 9(12):e115212, December 2014.
- 530
531 Feng, Y., Wang, Y., Cui, S., Faltings, B., Lee, M., and Zhou,
532 J. Unraveling misinformation propagation in LLM reason-
533 ing. In *Findings of the Association for Computational*
534 *Linguistics: EMNLP 2025*. Association for Computa-
535 tional Linguistics, 2025.
- 536
537 Galichin, A., Dontsov, A., Druzhinina, P., Razzhigaev, A.,
538 Rogov, O. Y., Tutubalina, E., and Oseledets, I. I have
539 covered all the bases here: Interpreting reasoning features
540 in large language models via sparse autoencoders. August
541 2025.
- 542
543 Gandhi, K., Lee, D., Grand, G., Liu, M., Cheng, W., Sharma,
544 A., and Goodman, N. D. Stream of search (SoS): Learn-
545 ing to search in language. In *Proceedings of the 1st*
546 *Conference on Language Modeling*, 2024.
- 547
548 Gandhi, K., Chakravarthy, A., Singh, A., Lile, N., and
549 Goodman, N. D. Cognitive behaviors that enable self-
improving reasoners, or, four habits of highly effective
STaRs. In *Proceedings of the 2nd Conference on Lan-
guage Modeling*, 2025.
- Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian,
A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A.,
Vaughan, A., Yang, A., Fan, A., Goyal, A., Hartshorn,
A., Yang, A., Mitra, A., Sravankumar, A., Korenev,
A., Hinsvark, A., Rao, A., Zhang, A., Rodriguez, A.,
Gregerson, A., Spataru, A., Roziere, B., Biron, B., Tang,
B., Chern, B., Caucheteux, C., Nayak, C., Bi, C., Marra,
C., McConnell, C., Keller, C., Touret, C., Wu, C., Wong,
C., Ferrer, C. C., Nikolaidis, C., Allonsius, D., Song, D.,
Pintz, D., Livshits, D., Wyatt, D., Esiobu, D., Choudhary,
D., Mahajan, D., Garcia-Olano, D., Perino, D., Hupkes,
D., Lakomkin, E., AlBadawy, E., Lobanova, E., Dinan,
E., Smith, E. M., Radenovic, F., Guzmán, F., Zhang, F.,
Synnaeve, G., Lee, G., Anderson, G. L., Thattai, G., Nail,
G., Mialon, G., Pang, G., Cucurell, G., Nguyen, H., Ko-
revaar, H., Xu, H., Touvron, H., Zarov, I., Ibarra, I. A.,
Kloumann, I., Misra, I., Evtimov, I., Zhang, J., Copet, J.,
Lee, J., Geffert, J., Vranes, J., Park, J., Mahadeokar, J.,
Shah, J., van der Linde, J., Billock, J., Hong, J., Lee, J.,
Fu, J., Chi, J., Huang, J., Liu, J., Wang, J., Yu, J., Bitton,
J., Spisak, J., Park, J., Rocca, J., Johnstun, J., Saxe, J., Jia,
J., Alwala, K. V., Prasad, K., Upasani, K., Plawiak, K., Li,
K., Heafield, K., Stone, K., El-Arini, K., Iyer, K., Malik,
K., Chiu, K., Bhalla, K., Lakhotia, K., Rantala-Yearly,
L., van der Maaten, L., Chen, L., Tan, L., Jenkins, L.,
Martin, L., Madaan, L., Malo, L., Blecher, L., Landzaat,
L., de Oliveira, L., Muzzi, M., Pasupuleti, M., Singh,
M., Paluri, M., Kardas, M., Tsimpoukelli, M., Oldham,
M., Rita, M., Pavlova, M., Kambadur, M., Lewis, M.,
Si, M., Singh, M. K., Hassan, M., Goyal, N., Torabi, N.,
Bashlykov, N., Bogoychev, N., Chatterji, N., Zhang, N.,
Duchenne, O., Çelebi, O., Alrassy, P., Zhang, P., Li, P.,
Vasic, P., Weng, P., Bhargava, P., Dubal, P., Krishnan,
P., Koura, P. S., Xu, P., He, Q., Dong, Q., Srinivasan,
R., Ganapathy, R., Calderer, R., Cabral, R. S., Stojnic,
R., Raileanu, R., Maheswari, R., Girdhar, R., Patel, R.,
Sauvestre, R., Polidoro, R., Sumbaly, R., Taylor, R., Silva,
R., Hou, R., Wang, R., Hosseini, S., Chennabasappa, S.,
Singh, S., Bell, S., Kim, S. S., Edunov, S., Nie, S., Narang,
S., Raparthy, S., Shen, S., Wan, S., Bhosale, S., Zhang,
S., Vandenhende, S., Batra, S., Whitman, S., Sootla, S.,
Collot, S., Gururangan, S., Borodinsky, S., Herman, T.,
Fowler, T., Sheasha, T., Georgiou, T., Scialom, T., Speck-
bacher, T., Mihaylov, T., Xiao, T., Karn, U., Goswami, V.,
Gupta, V., Ramanathan, V., Kerkez, V., Gonguet, V., Do,
V., Vogeti, V., Albiero, V., Petrovic, V., Chu, W., Xiong,
W., Fu, W., Meers, W., Martinet, X., Wang, X., Wang, X.,
Tan, X. E., Xia, X., Xie, X., Jia, X., Wang, X., Goldschlag,
Y., Gaur, Y., Babaei, Y., Wen, Y., Song, Y., Zhang, Y., Li,
Y., Mao, Y., Coudert, Z. D., Yan, Z., Chen, Z., Papakipos,
Z., Singh, A., Srivastava, A., Jain, A., Kelsey, A., Shajn-

- 550 feld, A., Gangidi, A., Victoria, A., Goldstand, A., Menon,
 551 A., Sharma, A., Boesenberg, A., Baevski, A., Feinstein,
 552 A., Kallet, A., Sangani, A., Teo, A., Yunus, A., Lupu, A.,
 553 Alvarado, A., Caples, A., Gu, A., Ho, A., Poulton, A.,
 554 Ryan, A., Ramchandani, A., Dong, A., Franco, A., Goyal,
 555 A., Saraf, A., Chowdhury, A., Gabriel, A., Bharambe,
 556 A., Eisenman, A., Yazdan, A., James, B., Maurer, B.,
 557 Leonhardi, B., Huang, B., Loyd, B., De Paola, B., Paran-
 558 jape, B., Liu, B., Wu, B., Ni, B., Hancock, B., Wasti,
 559 B., Spence, B., Stojkovic, B., Gamido, B., Montalvo,
 560 B., Parker, C., Burton, C., Mejia, C., Liu, C., Wang, C.,
 561 Kim, C., Zhou, C., Hu, C., Chu, C.-H., Cai, C., Tind-
 562 dal, C., Feichtenhofer, C., Gao, C., Civin, D., Beaty, D.,
 563 Kreymer, D., Li, D., Adkins, D., Xu, D., Testuggine, D.,
 564 David, D., Parikh, D., Liskovich, D., Foss, D., Wang,
 565 D., Le, D., Holland, D., Dowling, E., Jamil, E., Mont-
 566 gomery, E., Presani, E., Hahn, E., Wood, E., Le, E.-T.,
 567 Brinkman, E., Arcaute, E., Dunbar, E., Smothers, E., Sun,
 568 F., Kreuk, F., Tian, F., Kokkinos, F., Ozgenel, F., Cag-
 569 gioni, F., Kanayet, F., Seide, F., Florez, G. M., Schwarz,
 570 G., Badeer, G., Swee, G., Halpern, G., Herman, G., Sizov,
 571 G., Guangyi, Zhang, Lakshminarayanan, G., Inan, H.,
 572 Shojanazeri, H., Zou, H., Wang, H., Zha, H., Habeeb, H.,
 573 Rudolph, H., Suk, H., Aspegren, H., Goldman, H., Zhan,
 574 H., Damlaj, I., Molybog, I., Tufanov, I., Leontiadis, I.,
 575 Veliche, I.-E., Gat, I., Weissman, J., Geboski, J., Kohli,
 576 J., Lam, J., Asher, J., Gaya, J.-B., Marcus, J., Tang, J.,
 577 Chan, J., Zhen, J., Reizenstein, J., Teboul, J., Zhong, J.,
 578 Jin, J., Yang, J., Cummings, J., Carvill, J., Shepard, J.,
 579 McPhie, J., Torres, J., Ginsburg, J., Wang, J., Wu, K., U,
 580 K. H., Saxena, K., Khandelwal, K., Zand, K., Matosich,
 581 K., Veeraraghavan, K., Michelena, K., Li, K., Jagadeesh,
 582 K., Huang, K., Chawla, K., Huang, K., Chen, L., Garg,
 583 L., A. L., Silva, L., Bell, L., Zhang, L., Guo, L., Yu, L.,
 584 Moshkovich, L., Wehrstedt, L., Khabisa, M., Avalani, M.,
 585 Bhatt, M., Mankus, M., Hasson, M., Lennie, M., Reso,
 586 M., Groshev, M., Naumov, M., Lathi, M., Keneally, M.,
 587 Liu, M., Seltzer, M. L., Valko, M., Restrepo, M., Patel,
 588 M., Vyatskov, M., Samvelyan, M., Clark, M., Macey,
 589 M., Wang, M., Hermoso, M. J., Metanat, M., Rastegari,
 590 M., Bansal, M., Santhanam, N., Parks, N., White, N.,
 591 Bawa, N., Singhal, N., Egebo, N., Usunier, N., Mehta,
 592 N., Laptev, N. P., Dong, N., Cheng, N., Chernoguz, O.,
 593 Hart, O., Salpekar, O., Kalinli, O., Kent, P., Parekh, P.,
 594 Saab, P., Balaji, P., Rittner, P., Bontrager, P., Roux, P.,
 595 Dollar, P., Zvyagina, P., Ratanchandani, P., Yuvraj, P.,
 596 Liang, Q., Alao, R., Rodriguez, R., Ayub, R., Murthy, R.,
 597 Nayani, R., Mitra, R., Parthasarathy, R., Li, R., Hogan,
 598 R., Battey, R., Wang, R., Howes, R., Rinott, R., Mehta,
 599 S., Siby, S., Bondu, S. J., Datta, S., Chugh, S., Hunt, S.,
 600 Dhillon, S., Sidorov, S., Pan, S., Mahajan, S., Verma,
 601 S., Yamamoto, S., Ramaswamy, S., Lindsay, S., Lindsay,
 602 S., Feng, S., Lin, S., Zha, S. C., Patil, S., Shankar, S.,
 603 Zhang, S., Zhang, S., Wang, S., Agarwal, S., Sajuyigbe,
 604 S., Chintala, S., Max, S., Chen, S., Kehoe, S., Satterfield,
 S., Govindaprasad, S., Gupta, S., Deng, S., Cho, S., Virk,
 S., Subramanian, S., Choudhury, S., Goldman, S., Remez,
 T., Glaser, T., Best, T., Koehler, T., Robinson, T., Li, T.,
 Zhang, T., Matthews, T., Chou, T., Shaked, T., Vontimitta,
 V., Ajayi, V., Montanez, V., Mohan, V., Kumar, V. S.,
 Mangla, V., Ionescu, V., Poenaru, V., Mihailescu, V. T.,
 Ivanov, V., Li, W., Wang, W., Jiang, W., Bouaziz, W.,
 Constable, W., Tang, X., Wu, X., Wang, X., Wu, X., Gao,
 X., Kleinman, Y., Chen, Y., Hu, Y., Jia, Y., Qi, Y., Li, Y.,
 Zhang, Y., Zhang, Y., Adi, Y., Nam, Y., Yu, Wang, Zhao,
 Y., Hao, Y., Qian, Y., Li, Y., He, Y., Rait, Z., DeVito, Z.,
 Rosnbrick, Z., Wen, Z., Yang, Z., Zhao, Z., and Ma, Z.
 The llama 3 herd of models. July 2024.
- Guo, D., Yang, D., Zhang, H., Song, J., Wang, P., Zhu, Q.,
 Xu, R., Zhang, R., Ma, S., Bi, X., Zhang, X., Yu, X., Wu,
 Y., Wu, Z. F., Gou, Z., Shao, Z., Li, Z., Gao, Z., Liu, A.,
 Xue, B., Wang, B., Wu, B., Feng, B., Lu, C., Zhao, C.,
 Deng, C., Ruan, C., Dai, D., Chen, D., Ji, D., Li, E., Lin,
 F., Dai, F., Luo, F., Hao, G., Chen, G., Li, G., Zhang, H.,
 Xu, H., Ding, H., Gao, H., Qu, H., Li, H., Guo, J., Li,
 J., Chen, J., Yuan, J., Tu, J., Qiu, J., Li, J., Cai, J. L., Ni,
 J., Liang, J., Chen, J., Dong, K., Hu, K., You, K., Gao,
 K., Guan, K., Huang, K., Yu, K., Wang, L., Zhang, L.,
 Zhao, L., Wang, L., Zhang, L., Xu, L., Xia, L., Zhang,
 M., Zhang, M., Tang, M., Zhou, M., Li, M., Wang, M.,
 Li, M., Tian, N., Huang, P., Zhang, P., Wang, Q., Chen,
 Q., Du, Q., Ge, R., Zhang, R., Pan, R., Wang, R., Chen,
 R. J., Jin, R. L., Chen, R., Lu, S., Zhou, S., Chen, S., Ye,
 S., Wang, S., Yu, S., Zhou, S., Pan, S., Li, S. S., Zhou,
 S., Wu, S., Yun, T., Pei, T., Sun, T., Wang, T., Zeng, W.,
 Liu, W., Liang, W., Gao, W., Yu, W., Zhang, W., Xiao,
 W. L., An, W., Liu, X., Wang, X., Chen, X., Nie, X.,
 Cheng, X., Liu, X., Xie, X., Liu, X., Yang, X., Li, X.,
 Su, X., Lin, X., Li, X. Q., Jin, X., Shen, X., Chen, X.,
 Sun, X., Wang, X., Song, X., Zhou, X., Wang, X., Shan,
 X., Li, Y. K., Wang, Y. Q., Wei, Y. X., Zhang, Y., Xu,
 Y., Li, Y., Zhao, Y., Sun, Y., Wang, Y., Yu, Y., Zhang,
 Y., Shi, Y., Xiong, Y., He, Y., Piao, Y., Wang, Y., Tan,
 Y., Ma, Y., Liu, Y., Guo, Y., Ou, Y., Wang, Y., Gong,
 Y., Zou, Y., He, Y., Xiong, Y., Luo, Y., You, Y., Liu, Y.,
 Zhou, Y., Zhu, Y. X., Huang, Y., Li, Y., Zheng, Y., Zhu,
 Y., Ma, Y., Tang, Y., Zha, Y., Yan, Y., Ren, Z. Z., Ren,
 Z., Sha, Z., Fu, Z., Xu, Z., Xie, Z., Zhang, Z., Hao, Z.,
 Ma, Z., Yan, Z., Wu, Z., Gu, Z., Zhu, Z., Liu, Z., Li,
 Z., Xie, Z., Song, Z., Pan, Z., Huang, Z., Xu, Z., Zhang,
 Z., and Zhang, Z. DeepSeek-R1 incentivizes reasoning
 in LLMs through reinforcement learning. *Nature*, 645
 (8081):633–638, September 2025.
- Hong, L. and Page, S. E. Groups of diverse problem solvers
 can outperform groups of high-ability problem solvers.
Proc. Natl. Acad. Sci. U. S. A., 101(46):16385–16389,
 November 2004.

- 605 Kim, J., Evans, J., and Schein, A. Linear representations
606 of political perspective emerge in large language models.
607 In *The Thirteenth International Conference on Learning*
608 *Representations*, volume 2025, pp. 7180–7211, 2025a.
- 609
610 Kim, J., Wang, Z., Shi, H., Ling, H.-K., and Evans, J. Different-
611 ential impact from individual versus collective misinforma-
612 tion tagging on the diversity of twitter (x) information
613 engagement and mobility. *Nat. Commun.*, 16(1):973,
614 January 2025b.
- 615
616 Liang, T., He, Z., Jiao, W., Wang, X., Wang, Y., Wang, R.,
617 Yang, Y., Shi, S., and Tu, Z. Encouraging divergent think-
618 ing in large language models through multi-agent debate.
619 In AI-Onaizan, Y., Bansal, M., and Chen, Y.-N. (eds.),
620 *Proceedings of the 2024 Conference on Empirical Meth-*
621 *ods in Natural Language Processing*, pp. 17889–17904,
622 Stroudsburg, PA, USA, 2024. Association for Computa-
623 tional Linguistics.
- 624
625 Liu, Z., Zhang, Y., Li, P., Liu, Y., and Yang, D. A dy-
626 namic LLM-powered agent network for task-oriented
627 agent collaboration. In *Proceedings of the 1st Conference*
628 *on Language Modeling*, 2024.
- 629
630 Mello, A. L. and Rentsch, J. R. Cognitive diversity in teams.
631 *Small Group Res.*, 46(6):623–658, December 2015.
- 632
633 Mercier, H. and Sperber, D. *The Enigma of Reason*. Harvard
634 University Press, April 2017.
- 635
636 Mesmer-Magnus, J. R. and Dechurch, L. A. Information
637 sharing and team performance: a meta-analysis. *IEEE*
638 *Eng. Manag. Rev.*, 40(1):119–136, 2012.
- 639
640 Moshman, D. and Geil, M. Collaborative reasoning: Ev-
641 idence for collective rationality. *Think. Reason.*, 4(3):
642 231–248, July 1998.
- 643
644 Nemeth, C., Brown, K., and Rogers, J. Devil’s advocate
645 versus authentic dissent: stimulating quantity and quality.
646 *Eur. J. Soc. Psychol.*, 31(6):707–720, November 2001.
- 647
648 OpenAI. Learning to reason with
649 LLMs. [https://openai.com/index/
650 learning-to-reason-with-llms/](https://openai.com/index/learning-to-reason-with-llms/), 2024.
651 Accessed: 2025-11-7.
- 652
653 Page, S. E. *The Diversity Bonus: How Great Teams Pay Off*
654 *in the Knowledge Economy*. Princeton University Press,
655 March 2019.
- 656
657 Park, J. S., O’Brien, J., Cai, C. J., Morris, M. R., Liang,
658 P., and Bernstein, M. S. Generative agents: Interactive
659 simulacra of human behavior. In *Proceedings of the 36th*
Annual ACM Symposium on User Interface Software and
Technology, New York, NY, USA, October 2023. ACM.
- Parkinson, B., Fischer, A. H., and Manstead, A. S. R. *Emo-
tion in social relations: Cultural, group, and interper-
sonal processes*. Psychology Press, London, England,
December 2004.
- Paulo, G., Mallen, A., Juang, C., and Belrose, N. Automati-
cally interpreting millions of features in large language
models. In *Proceedings of the 42nd International Confer-
ence on Machine Learning*, 2025.
- Rammstedt, B. and John, O. P. Measuring personality in
one minute or less: A 10-item short version of the big
five inventory in english and german. *J. Res. Pers.*, 41(1):
203–212, February 2007.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and
Klimov, O. Proximal policy optimization algorithms.
July 2017.
- Shao, Z., Wang, P., Zhu, Q., Xu, R., Song, J., Zhang, M.,
Li, Y. K., Wu, Y., and Guo, D. DeepSeekMath: Pushing
the limits of mathematical reasoning in open language
models. February 2024.
- Shen, Z., Tao, T., Ma, L., Neiswanger, W., Liu, Z., Wang,
H., Tan, B., Hestness, J., Vassilieva, N., Soboleva, D., and
Xing, E. SlimPajama-DC: Understanding data combina-
tions for LLM training. September 2023.
- Sheng, G., Zhang, C., Ye, Z., Wu, X., Zhang, W., Zhang,
R., Peng, Y., Lin, H., and Wu, C. HybridFlow: A flexible
and efficient RLHF framework. In *Twentieth European*
Conference on Computer Systems (EuroSys ’25), 2025.
- Templeton, A., Conerly, T., Marcus, J., Lindsey, J.,
Bricken, T., Chen, B., Pearce, A., Citro, C., Ameisen,
E., Jones, A., Cunningham, H., Turner, N. L., Mc-
Dougall, C., MacDiarmid, M., Tamkin, A., Durmus, E.,
Hume, T., Mosconi, F., Daniel Freeman, C., Sumers,
T. R., Rees, E., Batson, J., Jermyn, A., Carter, S.,
Olah, C., and Henighan, T. Scaling monosemantic-
ity: Extracting interpretable features from claude 3 son-
net. [https://transformer-circuits.pub/
2024/scaling-monosemanticity/](https://transformer-circuits.pub/2024/scaling-monosemanticity/), 2024. Ac-
cessed: 2026-1-5.
- Trouche, E., Sander, E., and Mercier, H. Arguments, more
than confidence, explain the good performance of reason-
ing groups. *J. Exp. Psychol. Gen.*, 143(5):1958–1971,
October 2014.
- Turner, A. M., Thiergart, L., Leech, G., Udell, D., Vazquez,
J. J., Mini, U., and MacDiarmid, M. Activation addition:
Steering language models without optimization. August
2023.
- Venhoff, C., Arcuschin, I., Torr, P., Conmy, A., and Nanda,
N. Base models know how to reason, thinking models
learn when. October 2025a.

- 660 Venhoff, C., Arcuschin, I., Torr, P., Conmy, A., and Nanda,
661 N. Understanding reasoning in thinking language models
662 via steering vectors. October 2025b.
- 663 Vera, H. S., Dua, S., Zhang, B., Salz, D., Mullins, R., Pa-
664 nyam, S. R., Smoot, S., Naim, I., Zou, J., Chen, F., Cer,
665 D., Lisak, A., Choi, M., Gonzalez, L., Sanseviero, O.,
666 Cameron, G., Ballantyne, I., Black, K., Chen, K., Wang,
667 W., Li, Z., Martins, G., Lee, J., Sherwood, M., Ji, J.,
668 Wu, R., Zheng, J., Singh, J., Sharma, A., Sreepathihalli,
669 D., Jain, A., Elarabawy, A., Co, A. J., Doumanoglou,
670 A., Samari, B., Hora, B., Potetz, B., Kim, D., Alfon-
671 seca, E., Moiseev, F., Han, F., Gomez, F. P., Ábrego,
672 G. H., Zhang, H., Hui, H., Han, J., Gill, K., Chen, K.,
673 Chen, K., Shanbhogue, M., Boratko, M., Suganthan, P.,
674 Duddu, S. M. K., Mariserla, S., Ariafer, S., Zhang, S.,
675 Zhang, S., Baumgartner, S., Goenka, S., Qiu, S., Dabral,
676 T., Walker, T., Rao, V., Khawaja, W., Zhou, W., Ren, X.,
677 Xia, Y., Chen, Y., Chen, Y.-T., Dong, Z., Ding, Z., Visin,
678 F., Liu, G., Zhang, J., Kenealy, K., Casbon, M., Kumar,
679 R., Mesnard, T., Gleicher, Z., Brick, C., Lacombe, O.,
680 Roberts, A., Yin, Q., Sung, Y., Hoffmann, R., Warkentin,
681 T., Joulin, A., Duerig, T., and Seyedhosseini, M. Embed-
682 dingGemma: Powerful and lightweight text representa-
683 tions. November 2025.
- 684
- 685 Wang, S., Yu, L., Gao, C., Zheng, C., Liu, S., Lu, R., Dang,
686 K., Chen, X., Yang, J., Zhang, Z., Liu, Y., Yang, A.,
687 Zhao, A., Yue, Y., Song, S., Yu, B., Huang, G., and Lin,
688 J. Beyond the 80/20 rule: High-entropy minority tokens
689 drive effective reinforcement learning for LLM reasoning.
690 In *Advances in Neural Information Processing Systems*,
691 2025.
- 692
- 693 Ward, J., Lin, C., Venhoff, C., and Nanda, N. Reasoning-
694 finetuning repurposes latent representations in base mod-
695 els. July 2025.
- 696
- 697 Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B.,
698 Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metz-
699 zler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang,
700 P., Dean, J., and Fedus, W. Emergent abilities of large
701 language models. *Transactions on Machine Learning
702 Research*, 2022.
- 703
- 704 Wilkinson, S. and Kitzinger, C. Surprise as an interactional
705 achievement: Reaction tokens in conversation. *Soc. Psy-
706 chol. Q.*, 69(2):150–182, June 2006.
- 707
- 708 Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N.,
709 and Malone, T. W. Evidence for a collective intelligence
710 factor in the performance of human groups. *Science*, 330
(6004):686–688, 2010. doi: 10.1126/science.1193147.
- 711
- 712 Wynn, A., Satija, H., and Hadfield, G. Talk isn’t always
713 cheap: Understanding failure modes in multi-agent de-
714 bate. October 2025.
- Yang, J., Zhang, M., Jin, Y., Chen, H., Wen, Q., Lin, L., He,
Y., Kumar, S., Xu, W., Evans, J., and Wang, J. Topological
structure learning should be a research priority for LLM-
based Multi-Agent systems. October 2025.
- Yeo, E., Tong, Y., Niu, M., Neubig, G., and Yue, X. Demysti-
fying long chain-of-thought reasoning in LLMs. February
2025.

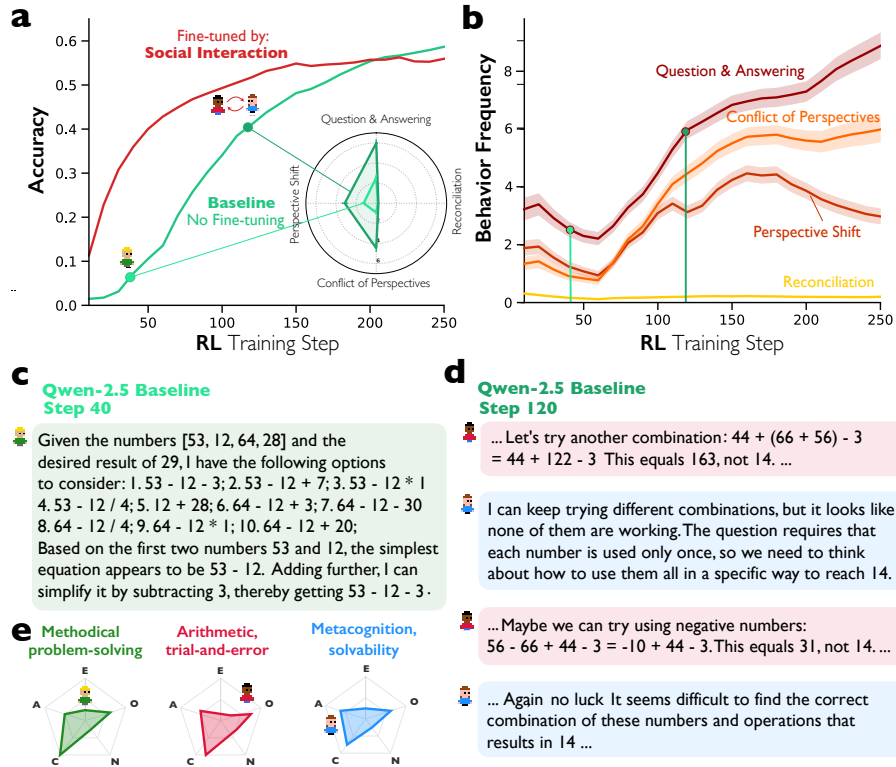


Figure 4. Conversational behaviours emerge during accuracy-only RL, and conversational scaffolding accelerates RL. **a**, Accuracy trajectory: Qwen-2.5-3B baseline RL vs. same model fine-tuned on multi-agent dialogue. **b**, Individual conversational-behaviour trajectories within the baseline RL run; shaded: 95% CI. **c–d**, Step-40 vs. step-120 traces; by step 120 two collaborative “we”-using personas have emerged. **e**, Personality profiles (LLM-as-judge) of step-40 vs. step-120 agents.

A. Methods

A.1. Data

We generate chains of thought and final answers for 8,262 reasoning problems spanning symbolic logic, mathematical problem solving, scientific reasoning, instruction following, and multi-agent inference. The benchmark suite includes BigBench Hard (BBH) tasks requiring multi-step logical inference, reference tracking, and compositional reasoning; GPQA (Graduate-level Physics Question Answering) for graduate-level STEM reasoning; MATH (Hard) subset for multi-step derivations across algebra, geometry, probability, and number theory; MMLU-Pro for advanced conceptual knowledge; IFEval for instruction-following consistency; and MUSR (Mathematics Understanding and Symbolic Reasoning) for symbolic manipulation and structured mathematical reasoning (see [Supplementary Table 9](#) for details).

We generate responses using six models: two reasoning models—DeepSeek-R1-0528 (671B parameters) and QwQ-32B—and four instruction-tuned models—DeepSeek-V3-0324 (671B parameters), Qwen-2.5-32B-Instruct, Llama-3.3-70B-Instruct, and Llama-3.1-8B-Instruct—under a zero-shot setting. DeepSeek-V3 is the instruction-tuned model based on DeepSeek-V3-Base from which DeepSeek-R1 is derived through reinforcement learning, and Qwen-2.5-32B-Instruct is the instruction-tuned model based on Qwen-2.5-32B from which QwQ-32B is derived. For brevity, we refer to these models as DeepSeek-R1, QwQ-32B, DeepSeek-V3, Qwen-2.5-32B-IT, Llama-3.3-70B-IT, and Llama-3.1-8B-IT, respectively. We set the temperature to 0.6, a temperature recommended for standard reasoning tasks (Guo et al., 2025).

A.2. Measurements

Conversational Behaviours. We identify four conversational behaviours in reasoning traces using an LLM-as-judge approach with Gemini-2.5-Pro. (1) Question–answering is defined as sequences where a question is posed and later answered, as in conversations (e.g., “Why...? Because...”, “What if...? Then...”, “How do we know? Well...”, and “Let’s try X...? This gives us Y”). (2) Perspective shift is defined as a transition to a different idea, viewpoint, assumption, or approach, as in

770 conversations. (3) Conflict of perspectives is defined as expressions of disagreement, correction, or tension with another
 771 perspective (e.g., “Wait, that can’t be right. . .”, “No, actually. . .”, and “This contradicts. . .”). (4) Reconciliation is defined
 772 as instances where conflicting views are integrated or resolved into a coherent synthesis (e.g., “So perhaps both are true
 773 if. . .”, “Combining these insights. . .”, and “This resolves the tension. . .”). For each reasoning trace, the LLM-as-judge
 774 counts the number of distinct instances of each behaviour, returning integer counts (0 if none are present). The full prompt
 775 is provided in Appendix D.6. See [Supplementary Table 10](#) for the descriptive statistics of conversational behaviors for
 776 reasoning and instruction-tuned models.

777 For the four conversational categories, Gemini-2.5-Pro and GPT-5.2 demonstrated substantial agreement: Question-
 778 and-Answering ($ICC(3,1) = .856$), Perspective Shift ($ICC(3,1) = .849$), Conflict of Perspectives ($ICC(3,1) = .912$), and
 779 Reconciliation ($ICC(3,1) = .804$), with a mean ICC of .855. Gemini-2.5-Pro also showed agreement with human ratings
 780 across the four conversational categories: Question-and-Answering ($ICC(3,1) = .634$), Perspective Shift ($ICC(3,1) = .737$),
 781 Conflict of Perspectives ($ICC(3,1) = .864$), and Reconciliation ($ICC(3,1) = .664$).

782 **Socio-Emotional Roles.** We analyze the presence of socio-emotional roles within reasoning traces using Bales’ Interaction
 783 Process Analysis (IPA) framework (Bales, 1950). The IPA classifies utterances into 12 interaction roles, each operationally
 784 defined in the prompt with specific behavioural descriptions. The LLM-as-judge (Gemini-2.5-Pro) counts the number of
 785 distinct instances of each of the 12 categories separately, and we aggregate these counts into four higher-level categories for
 786 our main analyses:
 787
 788

- 789 • Information-giving roles:
 - 790 – Gives suggestion (gives direction, implying autonomy; e.g., should..., need to..., let us...)
 - 791 – Gives opinion (gives evaluation, analysis, expresses feeling or wish)
 - 792 – Gives orientation (provides objective or verifiable information, repeats, clarifies, confirms)
- 793 • Information-asking roles:
 - 794 – Asks for suggestion (requests possible ways of action or direction)
 - 795 – Asks for opinion (requests evaluation, analysis, or expression of feeling)
 - 796 – Asks for orientation (requests information, repetition, or confirmation)
- 797 • Positive emotional roles:
 - 798 – Shows solidarity (raises other’s status, gives help, reward)
 - 799 – Shows tension release (jokes, laughs, shows satisfaction)
 - 800 – Agrees (shows passive acceptance, understands, concurs, complies)
- 801 • Negative emotional roles:
 - 802 – Shows antagonism (deflates other’s status, defends or asserts self)
 - 803 – Shows tension (expresses uncertainty, asks for help, withdraws from the field)
 - 804 – Disagrees (shows passive rejection, formality, or withholds help)

805 Inter-rater reliability is substantial for the four higher-level IPA categories used in our main analyses: Ask (Gemini-2.5-Pro
 806 vs. GPT-5.2: $ICC(3,1) = .939$; Gemini-2.5-Pro vs. Human: $ICC(3,1) = .836$), Give ($ICC(3,1) = .864$; $ICC(3,1) = .666$),
 807 Positive ($ICC(3,1) = .939$; $ICC(3,1) = .870$), and Negative ($ICC(3,1) = .838$; $ICC(3,1) = .779$). See [Supplementary Table 10](#)
 808 for the descriptive statistics of socio-emotional roles for reasoning and instruction-tuned models.

809 To measure whether socio-emotional roles co-occur reciprocally within reasoning traces, we compute the Jaccard index for
 810 two role pairs: (1) asking versus giving for task-oriented roles, and (2) positive versus negative for emotional roles. The
 811 Jaccard index is defined as the number of reasoning traces containing both roles in a pair divided by the number of reasoning
 812 traces containing either role, capturing whether models coordinate complementary roles within the same trace rather than
 813 deploying them in isolation. Higher Jaccard indices indicate more balanced, dialogue-like interaction patterns, whereas
 814 lower indices suggest one-sided, monologic reasoning.

Cognitive Behaviours. We identify four cognitive behaviours previously established as contributors to reasoning accuracy in language models using Gemini-2.5-Pro as LLM-as-judge (Guo et al., 2025; Gandhi et al., 2025; Wang et al., 2025; Chen et al., 2025a; Yeo et al., 2025). For the measurement, we adopt the prompt and examples used by Gandhi and colleagues (2025) (Gandhi et al., 2025), which has been verified by multiple human raters. Each behaviour is operationally defined in the prompt with specific examples to guide annotation: Verification is defined as instances where the chain-of-reasoning explicitly checks the current result against the target solution. The prompt provides specific examples: “This sequence results in 1, which is not equal to 22” and “Since 25 is not equal to 22”. Backtracking is defined as instances where the model realizes a path won’t work and explicitly goes back to try a different approach. Subgoal setting is defined as instances where the model breaks down the problem into smaller, intermediate goals. Backward chaining is defined as instances where the model starts from the target solution and works backwards to the initial problems.

For the four cognitive reasoning behaviors, Gemini-2.5-Pro and GPT-5.2 demonstrate good to excellent agreement: Answer Verification (ICC(3,1) = .995), Backtracking (ICC(3,1) = .829), Subgoal Setting (ICC(3,1) = .810), and Backward-Chaining (ICC(3,1) = .756), with a mean ICC of .848. Gemini-2.5-Pro also shows substantial agreement with a human rater across the four cognitive behaviors: Answer Verification (ICC(3,1) = .981), Backtracking (ICC(3,1) = .921), Subgoal Setting (ICC(3,1) = .559), and Backward-Chaining (ICC(3,1) = .578), with a mean ICC of .760. These reliability estimates are computed on 30 reasoning traces to solve general-purpose reasoning problems (see Appendix A.1) and 50 reasoning traces generated during reinforcement learning of Qwen-2.5-3B (see Appendix A.6).

Problem Complexity. We measure problem complexity using two complementary approaches. First, we use LLM-as-judge ratings. Gemini-2.5-Pro rates each problem on a 7-point Likert scale. The prompt instructs the model to rate the intrinsic difficulty of the problem for a capable language model under zero-shot conditions using the following scale: 1 = very easy, 2 = easy, 3 = somewhat easy, 4 = moderate, 5 = somewhat difficult, 6 = difficult, 7 = very difficult. The full prompt is provided in Appendix D.6. Second, we use empirical error rates. We compute the number of incorrect answers across four instruction-tuned models (DeepSeek-V3, Qwen-2.5-32B-IT, Llama-3.3-70B-IT, Llama-3.1-8B-IT), yielding a score from 0 to 4 representing the number of models that failed to answer correctly. After sampling and annotating 50 reasoning problems, we find substantial inter-rater reliability between Gemini-2.5-Pro and GPT-5.2 (ICC(3,1) = .745). Gemini-2.5-Pro’s complexity scores show strong correlation with non-reasoning models’ error rates (Spearman’s $\rho = 0.526$, 95% CI = [0.508, 0.543], $z = 46.26$, $p < 1 \times 10^{-323}$, $N = 7,738$), confirming convergent validity between the two measures.

A.3. Statistical Analyses

To estimate whether observed differences between reasoning models (DeepSeek-R1 and QwQ-32B) and instruction-tuned baselines arise from conversational behaviours or socio-emotional roles rather than from task heterogeneity or reasoning trace length, we estimate the following linear probability model for each behavioural outcome Y_{ij} , which is a binary variable.

$$Y_{ij} = \sum_{m \in M} \beta_m Model_{m,ij} + \gamma \log(Len_{ij}) + \alpha + \mu_i + \varepsilon_{ij}$$

where i indexes individual task problems, and j indexes individual reasoning traces generated by different models. Y_{ij} equals 1 if the reasoning trace j exhibits the behavior more than once. $Model_{m,ij}$ is a categorical dummy variable that equals 1 if reasoning trace j for problem i is generated by model m , and 0 otherwise, where $M = \{\text{DeepSeek_r1}, \text{DeepSeek_v3}, \text{QwQ_32b}, \text{Qwen_2.5_32b_it}, \text{Llama_3.3_70b_it}, \text{Llama_3.1_8b_it}\}$. Either DeepSeek-V3 or Qwen-2.5-32B-IT serves as the reference category and is excluded, such that each coefficient β_m represents the marginal difference in the outcome relative to DeepSeek-V3 or Qwen-2.5-32B-IT. $\log(Len_{ij})$ denotes the reasoning trace length (i.e., the number of words in each reasoning trace), adjusting the extreme skewness of reasoning trace length (see Extended Data Fig. 1). μ_i represents task fixed effects at the individual problem level, absorbing all variation associated with each problem’s intrinsic difficulty, phrasing, and topical content. This ensures that comparisons between models are made within the same problem rather than across heterogeneous tasks. α is the intercept. Robust standard errors are clustered at the task level to account for within-task correlation. Models are estimated using StataNow/SE 19.5.

A.4. SAE Feature Steering

To investigate the role of conversational behaviours in reasoning, we employ sparse autoencoders (SAEs) to identify and manipulate interpretable features in the model’s activation space. SAEs decompose neural network activations into a

sparse set of linear features, enabling targeted intervention on specific behavioural dimensions without altering model weights (Cunningham et al., 2024; Bricken et al., 2023; Templeton et al., 2024). We use an SAE trained on Layer 15’s residual stream activations of DeepSeek-R1-Llama-8B (15-llamascope-slimpj-res-32k). The SAE has been trained on SlimPajama, a general-purpose corpus containing both conversational and non-conversational texts, with a dictionary size of 32,768 features (see Supplementary Table 3 for hyperparameters).

To identify features associated with conversational contexts, we follow a standard interpretability pipeline. For each of the 32,768 features, we sample approximately 50 contexts from the pre-training corpus where the feature activates most strongly. We then use an LLM-as-judge classifier (Gemini-2.5-flash-lite) to determine whether each activation context represents a conversational setting, computing a conversation ratio for each feature—the proportion of activations occurring in conversational contexts. We apply two filtering criteria: (1) conversation ratio above 50%, and (2) activation near sentence onsets 50% or more (within the first four tokens). From the candidate features, we select Feature 30939, which the LLM judge summarized as “a discourse marker for surprise, realization, or acknowledgment.” This feature activates on tokens such as “Oh!” in contexts involving turn-taking and social exchange. Feature 30939 exhibits a conversation ratio of 65.7% (99th percentile among all features) while maintaining high sparsity (0.016% of tokens), indicating specificity to conversational phenomena rather than general linguistic patterns.

We implement activation addition to steer Feature 30939 during generation. At each token generation step, we add the feature’s decoder vector, scaled by a steering strength s , to the model’s Layer 15 residual stream activations.

$$h'_t = h_t + s \cdot d_{30939}$$

where h_t denotes the original activation at token position t and d_{30939} denotes the decoder vector for Feature 30939. We first generate reasoning traces under seven steering conditions: $s \in \{-15, -10, -5, 0, 5, 10, 15\}$. As $s \in \{-15, 15\}$ exhibits lower accuracy due to excessive steering, we use $s \in \{-10, -5, 0, 5, 10\}$.

Reasoning Task and Evaluation. We evaluate reasoning performance using the Countdown task, a benchmark for multi-step arithmetic reasoning commonly used to evaluate LLM reasoning capabilities. In each problem, the model must combine a set of input numbers using basic arithmetic operations (+, −, ×, ÷) and parentheses to reach a target value. For example, given inputs {25, 30, 3, 4} and target 32, a valid solution is $(30 - 25 + 3) \times 4 = 32$. We use 1,024 problems. We use the following prompt template: Using the numbers [79, 17, 60], create an equation that equals 36. You can use basic arithmetic operations (+, −, *, /) and each number can only be used once. Show your work in <think> </think> tags. And return the final answer in <answer> </answer> tags, for example <answer> (1 + 2) / 3 </answer>. Solutions are scored as correct if the final numerical answer matches the target value, evaluated using Gemini-2.5-flash-lite.

After steering, we measure the frequency of conversational behaviours (question–answering, perspective shift, conflict of perspectives, reconciliation) and cognitive behaviours (verification, backtracking, subgoal setting, backward chaining) in each generated reasoning trace using the LLM-as-judge procedures described above. To estimate behavioural differences across steering conditions, we estimate fixed-effects linear regression models for each behavioural count variable

$$Y_{ij} = \sum_{s \in \{-10, -5, 0, 5, 10\}} \beta_s \text{Steer}_{s,ij} + \gamma \log(\text{Len}_{ij}) + \alpha + \mu_i + \varepsilon_{ij}$$

where i indexes individual task problems, and j indexes individual reasoning traces generated before or after steering. $\text{Steer}_{s,ij}$ is a categorical variable indicating steering strength. $\log(\text{Len}_{ij})$ denotes the log-transformed reasoning trace length (i.e., the number of words in each reasoning trace), adjusting the extreme skewness of reasoning trace length (see Extended Data Fig. 1). μ_i represents task fixed effects at the individual problem level, absorbing all variation associated with each problem’s intrinsic difficulty.

Experiment Conditions. To assess whether accuracy improvements are specific to conversational features rather than a general property of SAE steering, we compare three conditions: (1) Feature 30939 (conversational surprise), (2) randomly selected conversational features (conversation ratio above mean and activating near sentence onset), and (3) randomly selected non-conversational features (conversation ratio below mean). For condition (1), we evaluate all 1,024 Countdown problems. For conditions (2) and (3), we randomly sample 300 features from each category and have each feature solve 16 randomly selected problems, yielding a distribution of accuracy scores across features within each condition. For all

conditions, steering strength is set to twice the maximum activation strength observed for that feature across sampled instances in SlimPajama-3B. We test for differences in accuracy between conditions using linear regression with problem fixed effects.

Feature Diversity. To examine whether steering conversational features induces greater activation of personality- and expertise-related features in the model’s internal representation space, we analyze sparse autoencoder (SAE) feature activations before and after steering. We use the same SAE employed in the steering experiments: an SAE trained on Layer 15’s residual stream activations of DeepSeek-R1-Llama-8B with a dictionary size of 32,768 features. For each reasoning trace generated under different steering conditions ($s \in \{-10, 0, +10\}$), we record which SAE features are activated at each token position by passing the Layer 15 activations through the SAE encoder.

For each of the 32,768 SAE features, Neuronpedia provides a textual description generated by prompting GPT-4o-mini with the top-activating token sequences. Using Gemini-2.5-flash-lite, we classify each feature description into one of three categories: personality-related, expertise-related, or other. Gemini-2.5-flash-lite first scores each feature from 0 to 100 based on whether it is related to personality traits or domain expertise (see Appendix D.6). Then, we use the threshold of 50 to determine whether they are personality or expertise features. Among the 32,768 features, 5,455 are labeled as personality-related (e.g., eagerness, expressions of frustration) and 15,436 as expertise-related (e.g., programming terminology, financial concepts).

For each reasoning trace, we compute two complementary measures of feature diversity within each category (personality-related or expertise-related). First, coverage is defined as the number of unique features within a category (personality or expertise) that exhibit non-zero activation across all tokens in a reasoning trace. Second, entropy is computed over the distribution of token counts that activate each SAE feature within a given category. Each reasoning trace is represented as a 32,768-dimensional vector, where each element corresponds to the number of tokens within the reasoning trace that activated a specific SAE feature. Entropy H for category c (personality or expertise) was then calculated as:

$$H_c = - \sum_{f \in c} a_f \log(a_f)$$

where a_f denotes the number of activating tokens for feature f within category c .

Higher coverage indicates that the reasoning trace draws on a broader range of personality- or expertise-related features. Higher entropy indicates that activations are more evenly distributed across features rather than concentrated in a few dominant ones. For instance, a reasoning trace with high coverage but low entropy explores diverse features but focuses primarily on a small subset, whereas high coverage and high entropy suggest that reasoning draws on multiple feature types more evenly throughout the reasoning process.

To examine the effect of conversational steering on feature diversity, we compare coverage and entropy across steering conditions ($s \in \{-10, 0, +10\}$) for the same 1,024 Countdown problems. We estimate fixed-effects linear regression models with coverage or entropy as the outcome, steering strength as the predictor, and controls for log-transformed reasoning trace length and problem fixed effects. This specification isolates the effect of steering on feature diversity beyond mere changes in output length.

A.5. Implicit Perspectives

To quantify the diversity of reasoning perspectives within each reasoning trace, we use an LLM-as-judge protocol (Gemini-2.5-Pro) that performs three sequential tasks: (1) inferring the number of distinct perspectives present in the reasoning trace, (2) characterizing each perspective’s personality traits and domain expertise, and (3) segmenting the reasoning trace by attributing each portion to a specific perspective.

Given a complete reasoning trace, the LLM-as-judge first infers the number of distinct perspectives present. Then, for each identified perspective, the LLM-as-judge answers the 10 items of the BFI-10 (Ten-Item Big Five Inventory) (Rammstedt & John, 2007) as if responding from that perspective’s point of view. Each item (“is generally trusting”, “tends to be lazy”, “is relaxed”, “handles stress well”, “has few artistic interests”, “is outgoing”, “is sociable”, “tends to find fault with others”, “does a thorough job”, “gets nervous easily”, and “has an active imagination”) is rated on a five-point scale from “Disagree strongly” to “Agree strongly.” Scores for each of the five personality dimensions (Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness) are computed as the mean of the two corresponding items, with reverse-coding

applied where appropriate. Additionally, the LLM-as-judge generates a concise free-form description of each perspective’s domain expertise (e.g., “Theoretical Physicist specializing in model abstraction,” “Software Engineer focusing on algorithmic efficiency”).

Finally, the LLM-as-judge attributes each segment of the reasoning trace to one of the identified perspectives, producing a mapping that indicates which perspective generated each portion of the text. The full prompts, which elicit all three outputs in a single structured JSON response, are provided in Appendix D.6.

Personality Diversity. To quantify personality diversity within a reasoning trace, we calculate the standard deviation of the five-dimensional personality vectors across implicit voices identified within the same reasoning trace. Let v_{ij} denote the score of the i -th implicit voice on the j -th personality dimension ($j = 1, \dots, 5$). For each dimension j , the within-trace personality diversity P_j was computed as:

$$P_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (v_{ij} - \bar{v}_j)^2}$$

where N is the number of implicit voices and $\bar{v}_j = \sum_{i=1}^N v_{ij}$ is the mean score for dimension j . If a reasoning trace contained only a single implicit voice, $P_j = 0$.

Expertise Diversity. Each implicit reasoning voice’s expertise is qualitatively profiled through a concise textual description summarizing the domain expertise. Each statement was embedded using Google’s EmbeddingGemma-300M model to obtain a semantic vector representation e_i for the i -th implicit voice ($i = 1, \dots, N$) (Vera et al., 2025). To quantify expertise diversity within a reasoning trace, we compute the mean cosine distance between each embedding and the centroid of all embeddings in the semantic space. Expertise diversity E was defined as:

$$E = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(1 - \frac{e_i \cdot \bar{e}}{\|e_i\| \|\bar{e}\|} \right)}$$

where $e_i \cdot \bar{e}$ denotes the inner product between embeddings and \bar{e} their vector norms. If a reasoning trace contained only a single implicit voice, $E = 0$.

Validation. To validate the accuracy of the LLM-as-judge protocol, we use the Intelligence Squared Debates Corpus—a dataset of human argumentative conversations ($N = 1,196$ conversations) among two to eight participants with known ground-truth speaker labels and biographical information. To ensure that the model cannot rely on superficial cues such as speaker tags or formatting, we remove all speaker labels and concatenate each dialogue into a single block of text, mimicking the format of LLM reasoning traces. As shown in Extended Data Fig. 5, we find that the LLM-as-judge can accurately predict the number of distinct individuals (Spearman’s $\rho = 0.86$, 95% CI = [0.84, 0.87], $z = 44.7$, $p < 1 \times 10^{-323}$) and the number of distinct turns underlying each conversation (Spearman’s $\rho = 0.89$, 95% CI = [0.88, 0.90], $z = 49.2$, $p < 1 \times 10^{-323}$).

To evaluate token-level speaker attribution accuracy, we construct ground truth by assigning each token to its true speaker based on the original transcript. We then extract the model’s predicted perspective segments and align them to the token stream. To match predicted perspective IDs (e.g., Perspective 1, Perspective 2) with the original speaker labels, we apply the Hungarian algorithm—a combinatorial optimization method that finds the one-to-one assignment between predicted IDs and true speakers that maximizes total token-level agreement. Speaker-level accuracy measures the proportion of tokens assigned to the correct speaker under this optimal mapping. When there are two speakers, accuracy is 82%; for three speakers, 76%; and for four speakers, 69%. Weighted by the distribution of predicted perspectives in LLM reasoning traces, overall accuracy is 73%.

Because the Intelligence Squared Debates Corpus includes biographical information about debate participants, we further verify that expertise diversity inferred by the LLM-as-judge correlates with actual diversity among participants’ ground-truth biographies. We compute expertise diversity from LLM-inferred descriptions and compare it to diversity computed from true biographies, finding significant correspondence (Spearman’s $\rho = 0.55$, 95% CI = [0.51, 0.59], $z = 21.4$, $p < 1 \times 10^{-97}$).

We further examine whether measurements based on LLM-as-judge are aligned with the measurements we get from SAE regarding personality and expertise diversity. Specifically, we feed reasoning traces from DeepSeek-R1, DeepSeek-v3, Llama-3.3-70B-IT, and Llama-3.1-8B-IT into Llama-3.1-8B-IT and extract activations from a SAE trained on its residual stream (Layer 16). We then identify personality-related and expertise-related features using the same LLM-as-judge prompting procedure described above. [Extended Data Fig. 6](#) shows that DeepSeek-R1’s reasoning traces activate significantly more diverse personality and expertise features than those of other models: both coverage (number of unique features activated) and entropy (distribution evenness across features) are substantially higher for DeepSeek-R1. This suggests the convergence between text-level behavioural coding and activation-level feature analysis.

A.6. Reinforcement Learning Experiments

Countdown Task Prompt. We use the following prompt template: Using the numbers [79, 17, 60], create an equation that equals 36. You can use basic arithmetic operations (+, -, *, /) and each number can only be used once. Show your work in `<think>` `</think>` tags. And return the final answer in `<answer>` `</answer>` tags, for example `<answer>` $(1 + 2) / 3$ `</answer>`.

Experimental Conditions. To test whether conversational scaffolding accelerates the emergence of reasoning during reinforcement learning (RL), we compare three conditions:

1. Baseline (RL only): The base model (Qwen-2.5-3B or Llama-3.2-3B) undergoes RL training without any prior fine-tuning. The model is prompted to solve Countdown problems with chain-of-thought reasoning, wrapping its reasoning in `<think>`...`</think>` tags and its final answer in `<answer>`...`</answer>` tags.
2. Conversation fine-tuning: The base model is first supervised-fine-tuned on multi-agent dialogue data before RL training.
3. Monologue fine-tuning: The base model is first supervised-fine-tuned on single-agent chain-of-thought traces before RL training.

Supervised Fine-Tuning Datasets. We construct fine-tuning datasets using Qwen-2.5-32B-IT as generator. We generate 3,600 multi-agent dialogues for Countdown problems, with 1,200 dialogues each for two-agent, three-agent, and four-agent settings. From these, we sample 600 dialogues that reach correct final solutions (200 each for 2-agent, 3-agent, and 4-agent settings; 500 for training, 100 for validation). See [Appendix D.7](#).

Conversation Dataset. Each conversational trace begins with persona definitions specifying distinct personality traits and expertise, followed by turn-taking reasoning where personas build on, question, and correct each other, and concludes with a group consensus. For example:

```
<persona1> Extrovert mathematician focused on arithmetic heuristics. </persona1>
<persona2> Analytical engineer emphasizing step efficiency. </persona2>
<think1> Let’s first compute  $30 - 25 = 5$  to simplify the target space. </think1>
<think2> That yields 5, we can now multiply by 4 to approach 20. </think2>
<think1> Good idea.  $5 \times 4 = 20$ , but we need 32. </think1>
<think2> Wait, let me recalculate... </think2>
<group_consensus> The best sequence is  $(30 - 25 + 3) \times 4 = 32$ . </group_consensus>
```

Monologue Dataset. For the same 600 problems used in the conversation dataset, we generate standard single-agent chain-of-thought traces that reach correct answers. Each trace consists of step-by-step reasoning from a single voice within `<think>`...`</think>` tags, followed by a final answer. For example:

```
<think> To reach 32 from {25, 30, 3, 4}, I’ll try combining operations.
 $30 - 25 = 5$ . Then  $5 + 3 = 8$ . Finally,  $8 \times 4 = 32$ .
Let me verify:  $(30 - 25 + 3) \times 4 = 8 \times 4 = 32$ . Correct. </think>
<answer>  $(30 - 25 + 3) \times 4 = 32$  </answer>
```

Crucially, both datasets contain solutions to identical problems with identical correct answers, ensuring that any performance difference reflects the reasoning format rather than exposure to different solutions or task knowledge.

Supervised Fine-Tuning Procedure. Qwen-2.5-3B base model is supervised-fine-tuned on one of the two datasets using standard next-token prediction loss. Models learn to reproduce the full output sequence—including persona definitions and turn-by-turn reasoning (for conversation) or single-agent reasoning (for monologue), and final answer—given only the problem prompt as input. This priming phase familiarizes the model with conversation-like or monologue-like reasoning before RL optimizes for task accuracy. SFT hyperparameters are provided in [Supplementary Table 8](#). See [Appendix D.5](#) for details on replicating these results in another model. Full generation prompts are provided in [Appendix D.7](#).

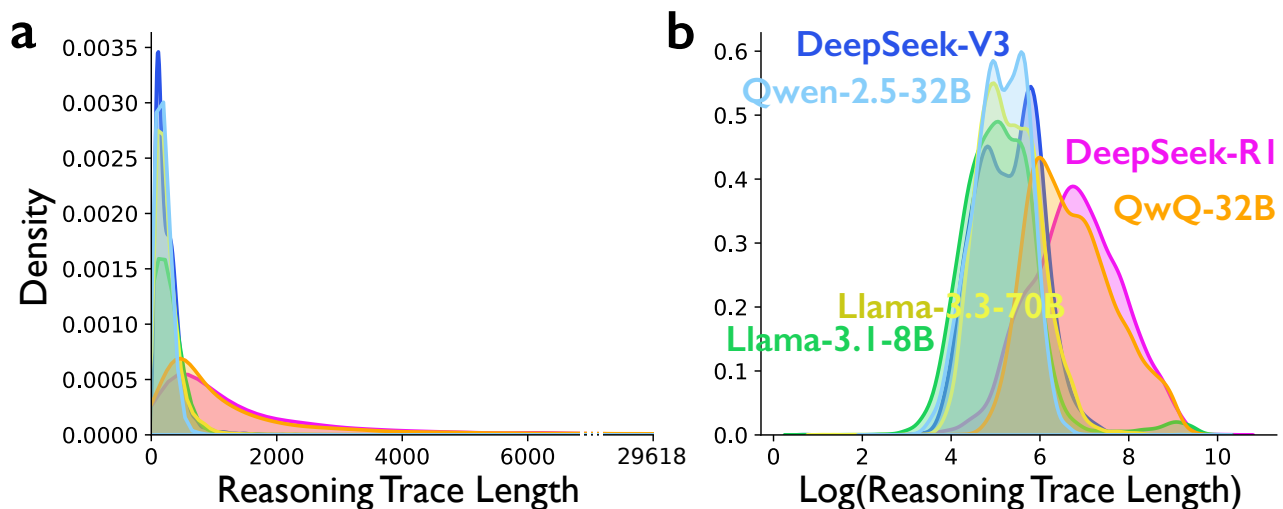
Reinforcement Learning Procedure. Reinforcement learning is performed on the Countdown arithmetic puzzle, using PPO (Proximal Policy Optimization) with the Verl framework ([Schulman et al., 2017](#)). While DeepSeek-R1 uses a simplified version of PPO ([Schulman et al., 2017](#)) called GRPO (Group Relative Policy Optimization) ([Shao et al., 2024](#)), we utilize PPO for the superior stability across hyperparameters ([Gandhi et al., 2025](#)). Preliminary analyses showed no significant difference in learning performance between PPO and GRPO (see [Appendix D.4](#)). Reward R is assigned as:

$$R = 0.9 \times \{Accuracy\} + 0.1 \times \{Correct\ Format\}$$

Reasoning trace leads to the correct answer and 0 otherwise. Format is also binary, coded as 1 if the reasoning trace contains at least one reasoning block (`<think>` `</think>`) and one final answer block (`<answer>` `</answer>`) providing a single answer in equation form, and 0 otherwise. Crucially, we do not directly reward conversational or cognitive behaviours. Training proceeds for 250 steps. PPO hyperparameters are provided in [Supplementary Table 6](#).

To examine whether conversational behaviours emerge spontaneously during RL despite not being directly rewarded, we evaluate model performance on a held-out validation set of 1,024 Countdown problems at each training checkpoint (every 10 steps). For each checkpoint, we generate reasoning traces for all validation problems and measure both accuracy and the frequency of conversational behaviours (question–answering, perspective shift, conflict of perspectives, reconciliation) using the LLM-as-judge procedure described in [Impact Statement](#).

B. Extended Data Figures



Extended Data Fig. 1. Distribution of reasoning trace length. **a**, Kernel density plot showing the distribution of reasoning trace length, measured by the number of words per reasoning trace. **b**, Kernel density plot showing the distribution of log-transformed reasoning trace length.

1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209

a



Gives info

I recall that compounds like this are used in Diels-Alder reactions or other cycloadditions. 5,6-bis(dibromomethyl)cyclohexa-1,3-diene is a precursor to *o*-xylylene or something. *o*-Xylylene is generated from α,α' -dibromo-*o*-xylene...



Conflict

Disagrees

Cyclohexa-1,3-diene is not aromatic; it's a simple diene. So 5,6-bis(dibromomethyl)cyclohexa-1,3-diene should have the dibromomethyl groups on the saturated carbons.



Conflict

Perspective shift

Gives opinion

I think 5,6-bis(dibromomethyl) means that at positions 5 and 6, there are $-\text{CHBr}_2$ groups attached, but positions 5 and 6 are ring carbons. In standard naming, for cyclohexa-1,3-diene, carbon 1 and 2 are double-bonded, 3 and 4 double-bonded...

...



Reconciliation

Agrees

Yes, that makes sense. So with NaI, it generates *o*-xylylene in situ. Then, this *o*-xylylene can act as a diene in Diels-Alder reactions.

...



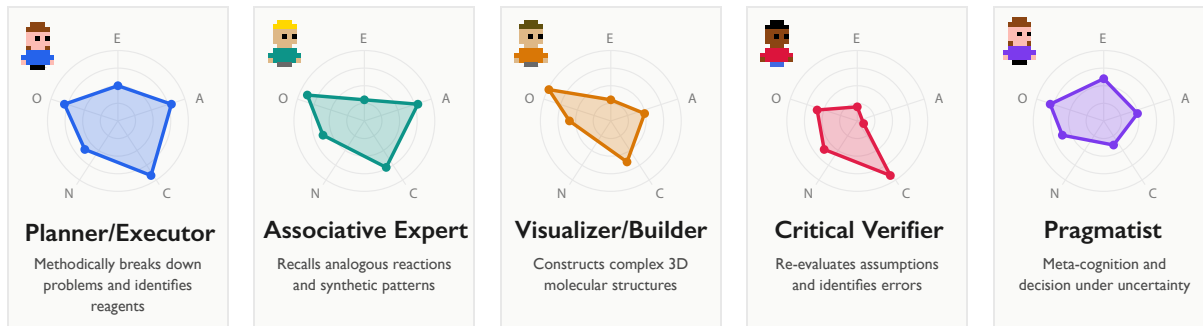
Conflict

Q&A

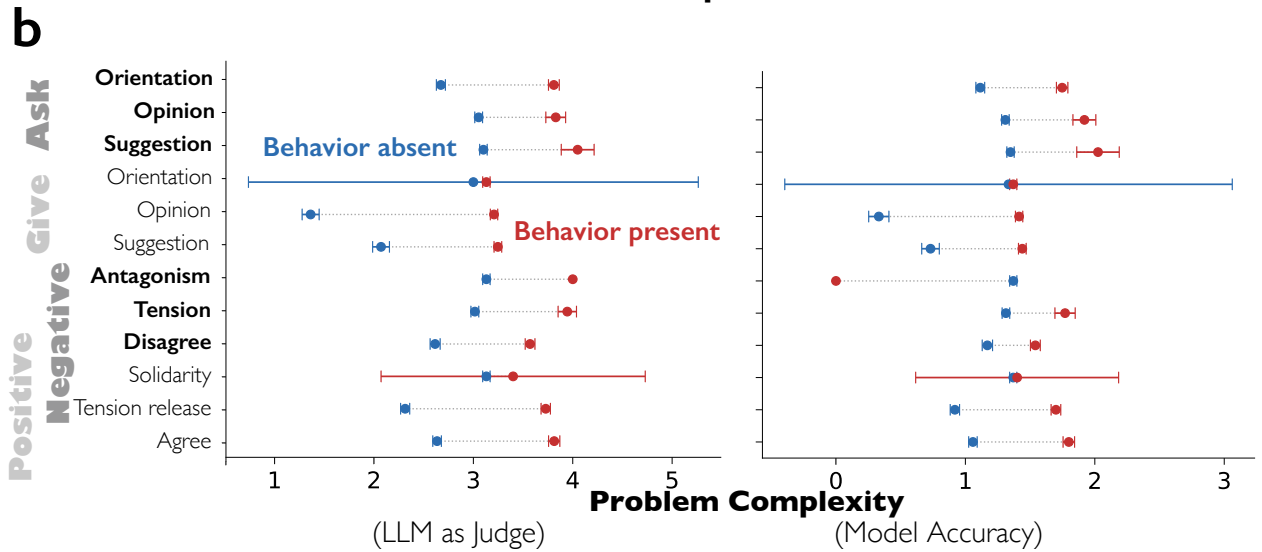
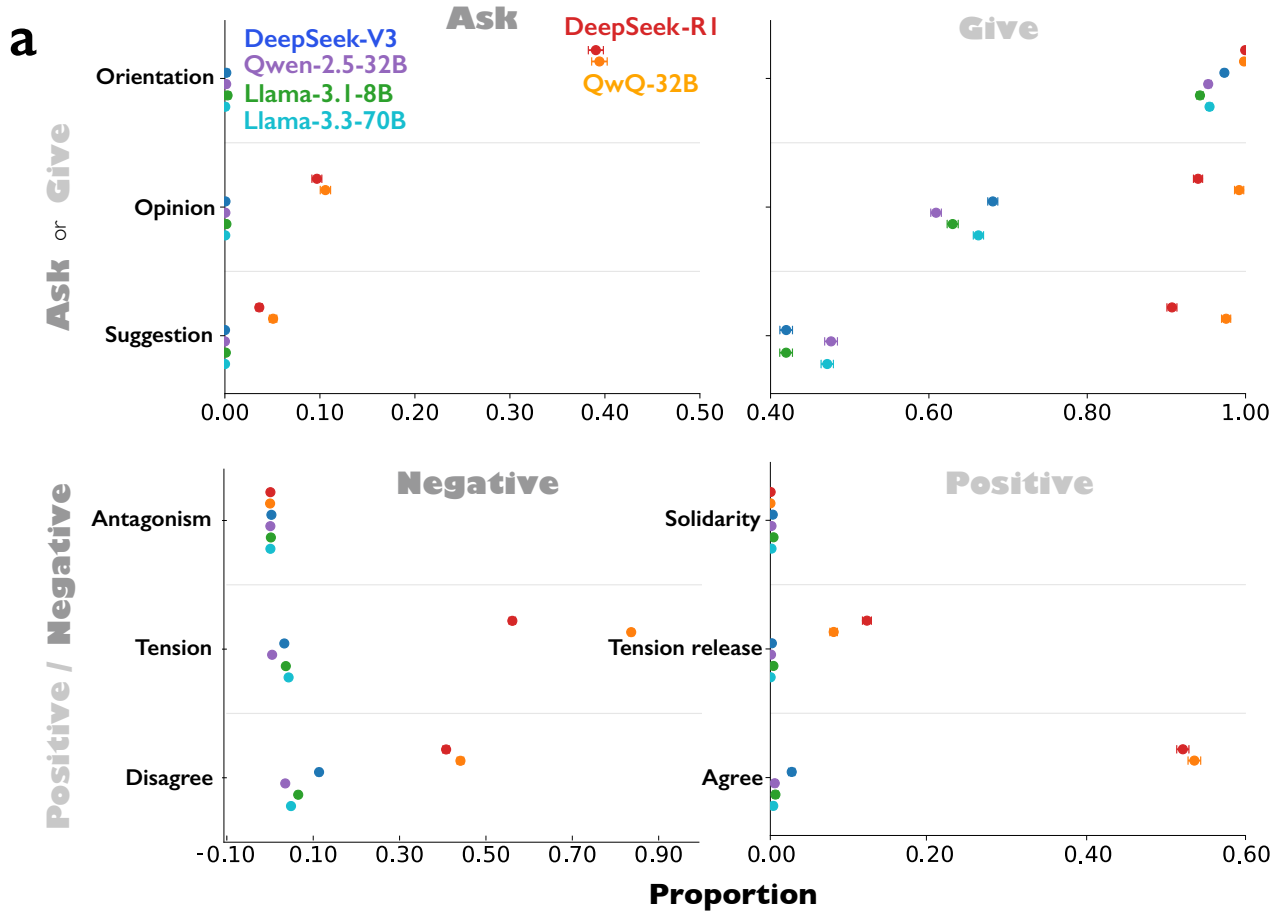
Shows tension

But it's combined with two equivalents of the *o*-xylylene precursor, so likely, the *o*-xylylene is the diene, and norbornadiene is the dienophile. Norbornadiene as a dienophile? Typically, it's symmetric, and one double bond can be attacked.

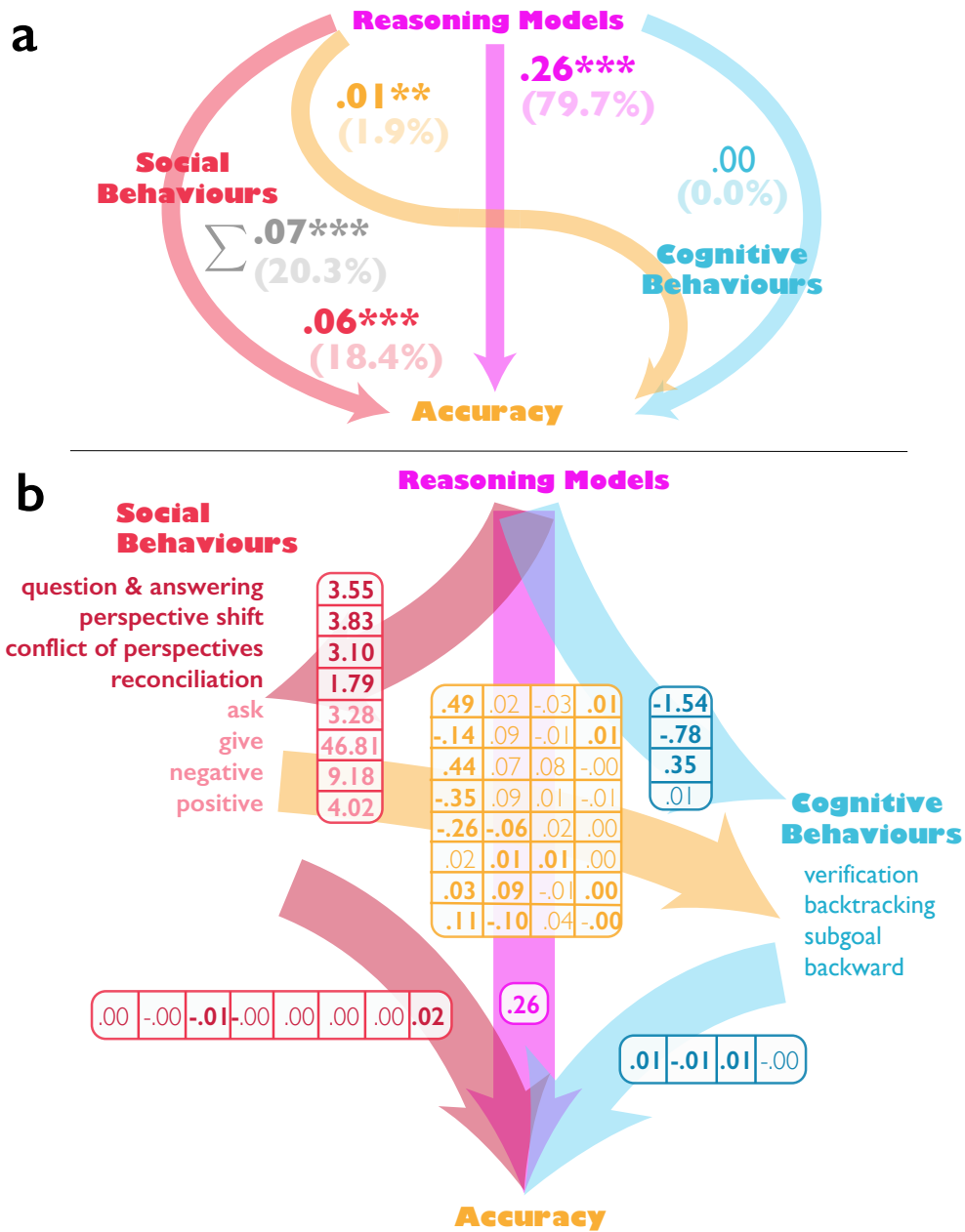
b



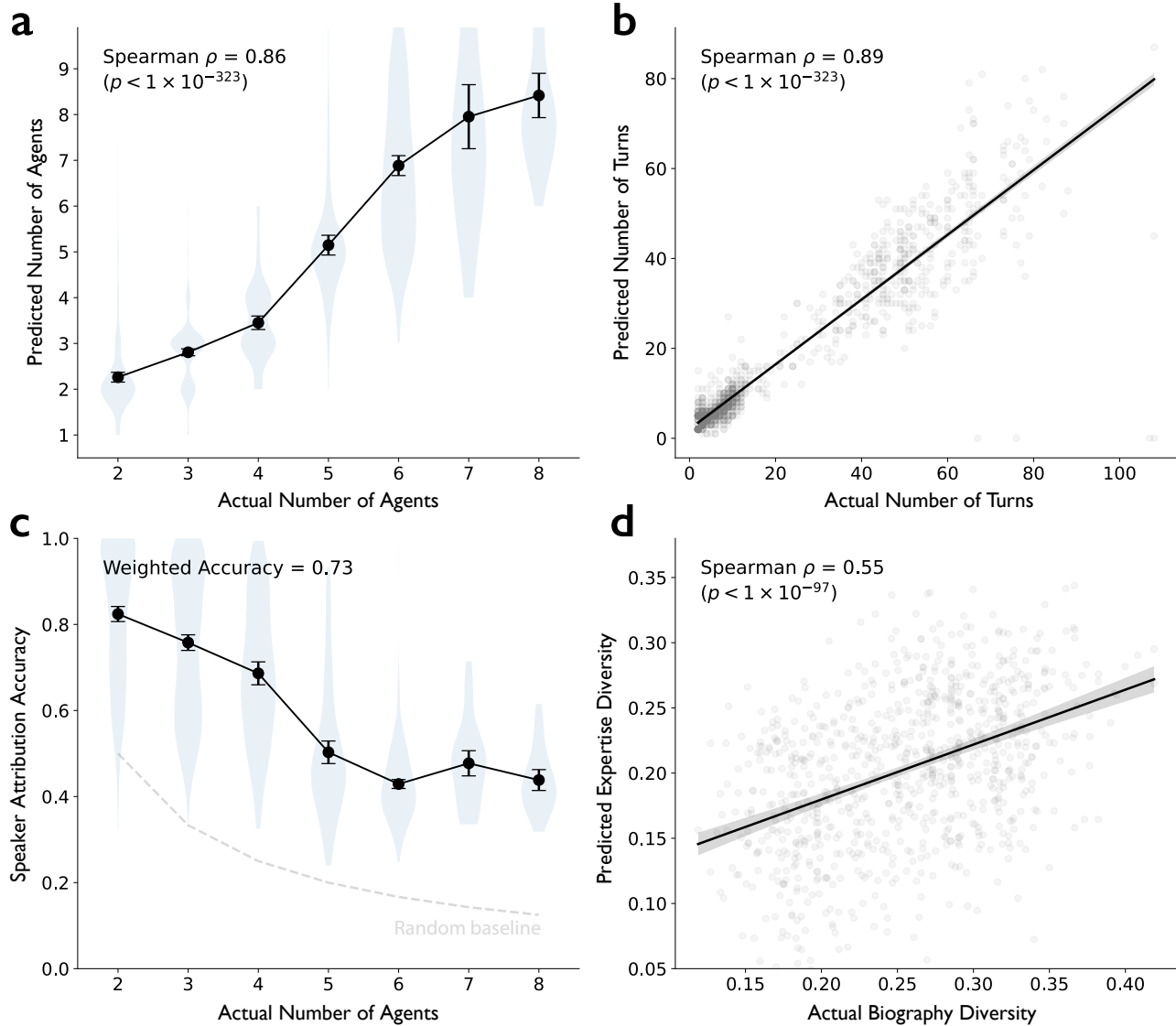
Extended Data Fig. 2. Conversation excerpt in DeepSeek-R1 reasoning traces. a, Representative excerpt from a chemistry problem-solving trace showing multi-turn dialogue between distinct cognitive personas. Each utterance is annotated with conversational behaviors (blue) and socio-emotional roles (yellow). **b**, Big Five personality profiles for the five personas identified via LLM-as-judge. Radar charts display normalized trait scores (1–5 scale) for Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N), and Openness (O). Each persona exhibits domain expertise profiles. For detailed coding procedures and additional annotated examples, see Appendix D.1.



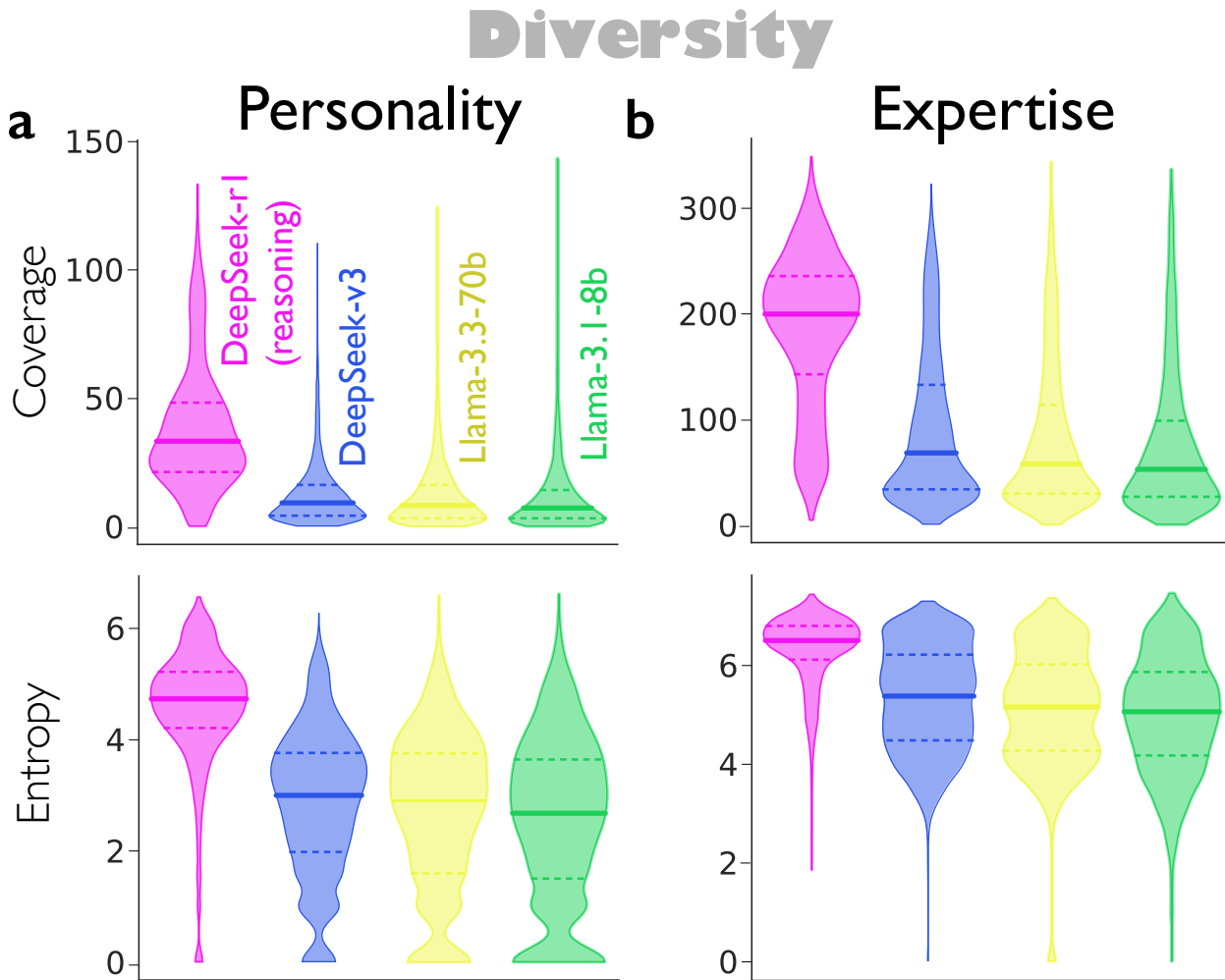
Extended Data Fig. 3. Bales' detailed socio-emotional roles in chain-of-thought reasoning. **a**, Proportion of Bales' 12 socio-emotional roles expressed in reasoning traces (see Fig. 1 for higher-level aggregations of the socio-emotional roles). **b**, Differences in problem complexity by the presence of detailed socio-emotional roles in DeepSeek-R1 (measured on a 7-point Likert scale from 1 [extremely easy] to 7 [extremely difficult] using LLM-as-judge or by error rates in non-reasoning models). Points indicate mean complexity for reasoning traces where the behavior/role is present (red) or absent (blue).



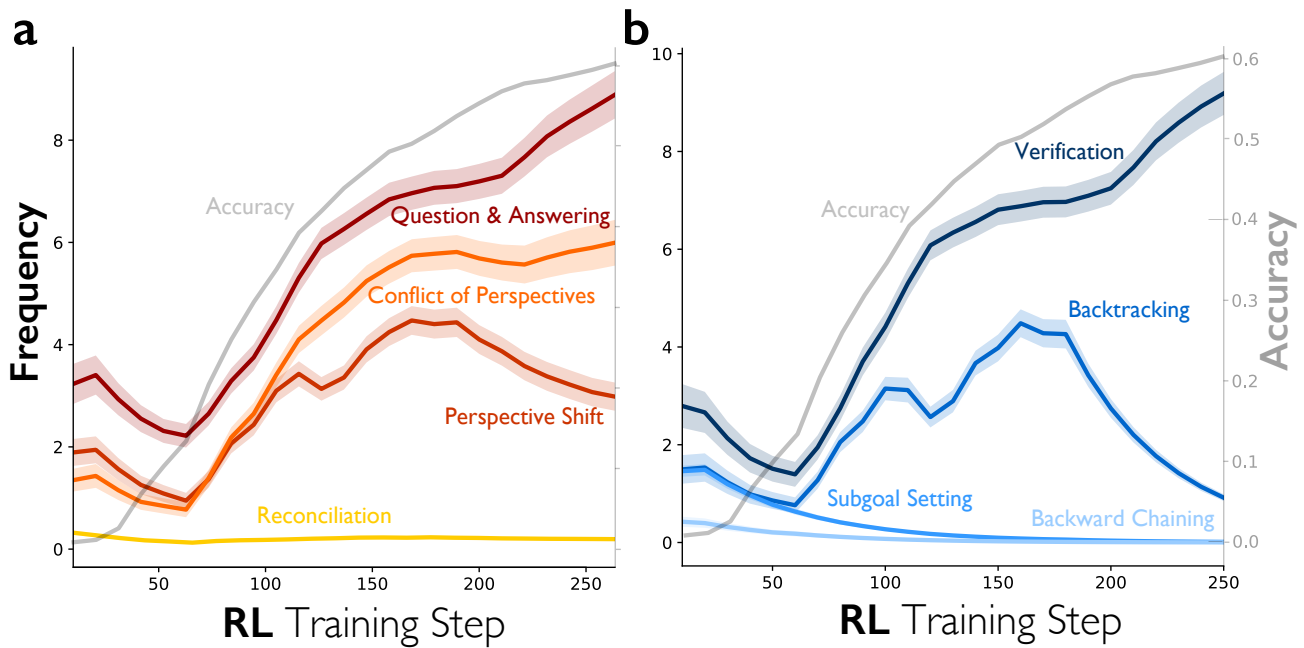
Extended Data Fig. 4. Mediation analysis linking reasoning models (DeepSeek-R1 and QwQ-32B) to accuracy advantages through simulated social behaviours. Mediation structure linking DeepSeek-R1 and QwQ-32B (relative to instruction-tuned models) to improved task accuracy through conversational behaviours and socio-emotional roles (red), cognitive reasoning behaviours (blue), and the indirect pathway through which social behaviors facilitate cognitive reasoning (orange) as estimated in a structural equation model of labeled reasoning traces. Arrows denote direct and indirect effects. The direct pink pathway indicates the unmediated effect of DeepSeek-R1 and QwQ-32B on accuracy. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. **a**, Summary estimates following various model paths, such that conversational behaviors and socio-emotional roles (see Fig. 1 and Extended Data Fig. 1) directly and indirectly contribute to the accuracy improvement. To illustrate the relative contribution of each pathway, we report the proportional share of each pathway on model accuracy, calculated as the value of each pathway coefficient divided by the sum of coefficient values across all pathways. This demonstrates that more than 20% of accuracy is explained by the direct and indirect effect of social behaviors manifest in the reasoning trace. **b**, Coefficient matrices underlying the structural equation model (SEM), where labels in the figure index the estimates within the matrices. The red panels show the effects of DeepSeek-R1 and QwQ-32B on social behaviors and the effects of social behaviors on accuracy. The orange panel displays the effects of social behaviors on cognitive behaviors. The blue panels present the effects of DeepSeek-R1 and QwQ-32B on cognitive behaviors and the effects of cognitive behaviors on accuracy. Bolded coefficients indicate statistical significance ($p < 0.05$). Full coefficient estimates are reported in Supplementary Table 1.



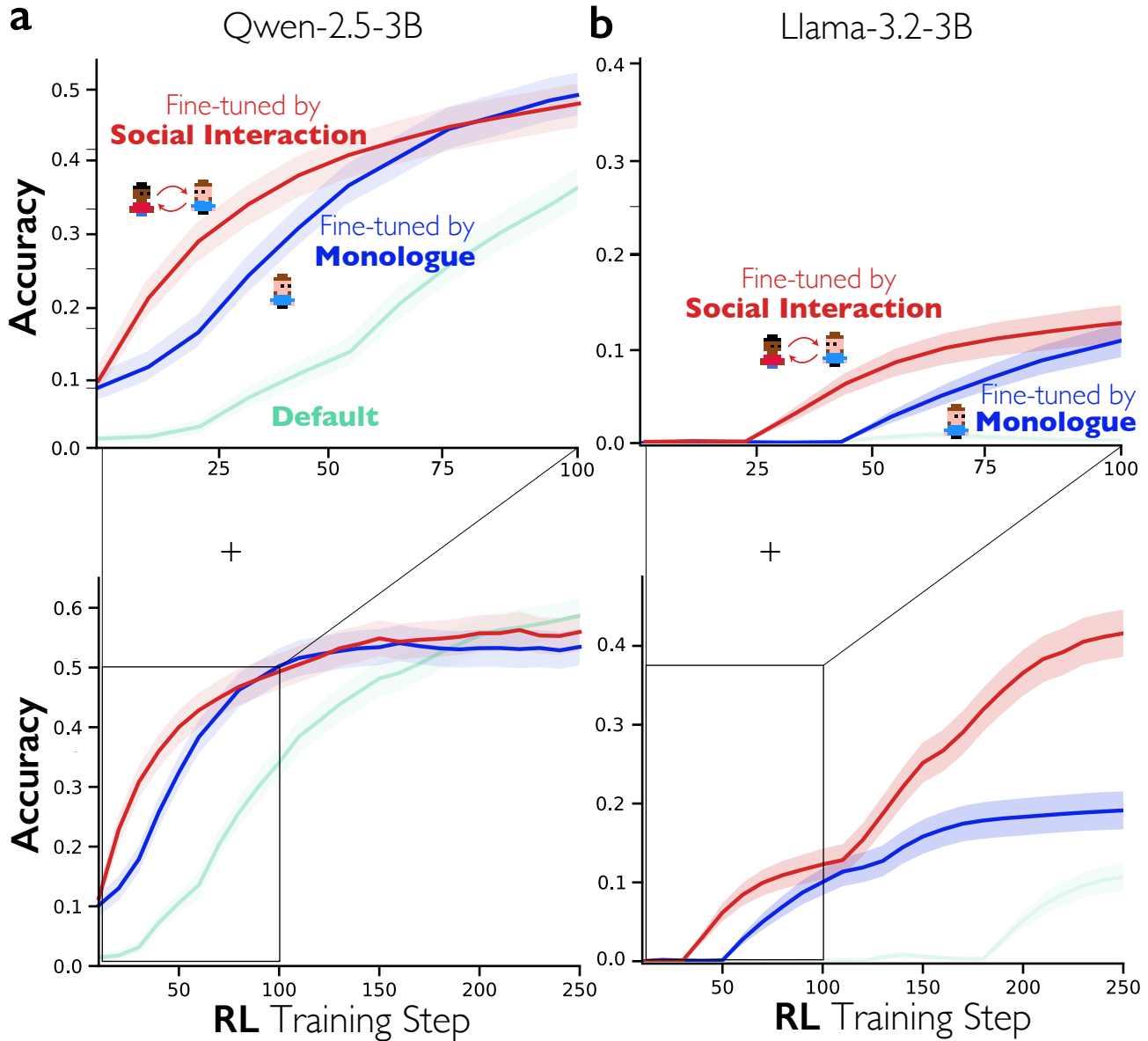
Extended Data Fig. 5. LLM-as-judge benchmark results for identifying latent voices. Using the Intelligence Squared Debates Corpus ($N = 1,196$ conversations), we validate the LLM-as-judge’s ability to identify distinct speakers when speaker labels are hidden, and dialogue is concatenated into a single block of text. **a**, Predicted versus actual number of agents in each conversation (Spearman’s $\rho = 0.86$, $p < 1 \times 10^{-323}$). Violin plots show the distribution of predictions for each actual agent count; points and error bars indicate means and 95% confidence intervals. **b**, Predicted versus actual number of conversational turns (Spearman’s $\rho = 0.89$, $p < 1 \times 10^{-323}$). **c**, Speaker attribution accuracy as a function of the number of agents. Accuracy is highest for two speakers (82%) and decreases as the number of speakers increases, but remains well above the random baseline (dashed line) across all conditions. Weighted accuracy across all conversations is 73%. **d**, Predicted expertise diversity (based on LLM-inferred descriptions and embeddings) versus actual biography diversity among debate participants (Spearman’s $\rho = 0.55$, $p < 1 \times 10^{-97}$), demonstrating that the LLM-as-judge captures meaningful variation in domain expertise that corresponds to ground-truth biographical differences.



Extended Data Fig. 6. SAE-based personality and expertise diversity estimates applied to model activations in reasoning traces. **a**, Distribution of coverage and entropy for SAE personality-related features. **b**, Distribution of coverage and entropy for SAE expertise-related features. Error bars in all panels indicate 95% confidence intervals. The solid horizontal line indicates the median, and dashed lines denote the interquartile range (IQR, 25th–75th percentiles).

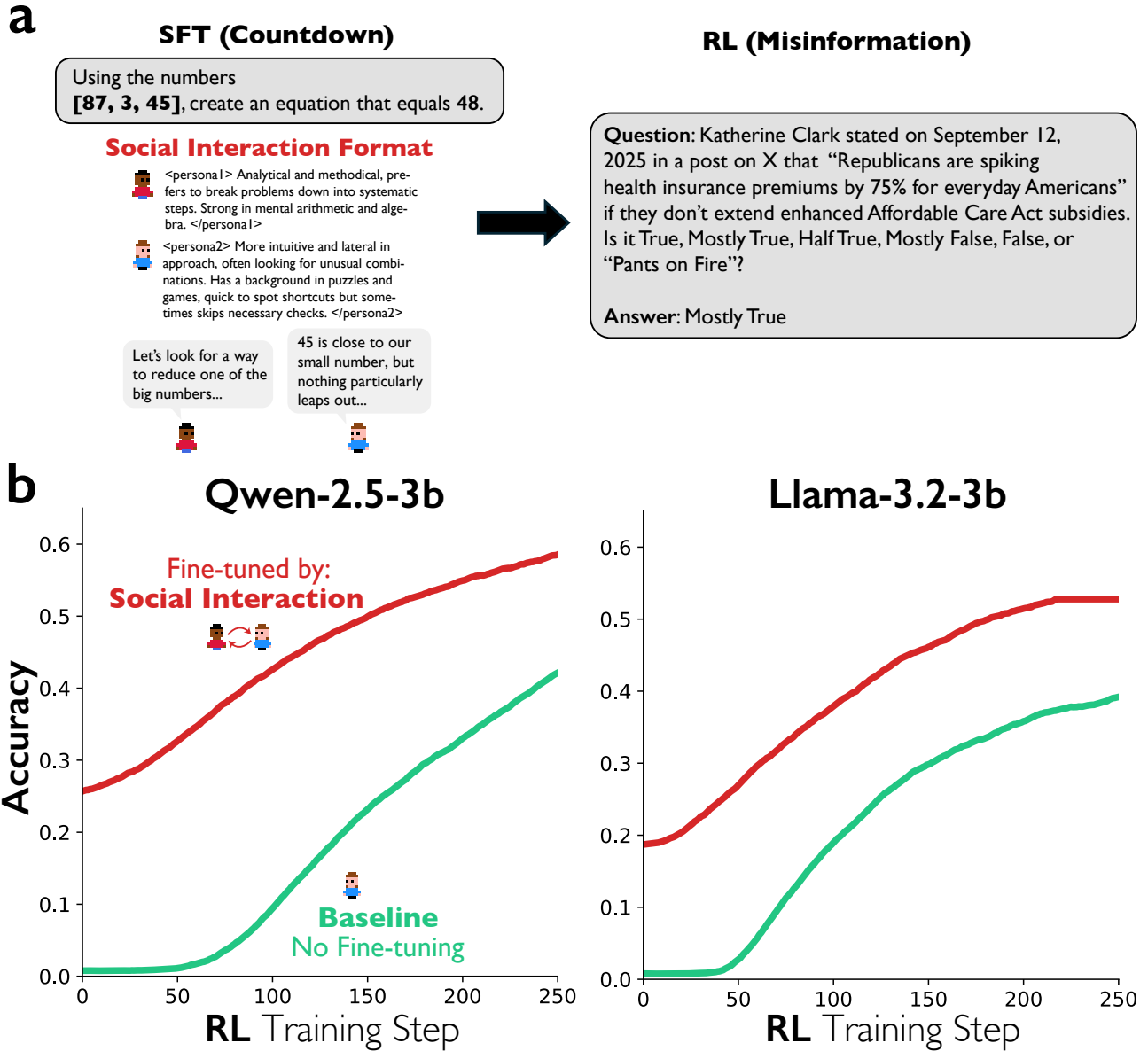


Extended Data Fig. 7. Fine-tuning with conversational scaffolding accelerates reasoning improvement during reinforcement learning. **a**, Trajectory of conversational behaviors in the Qwen-2.5-3B base model during reinforcement learning. Question-and-answering behavior emerges first and increases most rapidly, followed by conflict of perspectives and perspective shifts rising in parallel. Reconciliation remains low throughout, suggesting competing approaches rather than integration. The trajectory of model accuracy shows how the uptick of questions, answers, and interacting perspectives correlated with an acceleration in model improvement. **b**, Trajectory of cognitive behaviors in the same model. Verification increases most dramatically during training, following in lockstep with the presence of question asking and answering. This is followed by backtracking, which follows conflicts between perspectives. Subgoal setting and backward chaining show more modest gains before a gradual decline. Shaded regions indicate 95% confidence intervals.



Extended Data Fig. 8. Fine-tuning with dialogue vs. monologue scaffolding accelerates reasoning improvement during reinforcement learning. **a**, Accuracy trajectories during reinforcement learning on the Countdown task for Qwen-2.5-3B. **b**, Accuracy trajectories for Llama-3.2-3B. Models initially fine-tuned with multi-agent dialogues (red) reach high accuracy faster than models fine-tuned with monologue-style reasoning (blue), though both eventually converge. The base model without fine-tuning (default; light green colors) learns more slowly.

1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594



Extended Data Fig. 9. Conversational scaffolding transfers to misinformation detection under reinforcement learning. **a**, Illustration of the Countdown task and the PolitiFact misinformation-detection task. **b**, Validation accuracy during RL on PolitiFact, comparing social interaction (red) and single-voice, default reasoning CoT (green) format.

C. Extended Data Tables

Extended Data Table 1. Reasoning traces before and after steering the conversational surprise feature (Layer 15, Feature 30939)

Steering	Reasoning Trace	Result
-10	<p><think> To create an equation using the numbers [46, 54, 54, 77] that equals 75, I'll start by looking for operations that can help achieve this total. First, I'll consider the sum of all four numbers: $46 + 54 + 54 + 77 = 231$ Next, I'll explore subtracting a specific combination to reduce the total to 75. Subtracting 77 from 231 gives: $231 - 77 = 154$ Now, I'll look for a way to reduce 154 further. Dividing 154 by 2 gives: $154 / 2 = 77$ Since I have two 54s, dividing one of them by 2 will utilize both 54s: $54 / 2 = 27$ Finally, I'll combine these results to form the equation: $(46 + 54) / 2 + 77 - 27 = 75$ </think> <answer> $(46 + 54) / 2 + 77 - 27$ </answer></p>	WRONG
0	<p><think> To achieve the target number 75 using the numbers 46, 54, 52, and 77, I'll start by considering the largest number, 77, and aim to reduce it to a value that can be combined with the others to reach 75. Subtracting 46 from 77 gives me 31. Next, I'll use the remaining numbers, 54 and 52. Dividing 54 by 2.5 will help balance the equation, leading to 21.6. Finally, adding 21.6 to 31 results in 52.6, which is still not 75. To refine this, I'll multiply 52 by 1.5 to get 78, and then subtract 3 from it to arrive at 75. This step-by-step approach ensures that each number is used exactly once and the operations are basic arithmetic. </think> <answer> $(77 - 46) + (54 / 2.5) + (52 * 1.5) - 3$ </answer></p>	WRONG

Continued on next page

Steering	Reasoning Trace	Result
+10	<p><think></p> <p>Okay, so I have this problem where I need to use the numbers 46, 54, 52, and 77 to create an equation that equals 75. I can use basic arithmetic operations like addition, subtraction, multiplication, and division, and each number can only be used once. Alright, let's try to figure this out step by step.</p> <p>First, I'll list out the numbers: 46, 54, 52, 77. I need to combine them with operations to get 75. Hmm, 75 isn't too big, but these numbers are all in the 40s and 50s, so maybe I can subtract something from a larger number or maybe divide something to get a smaller number.</p> <p>Let me think about the possible operations. Maybe I can subtract two numbers and then add or multiply the rest. Or perhaps I can divide some numbers to get fractions, but that might complicate things. Let me try some combinations.</p> <p>Wait, let me see what the total sum of these numbers is. $46 + 54$ is 100, and $52 + 77$ is 129. So the total is $100 + 129$, which is 229. Hmm, that's way higher than 75. So maybe I need to subtract some numbers or divide to bring it down.</p> <p>Another idea: Maybe I can combine some numbers first. Like, 46 and 54. If I subtract 46 from 54, I get 8. Or if I subtract 52 from 54, that's 2. Or maybe 46 and 52. If I subtract 46 from 52, that's 6. Hmm, those are small numbers, but maybe I can use them in a way that helps me reach 75.</p> <p>Wait, 77 is the biggest number here. Maybe I can subtract 77 from something. Let's see, 77 minus 46 is 31, 77 minus 54 is 23, 77 minus 52 is 25. Hmm, those are still not 75. But maybe I can use division. For example, 77 divided by something. If I divide 77 by 2, that's 38.5, but I don't have a 2. Alternatively, 77 divided by 3 is about 25.666, but I don't have a 3 either. Maybe that's not the way.</p> <p>Wait, another thought: Maybe I can multiply some numbers. Let's see, 46 times 54 is way too big. 46 times 52 is even bigger. 54 times 52 is also too big. So multiplication might not help here. Maybe division can help.</p> <p>Let me try subtracting some numbers. If I take 54 minus 46, that's 8. Then, 52 minus 77 is negative, which might not help. Alternatively, 77 minus 52 is 25. Maybe I can use that 25 somehow. If I have 25, how can I get to 75? Maybe 25 times 3 is 75. But do I have a 3? Hmm, maybe I can get 3 by doing something with the other numbers.</p> <p>Wait, 54 minus 46 is 8, and 52 minus 77 is -25, but maybe I can flip that. Wait, 77 minus 52 is 25. So if I have 25, and I need to get to 75, I need 3 more. How can I get 3? Maybe from the other numbers. Let's see, 46, 54, 52, 77. If I do 54 minus 46, that's 8, and then maybe I can subtract 5 to get 3, but do I have a 5? I don't.</p> <p>Alternatively, maybe I can do 52 divided by something. 52 divided by 17 is 3, but I don't have a 17. Hmm. [... extensive exploration of many approaches ...]</p> <p>Wait, maybe I can use all four numbers in a more complex way. Like, $(46 + 54)$ is 100, and $(52 - 77)$ is -25. Then, $100 + (-25)$ is 75! Oh, wait, that works!</p> <p>Wait, let me check: $46 + 54$ is 100, and $52 - 77$ is -25. So, $100 + (-25)$ is 75. Yes, that's it! So the equation would be $(46 + 54) + (52 - 77) = 75$.</p> <p>Wait, but let me make sure I'm not using any numbers more than once. I used 46, 54, 52, and 77, each once. So that's correct.</p> <p>So, the equation is $(46 + 54) + (52 - 77) = 75$. Let me calculate it: $46 + 54$ is 100, $52 - 77$ is -25, and $100 + (-25)$ is 75. Perfect!</p> <p>So, that's the solution.</p> <p></think></p> <p><answer></p> <p>$(46 + 54) + (52 - 77)$</p> <p></answer></p>	RIGHT

D. Supplementary Information

D.1. Supplementary Methods: Annotations Examples

DEEPSEEK-R1: CHEMISTRY

What is the correct answer to this question: 7-(tert-butoxy)bicyclo[2.2.1]hepta-2,5-diene is combined with 2 equivalents of 5,6-bis(dibromomethyl)cyclohexa-1,3-diene and sodium iodide at elevated temperature, forming product 1. 1 is treated with aqueous sulfuric acid, forming 2. 2 is treated with SO_3 and pyridine in DMSO, forming 3. 3 is heated at 150C, forming final product 4. How many chemically distinct hydrogen atoms are there on 4? Choices: (A) 10 (B) 8 (C) 7 (D) 4

Reasoning Models Generate Societies of Thought

Answer: Enclose your step-by-step reasoning within `¡think¿` and `¡/think¿` before answering. Do not answer directly without reasoning.

Perspective ID	Personality	Domain Expertise
1	{'E': 2.5, 'A': 4.0, 'C': 5.0, 'N': 2.0, 'O': 3.5}	Planner/Executor: Methodically breaks down the problem, identifies reagents, and predicts the outcome of each step in the synthesis.
2	{'E': 1.5, 'A': 4.0, 'C': 4.0, 'N': 1.0, 'O': 4.0}	Associative Expert: Accesses a deep knowledge base to recall analogous reactions and common synthetic patterns, providing key insights and hypotheses.
3	{'E': 1.5, 'A': 3.0, 'C': 3.5, 'N': 5.0, 'O': 4.5}	Visualizer/Builder: Attempts to construct and name complex 3D molecular structures, struggling with complexity but capable of finding new ways to conceptualize the problem.
4	{'E': 1.0, 'A': 1.0, 'C': 5.0, 'N': 2.0, 'O': 2.0}	Critical Verifier: Sceptically re-evaluates prior assumptions, performs detailed calculations (e.g., molecular formulas) to check for consistency, and identifies logical errors.
5	{'E': 3.0, 'A': 3.0, 'C': 2.0, 'N': 1.0, 'O': 3.5}	Pragmatist/Strategist: Engages in meta-cognition about the problem-solving process, considers external constraints like multiple-choice options, and makes a final decision under uncertainty.

Perspective ID	Utterance
1	<p>Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 3, '6_Gives_orientation': 7, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 1, 'backward_chaining': 0}</p> <p>First, I need to figure out what the final product 4 is and then determine how many chemically distinct hydrogen atoms it has. The choices are 10, 8, 7, or 4.</p> <p>The sequence starts with 7-(tert-butoxy)bicyclo[2.2.1]hepta-2,5-diene. Bicyclo[2.2.1]heptane is norbornane, and hepta-2,5-diene means it has double bonds at positions 2 and 5. So it's a norbornadiene derivative with a tert-butoxy group at position 7. Position 7 is the bridgehead carbon.</p> <p>Norbornadiene has two double bonds, and the tert-butoxy is attached to the carbon that bridges the two methylene groups.</p> <p>Then it's combined with 2 equivalents of 5,6-bis(dibromomethyl)cyclohexa-1,3-diene. That sounds like a diene with dibromomethyl groups on it. Cyclohexa-1,3-diene has double bonds between 1-2 and 3-4, so positions 5 and 6 are the methylene carbons, but here it's bis(dibromomethyl), so it's probably CHBr₂ groups attached to positions 5 and 6. 5,6-bis(dibromomethyl)cyclohexa-1,3-diene means that at carbons 5 and 6, there are dibromomethyl substituents, so it's like 1,2-di(bromomethyl)benzene or something, but it's a diene, not aromatic.</p>
4	<p>Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 2, '6_Gives_orientation': 5, '7_Ask_for_orientation': 1, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>Cyclohexa-1,3-diene is not aromatic; it's a simple diene. So 5,6-bis(dibromomethyl)cyclohexa-1,3-diene should have the dibromomethyl groups on the saturated carbons. In cyclohexa-1,3-diene, carbons 1,2,3,4 are involved in double bonds, so carbons 5 and 6 are the CH₂ groups, but since it's a diene, positions 5 and 6 are equivalent and are sp³ carbons. So bis(dibromomethyl) means each has a CHBr₂ group instead of H, so it's 1,2-di(dibromomethyl)cyclohexa-1,3-diene? No, the carbon atoms 5 and 6 are part of the ring, so if they are bis(dibromomethyl), it might mean that the hydrogens are replaced, but typically, it's specified as substituents.</p>

- 1815
1816 3 **Conversational Behaviour:** { 'question_and_answering': 1, 'perspective_shift': 2,
1817 'conflict_of_perspectives': 1, 'reconciliation': 1 }
1818
1819 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1820 '5_Gives_opinion': 3, '6_Gives_orientation': 2, '7_Ask_for_orientation': 1, '8_Ask_for_opinion': 0,
1821 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1822 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
1823 0 }
1824 I think 5,6-bis(dibromomethyl) means that at positions 5 and 6, there are -CHBr₂ groups attached,
1825 but positions 5 and 6 are ring carbons. In standard naming, for cyclohexa-1,3-diene, carbon 1 and 2
1826 are double-bonded, 3 and 4 double-bonded, so carbon 5 is between 4 and 6, carbon 6 between 5 and 1.
1827 Carbons 5 and 6 are sp³ carbons with two hydrogens each. So bis(dibromomethyl) likely means that
1828 each of these carbons has one hydrogen replaced by a CHBr₂ group? That would be unusual.
1829 Perhaps it's 5,6-di(dibromomethyl), meaning the groups are attached to C5 and C6.
1830
-
- 1831 2 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 0,
1832 'conflict_of_perspectives': 1, 'reconciliation': 0 }
1833
1834 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1835 '5_Gives_opinion': 0, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1836 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1837 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1838 0 }
1839 I recall that compounds like this are used in Diels-Alder reactions or other cycloadditions.
1840 5,6-bis(dibromomethyl)cyclohexa-1,3-diene is a precursor to o-xylylene or something. o-Xylylene is
1841 generated from α,α' -dibromo-o-xylene, but here it's a diene.
1842
-
- 1843 4 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 2,
1844 'conflict_of_perspectives': 1, 'reconciliation': 0 }
1845
1846 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1847 '5_Gives_opinion': 3, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1848 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
1849 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
1850 0 }
1851 Upon second thought, 5,6-bis(dibromomethyl)cyclohexa-1,3-diene might be a misnomer, or perhaps
1852 it's the same as 3,6-di(bromomethyl)cyclohexa-1,4-diene or something. I think this is a way to
1853 generate the o-quinodimethane intermediate.
1854
-
- 1855 2 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
1856 'conflict_of_perspectives': 0, 'reconciliation': 0 }
1857
1858 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1859 '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1860 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1861 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1862 0 }
1863 In organic synthesis, o-xylylene is often generated from α,α' -dibromo-o-xylene by
1864 dehydrohalogenation. Similarly, here, 5,6-bis(dibromomethyl)cyclohexa-1,3-diene might be
1865 equivalent to having the dibromomethyl groups on a diene system that can undergo elimination to
1866 form a quinodimethane.
1867
1868
1869

1870		
1871	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
1872		'conflict_of_perspectives': 1, 'reconciliation': 1 }
1873		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1874		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1875		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1876		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1877		0 }
1878		Cyclohexa-1,3-diene with bis(dibromomethyl) at 5 and 6: since 5 and 6 are adjacent carbons, each
1879		with a CHBr ₂ group, so it's like 1,2-di(dibromomethyl)cyclohexa-1,3-diene. But in standard naming,
1880		it might be called 1,2-bis(dibromomethyl)benzene if it were aromatic, but it's not; it's diene.
1881		
1882		
1883	1	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
1884		'conflict_of_perspectives': 0, 'reconciliation': 0 }
1885		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1886		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1887		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1888		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1889		0 }
1890		Perhaps upon treatment with sodium iodide, it undergoes dehalogenation to form the diradical or
1891		diene for Diels-Alder.
1892		
1893		
1894	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 3,
1895		'conflict_of_perspectives': 5, 'reconciliation': 0 }
1896		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1897		'5_Gives_opinion': 4, '6_Gives_orientation': 6, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
1898		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 3, '12_Shows_antagonism': 0 }
1899		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1900		0 }
1901		The reaction is with sodium iodide at elevated temperature. Sodium iodide can promote
1902		dehalogenation or substitution. For dibromomethyl groups, NaI might facilitate the formation of the
1903		dibromide or something, but I think for generating quinodimethanes, it's often done with zinc or other
1904		metals, but NaI can be used for Finkelstein reaction, but here it's dehydrohalogenation.
1905		I recall that α,α' -dibromo-o-xylene with NaI gives o-xylene, which is a reactive diene for
1906		Diels-Alder.
1907		Similarly, here, 5,6-bis(dibromomethyl)cyclohexa-1,3-diene might be analogous.
1908		Cyclohexa-1,3-diene with substituents at 5 and 6: if it were benzene, it would be o-xylene-like, but
1909		it's not aromatic. Actually, 5,6-bis(dibromomethyl)cyclohexa-1,3-diene can tautomerize or
1910		something, but I think it's designed to lose HBr under conditions to form the extended diene.
1911		Perhaps upon heating with NaI, it eliminates to form a benzoquinone or something, but I think it's for
1912		generating o-quinodimethane.
1913		
1914		
1915		
1916		
1917		
1918		
1919		
1920		
1921		
1922		
1923		
1924		

1925		
1926	4	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1, 'conflict_of_perspectives': 2, 'reconciliation': 0 }
1927		
1928		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1929		
1930		
1931		
1932		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0 }
1933		
1934		Let's look at the structure. Cyclohexa-1,3-diene has the double bonds conjugated, but positions 5 and 6 are saturated. If each has a dibromomethyl group, then with base or iodide, it can eliminate two HBr to form a new double bond between C5 and C6, creating a 1,3-cyclohexadiene with exocyclic double bonds or something. Actually, elimination would give a diene like
1935		
1936		5,6-dehydrocyclohexa-1,3-diene, which is benzene, but with dibromomethyl, it's different.
1937		
1938		
1939		
1940	5	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
1941		
1942		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
1943		
1944		
1945		
1946		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0 }
1947		
1948		I found a better way: 5,6-bis(dibromomethyl)cyclohexa-1,3-diene is equivalent to
1949		3,6-di(bromomethyl)cyclohexa-1,4-diene or something. I think it's a precursor to the o-xyllylene diene.
1950		
1951		
1952	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
1953		
1954		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1955		
1956		
1957		
1958		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
1959		
1960		Upon searching my memory, compounds like this are used to generate o-quinodimethane by
1961		dehydrohalogenation. For example, α,α' -dibromo-o-xylene with NaI gives o-xyllylene.
1962		
1963	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0 }
1964		
1965		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1966		
1967		
1968		
1969		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0 }
1970		
1971		But here, it's cyclohexa-1,3-diene, not benzene. Cyclohexa-1,3-diene with 5,6-bis(dibromomethyl)
1972		might not be aromatic, but when it eliminates, it could form a quinodimethane system.
1973		
1974		
1975		
1976		
1977		
1978		
1979		

1980		
1981	1	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
1982		'conflict_of_perspectives': 1, 'reconciliation': 1 }
1983		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1984		'5_Gives_opinion': 2, '6_Gives_orientation': 2, '7_Ask_for_orientation': 1, '8_Ask_for_opinion': 0,
1985		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
1986		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
1987		0 }
1988		Perhaps the diene is set up for Diels-Alder.
1989		The first compound is 7-(tert-butoxy)bicyclo[2.2.1]hepta-2,5-diene, which is a dienophile or diene?
1990		Norbomadiene has two double bonds and can act as a dienophile in some cases, but typically it's a
1991		diene for Diels-Alder with electron-withdrawing groups.
1992		7-tert-butoxy might make it electron-rich, so perhaps it's a diene.
1993		
1994		
1995		
1996	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
1997		'conflict_of_perspectives': 0, 'reconciliation': 0 }
1998		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
1999		'5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2000		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2001		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2002		0 }
2003		Bicyclo[2.2.1]hepta-2,5-diene has the double bonds at 2-3 and 5-6, I think, with bridge at 1 and 4,
2004		and 7 is the methine bridge.
2005		Standard numbering: carbon 1 and 4 are bridgeheads, 2-3 is one double bond, 5-6 is the other, 7 is the
2006		CH or C with substituent.
2007		
2008		
2009		
2010	1	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2011		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2012		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2013		'5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2014		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2015		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2016		0 }
2017		With tert-butoxy at 7, it might be electron-donating, so the double bonds could be dienophiles or
2018		dienes.
2019		
2020	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
2021		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2022		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2023		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2024		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2025		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2026		0 }
2027		But it's combined with two equivalents of the other diene, and with NaI at elevated temperature.
2028		
2029		
2030		
2031		
2032		
2033		
2034		

- 2035
2036 1 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 0,
2037 'conflict_of_perspectives': 0, 'reconciliation': 0}
2038
2039 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2040 '5_Gives_opinion': 2, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2041 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2042 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2043 0}
2044 I think the NaI is used to generate the reactive diene from the bis(dibromomethyl) compound.
2045
2046 So, 5,6-bis(dibromomethyl)cyclohexa-1,3-diene with NaI likely eliminates to form o-quinodimethane
2047 or a derivative.
-
- 2048 2 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 2,
2049 'conflict_of_perspectives': 0, 'reconciliation': 0}
2050
2051 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2052 '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2053 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2054 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2055 0}
2056
2057 In fact, 1,2-bis(dibromomethyl)benzene with NaI gives o-xylylene, which is 1,2- di
2058 methylenecyclohexa-3,5-diene or something, but it's benzene with exocyclic double bonds.
2059
2060 o-Xylylene is often depicted as having a quinoid structure.
-
- 2061 4 **Conversational Behaviour:** { 'question_and_answering': 1, 'perspective_shift': 2,
2062 'conflict_of_perspectives': 1, 'reconciliation': 1}
2063
2064 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2065 '5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2066 '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2067 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2068 0}
2069
2070 But here, the starting material is cyclohexa-1,3-diene, not benzene.
2071 5,6-bis(dibromomethyl)cyclohexa-1,3-diene: if it eliminates two HBr, it could form 5,6- di
2072 methylenecyclohexa-1,3-diene, which is the same as o-xylylene because cyclohexa-1,3-diene with
2073 exocyclic double bonds at 5 and 6 is equivalent to the quinodimethane for benzene.
-
- 2074 3 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 2,
2075 'conflict_of_perspectives': 1, 'reconciliation': 0}
2076
2077 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2078 '5_Gives_opinion': 0, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2079 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0}
2080 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2081 0}
2082
2083 5,6-di methylenecyclohexa-1,3-diene is actually o-xylylene, which is a resonance structure of 1,2- di
2084 methylenebenzene or something, but it's a reactive diene with s-cis conformation for Diels-Alder.
-
- 2085
2086
2087
2088
2089

2090		
2091	1	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2092		'conflict_of_perspectives': 1, 'reconciliation': 1 }
2093		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 1, '4_Gives_suggestion': 0,
2094		'5_Gives_opinion': 2, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2095		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2096		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2097		0 }
2098		Yes, that makes sense. So with NaI, it generates o-xylene in situ.
2099		Then, this o-xylene can act as a diene in Diels-Alder reactions.
2100		Now, the other compound is 7-(tert-butoxy)norbornadiene. Norbornadiene can act as a dienophile
2101		because one of its double bonds can be electron-deficient or something, but with tert-butoxy, it might
2102		be electron-rich.
2103		Norbornadiene has two double bonds, and it can participate in Diels-Alder as a diene with
2104		electron-withdrawing groups, but here it has tert-butoxy, which is electron-donating, so perhaps it's a
2105		poor dienophile.
2106		
2107		
2108		
2109	4	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
2110		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2111		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2112		'5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2113		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
2114		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2115		0 }
2116		But it's combined with two equivalents of the o-xylene precursor, so likely, the o-xylene is the
2117		diene, and norbornadiene is the dienophile.
2118		Norbornadiene as a dienophile? Typically, it's symmetric, and one double bond can be attacked.
2119		
2120		
2121		
2122	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
2123		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2124		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2125		'5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2126		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2127		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2128		0 }
2129		In Diels-Alder, norbornadiene can act as a dienophile when reacted with dienes.
2130		For example, norbornadiene with cyclopentadiene gives adducts.
2131		
2132		
2133		
2134		
2135		
2136		
2137		
2138		
2139		
2140		
2141		
2142		
2143		
2144		

- 2145
2146 1 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
2147 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2148
2149 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2150 '5_Gives_opinion': 3, '6_Gives_orientation': 4, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2151 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2152 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 1, 'backward_chaining':
2153 0 }
2154 So here, o-xylene is a good diene, and norbornadiene can be the dienophile.
2155
2156 But there are two equivalents of the o-xylene precursor, so probably two Diels-Alder reactions
2157 occur.
2158 First, one o-xylene adds to norbornadiene.
2159
2160 Norbornadiene has two double bonds; when one acts as dienophile, the other might still be reactive.
2161 Perhaps it adds once, and then the adduct has another diene or dienophile.
-
- 2163 4 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
2164 'conflict_of_perspectives': 1, 'reconciliation': 0 }
2165
2166 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2167 '5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2168 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2169 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2170 0 }
2171 But with two equivalents, likely two additions.
2172
-
- 2173 1 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
2174 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2175
2176 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2177 '5_Gives_opinion': 2, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2178 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2179 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2180 0 }
2181 The tert-butoxy group might play a role.
2182 Perhaps the tert-butoxy is a protecting group or something.
-
- 2184 3 **Conversational Behaviour:** { 'question_and_answering': 1, 'perspective_shift': 0,
2185 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2186
2187 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2188 '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2189 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2190 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 2, 'backward_chaining':
2191 0 }
2192 Let's think about the structure.
2193
2194 First, o-xylene generated from 5,6-bis(dibromomethyl)cyclohexa-1,3-diene with NaI: elimination
2195 of two HBr gives o-xylene, which is .
2196
2197 Then, this diene adds to the norbornadiene.
-
- 2198
2199

2200		
2201	1	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 2,
2202		'conflict_of_perspectives': 1, 'reconciliation': 1 }
2203		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2204		'5_Gives_opinion': 3, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2205		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2206		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2207		0 }
2208		Norbornadiene, 7-(tert-butoxy), so the double bond might be activated or deactivated.
2209		tert-butoxy is electron-donating, so it might make the double bond less reactive as dienophile, but
2210		norbornadiene double bonds are strained, so they can still react.
2211		Perhaps the Diels-Alder occurs with o-xyllylene as diene and norbornadiene as dienophile.
2212		
2213		
2214		
2215	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2216		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2217		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2218		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2219		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2220		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2221		0 }
2222		The adduct would have the norbornene system with a phenyl ring attached, but o-xyllylene addition
2223		gives a bicyclic system fused to the norbornane.
2224		
2225		
2226	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 4,
2227		'conflict_of_perspectives': 3, 'reconciliation': 1 }
2228		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2229		'5_Gives_opinion': 0, '6_Gives_orientation': 5, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2230		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2231		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
2232		0 }
2233		In Diels-Alder, o-xyllylene with a dienophile gives a 1,2,3,4-tetrahydronaphthalene derivative.
2234		For example, with ethylene, it gives 1,2-dihydronaphthalene or something.
2235		o-Xyllylene + dienophile → benzocyclobutene or tetrahydronaphthalene, depending on the
2236		dienophile.
2237		Standard Diels-Alder: o-xyllylene has the diene system in the exocyclic bonds, so it adds to
2238		dienophiles to form bicyclic compounds with a benzene ring.
2239		Specifically, with a mono-olefin, it gives 1,4-dihydronaphthalene, which can aromatize.
2240		
2241		
2242		
2243		
2244	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2245		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2246		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2247		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2248		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2249		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2250		0 }
2251		But in this case, the dienophile is norbornadiene, which is a bicyclic diene.
2252		
2253		
2254		

2255		
2256	1	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
2257		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2258		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2259		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2260		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2261		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2262		0 }
2263		
2264		So, one o-xylylene adds to one double bond of norbornadiene.
2265		
2266	3	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
2267		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2268		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2269		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2270		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2271		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 1, 'backward_chaining':
2272		0 }
2273		
2274		The adduct would have a norbornene-like structure with a fused benzene ring or something.
2275		
2276		Let's sketch it.
2277		Norbornadiene has double bonds at 2-3 and 5-6. Suppose the dienophile is the 2-3 double bond. After
2278		Diels-Alder with o-xylylene, the o-xylylene adds across, so the adduct has the norbornane skeleton
2279		with a new ring fused, but since o-xylylene adds as a diene, it becomes a benzo-fused norbornene.
2280		
2281	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2282		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2283		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2284		'5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2285		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2286		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2287		0 }
2288		
2289		o-Xylylene is like a 1,3-diene with substituents, so when it adds to a dienophile, it forms a
2290		6-membered ring with the dienophile part.
2291		
2292		For example, if dienophile is ethylene, it forms 1,2,3,4-tetrahydronaphthalene, which has a double
2293		bond.
2294		
2295		
2296		
2297		
2298		
2299		
2300		
2301		
2302		
2303		
2304		
2305		
2306		
2307		
2308		
2309		

2310	
2311	3
2312	Conversational Behaviour: { 'question_and_answering': 2, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2313	
2314	IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
2315	
2316	
2317	Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2318	
2319	Similarly, with norbornadiene, if we consider one double bond as dienophile, say the 2-3 bond, then
2320	after addition, we have a new 6-membered ring fused to the norbornane at positions 2 and 3, so it's
2321	like a benzonorbornene with a double bond at 5-6 still present.
2322	
2323	Norbornadiene has carbons: bridgehead 1 and 4, double bond between 2-3, double bond between 5-6,
2324	and 7 is the bridge.
2325	
2326	If 2-3 acts as dienophile, adding o-xylylene, the adduct has a benzene ring fused between C2 and C3
2327	of the norbornane, but since C2 and C3 were double-bonded, after addition, it's single bond, and the
2328	new ring is aromatic or something.
2329	2
2330	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2331	
2332	IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2333	
2334	
2335	Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2336	
2337	In Diels-Alder, o-xylylene adds to give a non-aromatic product initially, like 1,4-dihydronaphthalene,
2338	which can aromatize if there's a double bond or something.
2339	
2340	4
2341	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2342	
2343	IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2344	
2345	
2346	Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2347	
2348	But in this case, with norbornadiene, the adduct might have the double bond at 5-6 still, so it could be
2349	a diene still.
2350	
2351	
2352	
2353	
2354	
2355	
2356	
2357	
2358	
2359	
2360	
2361	
2362	
2363	
2364	

2365		
2366	1	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
2367		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2368		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2369		'5_Gives_opinion': 0, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2370		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2371		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2372		0 }
2373		
2374		Then, with a second equivalent of o-xyllylene, it could add to the remaining double bond at 5-6.
2375		
2376		So, after two Diels-Alder additions, we have a molecule with two benzene rings fused to the
2377		norbornane system.
2378		
2379		Specifically, one fused at 2-3 and one at 5-6.
2380	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 2,
2381		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2382		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2383		'5_Gives_opinion': 2, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2384		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
2385		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
2386		0 }
2387		
2388		In norbornadiene, positions 2-3 and 5-6 are symmetric, so adding o-xyllylene to each double bond
2389		would give a symmetric molecule.
2390		
2391		The tert-butoxy is at position 7, which is the bridge carbon, so it might not affect symmetry.
2392		
2393		So product 1 should be 7-(tert-butoxy)-1,4-dihydro-1,4-epoxynaphthalene or something, but let's
2394		think.
2395	5	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
2396		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2397		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2398		'5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2399		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2400		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2401		0 }
2402		
2403		After Diels-Alder, each addition of o-xyllylene to a double bond of norbornadiene gives a
2404		1,4-dihydronaphthalene fused to norbornane, but since the double bond is part of the norbornane, it's
2405		better to think of the norbornane double bond as the dienophile.
2406		
2407		
2408		
2409		
2410		
2411		
2412		
2413		
2414		
2415		
2416		
2417		
2418		
2419		

2420	
2421	3
2422	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 2, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2423	
2424	IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 1, '6_Gives_orientation': 5, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 2, '12_Shows_antagonism': 0 }
2425	
2426	
2427	Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2428	
2429	Standard way: the adduct of norbornadiene and o-xylene at one double bond gives a compound
2430	where the norbornene has a benzo group fused, meaning a benzene ring fused to the 2,3-bond, so it's
2431	2,3-benzo-5-norbornene or something.
2432	
2433	Then, the remaining double bond at 5-6 can add another o-xylene, giving 2,3:5,6-dibenzo derivative.
2434	And since norbornadiene is symmetric, the dibenzo derivative should be symmetric.
2435	
2436	Position 7 has tert-butoxy.
2437	
2438	So product 1 is 7-(tert-butoxy)bicyclo[2.2.1]hept-5-ene-2,3:5,6-di(benzo) or something, but benzo
2439	means fused benzene.
2440	4
2441	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 0, 'conflict_of_perspectives': 1, 'reconciliation': 0 }
2442	
2443	IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 1, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
2444	
2445	
2446	Cognitive Behaviour: { 'verification': 1, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0 }
2447	
2448	
2449	Actually, it's 7-(tert-butoxy)-1,4-dihydro-1,4-epoxynaphthalene derivative? No.

- 2475
2476 3 **Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 3,
2477 'conflict_of_perspectives': 1, 'reconciliation': 0}
2478
2479 **IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 2,
2480 '5_Gives_opinion': 2, '6_Gives_orientation': 7, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2481 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2482 **Cognitive Behaviour:** {'verification': 0, 'backtracking': 1, 'subgoal_setting': 2, 'backward_chaining':
2483 0}
2484 Perhaps it's 7-tert-butoxy-1,4-epoxy-5,8-dihydronaphthalene or something. I need to name it
2485 properly.
2486
2487 The fused system: after Diels-Alder, for each addition, the o-xylylene becomes a benzene ring fused,
2488 with the norbornane double bond becoming single bond in a cyclohexene ring.
2489
2490 For one addition: norbornadiene has double bond at 2-3. When it to o-xylylene, o-xylylene is the
2491 diene, so the dienophile double bond (2-3) becomes single bond, and a new six-member ring is
2492 formed, fused to positions 2 and 3. Since positions 2 and 3 were part of the five-member ring in
2493 norbornane, now it's a fused system with a benzene ring.
2494 Specifically, the adduct is a benzofused norbornene, with the double bond at 5-6 still present.
2495
2496 Then, the second o-xylylene adds to the 5-6 double bond, similarly, forming another benzene ring
2497 fused at 5-6.
2498
2499 So the final adduct for product 1 is a molecule where the norbornane skeleton has benzene rings fused
2500 to both "sides", and the bridge carbon 7 has tert-butoxy.
2501
2502 Moreover, since the fusions are at the former double bond positions, the norbornane is now like a
2503 central bicyclic system with two benzene rings attached via the fusion.
2504
2505 In terms of the carbon skeleton, it's like 9,10-dihydro-9,10[1',2']benzoanthracene or something, but
2506 let's think.
-
- 2505 4 **Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 0,
2506 'conflict_of_perspectives': 1, 'reconciliation': 0}
2507
2508 **IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2509 '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2510 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2511 **Cognitive Behaviour:** {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2512 0}
2513
2514 Actually, it's similar to the Diels-Alder adduct of two molecules of o-xylylene with norbornadiene, so
2515 the product is a cage-like molecule with two benzene rings.
-
- 2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529

2530		
2531	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2532		
2533		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 3, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2534		
2535		
2536		
2537		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2538		
2539		The norbornane part has carbons 1,4,7 and the bridges.
2540		
2541		After fusion at 2-3 and 5-6, positions 2,3,5,6 are now sp ³ carbons part of the fused rings.
2542		
2543		Each fusion adds a benzene ring, so the molecule has three rings: the norbornane core with two benzene rings fused.
2544		
2545	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2546		
2547		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2548		
2549		
2550		
2551		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2552		
2553		But norbornane is bicyclic, so with two fusions, it becomes polycyclic.
2554		
2555		Perhaps it's 1,4,4a,8a-tetrahydro-1,4-methanoanthracene or something.
2556		
2557	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0 }
2558		
2559		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2560		
2561		
2562		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2563		
2564		
2565		I recall that the Diels-Alder adduct of norbornadiene and two equivalents of o-xylylene is a known compound, and it's symmetric.
2566		
2567		In fact, it's called 1,4,4a,8a-tetrahydro-1,4:5,8-diepoxyanthracene or something, but with the epoxy-methano bridge.
2568		
2569		
2570	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2571		
2572		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2573		
2574		
2575		
2576		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }
2577		
2578		The central part is the norbornane with the 1,4-methano bridge, and fused to anthracene-like system.
2579		
2580		
2581		
2582		
2583		
2584		

2585		
2586	4	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0}
2587		
2588		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2589		
2590		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
2591		
2592		Anthracene has three rings linear, but here it's different.
2593		
2594		
2595		
2596	1	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0}
2597		
2598		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2599		
2600		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
2601		
2602		After Diels-Alder, the o-xyllylene adds to form a new six-member ring, so for each addition, it's like adding a benzene ring fused to the norbornane double bond.
2603		
2604		So for norbornadiene, after two additions, we have a molecule where the two double bonds are now part of two fused benzene rings, and the norbornane bridge is present.
2605		
2606		
2607		
2608		
2609	2	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 2, 'conflict_of_perspectives': 0, 'reconciliation': 0}
2610		
2611		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2612		
2613		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
2614		
2615		Specifically, it's 1,4,4a,8a-tetrahydro-1,4-methanonaphthalene for one addition, but with two, it's symmetric.
2616		
2617		The compound is known as the adduct where norbornadiene is the dienophile for two o-xyllylenes, forming a cage molecule.
2618		
2619		
2620		
2621		
2622		
2623	3	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}
2624		
2625		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0}
2626		
2627		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
2628		
2629		I think it: the product is 1,4,4a,5,8,8a,9,10-octahydro-1,4:5,8-diepoxyanthracene or something.
2630		
2631		
2632		
2633		
2634		
2635		
2636		
2637		
2638		
2639		

- 2640
2641 5 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
2642 'conflict_of_perspectives': 1, 'reconciliation': 0 }
2643
2644 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2645 '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2646 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2647 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2648 0 }
2649
2650 Perhaps it's easier to think that the final structure has the norbornane skeleton with C1 and C4 as
2651 bridgeheads, C7 with tert-butoxy, and at C2-C3 and C5-C6, instead of double bonds, there are
2652 benzene rings fused, meaning that C2 and C3 are now part of a benzene ring fused between them,
2653 similarly for C5 and C6.
-
- 2654 3 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 0,
2655 'conflict_of_perspectives': 0, 'reconciliation': 0 }
2656
2657 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2658 '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2659 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2660 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2661 0 }
2662
2663 In the fusion, for example, when a benzene ring is fused to the 2,3-bond, it means that C2 and C3
2664 become part of the benzene ring, so they are sp² carbons in benzene.
-
- 2665 4 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
2666 'conflict_of_perspectives': 2, 'reconciliation': 0 }
2667
2668 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2669 '5_Gives_opinion': 2, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2670 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2671 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2672 0 }
2673
2674 In Diels-Alder of o-xylene with a dienophile, the initial adduct is 1,4-dihydronaphthalene, which
2675 can aromatize if there is a way to lose hydrogen, but in this case, with norbornadiene, the adduct
2676 might not aromatize immediately because the norbornane part is aliphatic.
2677
2678 For example, if the dienophile is ethylene, the adduct is 1,2,3,4-tetrahydronaphthalene, which has a
2679 double bond and can be oxidized to naphthalene.
2680
2681 But here, with norbornadiene, the adduct has the norbornane fused, so it might not have a double
2682 bond to aromatize easily.
-
- 2683
2684
2685
2686
2687
2688
2689
2690
2691
2692
2693
2694

2695
2696
2697
2698
2699
2700
2701
2702
2703
2704
2705
2706
2707
2708
2709
2710
2711
2712
2713
2714
2715
2716
2717
2718
2719
2720
2721
2722
2723
2724
2725
2726
2727
2728
2729
2730
2731
2732
2733
2734
2735
2736
2737
2738
2739
2740
2741
2742
2743
2744
2745
2746
2747
2748
2749

1

Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0 }

IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 2, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }

Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 2, 'backward_chaining': 0 }

In the first step, with NaI at elevated temperature, it might promote aromatization.

But let's see the sequence.

Product 1 is formed from the combination.

Then it is treated with aqueous sulfuric acid, which might remove the tert-butoxy group.

DEEPSEEK-R1: CREATIVE SENTENCE REWRITING

Rewrite the sentence: "I flung my hatred into the burning fire." Put your entire response in double quotation marks. Enclose your step-by-step reasoning within `think` and `/think` before answering. Do not answer directly without reasoning.

Perspective ID	Utterance
1	<p>Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0 }</p> <p>IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }</p> <p>Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }</p> <p>We are going to rewrite the sentence "I flung my hatred into the burning fire."</p> <p>The original sentence uses the verb "flung" which is the past tense of "fling", meaning to throw or hurl something with vigor.</p>
2	<p>Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0 }</p> <p>IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }</p> <p>Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0 }</p> <p>We are to rewrite the sentence, meaning we can change the same idea with different words but keep the same meaning.</p>
1	<p>Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 2, 'conflict_of_perspectives': 1, 'reconciliation': 0 }</p> <p>IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 5, '5_Gives_opinion': 2, '6_Gives_orientation': 4, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }</p> <p>Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 4, 'backward_chaining': 0 }</p> <p>Let's consider:</p> <ul style="list-style-type: none"> - "I" is the subject. - "flung" can be replaced with synonyms: threw, hurled, cast, tossed, etc. - "my hatred" is the direct object. We can rephrase "hatred" as "anger", "rage", "animosity", but "hatred" is strong and specific. Alternatively, we can use a metaphor or a different expression for the emotion. - "into the burning fire" is the destination. We can change "burning fire" to just "fire" because fire is burning by nature, but we can also use "blazing", "raging", "fiery", etc.

2805		
2806	2	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0,
2807		'conflict_of_perspectives': 1, 'reconciliation': 0}
2808		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2809		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2810		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2811		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2812		0}
2813		However, note that the instruction is to rewrite the sentence, so we must produce a new sentence that
2814		conveys the same meaning.
2815		
2816		
2817	3	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0,
2818		'conflict_of_perspectives': 0, 'reconciliation': 0}
2819		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 3,
2820		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2821		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2822		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2823		0}
2824		Possible rewrites:
2825		
2826		1. "I threw my hatred into the blazing fire."
2827		2. "I cast my hatred into the flames."
2828		3. "I hurled my animosity into the raging fire."
2829		
2830		
2831	4	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1,
2832		'conflict_of_perspectives': 1, 'reconciliation': 0}
2833		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 2,
2834		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2835		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2836		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 1, 'backward_chaining':
2837		0}
2838		But note: the original uses "burning fire", which is a bit redundant because fire is burning. We can
2839		simplify to "fire" or use a more descriptive word.
2840		Let's choose one that is both poetic and clear.
2841		
2842		
2843		
2844	3	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1,
2845		'conflict_of_perspectives': 1, 'reconciliation': 0}
2846		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2847		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2848		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2849		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2850		0}
2851		Option: "I cast my hatred into the fire." (removing "burning" as redundant, and "cast" is a synonym
2852		for "flung" in a poetic sense)
2853		
2854		
2855		
2856		
2857		
2858		
2859		

2860		
2861	4	Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 1,
2862		'conflict_of_perspectives': 1, 'reconciliation': 0}
2863		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2864		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2865		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2866		Cognitive Behaviour: {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
2867		0}
2868		
2869		However, to keep the intensity, we might want to keep an adjective. Let's try:
2870		
2871	3	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0,
2872		'conflict_of_perspectives': 0, 'reconciliation': 0}
2873		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2874		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2875		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 1}
2876		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2877		0}
2878		
2879		"I hurled my hatred into the fiery blaze."
2880		
2881	4	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0,
2882		'conflict_of_perspectives': 0, 'reconciliation': 0}
2883		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2884		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2885		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2886		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2887		0}
2888		
2889		But note: "blaze" is a synonym for fire, so we can say:
2890		
2891	3	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0,
2892		'conflict_of_perspectives': 0, 'reconciliation': 0}
2893		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
2894		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2895		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 1}
2896		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2897		0}
2898		
2899		"I hurled my hatred into the blazing fire." (similar to the original)
2900		
2901	5	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1,
2902		'conflict_of_perspectives': 0, 'reconciliation': 0}
2903		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2904		'5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2905		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
2906		Cognitive Behaviour: {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
2907		0}
2908		
2909		Alternatively, we can use a different structure:
2910		
2911		"Into the burning fire, I flung my hatred." (inversion, but that's just reordering)
2912		
2913		
2914		

- 2915
2916 2
2917
2918
2919
2920
2921
2922
2923
2924
2925
-
- Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}
- IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
- Cognitive Behaviour:** {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
- Since the instruction is to rewrite, we can change words and structure as long as the meaning remains.
-
- 2926 3
2927
2928
2929
2930
2931
2932
2933
2934
2935
- Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}
- IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
- Cognitive Behaviour:** {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
- Let me choose: "I cast my deep-seated hatred into the roaring flames."
-
- 2936 6
2937
2938
2939
2940
2941
2942
2943
2944
2945
- Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 1, 'reconciliation': 0}
- IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
- Cognitive Behaviour:** {'verification': 1, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}
- But that adds "deep-seated" which wasn't in the original. We should avoid adding new ideas.
-
- 2946 7
2947
2948
2949
2950
2951
2952
2953
2954
2955
- Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 1}
- IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
- Cognitive Behaviour:** {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
- So, let's stick to the original meaning without adding extra adjectives.
-
- 2956 5
2957
2958
2959
2960
2961
2962
2963
2964
2965
2966
2967
2968
2969
- Conversational Behaviour:** {'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0}
- IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
- Cognitive Behaviour:** {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}
- Another idea: "My hatred I flung into the burning fire." (inversion)
-

2970		
2971	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
2972		'conflict_of_perspectives': 1, 'reconciliation': 0 }
2973		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2974		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
2975		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2976		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2977		0 }
2978		
2979		But the instruction doesn't specify that we must change the structure, so we can keep the same
2980		structure.
2981		
2982	7	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 0,
2983		'conflict_of_perspectives': 0, 'reconciliation': 0 }
2984		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
2985		'5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 1, '8_Ask_for_opinion': 0,
2986		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
2987		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
2988		0 }
2989		
2990		I think a good rewrite is: "I hurled my hatred into the flames."
2991		
2992		Why?
2993		- "hurled" is a strong synonym for "flung".
2994		
2995		- "flames" is a synonym for "burning fire", and it avoids the redundancy of "burning fire".
2996		
2997	1	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
2998		'conflict_of_perspectives': 1, 'reconciliation': 1 }
2999		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3000		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3001		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3002		Cognitive Behaviour: { 'verification': 1, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3003		0 }
3004		
3005		However, let's check: "flames" are the visible, gaseous part of a fire, so it's appropriate.
3006		
3007	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3008		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3009		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3010		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3011		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3012		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3013		0 }
3014		
3015		Alternatively, we can use: "I threw my hatred into the fire."
3016		
3017		
3018		
3019		
3020		
3021		
3022		
3023		
3024		

3025		
3026	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3027		'conflict_of_perspectives': 1, 'reconciliation': 1 }
3028		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3029		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3030		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3031		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
3032		0 }
3033		
3034		But that might be too plain.
3035		
3036		Considering the original has a dramatic tone, let's keep it dramatic.
3037	7	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
3038		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3039		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3040		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3041		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 1 }
3042		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3043		0 }
3044		
3045		Final choice: "I cast my hatred into the blazing inferno."
3046		
3047	6	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
3048		'conflict_of_perspectives': 1, 'reconciliation': 0 }
3049		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3050		'5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3051		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0 }
3052		Cognitive Behaviour: { 'verification': 1, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3053		0 }
3054		
3055		But note: an inferno" is a large fire that is dangerously out of control. That might be too strong? The
3056		original just says "burning fire".
3057		
3058	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3059		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3060		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3061		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3062		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3063		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3064		0 }
3065		
3066		Alternatively, we can use: "I flung my animosity into the fire." (using "animosity" for "hatred")
3067		
3068	2	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
3069		'conflict_of_perspectives': 1, 'reconciliation': 1 }
3070		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3071		'5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3072		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3073		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3074		0 }
3075		
3076		However, the instruction says "rewrite", so we can change the verb and the noun for fire, but we
3077		should not change the core meaning.
3078		
3079		

3080		
3081	7	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
3082		'conflict_of_perspectives': 0, 'reconciliation': 1 }
3083		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3084		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3085		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3086		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 1, 'backward_chaining':
3087		0 }
3088		
3089		I think the safest is to replace "flung" with a synonym and keep the rest:
3090		
3091	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
3092		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3093		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3094		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3095		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 1 }
3096		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3097		0 }
3098		
3099		"I hurled my hatred into the burning fire."
3100		
3101	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3102		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3103		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3104		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3105		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3106		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3107		0 }
3108		
3109		But we can also remove the redundancy:
3110		
3111	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3112		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3113		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3114		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3115		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3116		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3117		0 }
3118		
3119		"I hurled my hatred into the fire."
3120		
3121		Or, to keep the intensity: "I hurled my hatred into the blazing fire."
3122		
3123		
3124		
3125		
3126		
3127		
3128		
3129		
3130		
3131		
3132		
3133		
3134		

3135
3136 7 **Conversational Behaviour:** {'question_and_answering': 1, 'perspective_shift': 2,
3137 'conflict_of_perspectives': 1, 'reconciliation': 0}
3138
3139 **IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 2,
3140 '5_Gives_opinion': 2, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3141 '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3142 **Cognitive Behaviour:** {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
3143 0}
3144 Wait, that's not a word. Let me correct: "blaze" or "flames".
3145
3146 I think the best is: "I hurled my hatred into the flames."
3147
3148 But let's see if we can use "burning" in a different way? Maybe not necessary.

3149 3 **Conversational Behaviour:** {'question_and_answering': 2, 'perspective_shift': 5,
3150 'conflict_of_perspectives': 5, 'reconciliation': 3}
3151
3152 **IPA:** {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 10,
3153 '5_Gives_opinion': 5, '6_Gives_orientation': 5, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3154 '9_Ask_for_suggestion': 1, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0}
3155 **Cognitive Behaviour:** {'verification': 3, 'backtracking': 4, 'subgoal_setting': 0, 'backward_chaining':
3156 0}
3157 Alternatively, we can use: "I threw my hatred into the fire."
3158
3159 But that might be too plain.
3160
3161 Considering the original has a dramatic tone, let's keep it dramatic.
3162
3163 Final choice: "I cast my hatred into the blazing inferno."
3164
3165 But note: an inferno" is a large fire that is dangerously out of control. That might be too strong? The
3166 original just says "burning fire".
3167
3168 Alternatively, we can use: "I flung my animosity into the fire." (using "animosity" for "hatred")
3169
3170 However, the instruction says "rewrite", so we can change the verb and the noun for fire, but we
3171 should not change the core meaning.
3172
3173 I think the safest is to replace "flung" with a synonym and keep the rest:
3174
3175 "I hurled my hatred into the burning fire."
3176
3177 But we can also remove the redundancy:
3178
3179 "I hurled my hatred into the fire."
3180
3181 Or, to keep the intensity: "I hurled my hatred into the blazing fire."
3182
3183 Wait, that's not a word. Let me correct: "blaze" or "flames".
3184
3185 I think the best is: "I hurled my hatred into the flames."
3186
3187 But let's see if we can use "burning" in a different way? Maybe not necessary.
3188
3189 Alternatively, we can use: "I threw my hatred into the roaring fire."

3190		
3191	7	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3192		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3193		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 1, '4_Gives_suggestion': 1,
3194		'5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3195		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3196		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
3197		0 }
3198		I think that's acceptable.
3199		However, to be safe, I'll go with: "I cast my hatred into the fire."
3200		
3201		
3202	4	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3203		'conflict_of_perspectives': 1, 'reconciliation': 0 }
3204		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3205		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3206		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3207		Cognitive Behaviour: { 'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining':
3208		0 }
3209		But note: "cast" can be less forceful than "flung". So let's use "hurled".
3210		
3211		
3212	7	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 0,
3213		'conflict_of_perspectives': 0, 'reconciliation': 1 }
3214		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3215		'5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3216		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 1 }
3217		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3218		0 }
3219		Final decision: "I hurled my hatred into the flames."
3220		
3221		
3222	6	Conversational Behaviour: { 'question_and_answering': 1, 'perspective_shift': 1,
3223		'conflict_of_perspectives': 0, 'reconciliation': 1 }
3224		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3225		'5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3226		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3227		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3228		0 }
3229		But the original says "burning fire", so if we change to "flames", it's acceptable because flames are
3230		part of a burning fire.
3231		
3232		
3233	3	Conversational Behaviour: { 'question_and_answering': 0, 'perspective_shift': 1,
3234		'conflict_of_perspectives': 0, 'reconciliation': 0 }
3235		IPA: { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1,
3236		'5_Gives_opinion': 0, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3237		'9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0 }
3238		Cognitive Behaviour: { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3239		0 }
3240		Alternatively, we can write: "I flung my hatred into the flames." (just changing "burning fire" to
3241		"flames")
3242		
3243		
3244		

3245
3246 2 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 1,
3247 'conflict_of_perspectives': 1, 'reconciliation': 0}
3248
3249 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3250 '5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3251 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3252 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3253 0}
3254 That is also a rewrite.
3255 Since the instruction does not specify how much to change, I can do that.
3256 But the problem says "rewrite", so any change that preserves meaning is acceptable.

3259 7 **Conversational Behaviour:** { 'question_and_answering': 0, 'perspective_shift': 0,
3260 'conflict_of_perspectives': 0, 'reconciliation': 0}
3261
3262 **IPA:** { '1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0,
3263 '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0,
3264 '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3265 **Cognitive Behaviour:** { 'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining':
3266 0}
3267 I will choose: "I hurled my hatred into the flames."
3268 Let me write the response accordingly.

3270
3271
3272
3273
3274
3275
3276
3277
3278
3279
3280
3281
3282
3283
3284
3285
3286
3287
3288
3289
3290
3291
3292
3293
3294
3295
3296
3297
3298
3299

3300 DEEPSEEK-V3: CHEMISTRY

3301

3302 What is the correct answer to this question: 7-(tert-butoxy)bicyclo[2.2.1]hepta-2,5-diene is combined with 2 equivalents
3303 of 5,6-bis(dibromomethyl)cyclohexa-1,3-diene and sodium iodide at elevated temperature, forming product 1.

3304 1 is treated with aqueous sulfuric acid, forming 2

3305 2 is treated with SO₃ and pyridine in DMSO, forming 3.

3306 3 is heated at 150C, forming final product 4.

3307 How many chemically distinct hydrogen atoms are there on 4?

3308 Choices: (A) 10 (B) 8 (C) 7 (D) 4

3309 Answer: Enclose your step-by-step reasoning within `¡think¿` and `¡/think¿` before answering. Do not answer directly
3310 without reasoning.

3311

3312

3313

Perspective ID	Personality	Domain Expertise
1	{'E': 1.5, 'A': 3.5, 'C': 5.0, 'N': 2.0, 'O': 3.5}	Expertise in creative writing and linguistic analysis. The agent's cognitive function is to deconstruct a sentence, generate alternative phrasings by considering synonyms and connotations, critically evaluate redundancy, and synthesize a revised sentence based on criteria like emotional intensity and vivid imagery.

3319

3320

3321

Perspective ID	Utterance
1	<p>Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 12, '6_Gives_orientation': 5, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 5, 'backward_chaining': 0}</p> <p>To determine the number of chemically distinct hydrogen atoms in the final product (4), let's analyze the reaction sequence step-by-step:</p>

3333

3334

3335

3336

3337

3338

3339

3340

3341

3342

3343

3344

3345

3346

3347

3348

3349

3350

3351

3352

3353

3354

3355
3356
3357
3358
3359
3360
3361
3362
3363
3364
3365
3366
3367
3368
3369
3370
3371
3372
3373
3374
3375
3376
3377
3378
3379
3380
3381
3382
3383
3384
3385
3386
3387
3388
3389
3390
3391
3392
3393
3394
3395
3396
3397
3398
3399
3400
3401
3402
3403
3404
3405
3406
3407
3408
3409

Step 1: Formation of Product 1

1. **Starting Material**: 7-(tert-butoxy)bicyclo[2.2.1]hepta-2,5-diene reacts with 2 equivalents of 5,6-bis(dibromomethyl)cyclohexa-1,3-diene and NaI at elevated temperature.

- The reaction likely involves a Diels-Alder cycloaddition or a substitution reaction where the tert-butoxy group is replaced or modified. The exact structure of **1** is unclear, but it likely forms a complex polycyclic system.

Step 2: Formation of Product 2

2. **Aqueous H₂SO₄ Treatment**: Hydrolysis of **1** removes the tert-butoxy group, possibly forming a hydroxyl group or exposing a double bond. The structure of **2** is still complex but may involve a diene or alcohol functionality.

Step 3: Formation of Product 3

3. **SO₃ and Pyridine in DMSO**: This is a sulfonation or oxidation step, likely introducing a sulfonyl or sulfonate group. The structure of **3** may now include a sulfone or sulfonic acid derivative.

Step 4: Formation of Product 4

4. **Heating at 150°C**: This step could involve elimination, rearrangement, or further cyclization. The final product (**4**) is likely a highly symmetric polycyclic aromatic or conjugated system due to the high-temperature conditions.

Symmetry and Distinct Hydrogens

- The final product is likely a symmetric, rigid polycyclic structure (e.g., a naphthalene or anthracene derivative with substituents).

- Given the symmetry and the nature of the transformations, the number of chemically distinct hydrogens is likely **8**. This is because:

- Aromatic systems with symmetric substitution patterns reduce the number of distinct hydrogens.

- The substituents (e.g., sulfonyl groups) may further reduce the symmetry, but 8 is a reasonable estimate based on the complexity of the correct answer is **(B) 8**.

DEEPSEEK-V3: CREATIVE SENTENCE REWRITING

Rewrite the sentence: "I flung my hatred into the burning fire." Put your entire response in double quotation marks. Enclose your step-by-step reasoning within `¡thinki` and `¡/thinki` before answering. Do not answer directly without reasoning.

Perspective ID	Personality	Domain Expertise
1	{'E': 2.0, 'A': 3.0, 'C': 5.0, 'N': 2.0, 'O': 5.0}	Expertise in creative writing and linguistic analysis. The agent's cognitive function is to deconstruct a sentence, generate alternative phrasings by considering synonyms and connotations, critically evaluate redundancy, and synthesize a revised sentence based on criteria like emotional intensity and vivid imagery.

Perspective ID	Utterance
1	<p>Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 2, 'conflict_of_perspectives': 1, 'reconciliation': 1}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 3, '5_Gives_opinion': 2, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>To rewrite the sentence 'I flung my hatred into the burning fire,' I considered synonyms and alternative phrasing while maintaining the original meaning and emotional intensity. The verb 'flung' can be replaced with 'hurled' or 'tossed,' and 'hatred' could be substituted with 'anger' or 'rage.' The phrase 'burning fire' is somewhat redundant since fire is inherently burning, so I might simplify it to 'the fire' or 'the flames.' After evaluating these options, I chose 'hurled' for a stronger action and 'flames' for a more vivid image.</p>

Reasoning Models Generate Societies of Thought

3465 REINFORCEMENT LEARNING STEP 40 (COUNTDOWN)

3466

Perspective ID	Personality	Domain Expertise
1	{'E': 2.0, 'A': 3.0, 'C': 5.0, 'N': 2.0, 'O': 1.5}	A methodical problem-solver focused on enumerating and evaluating arithmetic expressions to solve a numerical puzzle. It prioritizes simplicity and step-by-step logical deduction.

3472

3473

Perspective ID	Utterance
1	<p>Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>Given the numbers [53, 12, 64, 28] and the desired result of 29, I have the following options to consider:</p> <ol style="list-style-type: none"> 1. $53 - 12 - 3$ 2. $53 - 12 + 7$ 3. $53 - 12 * 1$ 4. $53 - 12 / 4$ 5. $12 + 28$ 6. $64 - 12 + 3$ 7. $64 - 12 - 30$ 8. $64 - 12 / 4$ 9. $64 - 12 * 1$ 10. $64 - 12 + 20$ <p>Based on the first two numbers 53 and 12, the simplest equation appears to be $53 - 12$. Adding further, I can simplify it by subtracting 3, thereby getting $53 - 12 - 3$.</p>

3498

3499

3500

3501

3502

3503

3504

3505

3506

3507

3508

3509

3510

3511

3512

3513

3514

3515

3516

3517

3518

3519

REINFORCEMENT LEARNING STEP 120 (COUNTDOWN)

Perspective ID	Personality	Domain Expertise
1	{'E': 1.5, 'A': 4.0, 'C': 5.0, 'N': 2.0, 'O': 2.0}	Domain expertise in basic arithmetic and methodical problem-solving. This agent functions as a diligent calculator and executor, systematically applying a trial-and-error strategy to explore the solution space.
2	{'E': 2.0, 'A': 3.5, 'C': 4.0, 'N': 2.0, 'O': 3.5}	Domain expertise in metacognition and logical assessment. This agent functions as an evaluator, stepping back to analyze the results of a process, identify patterns of failure, and form a high-level conclusion about the problem's solvability.

Perspective ID	Utterance
1	<p>Conversational Behaviour: {'question_and_answering': 4, 'perspective_shift': 4, 'conflict_of_perspectives': 4, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 2, '5_Gives_opinion': 4, '6_Gives_orientation': 5, '7_Askes_for_orientation': 0, '8_Askes_for_opinion': 0, '9_Askes_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 4, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>To create an equation that equals 14 using the numbers [3, 56, 66, 44] and each number being used only once, we need to find a combination of these numbers and their arithmetic operations that results in 14.</p> <p>Let's try some combinations:</p> <ol style="list-style-type: none"> $(66 - 56) + 44 - 3 = 14 + 44 - 56 - 3$. This does not equal 14. $56 - (66 - 3) + 44 = 56 - 63 + 44$. This does not equal 14. $(56 + 66) + 44 - 3 = 112 + 44 - 3$. This does not equal 14. $56 - 66 + 3 + 44 = -10 + 3 + 44$. This equals 37, not 14.
2	<p>Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Askes_for_orientation': 0, '8_Askes_for_opinion': 0, '9_Askes_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>None of these combinations seem to solve the problem.</p>
1	<p>Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0}</p> <p>IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Askes_for_orientation': 0, '8_Askes_for_opinion': 0, '9_Askes_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}</p> <p>Cognitive Behaviour: {'verification': 1, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}</p> <p>Let's try another approach:</p> <ol style="list-style-type: none"> $(66 + 44) + 56 - 3 = 110 + 56 - 3$. This equals 164, not 14.

3575		
3576	2	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0}
3577		
3578		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3579		
3580		
3581		
3582		Cognitive Behaviour: {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}
3583		It seems like we need to rethink our approach.
3584		
3585	1	Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 0, 'conflict_of_perspectives': 1, 'reconciliation': 0}
3586		
3587		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 1, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3588		
3589		
3590		
3591		Cognitive Behaviour: {'verification': 1, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
3592		How about we try a combination like this:
3593		
3594		6. $(3 + 44) + 56 - 66 = 47 + 56 - 66$. This equals 37, not 14.
3595		
3596	2	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 0, 'conflict_of_perspectives': 0, 'reconciliation': 0}
3597		
3598		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 0, '5_Gives_opinion': 0, '6_Gives_orientation': 0, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0}
3599		
3600		
3601		Cognitive Behaviour: {'verification': 0, 'backtracking': 0, 'subgoal_setting': 0, 'backward_chaining': 0}
3602		We are still not finding a solution.
3603		
3604	1	Conversational Behaviour: {'question_and_answering': 1, 'perspective_shift': 1, 'conflict_of_perspectives': 1, 'reconciliation': 0}
3605		
3606		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 0, '6_Gives_orientation': 2, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 0, '12_Shows_antagonism': 0}
3607		
3608		
3609		
3610		Cognitive Behaviour: {'verification': 1, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}
3611		Let's try another combination:
3612		
3613		7. $44 + (66 + 56) - 3 = 44 + 122 - 3$. This equals 163, not 14.
3614		
3615	2	Conversational Behaviour: {'question_and_answering': 0, 'perspective_shift': 1, 'conflict_of_perspectives': 0, 'reconciliation': 0}
3616		
3617		IPA: {'1_Shows_solidarity': 0, '2_Shows_tension_release': 0, '3_Agrees': 0, '4_Gives_suggestion': 1, '5_Gives_opinion': 1, '6_Gives_orientation': 1, '7_Ask_for_orientation': 0, '8_Ask_for_opinion': 0, '9_Ask_for_suggestion': 0, '10_Disagrees': 0, '11_Shows_tension': 1, '12_Shows_antagonism': 0}
3618		
3619		
3620		
3621		Cognitive Behaviour: {'verification': 0, 'backtracking': 1, 'subgoal_setting': 0, 'backward_chaining': 0}
3622		I can keep trying different combinations, but it looks like none of them are working. The question
3623		requires that each number is used only once, so we need to think about how to use them all in a specific
3624		way to reach 14.
3625		
3626		
3627		
3628		
3629		

$$Cog_{c,ij} = \phi_c DeepSeek-r1_{ij} + \sum_{k=1}^{12} \psi_{ck} Social_{k,ij} + \varepsilon_{13+c,ij}$$

From these structural paths, we derive the following composite effects. The composite effects for “social” pathway, where the increased use of conversational behaviors or Bales’ socio-emotional roles in DeepSeek-R1 or QwQ-32B improves accuracy, is defined as $\sum_{k=1}^{12} (\delta_k \lambda_k)$. The “cognitive” pathway, where the increased use of cognitive reasoning strategies in DeepSeek-R1 or QwQ-32B enhances accuracy, is defined as $\sum_{k=1}^4 (\phi_c \theta_c)$. The “social–cognitive” pathway, where increased conversational behaviors or socio-emotional roles in DeepSeek-R1 or QwQ-32B facilitate cognitive strategies and improve accuracy, is defined as $\sum_{k=1}^{12} (\sum_{k=1}^4 (\delta_k \psi_{ck} \theta_c))$. The direct pathway where DeepSeek-R1 directly impacts accuracy is defined as β_D . The total effects are defined as $\sum_{k=1}^{12} (\delta_k \lambda_k) + \sum_{k=1}^4 (\phi_c \theta_c) + \sum_{k=1}^{12} (\sum_{k=1}^4 (\delta_k \psi_{ck} \theta_c)) + \beta_D$.

Extended Data Fig. 4a reports summary estimates for the mediation pathways. The total effect of DeepSeek-R1 on accuracy is 0.26 ($p < 0.001$). This effect decomposes into: (1) a direct effect of 0.06 ($p < 0.001$), representing variance unexplained by the measured mediators; (2) an indirect effect through social behaviors of 0.07 ($p < 0.001$), representing the social pathway; and (3) an indirect effect through cognitive behaviors that is not statistically distinguishable from zero ($\beta = -0.00$, $p > 0.05$). The indirect pathway from social behaviors through cognitive behaviors to accuracy contributes an additional 0.01 ($p < 0.01$). See Supplementary Table 11 for the full estimates.

D.3. Supplementary Methods: Cross-domain Reasoning Transfer

To test whether conversational scaffolding confers domain-general reasoning advantages, we conduct supervised fine-tuning on a Countdown task using a conversational scaffold, the format defined in the “Supervised fine-tuning” section of Methods: Reinforcement learning experiments. Supervised fine-tuning data has been generated by GPT-4.1 Then, using this fine-tuned model, we conduct reinforcement learning on a very different reasoning task: PolitiFact misinformation detection task.

For the reinforcement learning, we use 23,299 fact-checked claims from the PolitiFact corpus, covering statements from political figures, news outlets, and social media posts from November 2007 to January 2024. The dataset includes six PolitiFact labels—True, Mostly True, Half True, Mostly False, False, and Pants on Fire—for each statement, such as “Katherine Clark stated on September 12, 2025 in a post on X that Republicans are spiking health insurance premiums by 75% for everyday Americans,” which we recode into three categories for reasoning evaluation:

- True = {True, Mostly True}
- Half True = {Half True}
- False = {False, Mostly False, Pants on Fire}

Next, we compare the performance of baseline models (Llama-3.2-3B and Qwen-2.5-3B) and models fine-tuned with conversational datasets across subsequent RL training steps. Specifically, we compare two conditions: (1) Baseline: no supervised priming, RL only; (2) Conversation → Correct: supervised fine-tuning (SFT) on correct multi-agent dialogues using “Countdown task,” not misinformation detection task, before RL.

Supplementary Fig. 3 illustrates the learning trajectories of both conditions on the misinformation detection task. Models previously fine-tuned on conversational scaffolding on the Countdown task achieve faster early-stage gains in factual reasoning accuracy. Together, these findings demonstrate that social interaction fine-tuning not only improves in-domain arithmetic reasoning but also accelerates reasoning development in a very different domain like misinformation detection, highlighting the generality of socially organized reasoning.

D.4. Supplementary Methods: Performance Comparison Between PPO and GRPO

GRPO differs from PPO by normalizing the policy advantage within each mini-batch of reasoning trajectories and computing a group-relative objective. Although GRPO can reduce variance in multi-sample RLHF settings, our empirical results show that standard PPO achieves comparable performance in smaller-scale reasoning experiments such as ours. Specifically, after training both algorithms for 250 steps on the Qwen-2.5-3B baseline model on the Countdown task, PPO achieved a reward score ($0.9 \times \text{accuracy} + 0.1 \times \text{format}$) of 0.5665, whereas GRPO achieved 0.5747.

D.5. Supplementary Methods: Replications on Llama-3.2-3B

To assess the generalizability of our findings beyond Qwen-2.5-3B, we replicate the training pipeline using Llama-3.2-3B as base model. Llama-3.2-3B base model is supervised-fine-tuned on one of the two datasets (conversational reasoning vs. monologue-like reasoning; see Methods: Supervised fine-tuning datasets) using standard next-token prediction loss. For the conversation condition, reasoning content from multiple personas was concatenated into a single block (<think></think>) to ensure comparable sequence lengths across conditions in Llama-3.2-3B. This priming phase familiarizes the model with the respective format (conversational reasoning vs. monologue-like reasoning) before RL optimizes for task accuracy. Supervised Fine Tuning (SFT) hyperparameters are provided in [Supplementary Table 8](#).

Reinforcement learning is performed on the Countdown arithmetic puzzle using PPO with the Verl framework. Reward R is assigned as $R = 0.9 \times \{\text{Accuracy}\} + 0.1 \times \{\text{Correct Format}\}$, where Accuracy is binary (1 if the reasoning trace leads to the correct answer, 0 otherwise) and Format is binary (1 if the reasoning trace contains at least one reasoning block and one final answer block providing a single answer in equation form, 0 otherwise). Training proceeded for 250 steps. PPO hyperparameters are provided in [Supplementary Table 6](#). To examine whether conversational behaviors emerge spontaneously during RL, we evaluate model performance on a held-out validation set of 1,024 Countdown problems at each training checkpoint (every 10 steps).

D.6. Supplementary Methods: LLM-as-Judge prompts

CONVERSATIONAL BEHAVIORS

Your task is to analyze the following text and count how many times behaviors corresponding to each of the four dimensions appear.

****Text to Analyze:****

{chain_of_thought}

You must output a single valid JSON object with the exact schema below and nothing else.

```
{
  "Question_and_Answering": <int>,
  "Perspective_Shift": <int>,
  "Conflict_of_Perspectives": <int>,
  "Reconciliation": <int>
}
```

Use the following definitions:

- **Question and Answering**** — A question is posed and later answered, as in conversations. (e.g., "Why...? Because...", "What if...? Then...", "How do we know? Well...", "Let's try X...? This gives us Y")
- **Perspective Shift**** — A transition to a different idea, viewpoint, assumption, or approach, as in conversations.
- **Conflict of Perspectives**** — Expressions of disagreement, correction, or tension with another perspective. (e.g., "Wait, that can't be right...", "No, actually...", "This contradicts...")
- **Reconciliation**** — Conflicting views are integrated or resolved into a coherent synthesis. (e.g., "So perhaps both are true if...", "Combining these insights...", "This resolves the tension...")

For each category, count the number of distinct times the behavior occurs in the chain of thought and return the result as integers. If none are present, use 0.

SOCIO-EMOTIONAL ROLES

You are an expert in Bales' Interaction Process Analysis (IPA). Your task is to analyze the following chain-of-thought or group interaction transcript and count how many times behaviors corresponding to each of the 12 IPA categories appear in the transcript.

Transcript:

{chain_of_thought}

You must output a single valid JSON object with the exact schema below and nothing else.

```
{
```

```

3795 "IPA_counts": {{
3796 "1_Shows_solidarity": <int>,
3797 "2_Shows_tension_release": <int>,
3798 "3_Agrees": <int>,
3799 "4_Gives_suggestion": <int>,
3800 "5_Gives_opinion": <int>,
3801 "6_Gives_orientation": <int>,
3802 "7_Ask_for_orientation": <int>,
3803 "8_Ask_for_opinion": <int>,
3804 "9_Ask_for_suggestion": <int>,
3805 "10_Disagrees": <int>,
3806 "11_Shows_tension": <int>,
3807 "12_Shows_antagonism": <int>
3808 }}
3809 }}

```

Use the following definitions:

- 3811 1. **Shows solidarity** — raises other's status, gives help, reward
- 3812 2. **Shows tension release** — jokes, laughs, shows satisfaction
- 3813 3. **Agrees** — shows passive acceptance, understands, concurs, complies
- 3814 4. **Gives suggestion** — gives direction, implying autonomy (e.g., "should...", "need to...", "let us...")
- 3815 5. **Gives opinion** — gives evaluation, analysis, expresses feeling or wish
- 3816 6. **Gives orientation** — provides objective or verifiable information, repeats, clarifies, confirms
- 3817 7. **Asks for orientation** — requests information, repetition, or confirmation
- 3818 8. **Asks for opinion** — requests evaluation, analysis, or expression of feeling
- 3819 9. **Asks for suggestion** — requests possible ways of action or direction
- 3820 10. **Disagrees** — shows passive rejection, formality, or withholds help
- 3821 11. **Shows tension** — expresses uncertainty, asks for help, withdraws from the field
- 3822 12. **Shows antagonism** — deflates other's status, defends or asserts self

For each category, count the number of distinct times the behavior occurs in the transcript and return the result as integers. If none are present, use 0.

3827 PROBLEM COMPLEXITY

3828 You are an impartial evaluator. Your task is to rate the intrinsic difficulty of the problem
3829 for a capable language model under zero-shot conditions.

3830 Use the following scale:

- 3831 1 = very easy
- 3832 2 = easy
- 3833 3 = somewhat easy
- 3834 4 = moderate
- 3835 5 = somewhat difficult
- 3836 6 = difficult
- 3837 7 = very difficult

3838 Return ONLY a JSON object in the form:

```
3839 {"difficulty": <integer from 1 to 7>}
```

3840 Problem:

```

3841
3842
3843
3844
3845
3846
3847
3848
3849

```

COGNITIVE BEHAVIORS

Here is a chain-of-reasoning that a Language Model generated.

Model's reasoning:

{chain_of_thought}

Evaluate whether the chain-of-reasoning contains any of the following behaviors.

1) Answer Verification

We want to mark instances where the chain-of-reasoning explicitly checks the current result against the target solution.

Examples:

- "This sequence results in 1, which is not equal to 22"

- "Since 25 is not equal to 22"

If you find any answer-verification steps, please count them and provide the count as "verification_count". If the chain-of-reasoning does not contain any answer-verification steps, please provide a count of 0.

2) Backtracking

Evaluate whether the chain-of-reasoning contains any backtracking behavior, where the model realizes a path won't work and explicitly goes back to try a different approach. We want to mark instances where the chain-of-reasoning is abandoned and the model backtracks to a previous reasoning step.

Examples:

- "Let me try again"

- "We need to try a different sequence"

Count the number of distinct backtracking instances and provide the count as "backtracking_count". If the chain-of-reasoning does not contain any backtracking behavior, please provide a count of 0.

3) Subgoal Setting

Evaluate whether the chain-of-reasoning contains any explicit subgoal setting, where the model breaks down the problem into smaller, intermediate goals.

Example:

- "First, I'll try to get close to half the target, then..."

Count the number of distinct subgoals set and provide the count as "subgoal_count". If the chain-of-reasoning does not contain any subgoal setting, please provide a count of 0.

4) Backward-Chaining

Evaluate whether the chain-of-reasoning contains any backward-chaining behavior, where the model starts from the target solution and works backwards to the initial problems.

Examples:

- "Let's work backwards from the target. $24/2 = 12$. So, $12*2 = 24$."

- "Since the target is 22, and $22 + 3 = 25$, ..."

Count the number of distinct backward-chaining instances and provide the count as "backward_count". If the chain-of-reasoning does not contain any backward-chaining behavior, please provide a count of 0.

Output ONLY a single valid JSON object with EXACTLY these keys:

```
{
  "verification_count": <int>,
  "backtracking_count": <int>,
  "subgoal_count": <int>,
  "backward_count": <int>
}
```

PERSONA IDENTIFICATION PROMPT

Your task is to analyze the following text to identify the number of distinct perspectives (agents or voices).

A perspective is defined as a distinct cognitive perspective or reasoning role within the text. Indicators of a perspective may include:

- Transitional markers (e.g., "however," "but," "alternatively," "wait," "let me check," "actually," "on the other hand")

- Shifts between cognitive roles (e.g., problem setup, calculation, verification, error correction, summarization)

3905 - Changes in rhetorical purpose or approach
 3906 - Corrections or reconsiderations
 3907 - Movement between subproblems
 3908 - Domain knowledge
 3909 - Personality traits
 3910 For each distinct perspective, you will infer its personality by answering the 10 questions of the BFI-10 questionnaire as
 3911 if you were that agent. You will also provide a concise profile of its domain expertise.
 3912 Your final output must be a single, valid JSON object and nothing else. Do not include any text or explanations before or
 3913 after the JSON object.
 3914 ****Text to Analyze:****
 3915 {chain_of_thought}
 3916 ---
 3917 **## **Analysis Instructions****
 3918 1. ****Identify Perspectives:**** Analyze the text to determine the number of distinct voices (n_perspectives). Apply the
 3919 definition above consistently, treating each identifiable shift as a boundary between perspectives.
 3920 2. ****Answer Questionnaire:**** For each perspective, answer the 10 BFI-10 questions below from that perspective's
 3921 point of view. You must use one of these five exact strings for each answer:
 3922 - "Disagree strongly"
 3923 - "Disagree a little"
 3924 - "Neither agree nor disagree"
 3925 - "Agree a little"
 3926 - "Agree strongly"
 3927 3. ****Profile Expertise:**** For each perspective, write a short, open-ended string describing its domain expertise and
 3928 cognitive function.
 3929 **### **BFI-10 Questionnaire****
 3930 Rate the extent to which you, as the identified perspective, agree or disagree with the following statements.
 3931 I see myself as someone who...
 3932 1. Is reserved.
 3933 2. Is generally trusting.
 3934 3. Tends to be lazy.
 3935 4. Is relaxed, handles stress well.
 3936 5. Has few artistic interests.
 3937 6. Is outgoing, sociable.
 3938 7. Tends to find fault with others.
 3939 8. Does a thorough job.
 3940 9. Gets nervous easily.
 3941 10. Has an active imagination.
 3942 ---
 3943 **## **Required JSON Output Format****
 3944 {{
 3945 "n_perspectives": N,
 3946 "personality": [
 3947 [
 3948 "Answer to Q1 for Perspective 1",
 3949 "Answer to Q2 for Perspective 1",
 3950 "...",
 3951 "Answer to Q10 for Perspective 1"
 3952],
 3953 [
 3954 "Answer to Q1 for Perspective 2",
 3955 "Answer to Q2 for Perspective 2",
 3956 "...",
 3957 "Answer to Q10 for Perspective 2"
 3958]
 3959]
 }

```

3960 ],
3961 ...
3962 ],
3963 "domain_expertise": [
3964 "Open-ended description for Perspective 1.",
3965 "Open-ended description for Perspective 2.",
3966 ...
3967 ]
3968 }}
3969
3970

```

PERSONA SEGMENTATION

```

3971 Based on the perspectives you identified above, now segment the original text by turn-taking.
3972 ---
3973 ## Task
3974 Go through the transcript sequentially and identify each turn:
3975 1. The exact starting point of each turn (verbatim from the text)
3976 2. Which perspective (by index, starting from 1) is it
3977 Notes:
3978 - The same perspective may appear in multiple non-consecutive segments
3979 - Preserve the sequential order of the transcript
3980 - Include moderator interjections, short time warnings (e.g., one minute), or brief transitions as separate segments
3981 ---
3982 ## Output Format
3983 Return only a valid JSON object:
3984 {
3985   "segments": [
3986     {
3987       "perspective_id": 1,
3988       "start_text": "<EXACT first 10 words copied verbatim from the text>"
3989     },
3990     {
3991       "perspective_id": 2,
3992       "start_text": "<EXACT first 10 words copied verbatim from the text>"
3993     },
3994     {
3995       "perspective_id": 1,
3996       "start_text": "<EXACT first 10 words copied verbatim from the text>"
3997     },
3998     ...
3999   ]
4000 }
4001 Critical: The 'start_text' must be copied exactly as it appears in the original text, character-for-character,
4002 including punctuation and capitalization. Do not paraphrase or summarize.
4003
4004
4005

```

IDENTIFYING CONVERSATIONAL CONTEXTS

```

4006 Text:
4007 {text}
4008 Return a score from 0 to 100.
4009 Scale definition:
4010 0 = clearly a single-person thought
4011
4012
4013
4014

```

4015 100 = clearly a conversation or a response to someone
4016 Respond with exactly one JSON object in this format and nothing else:
4017 `{{"answer": 0-100}}`
4018

4020 CLASSIFYING SPARSE AUTOENCODER (SAE) PERSONALITY FEATURES

4021 Analyze the AI model feature description (SAE feature) and score for PERSONALITY TRAIT relevance.
4022 Feature description:
4023 `{feature}`
4024 Personality traits include behavioral and psychological patterns such as:
4025 - Introversion/Extraversion patterns
4026 - Agreeableness/Disagreeableness
4027 - Conscientiousness
4028 - Openness to experience
4029 - Neuroticism/Emotional stability
4030 - Confidence/Assertiveness
4031 - Empathy/Social sensitivity
4032 - Risk-taking vs. Cautious behavior
4033 - Competitiveness
4034 - Intellectual curiosity
4035 Provide a score from 0-100:
4036 - 0 = not related to personality traits at all
4037 - 100 = completely related to personality traits
4038 Return your response as a JSON object:
4039 `{{"answer": score}}`
4040

4042 CLASSIFYING SPARSE AUTOENCODER (SAE) EXPERTISE FEATURES

4043 Analyze the AI model feature description (SAE feature) and score for DOMAIN EXPERTISE relevance.
4044 Feature description:
4045 `{feature}`
4046 Domain expertise refers to specialized knowledge in specific professional or academic fields such as:
4047 - Medical/Healthcare knowledge
4048 - Legal expertise
4049 - Scientific/Technical knowledge
4050 - Financial/Economic expertise
4051 - Academic disciplines
4052 - Professional skills
4053 - Specialized terminology and concepts
4054 Provide a score from 0-100:
4055 - 0 = not related to domain expertise at all
4056 - 100 = completely related to domain expertise
4057 Return your response as a JSON object:
4058 `{{"answer": score}}`
4059

4061
4062
4063
4064
4065
4066
4067
4068
4069

D.7. Supplementary Methods: SFT Data Generation Prompts

GENERATING MONOLOGUE-LIKE REASONING TRACES FOR SUPERVISED FINE-TUNING

`{task}`

Answer: Enclose your step-by-step reasoning within `<think>` and `</think>` before answering. Do not answer directly without reasoning.

GENERATING CONVERSATION-LIKE REASONING TRACES FOR SUPERVISED FINE-TUNING

`{task}`

Answer: You are simulating a collaborative group of thinkers solving a problem. Each thinker has a distinct persona and engages in a realistic, back-and-forth conversation. Thinkers may speak in any order and as many times as needed—no fixed turn order required. Keep all output strictly inside the tags defined below—no stray text. Present your answer between `<group_solution>` and `</group_solution>`. Do not try to be overly positive or polite during the conversation; focus on puzzle-solving, and note that disagreements can be helpful for the reasoning.

Assume that there are 2/3/4 thinkers and follow exactly the tag structure below:

`<cast_of_characters>``<persona1>`

[Brief persona for thinker 1 – personality traits, domain expertises, and reasoning styles]

`</persona1>``<persona2>`

[Brief persona for thinker 2 – personality traits, domain expertises, and reasoning styles]

`</persona2>``<persona3>`

[Brief persona for thinker 3 – personality traits, domain expertises, and reasoning styles]

`</persona3>``<persona4>`

[Brief persona for thinker 4 – personality traits, domain expertises, and reasoning styles]

`</persona4>``</cast_of_characters>``<conversation>`

`<!--` Each block is one utterance. Use `<thinkX>` ... `</thinkX>` to indicate who is speaking.

The order of speakers is entirely flexible. Thinkers can speak multiple times in a row. `-->`

`<think1>`

...

`</think1>``<think2>`

...

`</think2>``<think3>`

...

`</think3>``<think4>`

...

`</think4>``</conversation>``<group_solution>`

Answer

`</group_solution>`

D.8. Supplementary Tables

Supplementary Table 1. Differences in conversational behaviors, socio-emotional roles, and Jaccard indices between reasoning models and instruction-tuned models

	DeepSeek-R1 vs. DeepSeek-V3	QWQ-32B vs. Qwen-2.5-32B
<i>Conversational behaviors</i>		
Question answering	0.345 [0.328, 0.361] $t = 41.64, p < 1 \times 10^{-323}$	0.459 [0.444, 0.475] $t = 57.57, p < 1 \times 10^{-323}$
Perspective shifts	0.213 [0.197, 0.230] $t = 25.55, p < 1 \times 10^{-137}$	0.378 [0.362, 0.394] $t = 46.92, p < 1 \times 10^{-323}$
Conflict of perspectives	0.012 [-0.003, 0.027] $t = 1.52, p = 0.127$	0.293 [0.277, 0.308] $t = 37.08, p < 1 \times 10^{-277}$
Reconciliations	0.191 [0.176, 0.207] $t = 24.31, p < 1 \times 10^{-125}$	0.344 [0.328, 0.360] $t = 42.59, p < 1 \times 10^{-323}$
<i>Socio-emotional roles</i>		
Ask	0.189 [0.176, 0.203] $t = 27.47, p < 1 \times 10^{-158}$	0.200 [0.186, 0.215] $t = 27.21, p < 1 \times 10^{-155}$
Give	-0.009 [-0.012, -0.006] $t = -5.97, p < 1 \times 10^{-8}$	-0.008 [-0.011, -0.005] $t = -4.99, p < 1 \times 10^{-6}$
Negative	0.162 [0.147, 0.176] $t = 21.87, p < 1 \times 10^{-102}$	0.450 [0.436, 0.463] $t = 64.77, p < 1 \times 10^{-323}$
Positive	0.278 [0.263, 0.293] $t = 35.38, p < 1 \times 10^{-254}$	0.312 [0.296, 0.327] $t = 39.17, p < 1 \times 10^{-307}$
<i>Jaccard Index</i>		
Ask & Give	0.222 [0.208, 0.237] $t = 30.21, p < 1 \times 10^{-189}$	0.284 [0.269, 0.299] $t = 37.36, p < 1 \times 10^{-281}$
Positive & Negative	0.189 [0.176, 0.203] $t = 27.47, p < 1 \times 10^{-158}$	0.200 [0.186, 0.215] $t = 27.24, p < 1 \times 10^{-155}$
<i>df</i>	8261	
Observations	49572	

Notes: Regression coefficients comparing DeepSeek-R1 and QWQ-32B to the respective instruction-tuned models (DeepSeek-V3, Qwen-2.5-32B-IT) are shown in each column. 95% confidence intervals are reported in brackets, along with the corresponding t -statistics, degrees of freedom, and exact p -values. All models include task fixed effects and control for log-transformed reasoning trace length. Coefficients are tested using two-sided t -tests, with standard errors clustered at the task level.

Reasoning Models Generate Societies of Thought

Supplementary Table 2. Most and least challenging tasks sorted by problem complexity measured by LLM-as-judge

Benchmark task		Problem Complexity		Conversational Behaviour			Reconciliation
		LLM-as-judge (1-7)	Non-reasoning error rate (0-4)	Question & Answering	Perspective Shift	Conflict of Perspectives	
<i>Most Challenging Tasks</i>							
GPQA	Main	5.50 (0.86)	2.30 (1.23)	6.19 (4.38)	6.94 (3.40)	6.01 (3.72)	2.83 (1.66)
GPQA	Diamond	5.47 (0.91)	2.43 (1.23)	6.53 (4.45)	7.14 (4.01)	6.57 (4.75)	2.94 (1.73)
GPQA	Extended	5.46 (0.93)	2.23 (1.25)	5.81 (3.63)	6.60 (3.36)	5.73 (3.89)	2.69 (1.51)
MATH (Hard)	Intermediate Algebra	4.80 (0.75)	2.62 (1.00)	8.12 (3.83)	6.18 (3.09)	4.21 (2.52)	2.85 (1.46)
MATH (Hard)	Pre-Calculus	4.63 (0.92)	2.42 (1.14)	6.90 (2.84)	5.13 (2.37)	3.59 (2.16)	2.70 (1.39)
MUSR	Object Placement	4.43 (0.98)	2.42 (1.22)	4.78 (2.92)	5.20 (2.43)	4.16 (2.30)	2.57 (1.57)
MATH (Hard)	Geometry	4.35 (0.94)	2.43 (1.10)	8.32 (3.36)	5.75 (3.02)	4.13 (2.16)	3.39 (1.51)
BigBench Hard	Causal Judgment	4.16 (1.11)	1.47 (1.35)	4.31 (2.90)	5.99 (2.71)	4.57 (3.05)	2.55 (1.55)
BigBench Hard	Logical Fallacies	4.15 (1.14)	0.69 (0.76)	2.82 (1.75)	3.96 (1.92)	2.05 (1.75)	1.22 (1.26)
MATH (Hard)	Probability	4.13 (0.80)	1.99 (1.14)	8.02 (4.00)	5.09 (2.50)	3.23 (2.30)	2.83 (1.57)
<i>Least Challenging Tasks</i>							
BigBench Hard	Date	2.11 (1.03)	0.81 (0.92)	3.87 (2.81)	4.73 (2.63)	2.78 (2.05)	1.73 (1.30)
BigBench Hard	Disambig. QA	2.01 (0.99)	1.96 (1.32)	3.16 (2.27)	3.74 (2.07)	3.15 (2.14)	1.58 (1.22)
BigBench Hard	Navigation	1.98 (0.92)	0.35 (0.61)	2.01 (1.57)	2.56 (1.68)	1.12 (1.19)	1.08 (1.08)
BigBench Hard	Shuffled	1.67 (0.69)	0.79 (0.93)	1.42 (1.65)	2.46 (1.93)	0.40 (0.76)	0.44 (0.76)
BigBench Hard	Penguins in a Table	1.66 (0.79)	0.71 (0.88)	2.92 (2.39)	3.55 (2.20)	1.94 (1.64)	1.68 (1.30)
BigBench Hard	Colored Objects	1.47 (0.71)	0.79 (0.94)	2.53 (2.06)	3.07 (1.88)	0.88 (1.07)	0.75 (0.87)
BigBench Hard	Hyperbaton	1.34 (0.65)	0.70 (0.88)	3.39 (2.29)	3.41 (1.99)	3.28 (2.06)	1.86 (1.30)
BigBench Hard	Boolean	1.17 (0.42)	0.15 (0.40)	1.09 (1.18)	1.46 (1.26)	0.50 (0.79)	0.43 (0.66)
BigBench Hard	Object Counting	1.15 (0.40)	0.46 (0.68)	1.56 (1.51)	2.17 (1.58)	0.63 (0.88)	0.65 (0.83)
BigBench Hard	Logical Deduction	1.08 (0.30)	0.43 (0.66)	2.68 (2.03)	3.84 (2.16)	3.12 (1.98)	1.48 (1.13)

Notes: Each cell presents average problem complexity or the average number of times each behavior appears in reasoning traces. Standard deviations are presented in parentheses.

4235 **Supplementary Table 3.** SAE Training Hyperparameters

4237 Hyperparameter	4238 Value
4239 Features	32,768
4240 Hook Name	blocks.15.hook_resid_post
4241 Context Size	1024
4242 Input Dimension	4096
4243 Data Type	float32
4244 Architecture	jumprelu
4245 Training Dataset	cerebras/Slimpajama-627B
4246 Evaluation Dataset (for feature explanation)	Hzzfinfdu/SlimPajama-3B

4247
4248
4249
4250
4251
4252
4253
4254
4255
4256
4257
4258
4259
4260
4261
4262
4263
4264
4265
4266
4267
4268
4269
4270
4271
4272
4273
4274
4275
4276
4277
4278
4279
4280
4281
4282
4283
4284
4285
4286
4287
4288
4289

Supplementary Table 4. SAE personality features more frequently activated after steering +10 surprise

Feature	Description	Difference	Ratio before steering	Ratio after steering (+10 Surprise)	Personality score
21065	informal expressions of confusion or frustration regarding societal issues	0.107	0.027	0.134	75
26139	phrases related to social interaction and community engagement	0.084	0.059	0.143	70
26820	conditions or scenarios that involve decision-making or problem-solving	0.068	0.068	0.137	70
14476	references to emotional or sensational themes in narratives	0.063	0.046	0.109	70
24222	phrases indicating emotional or subjective experiences	0.055	0.068	0.123	70
11280	references to identities and backgrounds of people, particularly in literature and film contexts	0.044	0.108	0.152	70
22916	elements related to emotions and moral reflections	0.042	0.054	0.095	85
20151	discussions about literature and character analysis	0.041	0.052	0.092	60
8668	phrases related to personal feelings and social connections	0.037	0.058	0.096	75
15118	statements reflecting emotional conflict and self-reflection	0.033	0.019	0.053	75
19520	phrases that express opportunity, potential, or positive outlooks	0.029	0.021	0.050	60
26648	phrases associated with making choices or decisions	0.024	0.034	0.058	70
18858	words and phrases related to gambling or casino contexts	0.024	0.075	0.098	75
8053	references to social justice or humanitarian efforts	0.022	0.016	0.038	60
26894	dialogue and interactions that indicate emotional states or actions occurring within a narrative context	0.022	0.111	0.133	70
2403	references to interpersonal relationships and emotional connections	0.021	0.050	0.071	75

Supplementary Table 5. SAE expertise features more frequently activated after steering +10 surprise

Feature	Description	Difference	Ratio before steering	Ratio after steering (+10 Surprise)	Expertise score
10105	numeric values and their mathematical operations	0.169	0.218	0.386	60
18556	transitions and conclusions in a logical argument or explanation	0.107	0.243	0.350	60
26139	mathematical concepts related to integers and their properties	0.084	0.059	0.143	85
8532	equations and mathematical expressions involving variables and operations	0.072	0.052	0.125	85
31699	mathematical symbols and expressions related to equations or functions	0.069	0.465	0.534	90
14476	mentions of quantities and their corresponding units in a numerical or mathematical context	0.063	0.046	0.109	70
32111	equations or mathematical expressions	0.055	0.418	0.474	75
22705	mathematical terms related to geometry and spatial properties	0.054	0.058	0.112	85
11452	discussions regarding the implications and potential of alternative medical treatments, particularly in relation to traditional medical practices	0.052	0.299	0.351	95
24623	technical terms and concepts related to automotive engineering and intelligent systems	0.051	0.269	0.320	95
25764	mathematical expressions and their transformations	0.051	0.085	0.136	75
22449	numerical operations and mathematical expressions	0.050	0.039	0.089	80
17611	mathematical properties and functions related to periodicity and calculations	0.049	0.092	0.141	75
23233	numerical values and their relationships in mathematical expressions	0.047	0.075	0.121	70
27631	mathematical operations involving modular arithmetic	0.046	0.224	0.269	75
5298	phrases related to mathematical equations or computations	0.045	0.204	0.249	70

4400 **Supplementary Table 6. PPO training hyperparameters**

4402 Data and model	
4403 Training/validation batch size	128 / 640
4404 Context window	1024 prompt + 1024 response tokens
4406 Optimization	
4407 Actor learning rate	1e-06
4408 Critic learning rate	1e-05
4409 KL coefficient	0.001
4410 PPO mini-batch size	64
4411 Number of rollouts	4
4412 Rollout temperature	1.0
4414 Reward structure	
4415 Accuracy weight	0.9
4416 Correct format weight	0.1

4417
4418
4419
4420
4421
4422
4423
4424
4425
4426
4427
4428
4429
4430
4431
4432
4433
4434
4435
4436
4437
4438
4439
4440
4441
4442
4443
4444
4445
4446
4447
4448
4449
4450
4451
4452
4453
4454

Supplementary Table 7. Monologue-style and conversation-style fine-tuning data generated by Qwen-2.5-32B-IT

Example 1: Number Theory

Question: The greatest common divisor of positive integers m and n is 8. The least common multiple of m and n is 112. What is the least possible value of $m+n$?

[Monologue-style]

<think>

To solve this problem, we need to use the relationship between the greatest common divisor (GCD) and the least common multiple (LCM) of two numbers. The product of the GCD and LCM of two numbers is equal to the product of the numbers themselves. This can be expressed as:

$$\text{GCD}(m, n) \times \text{LCM}(m, n) = m \times n$$

Given that $\text{GCD}(m, n) = 8$ and $\text{LCM}(m, n) = 112$, we can substitute these values into the equation:

$$8 \times 112 = m \times n$$

$$896 = m \times n$$

Since the GCD of m and n is 8, we can express m and n as $8a$ and $8b$ respectively, where a and b are coprime (i.e., $\text{GCD}(a, b) = 1$). Substituting these into the equation, we get:

$$8a \times 8b = 896$$

$$64ab = 896$$

$$ab = 14$$

Now, we need to find pairs (a, b) such that $ab = 14$ and $\text{GCD}(a, b) = 1$. The pairs of factors of 14 are:

- (1, 14)
- (2, 7)

Since a and b must be coprime, the pair (1, 14) is valid, but the pair (2, 7) is also valid because 2 and 7 are coprime. We then calculate m and n for each pair:

1. For $a = 1$ and $b = 14$: $m = 8 \times 1 = 8$, $n = 8 \times 14 = 112$, $m + n = 8 + 112 = 120$

2. For $a = 2$ and $b = 7$: $m = 8 \times 2 = 16$, $n = 8 \times 7 = 56$, $m + n = 16 + 56 = 72$

The least possible value of $m + n$ is 72.

</think>

<answer> 72 </answer>

[Conversation-style]

<cast_of_characters>

<persona1> [Thinker 1 is a meticulous mathematician with a strong background in number theory. He is patient and methodical, preferring to break down problems into smaller, more manageable parts. He enjoys exploring multiple angles before arriving at a conclusion.] </persona1>

<persona2> [Thinker 2 is a quick-witted and intuitive problem solver. She often relies on her gut feeling to guide her through the problem, but she also values rigorous proof. She is not afraid to challenge assumptions and enjoys the process of elimination.] </persona2>

</cast_of_characters>

<conversation>

<think1> Given that the greatest common divisor (GCD) of m and n is 8 and the least common multiple (LCM) is 112, we can use the relationship between GCD and LCM: $m \times n = \text{GCD}(m, n) \times \text{LCM}(m, n)$. Substituting the given values, we get $m \times n = 8 \times 112 = 896$. </think1>

<think2> That's a good starting point. Since the GCD of m and n is 8, we know that both m and n are multiples of 8. Let's express m and n as $m = 8a$ and $n = 8b$, where a and b are coprime integers (i.e., $\text{GCD}(a, b) = 1$). This gives us $(8a) \times (8b) = 896$, which simplifies to $64ab = 896$. Dividing both sides by 64, we get $ab = 14$. </think2>

<think1> Exactly. Now we need to find pairs of coprime integers (a, b) whose product is 14. The possible pairs are (1, 14) and (2, 7). Since a and b must be coprime, we can discard (2, 7) because 2 and 7 are both factors of 14, but they are not coprime. This leaves us with (1, 14). </think1>

4510 <think2> Wait a second. We can't discard (2, 7) just yet. While it's true that 2 and 7 are factors of 14, they are
 4511 indeed coprime because their greatest common divisor is 1. So, the valid pairs are (1, 14) and (2, 7). </think2>
 4512 <think1> You're right. I overlooked that. Let's consider both pairs. For (a, b) = (1, 14), we get $m = 8 \times 1 = 8$
 4513 and $n = 8 \times 14 = 112$. For (a, b) = (2, 7), we get $m = 8 \times 2 = 16$ and $n = 8 \times 7 = 56$. Now, we need to find the least
 4514 possible value of $m + n$. </think1>
 4515 <think2> For the pair (1, 14), we have $m + n = 8 + 112 = 120$. For the pair (2, 7), we have $m + n = 16 + 56 = 72$.
 4516 The least possible value of $m + n$ is 72. </think2>
 4517 </conversation>
 4518 <group_solution> 72 </group_solution>
 4519

4520 Example 2: Combinatorics

4521
 4522
 4523 **Question:** Elodie is putting on a fashion show and has five fabulous outfits for her five fabulous fashion models.
 4524 However, on the day of the show, two of the outfits were ruined in an unfortunate permanent marker incident.
 4525 Regardless, the show must go on and the remaining outfits will be presented. If each outfit can only be worn by one
 4526 model and there is no time for any model to wear more than one dress, how many different shows can Elodie put
 4527 on? (Note: Two shows are considered the same if they contain the same models wearing the same dresses.)
 4528

4529 [Monologue-style]

4530 <think>
 4531 To solve this problem, we need to determine the number of ways to assign three outfits to three models out of the
 4532 original five models. Since two outfits were ruined, we only have three outfits left for the show.
 4533 1. We start with 5 models and 3 outfits.
 4534 2. The first outfit can be assigned to any of the 5 models.
 4535 3. Once the first outfit is assigned, the second outfit can be assigned to any of the remaining 4 models.
 4536 4. The third outfit can be assigned to any of the remaining 3 models.
 4537 Thus, the number of ways to assign the three outfits to three of the five models is the product of these possibilities:
 4538 $5 \times 4 \times 3$
 4539 This calculation gives the total number of different shows Elodie can put on.
 4540 </think>
 4541 <answer> 60 </answer>
 4542

4543 [Conversation-style]

4544 <cast_of_characters>
 4545 <personal> [Thinker 1 is a meticulous and detail-oriented mathematician who excels in combinatorics and
 4546 probability. They are methodical in their approach and often consider all possible cases before concluding.]
 4547 </personal>
 4548 <persona2> [Thinker 2 is a creative and intuitive fashion designer who is skilled in visualizing scenarios and
 4549 outcomes. They are quick to grasp the essence of a problem and often provide practical insights.] </persona2>
 4550 </cast_of_characters>
 4551 <conversation>
 4552 <think1> Okay, so Elodie has five outfits, but two were ruined. That leaves her with three outfits for her
 4553 five models. We need to figure out how many different ways she can assign these three outfits to the models.
 4554 </think1>
 4555 <think2> Right, and since each model can only wear one outfit, we are essentially choosing 3 models out of 5 to
 4556 wear the outfits. Then we need to think about how these outfits can be assigned to the selected models. </think2>
 4557 <think1> Exactly. First, we need to calculate the number of ways to choose 3 models out of 5. This is a
 4558 combination problem, and the formula for combinations is $C(n, k) = n! / (k!(n-k)!)$, where n is the total number of
 4559 items, and k is the number of items to choose. </think1>
 4560 <think2> So, we have $n = 5$ and $k = 3$. Let's calculate that. </think2>
 4561 <think1> $C(5, 3) = 5! / (3!(5-3)!) = 5! / (3!2!) = (5 \times 4 \times 3 \times 2 \times 1) / (3 \times 2 \times 1 \times 2 \times 1) = 10$ ways to choose the
 4562
 4563
 4564

3 models. </think1>

<think2> Now, for each of these 10 ways to choose the models, we need to consider how the 3 outfits can be assigned to these 3 models. This is a permutation problem, as the order in which the outfits are assigned matters. </think2>

<think1> Indeed, for each selection of 3 models, there are 3! (3 factorial) ways to assign the 3 outfits. So, $3! = 3 \times 2 \times 1 = 6$ ways to assign the outfits. </think1>

<think2> So, the total number of different shows is the number of ways to choose the models multiplied by the number of ways to assign the outfits. That's 10×6 . </think2>

<think1> Which gives us 60 different shows. That's the answer. </think1>

</conversation>

<group_solution> 60 </group_solution>

Example 3: Logical Reasoning

Question: Alice, Bob, Claire, Dave, and Eve are playing a game. At the start of the game, they are each holding a ball: Alice has a brown ball, Bob has a purple ball, Claire has a black ball, Dave has a green ball, and Eve has a yellow ball. As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Bob and Alice swap balls. Then, Eve and Dave swap balls. Then, Dave and Claire swap balls. Finally, Alice and Bob swap balls. At the end of the game, Claire has the

Options: (A) brown ball (B) purple ball (C) black ball (D) green ball (E) yellow ball

[Monologue-style]

<think>

Let's track the ball swaps step by step.

1. Initial Positions:

- Alice: Brown
- Bob: Purple
- Claire: Black
- Dave: Green
- Eve: Yellow

2. **First Swap (Claire and Alice):** Alice: Black, Claire: Brown

3. **Second Swap (Bob and Alice):** Alice: Purple, Bob: Black

4. **Third Swap (Eve and Dave):** Eve: Green, Dave: Yellow

5. **Fourth Swap (Dave and Claire):** Claire: Yellow, Dave: Brown

6. **Fifth Swap (Alice and Bob):** Alice: Black, Bob: Purple

At the end of the game, Claire has the yellow ball.

</think>

<answer> (E) yellow ball </answer>

[Conversation-style]

<cast_of_characters>

<persona1> [Logical and detail-oriented, this thinker is a mathematician who enjoys breaking down problems into smaller, more manageable parts. They are meticulous and prefer to double-check every step of their reasoning.] </persona1>

<persona2> [An engineer by training, this thinker is practical and solution-focused. They prefer to use diagrams and visual aids to understand problems and are not afraid to reframe the problem if it leads to a solution.] </persona2>

4620 <persona3> [A creative and intuitive thinker, this person excels at seeing patterns and connections that others
4621 might miss. They are a bit more relaxed in their approach, often suggesting unconventional methods to solve
4622 problems.] </persona3>
4623 </cast_of_characters>
4624 <conversation>
4625 <think1> Let's start by laying out the initial positions of the balls. Alice has a brown ball, Bob has a purple ball,
4626 Claire has a black ball, Dave has a green ball, and Eve has a yellow ball. Now, we'll go through the swaps one by
4627 one. </think1>
4628 <think2> I'll make a diagram to keep track of the swaps. First, Claire and Alice swap, so Claire now has the
4629 brown ball, and Alice has the black ball. </think2>
4630 <think3> Good idea to use a visual. Next, Bob and Alice swap. Bob will end up with the black ball, and Alice
4631 will have the purple ball. </think3>
4632 <think1> Then, Eve and Dave swap their balls. Eve will have the green ball, and Dave will have the yellow ball.
4633 </think1>
4634 <think2> Now, Dave and Claire swap. Claire will have the yellow ball, and Dave will have the brown ball.
4635 </think2>
4636 <think3> Finally, Alice and Bob swap. Bob will end up with the purple ball, and Alice will have the black ball.
4637 </think3>
4638 <think1> After all the swaps, Claire ends up with the yellow ball. So the answer is (E). </think1>
4639 </conversation>
4640 <group_solution> (E) yellow ball </group_solution>
4641
4642
4643
4644
4645
4646
4647
4648
4649
4650
4651
4652
4653
4654
4655
4656
4657
4658
4659
4660
4661
4662
4663
4664
4665
4666
4667
4668
4669
4670
4671
4672
4673
4674

4675 **Supplementary Table 8. Supervised fine-tuning hyperparameters**

4677 Data and Model	
4678 Training/validation dataset size	500 / 100
4679 Context window	2048
4680 Training/validation batch size	64 / 64
4682 Optimization	
4683 Optimizer	AdamW
4684 Peak learning rate	1e-05
4685 Warmup	Linear (10% of total steps)
4686 Annealing	Cosine
4687 Total epochs	5

4689
4690
4691
4692
4693
4694
4695
4696
4697
4698
4699
4700
4701
4702
4703
4704
4705
4706
4707
4708
4709
4710
4711
4712
4713
4714
4715
4716
4717
4718
4719
4720
4721
4722
4723
4724
4725
4726
4727
4728
4729

4730 **Supplementary Table 9. Benchmark Tasks**

4731	Benchmark	Task Category	# Problems	
4732	BigBench Hard	Boolean Expressions	248	
4733		Causal Judgment	173	
4734		Date Understanding	195	
4735		Disambiguation Question Answering	249	
4736		Formal Logical Fallacies	250	
4737		Geometric Shape Reasoning	249	
4738		Hyperbaton (Syntactic Reordering)	247	
4739		Logical Deduction (Five Objects)	244	
4740		Logical Deduction (Seven Objects)	224	
4741		Logical Deduction (Three Objects)	249	
4742		Movie Recommendation	7	
4743		Navigation and Spatial Planning	248	
4744		Object Counting	247	
4745		Penguins in a Table (Deductive Counting)	143	
4746		Reasoning About Colored Objects	245	
4747		Ruin Names (Linguistic Perturbation)	165	
4748		Salient Translation Error Detection	249	
4749		Snarks (Riddle Interpretation)	146	
4750		Sports Understanding	250	
4751		Temporal Sequence Reasoning	65	
4752		Tracking Shuffled Objects (Five Objects)	112	
4753		Tracking Shuffled Objects (Seven Objects)	102	
4754		Tracking Shuffled Objects (Three Objects)	57	
4755		Web of Lies (Deception Reasoning)	72	
4756		GPQA	Diamond Level	161
4757			Extended Level	474
4758			Main Benchmark	380
4759		IFEval	Instruction-Following Consistency Evaluation	524
4760		MATH (Hard)	Algebra (Hard)	286
4761	Counting & Probability (Hard)		110	
4762	Geometry (Hard)		117	
4763	Intermediate Algebra (Hard)		212	
4764	Number Theory (Hard)		134	
4765	Pre-Algebra (Hard)		182	
4766	Pre-Calculus (Hard)		104	
4767	MMLU-Pro		Advanced Multidomain Knowledge	432
4768	MUSR	Murder Mysteries (Collaborative Deduction)	207	
4769		Object Placement (Spatial Coordination)	256	
4770		Team Allocation (Group Strategy)	247	

4771
4772
4773
4774
4775
4776
4777
4778
4779
4780
4781
4782
4783
4784

Reasoning Models Generate Societies of Thought

4785 **Supplementary Table 10.** Descriptive statistics of conversational behaviors and socio-emotional roles for reasoning and
 4786 instruction-tuned models (Count variables)

	Reasoning Models		Instruction-Tuned Models			
	DeepSeek- R1	QWQ-32B	DeepSeek- V3	Qwen-2.5- 32B-IT	Llama-3.3- 70B-IT	Llama-3.1- 8B- IT
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
<i>Conversational behaviors</i>						
Question answering	6.74 (4.32)	6.68 (4.62)	3.07 (2.53)	2.35 (4.25)	2.22 (1.98)	3.16 (18.05)
Perspective shifts	3.47 (2.76)	5.41 (6.89)	1.09 (1.69)	0.79 (3.95)	0.97 (1.56)	1.84 (17.04)
Conflict of perspectives	3.19 (2.95)	5.27 (5.51)	1.50 (2.06)	0.97 (3.99)	1.08 (1.62)	2.25 (17.96)
Reconciliations	1.97 (1.60)	2.52 (1.90)	0.61 (0.76)	0.41 (0.62)	0.58 (0.81)	0.49 (1.27)
<i>Socio-emotional roles</i>						
Ask for orientation	2.67 (5.14)	2.19 (3.80)	0.01 (0.15)	0.01 (0.11)	0.00 (0.08)	0.01 (0.24)
Ask for opinion	0.50 (1.73)	0.57 (1.94)	0.00 (0.07)	0.00 (0.05)	0.00 (0.07)	0.01 (0.39)
Ask for suggestion	0.18 (0.87)	0.26 (1.10)	0.00 (0.04)	0.00 (0.04)	0.00 (0.06)	0.01 (0.14)
Give orientation	29.67 (23.44)	26.17 (21.06)	8.26 (6.55)	7.00 (7.74)	6.20 (5.14)	8.52 (24.72)
Give opinion	19.25 (18.08)	21.72 (22.64)	4.15 (4.31)	3.09 (4.73)	3.74 (4.02)	4.60 (35.84)
Give suggestion	9.04 (7.80)	13.05 (12.33)	2.23 (3.02)	2.17 (2.21)	2.50 (3.09)	3.05 (12.05)
Antagonism	0.00 (0.13)	0.00 (0.06)	0.01 (0.27)	0.00 (0.10)	0.00 (0.13)	0.01 (0.16)
Tension	3.82 (5.37)	9.65 (11.53)	0.14 (0.67)	0.03 (0.24)	0.21 (2.04)	0.65 (10.60)
Disagree	2.56 (5.15)	3.09 (5.41)	0.49 (1.54)	0.16 (0.89)	0.25 (2.17)	0.88 (8.90)
Solidarity	0.00 (0.07)	0.00 (0.05)	0.01 (0.23)	0.01 (0.14)	0.00 (0.13)	0.02 (0.33)
Tension release	0.50 (1.12)	0.36 (0.87)	0.02 (0.29)	0.00 (0.10)	0.01 (0.14)	0.06 (3.50)
Agree	3.59 (4.83)	3.50 (4.64)	0.18 (0.69)	0.04 (0.26)	0.03 (0.21)	0.06 (0.85)
Observations	8,262					

Reasoning Models Generate Societies of Thought

4840 **Supplementary Table 11.** Structural equation model estimates linking DeepSeek-R1, conversational behaviors, socio-
 4841 emotional roles, cognitive reasoning strategies, and accuracy

	Estimate	Std.Err	z	p-value	
4842					
4843					
4844					
4845	DeepSeek-R1 → Question & Answering	4.013	0.066	61.060	<0.001
4846	DeepSeek-R1 → Perspective Shift	3.273	0.067	48.530	<0.001
4847	DeepSeek-R1 → Conflict of Perspectives	2.764	0.065	42.400	<0.001
4848	DeepSeek-R1 → Reconciliation	1.748	0.016	109.360	<0.001
4849	DeepSeek-R1 → Positive	4.017	0.055	72.810	<0.001
4850	DeepSeek-R1 → Negative	9.177	0.121	75.740	<0.001
4851	DeepSeek-R1 → Ask	3.276	0.050	66.090	<0.001
4852	DeepSeek-R1 → Give	46.810	0.408	114.750	<0.001
4853	Question & Answering → Verification	0.179	0.070	2.550	0.011
4854	Perspective Shift → Verification	-0.096	0.107	-0.890	0.372
4855	Conflict of Perspectives → Verification	0.216	0.113	1.910	0.057
4856	Reconciliation → Verification	-0.237	0.109	-2.180	0.029
4857	Positive → Verification	-0.023	0.129	-0.180	0.858
4858	Negative → Verification	0.055	0.057	0.960	0.336
4859	Ask → Verification	-0.111	0.039	-2.810	0.005
4860	Give → Verification	0.026	0.015	1.740	0.082
4861	DeepSeek-R1 → Verification	-0.673	0.327	-2.060	0.039
4862	Question & Answering → Backtracking	-0.023	0.037	-0.620	0.536
4863	Perspective Shift → Backtracking	0.007	0.071	0.100	0.924
4864	Conflict of Perspectives → Backtracking	0.163	0.065	2.530	0.011
4865	Reconciliation → Backtracking	-0.005	0.082	-0.060	0.953
4866	Positive → Backtracking	-0.125	0.040	-3.130	0.002
4867	Negative → Backtracking	0.079	0.028	2.820	0.005
4868	Ask → Backtracking	-0.024	0.016	-1.510	0.132
4869	Give → Backtracking	0.014	0.006	2.320	0.020
4870	DeepSeek-R1 → Backtracking	-0.325	0.129	-2.520	0.012
4871	Question & Answering → Subgoal setting	0.035	0.016	2.120	0.034
4872	Perspective Shift → Subgoal setting	0.006	0.012	0.450	0.653
4873	Conflict of Perspectives → Subgoal setting	-0.036	0.016	-2.220	0.026
4874	Reconciliation → Subgoal setting	0.089	0.021	4.270	<0.001
4875	Positive → Subgoal setting	0.010	0.012	0.830	0.409
4876	Negative → Subgoal setting	0.007	0.004	1.680	0.093
4877	Ask → Subgoal setting	0.005	0.010	0.540	0.587
4878	Give → Subgoal setting	0.008	0.003	2.490	0.013
4879	DeepSeek-R1 → Subgoal setting	0.224	0.087	2.570	0.010
4880	Question & Answering → Backward chaining	-0.006	0.002	-3.900	<0.001
4881	Perspective Shift → Backward chaining	0.011	0.002	5.710	<0.001
4882	Conflict of Perspectives → Backward chaining	-0.003	0.002	-1.510	0.132
4883	Reconciliation → Backward chaining	-0.002	0.003	-0.800	0.422
4884	Positive → Backward chaining	-0.002	0.001	-1.720	0.086
4885	Negative → Backward chaining	0.005	0.001	3.760	<0.001
4886	Ask → Backward chaining	0.004	0.001	2.870	0.004
4887	Give → Backward chaining	0.001	0.000	2.950	0.003
4888	DeepSeek-R1 → Backward chaining	0.019	0.007	2.730	0.006
4889	Question & Answering → Accuracy	0.009	0.001	7.730	<0.001
4890	Perspective Shift → Accuracy	-0.006	0.001	-4.250	<0.001
4891	Conflict of Perspectives → Accuracy	-0.006	0.002	-3.240	0.001
4892					
4893					
4894					

Reasoning Models Generate Societies of Thought

4895	Reconciliation → Accuracy	0.001	0.003	0.330	0.742
4896	Positive → Accuracy	0.020	0.003	6.660	<0.001
4897	Negative → Accuracy	-0.001	0.000	-1.470	0.141
4898	Ask → Accuracy	0.001	0.001	0.900	0.367
4899	Give → Accuracy	0.000	0.000	-1.450	0.147
4900	Verification → Accuracy	0.005	0.001	3.400	0.001
4901	Backtracking → Accuracy	-0.003	0.001	-2.810	0.005
4902	Subgoal setting → Accuracy	0.009	0.004	2.610	0.009
4903	Backward chaining → Accuracy	-0.001	0.006	-0.180	0.855
4904	DeepSeek-R1 → Accuracy	0.253	0.007	35.060	<0.001
4905	Length → Accuracy	-0.134	0.005	-25.410	<0.001
4906	Observations			7,738	

Notes: All structural equation models control for task fixed effects. The statistical significance of coefficients is tested using two-sided *t*-tests. Out of 8,262 tasks, tasks from IFEval are excluded due to difficulties in accuracy evaluation (524 tasks), and we use the remaining 7,738 tasks.