Teaching LLMs Naturally: Pedagogical Strategies for Interactive Knowledge Acquisition

Sabrina Patania
University of Milan-Bicocca
Milan, Italy
sabrina.patania@unimib.it

Luca Annese
University of Milan-Bicocca
Milan, Italy
luca.annese1@unimib.it

Cansu Koyuturk
University of Milan-Bicocca
Milan, Italy
c.koyutuerk@campus.unimib.it

Dimitri Ognibene
University of Milan-Bicocca
Milan, Italy
ISTC-CNR
Rome, Italy
dimitri.ognibene@unimib.it

ABSTRACT

Large Language Models (LLMs) excel at learning from extensive offline datasets but face significant challenges in dynamically acquiring complex knowledge during online interactions. Traditional training paradigms, grounded in supervised or reinforcement learning, emphasize independent discovery but depend on vast data and sparse feedback, limiting adaptability.

In this study, we investigate whether pedagogically inspired forms of interaction can provide more natural ways of presenting information to LLMs and support the integration of new knowledge. Drawing on Vygotsky's sociocultural theory, we explore settings where a learner LLM acquires a synthetic taxonomy through structured exchanges with a teacher LLM. These interactions combine explicit explanations with opportunities for the learner to test and refine its understanding, creating a process that extends beyond conventional prompt engineering or internal reasoning heuristics. Unlike fine-tuning or few-shot learning, the approach introduces new knowledge without modifying model weights or relying solely on task-specific examples.

Evaluation focuses on learner performance in two complementary ways: first, by reconstructing the acquired ontology and aligning it with the original reference ontology, and second, by applying its knowledge in downstream tasks. These include playing the challenging and well-established 20 Questions Game, which requires efficient reasoning over incomplete information.

Results show that pedagogically structured learning yields better ontology reconstruction than non-pedagogical exposition. Top-down explanation enables near-complete, stable transfer, whereas Mixed Learner Questions produces more human-like dynamics but is less stable for deeper relational knowledge. Overall, combining structured pedagogy with targeted teacher verification may support knowledge transmission that is both reliable and natural for humans and language models. Code and data available here: https://github.com/DimNeuroLab/AI-Pedagogy

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/JGLB4831

KEYWORDS

in-context learning; social learning; LLMs; artificial pedagogy; ontology acquisition; mixed-initiative dialogue; natural AI training; interactive learning.

ACM Reference Format:
Sabrina Patania, Luca Annese, Cansu Koyuturk, and Dimitri Ognibene. 2026. Teaching LLMs Naturally: Pedagogical Strategies for Interactive Knowledge Acquisition. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 9 pages. https://doi.org/10.65109/JGLB4831

1 INTRODUCTION

Over the past decade, Large Language Models (LLMs) have reshaped natural language processing, showing remarkable competence across an expanding range of linguistic and reasoning tasks [5, 8]. Such achievements, however, are grounded in a fundamentally offline paradigm: large-scale pre-training on text corpora and, more recently, reinforcement learning from human feedback [31]. Once deployed, these models cannot truly learn from interaction. They can adapt behaviour through prompt conditioning, but cannot internalise a new, formally structured body of facts without costly fine-tuning. This creates a persistent mismatch: humans continually extend, reorganise, and verify their knowledge during conversation, whereas current LLMs lack mechanisms for dynamic, situated knowledge acquisition.

A second mismatch concerns the social nature of learning. Human expertise is largely transmitted interpersonally: teachers and peers scaffold understanding, highlight key distinctions, and monitor comprehension in real time. This Vygotskian perspective [6, 14, 46] contrasts sharply with the dominant assumption in machine learning that intelligence emerges from individual discovery. If we want artificial agents to collaborate meaningfully with humans—co-constructing knowledge, diagnosing problems, or learning new protocols on demand—we may need to endow them with more natural modes of being taught.

Educational psychology offers a rich foundation for such inquiry. Research on scaffolding [7, 48], worked examples [16], and the management of cognitive load [38] shows that explicit pedagogical guidance allows learners to acquire and integrate structured knowledge more efficiently and with fewer misconceptions. Translating these principles to artificial systems raises an intriguing question: can an LLM acquire new, verifiable conceptual structures through pedagogical interaction, rather than through exposure or reward optimisation alone?

Building upon previous evidence [32], we address this challenge through a controlled study of artificial pedagogy in ontology learning. We simulate teaching sessions in which a knowledgeable teacher LLM conveys the structure of a synthetic taxonomy of alien species to a naïve learner LLM, which must later reconstruct and apply this knowledge without further supervision. To guide this investigation, we formulate three distinct research questions:

RQ1 : Can an LLM acquire verifiable conceptual structures through pedagogical interaction (social scaffolding) rather than direct access to structured data?
RQ2 : How do different pedagogical strategies impact the accuracy of ontology reconstruction?
RQ3 : Does high fidelity in knowledge reconstruction translate to strategic application in a sequential decision-making task?

To address RQ1, we contrast pedagogical strategies with a non-pedagogical baseline—a glossary-style condition that exposes the learner to the same information in static textual form, without social exchange. This comparison isolates the contribution of social scaffolding from mere content exposure. To investigate RQ2, we systematically vary pedagogical strategies along three dimensions: conceptual framing (top-down vs. bottom-up), inquiry direction (teacher-led vs. learner-led), and interactional initiative (single vs. mixed). This yields a typology spanning both guided and exploratory forms of dialogue.

Finally, to answer RQ3, we look beyond immediate reconstruction. We evaluate the stability of learning across repeated sessions and replicate the evaluation of knowledge application through the 20 Questions Game. This dual focus enables us to test whether the knowledge acquired can be effectively applied in a sequential decision-making task.

Our results indicate that top-down explanation and mixed learner-question strategies are particularly informative: the former produces near-complete and stable knowledge acquisition, while the latter—though more natural—reveals greater variability. Notably, we find that structural understanding and reasoning performance do not always correlate, suggesting that high fidelity in reconstruction does not automatically guarantee strategic application (RQ3).

Beyond their technological implications, our findings outline a broader vision—interactive systems that can be taught by humans through conversation rather than data curation, bridging cognitive science and machine learning in the design of genuinely social AI.

2 RELATED WORKS

Developmental and educational psychology provides valuable insights into effective pedagogical strategies and knowledge acquisition paradigms, forming a theoretical basis for our proposal. Research in educational psychology has examined structured, expert-led instruction, such as direct explanation strategies where teachers lead the learning process by presenting information in a sequenced, hierarchical manner with minimal student interaction or autonomy [27]. These strategies reflect top-down knowledge delivery, also commonly employed in traditional educational settings to ensure coherence and expert framing [17]. In contrast, bottom-up instructional strategies, which emphasize localized knowledge construction, allow for responsiveness to learners' prior knowledge and context. However, such approaches may yield narrower or less integrative outcomes if not sufficiently guided [17]. Considering how learning places demands on memory, structured top-down approaches help reduce extraneous cognitive load for novice learners, while unguided or minimally guided bottom-up strategies may overload working memory if learners lack sufficient prior knowledge [27, 38].

Seminal theories by Piaget and Vygotsky both stress active learner engagement, but differ in how central they consider social interaction to learning. Piaget foregrounds independent exploration, where learners construct knowledge through direct interaction with the environment [35]. Teachers mainly facilitate by setting up conditions for self-directed problem solving and offering minimal, non-directive guidance [11]. In discovery-based learning, understanding emerges from self-generated inferences and trial-and-error exploration [13]; learner-led questioning likewise positions the learner as the initiator, with inquiry and hypothesis testing driving knowledge construction.

Vygotsky, by contrast, emphasises instruction and socially mediated interaction, where knowledge is co-constructed with more knowledgeable others [46]. Learning is supported by structured guidance rather than unguided exploration. A core mechanism is scaffolding: experts adapt support to the learner's current needs and progress, fostering increasing autonomy and mastery [4, 7]. Scaffolding provides feedback and elaborative support (e.g., modelling, explanation), helping learners form more accurate and coherent representations during complex tasks [45]. This also connects to cognitive load theory [37], as structured support can reduce early cognitive demands by organising information into manageable schemas.

Early developmental robotics [26] explored how robots might acquire cognitive functions, including language [28], via autonomous interaction that physically models child development. Although largely Piagetian in spirit, this line anticipated a role for language in supporting learning, later foregrounded in epigenetic robotics [2, 23, 24]. For instance, [9] studies forms of robot social learning without an explicit teacher, aimed at mitigating the exploration costs of solitary reinforcement learning [29]. Lockerd and Breazeal [25] propose an early pedagogy-inspired architecture: despite being hand-engineered and lacking open-ended language (hence not acquiring new ontologies or generic tasks), human tutoring enables a three-button sequence in 3–4 demonstrations versus roughly 45 trials for a Q-learning baseline, illustrating how minimal social feedback can sharply reduce data requirements.

The role of social teaching is central to Socially Guided Machine Learning [43, 44], which argues for principled accounts of how non-expert humans shape learning. These studies examine how teachers adapt strategies when training an RL-based simulated robot through a narrow channel—numeric rewards reflecting the system's state or aspects of it. De Greef and Belpaeme [18] similarly emphasise human-like social learning, highlighting unstructured interaction and coordination via social cues, and giving the robot an active role in signalling learning needs—though communication remains far more constrained than natural-language dialogue.

Although language as a means of improving (robot) learning and cognition has been widely explored [34, 36, 39], its integration into explicitly pedagogical robotics has received less attention, particularly since the advent of LLMs with broad knowledge and zero/few-shot capabilities [5]. LLMs are increasingly embedded in robotic architectures [1, 20, 40], where frozen models can generate curricula, high-level plans, and success predicates for manipulation, positioning language as a general teaching interface for embodied agents. Interest is growing in reasoning-orientated models that can further improve robustness and flexibility [15, 49], albeit with additional training costs [19, 21, 41, 50]. Our method extends this paradigm by enabling two LLMs to co-construct knowledge through mixed-direction dialogue, rather than relying on one-shot prompting that often requires iterative requests and corrections [42, 47], which can be expensive in robotic settings.

Our approach aligns with [30], which highlights how cooperative interaction allows LLMs to build shared representations of goals and environment through language. While that work emphasises plan execution in cooperative settings, we focus on acquiring structured knowledge, evaluated via reasoning.

A complementary lens is the 20 Questions game, which tests how well acquired knowledge supports inference and decision-making. Recent work frames each question–answer exchange as an active Bayesian update [12] and shows that models such as ChatGPT can spontaneously approximate near-optimal inquiry strategies [3, 33]. Accordingly, we use the same game as our test phase, linking pedagogical training regimes to quantitative evidence of online information acquisition.

3 METHOD

3.1 Experimental Design

The study investigates how different pedagogical strategies influence the ability of LLMs to acquire and operationalise new conceptual knowledge through interaction. Learning is modelled as a socially mediated process in which a learner agent constructs an ontology through guided dialogue with a teacher agent or by presenting a glossary-based static text as a non-pedagogical baseline.

The design systematically varies three main pedagogical dimensions to characterise instructional interaction:

- Conceptual framing: how knowledge is structured and presented — top-down (deductive, hierarchical explanation) vs bottom-up (inductive, exemplar-driven).
- Inquiry direction: who drives the learning process making questions — teacher-led (guiding or questioning) vs learner-led (asking or exploring).
- Initiative: distinguishes purely one-way initiative (an active subject and a reactive one) with mixed-initiative (both contribute reciprocally).

This factorial design yields a family of pedagogical strategies, each corresponding to a distinct combination of framing, inquiry direction, and initiative modality.

The experimental programme unfolds in two macro-phases:

(1) Strategy assessment: broad screening of conditions spanning significant combinations of framing, inquiry direction, and initiative.
(2) Extended evaluation: longitudinal testing of selected strategies and comparison with a non-pedagogical baseline.

Within each macro-phase, the procedure includes a learning stage—where the learner acquires the target ontology through pedagogical dialogue—and a testing stage comprising two complementary tasks: (a) knowledge reconstruction, in which the learner reproduces the ontology from memory to assess structural alignment, and (b) knowledge application, where the learner employs the acquired knowledge in a compact 20 Questions Game to evaluate generalisation and reasoning.

3.2 Materials

The knowledge domain consists of a synthetic ontology of alien species designed to support controlled analysis of conceptual learning. The ontology was initially generated with GPT-4o and subsequently refined through manual revision to ensure internal coherence and remove possible lexical or cultural biases associated with prior human or real-world knowledge. This process allowed us to preserve the creative variability of the model's output while adapting the content to the experimental context.

Each ontology consists of ten fictional species described through five categorical features, yielding a total of fifty feature–value associations. The feature space includes controlled overlap across species, allowing both top-down abstraction and bottom-up pattern discovery. Feature distributions were balanced to avoid dominant categories and ensure comparable learning difficulty across conditions.

Ontologies are stored as structured JSON objects and remain identical within each experimental phase to guarantee content equivalence. During the experiments, the teacher agent has full access to the ontology, while the learner starts with no prior information and can acquire knowledge solely through dialogue.

3.3 Pedagogical Strategies

Each condition is instantiated through a dedicated prompt controlling the teacher's behaviour:

- **Top-down (TD)**: Structured, declarative ontology explanation from high-level categories to examples.
- **Bottom-up (BU)**: Starts from species/features; learner infers categories inductively.
- **Learner Questions (LQ)**: Learner asks; teacher answers truthfully without guiding or reframing.
- **Teacher Questions (TQ)**: Teacher uses structured questions to prompt reasoning and pattern discovery.
- **Mix-TQ**: Explain → brief check question → feedback, then continue.
- **Mix-LQ**: Explain → invite learner questions on that part, then continue.
- **Mix-TD-TQ**: Top-down blocks alternating with comprehension questions.
- **Mix-BU-TQ**: Bottom-up examples plus short consolidation questions (inductive generalisation).

- **Mix-TD-LQ**: Top-down scaffold with learner-question pauses after each block.
- **Mix-BU-LQ**: Bottom-up examples with learner-question pauses to support pattern discovery.

Among these, we select the two best-performing strategies (Top-down and Mixed Learner Questions) and the two weakest (Bottom-up and Learner Questions) for extended evaluation, representing the extremes of the performance spectrum.

To separate pedagogical benefits from simple exposure, a glossary-style control is introduced: the learner received the ontology as a concise textual description without any interactive exchange, simulating a non-dialogic "book learning" condition.

3.4 Learning Phase

During training, a knowledgeable teacher LLM and a naïve learner LLM engage in short, fixed-length dialogues centred on a target alien ontology. Each session consists of a predetermined number of alternating question–answer exchanges (20), ensuring consistent interaction time across experimental conditions. The structure and control of the exchange depend on the assigned pedagogical or non-pedagogical strategy.

At the end of every teacher contribution, the learner is asked to summarise its current understanding in one or two sentences. These micro-summaries serve two purposes: they promote active consolidation for the learner, and they give us insights on the ability of the learner to acquire and structure knowledge.

Finally, we estimate the amount of information transmitted from teacher to learner during each dialogue. This yields a quantitative measure of information exposure, the cumulative number of factual elements made available to the learner over the course of the dialogue.

3.5 Testing Phase

To evaluate knowledge acquisition, each learner is first tested on its ability to reconstruct the underlying ontology that have been the focus of the teaching dialogue. The reconstructed ontology is then aligned with the reference ontology using an entity-feature-value matching procedure that measured structural accuracy and completeness. Performance is quantified over multiple runs as the proportion of correctly recovered (entity, feature, value) triples relative to the ground truth.

Moreover, each learner is evaluated in a standard 20-Questions Game, a classic parlor game. For each trial, the learner faces a set of eight candidate species drawn from the same ontology and attempts to identify a secretly chosen target via up to 20 yes/no questions. An oracle agent, fully informed of the ontology, provides truthful responses. At the end of each trial, the learner issues a final guess. We ran multiple trials per condition to sample performance across different targets and candidate sets. This interactive format provides a naturalistic way to assess an AI's ability to learn by asking targeted questions and integrating the received information to narrow down possibilities [12].

4 EXPERIMENTS

4.1 Experimental setup

We conduct all simulations with GPT-4o accessed through the OpenAI API.[1] The procedure consists of four stages: ontology generation, teacher initialisation, pedagogical interaction, knowledge reconstruction and knowledge application.

Ontology generation. A custom prompt requests GPT-4o to create a JSON ontology containing ten alien species, each described by five categorical features: *Diet*, *Habitat*, *Morphology*, *Locomotion*, and *Social Structure*. The resulting ontology is stored unchanged and used for both training and test.

Teacher initialisation. The teacher agent receives the structured ontology as part of its system prompt. Thus, for every subsequent turn, the teacher has perfect, verifiable knowledge of the reference ontology.

Pedagogical interaction. We create the 10 training conditions, each instantiated once and all paired with the same architecture (GPT-4o). Each dialogue comprises 40 alternating turns (20 per agent). After every teacher contribution, the learner produces a short summary of its current understanding, promoting active consolidation and providing interpretable intermediate states.

4.2 Quantifying Information Transmission

To evaluate how much knowledge was conveyed during training, each dialogue is analysed step by step, including both teacher and learner contributions. For each turn, the dialogue is parsed into discrete informational units, defined as unique (entity, feature, value) triplets present in the ontology. The triplets are extracted using GPT-4o, which had access to the full reference ontology. To improve reliability, extraction is performed multiple times per step, and the most frequently identified triplets are retained as the final set for that turn. The cumulative number of unique triplets throughout the dialogue provided a quantitative measure of information exposure for each session. This metric captures both the amount of structured knowledge presented and how it accumulates over the course of the interaction (see Figure 1).

4.3 Knowledge Reconstruction

After completing the training dialogues, each learner is tasked with reconstructing the target ontology. Specifically, the learner is prompted to reproduce all entities, their features, and associated values. To capture the variability and stability of learning, each learner performed five independent reconstruction attempts for each ontology.

The precision of reconstruction is then quantified by aligning each reconstructed ontology with the reference ontology. Alignment is performed by decomposing both ontologies into atomic triplets (entity, feature, value) and comparing them to the ground truth. Performance is measured as the proportion of correctly reproduced triplets relative to the reference ontology. Additional analyses consider partial matches at the feature or value level to capture more fine-grained patterns of knowledge retention.

[1]Temperature is fixed at 0.3 and `max_tokens` at 10000 for every call.

4.4 Knowledge Application: 20-Questions evaluation

After training, the learner is tested in an automated variant of the classic 20-Questions Game. We generate 20 candidate sets, each containing eight distinct species randomly drawn from the ontology. For each set an oracle with full ontology access secretly selects one target. The learner asks yes/no feature questions until it either identifies the target or reaches the maximum budget of 20 questions. We record the questions count and whether the final guess is correct.

To gauge an upper bound we evaluate an expert baseline, another GPT-4o instance whose system prompt contains the full ontology during the 20-Questions Game. Apart from that advantage the baseline plays exactly the same protocol as the learners.

5 RESULTS

Across all experimental conditions, pedagogically guided strategies outperform the non-interactive glossary baseline in the reconstruction of knowledge. This advantage persists across repetitions and ontologies, confirming that social structure, rather than simple content exposure, drives more consistent conceptual acquisition.

The reconstruction task assesses how accurately learners could reproduce the ontology after training. Top-down explanations yielded near-complete reconstruction from scratch, with minimal variance across ontologies and repetitions. This confirms that fully structured, deductive exposition remains the most reliable way to convey and stabilise hierarchical knowledge.

The Mixed Learner Questions condition achieved comparable average accuracy but displayed higher variability, particularly at the entity and triplet levels of the ontology (see Figure 3). This instability indicates that while mixed learner-led interactions promote flexibility and more natural conversational flow, they can also produce partial or inconsistent representations when the teacher does not verify the learner's understanding before proceeding.

The Learner Questions strategy consistently ranked among the weakest, suffering from this lack of teacher control and verification. Comparison across ontologies revealed limited variance overall, suggesting that the pedagogical mechanisms generalise across knowledge domains, although reconstruction of specific relational or compositional features remained more sensitive to ontology size and internal regularities.

Considering Figure 1, we observe that Learner-driven strategies were the ones that primarily , apart from Top-down, covered the most facts during the training dialogue. Correlation analyses confirm a positive relationship between information exchange level and reconstruction performance: the more efficiently information circulates, the better the learner's ontology aligns with the ground truth.

When tested across three distinct ontologies, overall performance remained stable, showing that pedagogical structure is more influential than the specific training domain. Still, entity- and triplet-level reconstructions exhibited slightly higher variance, likely due to differences in the internal connectivity and feature overlap among ontologies. These effects were most visible in the mixed learner-question strategy, where conversational flexibility occasionally led to local inconsistencies or omissions.

Table 1: Pearson correlation coefficients (r) between key performance measures across strategies. p-values are shown in parentheses. Time refers to the number of steps after which the teacher-dialogue learner reaches a plateau; Info is the maximum number of facts exchanged during the training dialogue.

	Reconstruction	20Q
Time	0.385 (p = 0.271)	-0.010 (p = 0.979)
Info	**0.890 (p = 0.001)**	-0.057 (p = 0.875)

We next evaluate the use of acquired knowledge in an applied context, as measured by the 20 Questions task. Figure 4 reports the number of questions per trial for each pedagogical strategy alongside the expert baseline. For each strategy, we ran two-sided Welch's t-tests against the expert and controlled the false discovery rate with Benjamini–Hochberg; significance is encoded under the x-axis as ns ($p_{\mathrm{FDR}} \geq .05$) or * ($p_{\mathrm{FDR}} < .05$). The magenta dashed line marks the expert median. Crucially, six pedagogical strategies are at least not significantly worse than the expert baseline—TD, LQ, TQ, mix-TD-LQ, mix-TD-TQ, and mix-LQ—indicating that, within sampling variability and after multiple-comparison control, these approaches match (and in some cases nominally improve upon) the expert reference in terms of questions required but with the advantage to be, in most natural settings, more flexible and human-friendly.

Combining results from dialogue exchange, knowledge acquisition, and applied performance, we compute correlations among information shared during learning, ontology reconstruction, and 20 Questions performance. No significant relationships emerge between 20 Questions performance and dialogue-level measures (information exchange or time to reach the maximum information), indicating that the amount or pace of information transmission does not directly affect performance in the applied task.

As shown in Table 1, Pearson's r confirms the absence of correlation between 20Q steps and both information exchange and the number of dialogue turns required to reach the plateau (see Figure 1). In contrast, information shared during the dialogue correlates positively with reconstruction accuracy, suggesting that structured exposition supports declarative learning.

However, reconstruction accuracy itself shows no association with 20Q performance ($r = -0.044$, $p = .903$), indicating that even well-acquired knowledge is not necessarily used more effectively in applied contexts.

This dissociation suggests that while structured dialogues effectively support knowledge acquisition, they do not automatically enhance the learner's ability to use that knowledge strategically. Reconstructing the ontology relies on memory and alignment with explicit structure, whereas success in the 20Q Game requires adaptive reasoning and uncertainty management. The findings therefore point to a gap between acquiring factual knowledge and deploying it for efficient information-seeking—an analogue to the distinction between "knowing that" and "knowing how" in human cognition.

Overall, results confirm that pedagogical framing plays a critical role in both the reconstruction and the application of structured

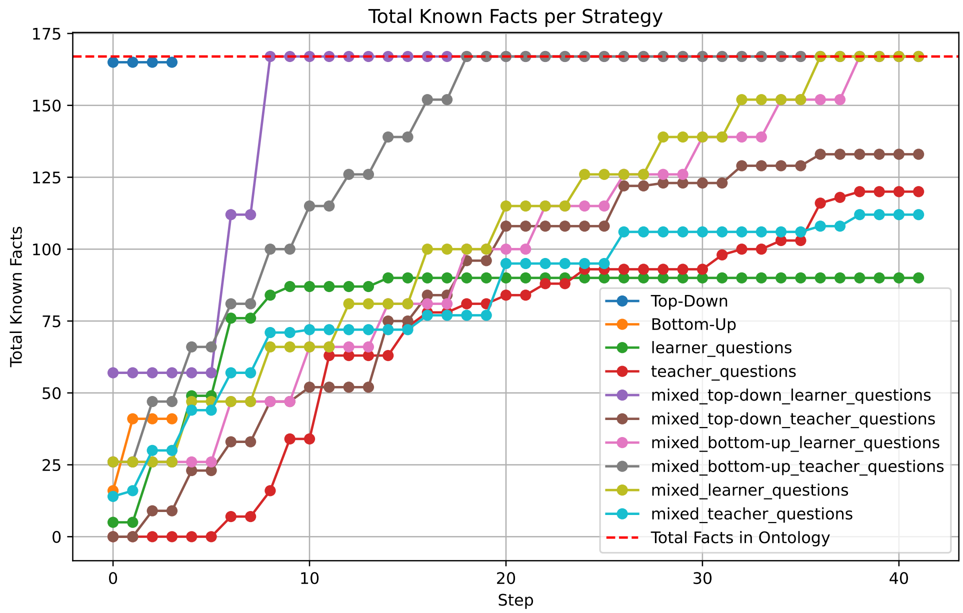

Figure 1: Cumulative knowledge exposure over dialogue steps. The curves quantify information transmission by tracking the total count of unique facts (extracted as entity-feature-value triplets) revealed to the learner over time. The red dashed line marks the total facts available in the ground-truth ontology. The data reveals distinct profiles of information flow: Top-Down strategies provide near-instantaneous coverage, whereas Learner-driven and Mixed approaches facilitate a steady, iterative discovery of facts.

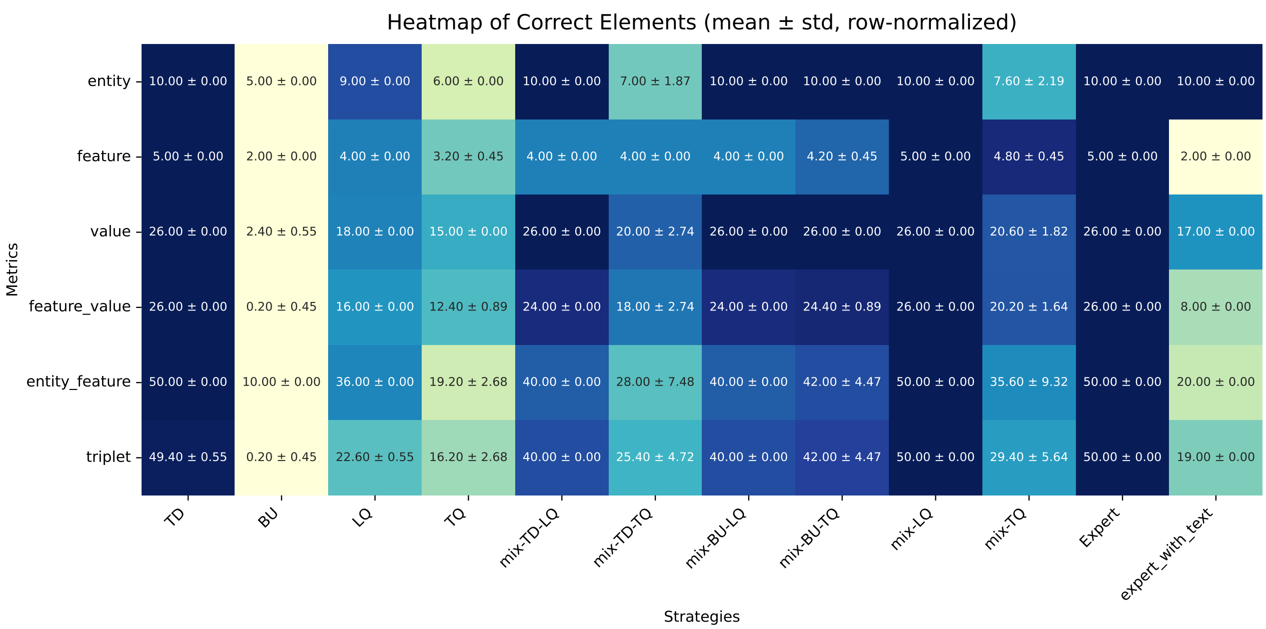

Figure 2: Heatmap of reconstruction accuracy across all feature dimensions and pedagogical strategies. Columns correspond to strategies, rows to features. TD and Mix-LQ configurations achieve the highest and most homogeneous coverage.

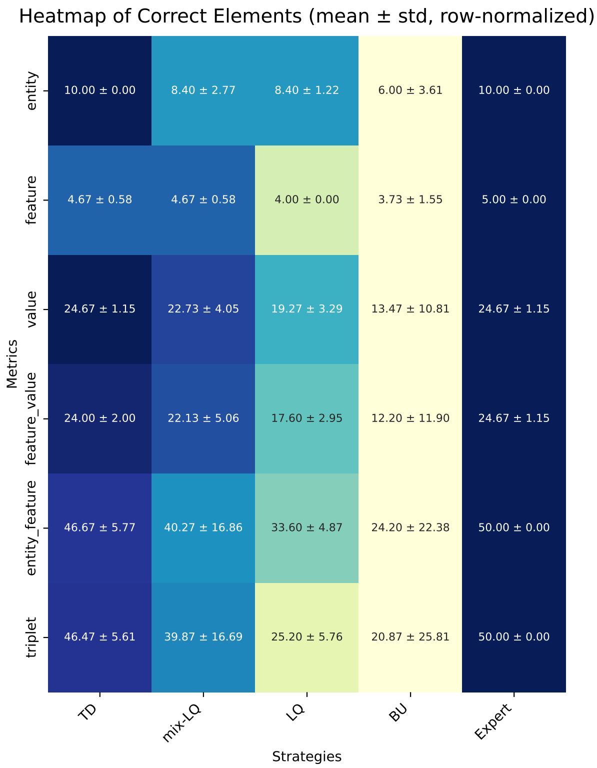

Figure 3: Reconstruction accuracy heatmap for the four strategies selected for the second phase. Mix-LQ shows promising performance but higher instability across repetitions compared to the robust TD condition.

knowledge. Top-down strategies consistently produce stable, nearly complete knowledge representations, validating their suitability for zero-knowledge teaching. In contrast, the mixed learner-question condition—while more natural and human-like—shows greater instability, particularly in deeper relational structures. This suggests that interactive but unverified learning can lead to partial conceptual understanding.

Introducing brief teacher verification phases, where the teacher tests the learner's understanding before moving to the application stage, could help identify and correct missing pieces of knowledge. Such mechanisms would integrate the strengths of mixed learner-question dialogue with the reliability of guided teaching, forming a hybrid strategy that balances naturalness, efficiency, and control.

6 DISCUSSION

In line with [32], our findings suggest that socially grounded instruction may provide a natural and complementary pathway to dataset-based fine-tuning. The advantage of pedagogical conditions over the glossary-style control indicates that interactive, socially grounded exchanges—not mere exposure to information—play a central role in supporting conceptual internalisation in LLMs, pointing toward more human-aligned forms of communication between teachers and artificial learners.

These results also relate to a broader challenge in current LLM architectures: their internal representations are opaque and non-symbolic, making it difficult to ensure that newly acquired information integrates coherently with existing knowledge. The structure of the knowledge presented to an LLM may therefore strongly influence, or at least differently affect, its ability to support more complex forms of inference compared to simpler ones. While retrieval-augmented generation (RAG) pipelines or memory-augmented variants of GPT-like systems aim to address these limitations, they require careful hand-tuning and often lack transparency and control over how information is stored and reused. By contrast, the present pedagogical approach offers an interpretable and socially grounded alternative, where the flow of knowledge can be explicitly modelled, monitored, and corrected through interaction.

A promising future direction is therefore to incorporate teacher verification steps—brief interactive tests that probe the learner's current model before knowledge application. Such methods, inspired by formative assessment in human pedagogy, may close the feedback loop and prevent error propagation.

Beyond their technological relevance, these results bear epistemological implications. If an LLM can be "taught" through explanation and dialogue rather than dataset engineering, knowledge transfer becomes more accessible and human-aligned. This vision of natural instruction allows users to convey domain knowledge directly through conversation, mirroring how humans teach each other. By incorporating rich language interactions and scaffolding techniques, it could be possible to model complex skill acquisition processes akin to those observed in natural social contexts. This progression represents a crucial step toward more sophisticated and human-like AI learning systems and the production of novel contributions of developmental and epigenetic robotics to developmental psychology, reinforcing the significance of embodied combined with socially situated learning [10].

Hence, this study opens the possibility of efficiently improving performance without the need for additional data or retraining required for reasoning [19, 21, 41, 50]. In addition, this interactive training approach, integrating learner question making to reduce uncertainty about the new concepts acquired [33], may lead to more reliable prompting [22, 42, 47] and an active zero-shot or few learning strategy[5] extending the flexibility of these methods with enhanced efficacy and accuracy.

The challenge ahead is to extend these strategies to human–AI teaching settings, exploring how people can intuitively use them to build shared conceptual spaces with artificial learners. Future research could extend this framework to human participants, exploring whether people are able to adopt these pedagogical strategies when teaching an AI partner. Another promising direction is to examine optimisation trade-offs—such as reducing dialogue length, regulating information flow, and analysing how stability scales with ontology size—to refine the approach toward more practical and adaptive forms of interactive AI education.

ACKNOWLEDGMENTS

This research was supported by the Italian Ministry of University and Research under Grant No. 2023-NAZ-0206, PsyFuture – Dipartimento di Eccellenza 2023-2027 and by Volkswagen Foundation

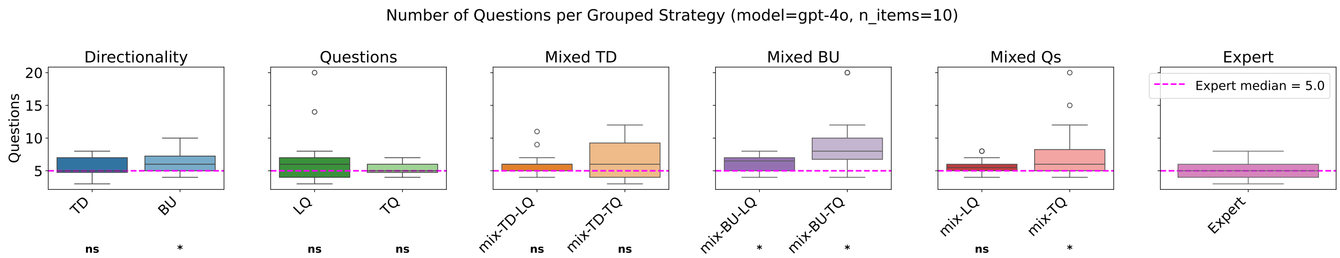

Figure 4: Distribution of questions required to solve the 20-Questions task. Box-plots show the strategy; the last box on the right is the expert baseline with full JSON-structured ontology access. Lower values indicate greater efficiency. Significance markers under x-axis labels report two-sided Welch's t-tests comparing each strategy with the expert baseline, with Benjamini–Hochberg FDR correction across comparisons (ns p≥.05 and * p<.05). The magenta dashed line denotes the expert median.

OpenUp Grant Ref. 9E530 Developing an Artificial Social Childhood (ASC).

REFERENCES

[1] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Chuyuan Fu, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, Daniel Ho, Jasmine Hsu, Julian Ibarz, Brian Ichter, Alex Irpan, Eric Jang, Rosario Jauregui Ruano, Kyle Jeffrey, Sally Jesmonth, Nikhil Joshi, Ryan Julian, Dmitry Kalashnikov, Yuheng Kuang, Kuang-Huei Lee, Sergey Levine, Yao Lu, Linda Luu, Carolina Parada, Peter Pastor, Jornell Quiambao, Kanishka Rao, Jarek Rettinghouse, Diego Reyes, Pierre Sermanet, Nicolas Sievers, Clayton Tan, Alexander Toshev, Vincent Vanhoucke, Fei Xia, Ted Xiao, Peng Xu, Sichun Xu, Mengyuan Yan, and Andy Zeng. 2022. Do As I Can and Not As I Say: Grounding Language in Robotic Affordances. In *arXiv preprint arXiv:2204.01691*.

[2] Luc Berthouze and Tom Ziemke. 2003. Epigenetic robotics—modelling cognitive development in robotic systems.

[3] Leonardo Bertolazzi, Davide Mazzaccara, Filippo Merlo, and Raffaella Bernardi. 2023. Chatgpt's information seeking strategy: Insights from the 20-questions game. In *Proceedings of the 16th International Natural Language Generation Conference*. 153–162.

[4] Parviz Birjandi and Sepideh Jazebi. 2014. A comparative analysis of teachers' scaffolding practices. *International Journal of Language and Linguistics* 2, 3 (2014), 154–164.

[5] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[6] Jerome Bruner. 1985. Child's talk: Learning to use language. *Child Language Teaching and Therapy* 1, 1 (1985), 111–114.

[7] Jerome Bruner. 1985. Vygotsky: A historical and conceptual perspective. *Culture, communication, and cognition: Vygotskian perspectives* 21 (1985), 34.

[8] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. 2023. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712* (2023).

[9] Maya Cakmak, Nick DePalma, Rosa I Arriaga, and Andrea L Thomaz. 2010. Exploiting social partners in robot learning. *Autonomous Robots* 29 (2010), 309–329.

[10] Angelo Cangelosi and Matthew Schlesinger. 2018. From babies to robots: the contribution of developmental robotics to developmental psychology. *Child Development Perspectives* 12, 3 (2018), 183–188.

[11] Susan Carey, Deborah Zaitchik, and Igor Bascandziev. 2015. Theories of development: In dialog with Jean Piaget. *Developmental Review* 38 (2015), 36–54.

[12] Yihong Chen, Bei Chen, Xuguang Duan, Jian-Guang Lou, Yue Wang, Wenwu Zhu, and Yong Cao. 2018. Learning-to-ask: Knowledge acquisition via 20 questions. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1216–1225.

[13] Michelene TH Chi. 2009. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science* 1, 1 (2009), 73–105.

[14] Michelene TH Chi and Ruth Wylie. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist* 49, 4 (2014), 219–243.

[15] Simone Colombani, Dimitri Ognibene, and Giuseppe Boccignone. 2024. One to rule them all: natural language to bind communication, perception and action. In *12th Italian Workshop on Planning and Scheduling (IPS-2024)*.

[16] Graham Cooper and John Sweller. 1987. Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of educational psychology* 79, 4 (1987), 347.

[17] J Bradley Cousins and Kenneth A Leithwood. 1993. Enhancing knowledge utilization as a strategy for school improvement. *Knowledge* 14, 3 (1993), 305–333.

[18] Joachim De Greeff and Tony Belpaeme. 2015. Why robots should be social: Enhancing machine learning through social human-robot interaction. *PLoS one* 10, 9 (2015), e0138061.

[19] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948* (2025).

[20] Huy Ha, Pete Florence, and Shuran Song. 2023. Scaling up and distilling down: Language-guided robot skill acquisition. In *Conference on Robot Learning*. PMLR, 3766–3777.

[21] Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, et al. 2024. Openai o1 system card. *arXiv preprint arXiv:2412.16720* (2024).

[22] Cansu Koyuturk, Emily Theophilou, Sabrina Patania, Gregor Donabauer, Andrea Martinenghi, Chiara Antico, Alessia Telari, Alessia Testa, Sathya Buršić, Franca Garzotto, et al. 2025. Understanding learner-LLM Chatbot interactions and the impact of prompting guidelines. In *International Conference on Artificial Intelligence in Education*. Springer, 364–377.

[23] Hideki Kozima and Jordan Zlatev. 2000. An epigenetic approach to human-robot communication. In *Proceedings 9th IEEE International Workshop on Robot and Human Interactive Communication. IEEE RO-MAN 2000 (Cat. No. 00TH8499)*. IEEE, 346–351.

[24] Jessica Lindblom and Tom Ziemke. 2003. Social situatedness of natural and artificial intelligence: Vygotsky and beyond. *Adaptive Behavior* 11, 2 (2003), 79–96.

[25] A. Lockerd and C. Breazeal. 2004. Tutelage and socially guided robot learning. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, Vol. 4. 3475–3480 vol.4. https://doi.org/10.1109/IROS.2004.1389954

[26] Max Lungarella, Giorgio Metta, Rolf Pfeifer, and Giulio Sandini. 2003. Developmental robotics: a survey. *Connection science* 15, 4 (2003), 151–190.

[27] Andrew J Martin. 2023. Integrating motivation and instruction: Towards a unified approach in educational psychology. *Educational Psychology Review* 35, 2 (2023), 54.

[28] Anthony F Morse and Angelo Cangelosi. 2017. Why are there developmental stages in language learning? A developmental robotics model of language development. *Cognitive Science* 41 (2017), 32–51.

[29] Dimitri Ognibene, Vincenzo G Fiore, and Xiaosi Gu. 2019. Addiction beyond pharmacological effects: The role of environment complexity and bounded rationality. *Neural Networks* 116 (2019), 269–278.

[30] Dimitri Ognibene, Sabrina Patania, Luca Annese, Cansu Koyuturk, Franca Garzotto, Giuseppe Vizzari, Azzurra Ruggeri, and Simone Colombani. 2025. SCOOP: A Framework for Proactive Collaboration and Social Continual Learning through Natural Language Interaction andCausal Reasoning. *arXiv preprint arXiv:2503.10241* (2025).

[31] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 35 (2022), 27730–27744.

[32] Sabrina Patania, Luca Annese, Cansu Koyuturk, Azzurra Ruggeri, and Dimitri Ognibene. 2025. AI Pedagogy: Dialogic Social Learning for Artificial Agents. *arXiv preprint arXiv:2507.21065* (2025).

[33] Sabrina Patania, Emanuele Masiero, Luca Brini, Valentyn Piskovskyi, Dimitri Ognibene, Gregor Donabauer, Udo Kruschwitz, et al. 2024. Large Language Models as an active Bayesian filter: information acquisition and integration. In *Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue.*

[34] Maxime Petit, Stéphane Lallée, Jean-David Boucher, Grégoire Pointeau, Pierrick Cheminade, Dimitri Ognibene, Eris Chinellato, Ugo Pattacini, Ilaria Gori, Uriel Martinez-Hernandez, et al. 2012. The coordinating role of language in real-time multimodal learning of cooperative tasks. *IEEE Transactions on Autonomous Mental Development* 5, 1 (2012), 3–17.

[35] Jean Piaget, Margaret Cook, et al. 1952. *The origins of intelligence in children.* Vol. 8. International universities press New York.

[36] Yuuya Sugita and Jun Tani. 2005. Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive behavior* 13, 1 (2005), 33–52.

[37] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.

[38] John Sweller. 2011. Cognitive load theory. In *Psychology of learning and motivation.* Vol. 55. Elsevier, 37–76.

[39] Akira Taniguchi, Yoshinobu Hagiwara, Tadahiro Taniguchi, and Tetsunari Inamura. 2020. Improved and scalable online learning of spatial concepts and language models with mapping. *Autonomous Robots* 44, 6 (2020), 927–946.

[40] Gemini Robotics Team, Saminda Abeyruwan, Joshua Ainslie, Jean-Baptiste Alayrac, Montserrat Gonzalez Arenas, Travis Armstrong, Ashwin Balakrishna, Robert Baruch, Maria Bauza, Michiel Blokzijl, et al. 2025. Gemini robotics: Bringing ai into the physical world. *arXiv preprint arXiv:2503.20020* (2025).

[41] Kimi Team, Angang Du, Bofei Gao, Bowei Xing, Changjiu Jiang, Cheng Chen, Cheng Li, Chenjun Xiao, Chenzhuang Du, Chonghua Liao, et al. 2025. Kimi k1. 5: Scaling reinforcement learning with llms. *arXiv preprint arXiv:2501.12599* (2025).

[42] Emily Theophilou, Cansu Koyutürk, Mona Yavari, Sathya Bursic, Gregor Donabauer, Alessia Telari, Alessia Testa, Raffaele Boiano, Davinia Hernandez-Leo, Martin Ruskov, et al. 2023. Learning to prompt in the classroom to understand AI limits: a pilot study. In *International conference of the Italian association for artificial intelligence.* Springer, 481–496.

[43] Andrea L Thomaz and Cynthia Breazeal. 2008. Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence* 172, 6-7 (2008), 716–737.

[44] Andrea L Thomaz, Guy Hoffman, and Cynthia Breazeal. 2006. Experiments in socially guided machine learning: understanding how humans teach. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction.* 359–360.

[45] Janneke Van de Pol, Monique Volman, and Jos Beishuizen. 2010. Scaffolding in teacher–student interaction: A decade of research. *Educational psychology review* 22 (2010), 271–296.

[46] Lev Semenovich Vygotsky and Michael Cole. 1978. *Mind in society: Development of higher psychological processes.* Harvard university press.

[47] Jules White, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C Schmidt. 2023. A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382* (2023).

[48] David Wood, Jerome S Bruner, and Gail Ross. 1976. The role of tutoring in problem solving. *Journal of child psychology and psychiatry* 17, 2 (1976), 89–100.

[49] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2023. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR).*

[50] Yang Yue, Zhiqi Chen, Rui Lu, Andrew Zhao, Zhaokai Wang, Shiji Song, and Gao Huang. 2025. Does Reinforcement Learning Really Incentivize Reasoning Capacity in LLMs Beyond the Base Model? *arXiv preprint arXiv:2504.13837* (2025).