geometric measurement instrument for quantifying how different interrogative forms structure uncertainty in language systems. The system models questions as spatial injections into a six-dimensional cubic field representing the interrogatives Who, What, When, Where, Why, and How. Each question perturbs the field deterministically, enabling reconstruction of an "interrogative entropy" profile independent of the stochastic behavior of a language model's generated responses. The instrument computes Interrogative Entropy (H⊠) using Shannon entropy over the cube's state distribution, allowing question sequences to be measured as structured entropy trajectories rather than treated as indistinguishable prompt text. Experimental trials demonstrate that identical interrogative sequences yield identical entropy values across repeated runs, confirming that the system's dynamics are deterministic and governed solely by the geometry of inquiry rather than probabilistic language model outputs. This approach provides a diagnostic capability for distinguishing interrogative structure from answer variability, enabling formal analysis of how specific question forms influence cognitive load, ambiguity distribution, and information geometry within AI systems. The instrument can be used as an evaluation tool for prompt design, interrogation strategy, and model behavior analysis, and offers a foundation for treating inquiry as a measurable system rather than an abstract linguistic artifact.

No existing literature implements the proposed geometric measurement instrument, as none of the 12 reviewed studies employ a six-dimensional cubic field representing classical interrogatives, deterministic spatial injections, or interrogative entropy computed over field states, indicating that this approach represents a novel contribution without direct empirical validation or precedent in current research on interrogative systems and uncertainty quantification.

Abstract

This systematic review of 12 sources examining interrogative systems and uncertainty quantification in language and AI contexts found no existing work implementing the specific geometric measurement instrument described in the research question. None of the included studies reference the six classical interrogatives (Who, What, When, Where, Why, How) as a structural framework for categorizing questions , nor do any employ spatial injections into a six-dimensional cubic field or compute "interrogative entropy" over field states rather than response distributions. While Shannon entropy and conditional entropy are widely used for uncertainty quantification, these measures operate over response spaces and grounding variables rather than interrogative form categories. The determinism requirement—identical interrogative sequences yielding identical entropy values—remains largely unaddressed, as

most systems rely fundamentally on probabilistic language model outputs .

Despite this gap, several findings provide indirect support for the instrument's underlying premises. Multiple studies demonstrate that strategic question selection significantly affects uncertainty reduction and downstream performance , with one finding that a one-standard-deviation increase in information gain reduces expected task length by 43% . Evidence that models fail to consistently retrieve correct answers under varied question forms supports the conceptual separation of interrogative structure from response variability. Rothe et al.'s treatment of questions as formal programs with deterministic outputs and Saparov et al.'s deterministic dataset generation from symbolic ontologies provide methodological precedents. The literature thus establishes that interrogative form carries measurable information affecting AI system behavior, but the specific geometric architecture and deterministic entropy measurement proposed in the research question represent a novel contribution without direct empirical validation in existing work.

Paper search

We performed a semantic search using the query ”geometric measurement instrument for quantifying how different interrogative forms structure uncertainty in language systems. The system models questions as spatial injections into a six-dimensional cubic field representing the interrogatives Who, What, When, Where, Why, and How. Each question perturbs the field deterministically, enabling reconstruction of an "interrogative entropy" profile independent of the stochastic behavior of a language model's generated responses. The instrument computes Interrogative Entropy (H_i) using Shannon entropy over the cube's state distribution, allowing question sequences to be measured as structured entropy trajectories rather than treated as indistinguishable prompt text. Experimental trials demonstrate that identical interrogative sequences yield identical entropy values across repeated runs, confirming that the system's dynamics are deterministic and governed solely by the geometry of inquiry rather than probabilistic language model outputs.

This approach provides a diagnostic capability for distinguishing interrogative structure from answer variability, enabling formal analysis of how specific question forms influence cognitive load, ambiguity distribution, and information geometry within AI systems. The instrument can be used as an evaluation tool for prompt design, interrogation strategy, and model behavior analysis, and offers a foundation for treating inquiry as a measurable system rather than an abstract linguistic artifact." across over 138 million academic papers from the Elicit search engine, which includes all of Semantic Scholar and OpenAlex.

We retrieved the 500 papers most relevant to the query.

Screening

We screened in sources that met these criteria:

- **Computational Interrogative Analysis**: Does the study involve computational measurement or quantification of interrogative structures in natural language processing systems?
- **Mathematical Modeling Approach**: Does the research employ geometric, spatial, or mathematical modeling approaches to analyze linguistic elements or question structures?
- **Information Theory Components**: Does the study investigate entropy, information theory, or uncertainty quantification in language systems or AI models?
- **Question-Processing Relationships**: Does the research examine the relationship between question types/forms and cognitive load, ambiguity, or information processing in computational systems?

- **AI Interrogation Strategies**: Does the study analyze prompt design, question sequencing, or interrogation strategies in AI systems?
- **Computational System Focus**: Does the study include computational or AI system components (rather than focusing solely on human linguistic behavior)?
- **Quantitative Measurement**: Does the research include quantitative measurement or modeling components (rather than being limited to traditional qualitative linguistic analysis only)?
- **Question Structure Analysis**: Does the study analyze question structure or interrogative forms (rather than examining only answer generation or response quality)?
- **Empirical Methodology**: Does the study employ empirical methodology with measurable outcomes (rather than being a case report, opinion piece, or purely theoretical paper)?

We considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Data extraction

We asked a large language model to extract each data column below from each paper. We gave the model the extraction instructions shown below for each column.

- **Interrogative System Design**:

  Extract details about the questioning/interrogative system studied including:

- Type of system (automated question generation, interactive dialogue, question selection algorithm, etc.)
- How questions are generated, selected, or structured
- Whether questions are mapped to specific dimensions or categories
- Any reference to the six classical interrogatives (Who, What, When, Where, Why, How)
- System architecture or framework used

- **Uncertainty Quantification Method**:

  Extract how uncertainty or entropy is measured including:

- Specific entropy measures used (Shannon entropy, conditional entropy, mutual information, etc.)
- Mathematical formulations or calculations described
- Whether uncertainty is measured over responses, question structure, or both
- Any mention of 'interrogative entropy' or similar concepts
- How uncertainty relates to information gain or reduction

- **Geometric/Spatial Modeling**:

  Extract any geometric, spatial, or dimensional modeling approaches including:

- Use of multi-dimensional spaces or fields
- Spatial representations of questions or uncertainty
- Geometric transformations or perturbations
- Field theory or topology applications
- Any mention of cubic fields, spatial injections, or dimensional structures
- Trajectory or path modeling in question spaces

- **Question Structure Framework**:

  Extract how questions are categorized, structured, or analyzed including:

- Question taxonomies or classification systems used
- Structural features of questions that are measured
- How question sequences or chains are modeled
- Whether questions are treated as discrete units or continuous structures
- Any analysis of question form vs. content
- Relationship between question type and uncertainty

- **Deterministic vs Stochastic**:

  Extract information about the deterministic or probabilistic nature of the approach including:

- Whether the system produces consistent/identical outputs for identical inputs
- Role of randomness or probability in the model
- How the approach handles or separates deterministic structure from stochastic behavior
- Whether results are reproducible across runs
- Any discussion of separating question structure from answer variability

- **Evaluation Metrics**:

  Extract all performance measures and evaluation criteria including:

- Quantitative metrics used (accuracy, efficiency, information gain, etc.)
- How question quality or effectiveness is assessed
- Measures of cognitive load, ambiguity, or complexity
- Comparison baselines or benchmarks
- Statistical significance tests
- Any novel evaluation frameworks introduced

- **Applications Tested**:

  Extract specific use cases and applications including:

- Domain areas tested (dialogue systems, recommendation, medical diagnosis, etc.)
- Whether used for prompt design, model evaluation, or interrogation strategy
- Interactive vs. non-interactive settings
- Human-AI vs. AI-AI interactions
- Real-world deployment or theoretical analysis
- Specific AI models or systems evaluated

- **Key Findings**:

  Extract main results about interrogative structure and entropy including:

- How question form affects uncertainty or information geometry
- Relationships between question type and cognitive load
- Evidence for separating question structure from answer variability
- Performance improvements or insights gained
- Any validation of theoretical predictions
- Comparative effectiveness of different questioning approaches
- Limitations or failure cases identified

Characteristics of Included Studies

The systematic search identified 12 sources examining interrogative systems and uncertainty quantification in language and AI contexts. None of the included studies directly addressed the specific geometric measurement instrument described in the research question—a six-dimensional cubic field representing the classical interrogatives (Who, What, When, Where, Why, How) with deterministic spatial injections. However, the literature provides substantial coverage of related concepts including information-theoretic question selection, entropy-based uncertainty quantification, and question structure frameworks.

Study	Full text retrieved?	Study Type	Domain	Primary Focus
Sachithra Hemachandra et al., 2014	No	Primary study	Robotics/spatial semantics	Entropy reduction through targeted questions
Jimmy Wang et al., 2025	Yes	Primary study	Adaptive assessment	Meta-learned language models for uncertainty quantification
Robin Deits et al., 2013	No	Primary study	Human-robot dialogue	Information-theoretic clarifying dialog
Dylan Hutson et al., 2025	No	Primary study	LLM evaluation	Strategic question-asking measurement
Hongxin Ding et al., 2025	No	Primary study	Medical diagnosis	Proactive questioning in medical LLMs
Yu Feng et al., 2024	Yes	Primary study	Uncertainty estimation	Multi-agent interaction for uncertainty
Stefanie Tellex et al., 2012	Yes	Primary study	Human-robot dialogue	Template-based clarifying questions
Kwan Ho Ryan Chan et al., 2025	No	Primary study	Interactive QA	Conformal prediction for question selection
Wasu Top Piriyakulkij et al., 2023	Yes	Primary study	Preference inference	Entropy-based question selection
Anselm Rothe et al., 2017	No	Primary study	Cognitive modeling	Questions as formal programs
Abulhair Saparov et al., 2022	Yes	Primary study	LLM reasoning	Chain-of-thought evaluation
Michael J.Q. Zhang et al., 2023	Yes	Primary study	Ambiguity resolution	Clarifying question frameworks

The studies span multiple domains including robotics and human-robot interaction , medical diagnosis , adaptive assessment , LLM evaluation , preference inference , and general NLP applications . All studies examined automated or semi-automated question generation or selection systems, though none employed the six-dimensional interrogative classification scheme central to the research question.

Interrogative System Design Approaches

The included studies employ diverse approaches to question generation and selection, though none implement a geometric model based on the six classical interrogatives.

Study	Question Generation Method	Reference to Classical Interrogatives	Architecture
Sachithra Hemachandra et al., 2014	Targeted questions to reduce entropy	Not mentioned	Optimization problem balancing information value with cost
Jimmy Wang et al., 2025	Meta-learned language model simulation	Not mentioned	Greedy selection and MCTS for question selection
Robin Deits et al., 2013	Information-theoretic selection maximizing entropy reduction	Not mentioned	G3 graphical model
Dylan Hutson et al., 2025	Free-form questions without predefined choices	Not mentioned	GuessingGame protocol with Bayesian and entropy metrics
Hongxin Ding et al., 2025	MCTS with Shapley Information Gain rewards	Not mentioned	Two-stage training pipeline
Yu Feng et al., 2024	Questions requiring same underlying information from different perspectives	Not mentioned	Multi-agent interaction
Stefanie Tellex et al., 2012	Template-based algorithm targeting high-uncertainty variables	Not mentioned	G3 framework
Kwan Ho Ryan Chan et al., 2025	Information Pursuit greedy algorithm	Not mentioned	Conformal prediction sets
Wasu Top Piriyakulkij et al., 2023	Sampling from proposal distribution with information gain optimization	Not mentioned	Probabilistic model with LLM-defined distributions
Anselm Rothe et al., 2017	Questions as formal programs with probability distribution	Not mentioned	Compositional program space
Abulhair Saparov et al., 2022	Not applicable—focus on reasoning evaluation	Not applicable	Synthetic ontology-based dataset
Michael J.Q. Zhang et al., 2023	Oracle method with few-shot prompting	Not mentioned	Three-subtask framework for ambiguity resolution

Notably, none of the 12 studies reference the six classical interrogatives (Who, What, When, Where, Why, How) as a structural framework for categorizing or modeling questions . The predominant approach treats questions as information-gathering instruments optimized for entropy reduction rather than as instances of distinct interrogative categories with unique geometric properties.

Uncertainty Quantification Methods

All studies that address uncertainty employ information-theoretic measures, with Shannon entropy and its variants being the predominant framework.

Study	Entropy Measure	Uncertainty Scope	Relation to Information Gain
Sachithra Hemachandra et al., 2014	Not specified	Responses to targeted questions	Uncertainty reduction improves map accuracy
Jimmy Wang et al., 2025	Conditional entropy	Responses	Expected Information Gain (EIG)
Robin Deits et al., 2013	Shannon entropy	Responses/command understanding	Maximize entropy reduction
Dylan Hutson et al., 2025	Not specified	Not specified	Higher IG predicts efficiency
Hongxin Ding et al., 2025	Not mentioned	Not mentioned	Shapley Information Gain
Yu Feng et al., 2024	Weighted entropy	Responses from multiple agents	Predictive entropy approximation
Stefanie Tellex et al., 2012	Shannon entropy	Grounding variables	Select questions targeting uncertain phrases
Kwan Ho Ryan Chan et al., 2025	Conditional entropy (indirect)	Prediction sets	Minimize uncertainty via information gain
Wasu Top Piriyakulkij et al., 2023	Expected entropy minimization	Responses to questions	Maximize expected information gain
Anselm Rothe et al., 2017	Not mentioned	Not mentioned	Not mentioned
Abulhair Saparov et al., 2022	Not mentioned	Not mentioned	Not applicable
Michael J.Q. Zhang et al., 2023	Entropy over user intents	User intents	Expected performance improvement

No study employed the specific concept of ”interrogative entropy” as defined in the research question—that is, entropy computed over a geometric field representing interrogative categories. The closest conceptual parallel appears in systems that compute Shannon entropy over grounding variables or conditional entropy over potential answers , but these measure uncertainty over content spaces rather than interrogative form spaces. The INTENT-SIM approach estimates entropy over user intents , which represents a semantic rather than structural categorization of questions.

Geometric and Spatial Modeling

The research question posits questions as spatial injections into a six-dimensional cubic field. The literature provides minimal support for such geometric approaches to modeling interrogative structure.

Study	Spatial/Geometric Approach	Dimensional Modeling
Sachithra Hemachandra et al., 2014	Hybrid metric, topological, and semantic representation	Not six-dimensional interrogative field
Stefanie Tellex et al., 2012	G3 framework maps language to spatial groundings including objects, places, paths	Hierarchical linguistic composition
All other studies	Not mentioned	Not applicable

Only two studies employ spatial representations. Hemachandra et al. use a hybrid metric-topological-semantic representation for spatial-semantic maps , while Tellex et al. employ the G3 framework to map language to external world groundings including objects, places, and paths . However, neither implements a cubic field based on interrogative categories, nor do they model questions as dimensional perturbations. The remaining ten studies do not discuss geometric or spatial modeling approaches .

Question Structure and Classification

While no study implements a six-dimensional interrogative taxonomy, several employ structural frameworks for analyzing or categorizing questions.

Study	Question Categorization	Structural Features Measured	Treatment of Questions
Yu Feng et al., 2024	Entity-centric QA, general QA, false assumption QA	Consistency across agents and varied questions	Discrete units answered by individual agents
Stefanie Tellex et al., 2012	By ability to disambiguate uncertain variables	Entropy reduction potential	Discrete units in iterative process
Wasu Top Piriyakulkij et al., 2023	Yes/no questions about single features	Information gain	Constrained binary structure
Anselm Rothe et al., 2017	Questions as formal programs	Conciseness, informativeness, complexity	Compositional program structures
Most other studies	Not mentioned	Not applicable	Not specified

The most sophisticated structural treatment appears in Rothe et al., which treats questions as formal programs operating over a compositional space . This approach shares conceptual similarities with the research question's deterministic modeling of interrogative form, as programs produce consistent outputs for consistent inputs. Tellex et al. standardize question form using templates ("What do the words X refer to?") , which represents a rudimentary form-content separation. Piriyakulkij et al. constrain questions to binary yes/no format about single features , limiting structural variation but enabling cleaner information-theoretic analysis.

Deterministic vs. Stochastic Properties

The research question emphasizes determinism—identical interrogative sequences yielding identical entropy values across runs. The literature presents a mixed picture.

Study	Deterministic Elements	Stochastic Elements	Reproducibility
Sachithra Hemachandra et al., 2014	Systematic entropy reduction through optimization	Implied probabilistic elements	Not explicitly addressed
Jimmy Wang et al., 2025	Training structure	Autoregressive simulation, stochastic gradient descent	Varies due to randomness
Yu Feng et al., 2024	Assumption of consistent recall for certain models	Weighted entropy, agent interactions	Not entirely reproducible
Stefanie Tellex et al., 2012	None identified	Probabilistic G3 model	Variable based on probabilistic factors
Kwan Ho Ryan Chan et al., 2025	Conformal prediction sets as distribution-free measurement	LLM probability variations	Not explicitly addressed
Wasu Top Piriyakulkij et al., 2023	None identified	Probabilistic sampling and model optimization	May not be reproducible
Anselm Rothe et al., 2017	None identified	Probability distribution over programs	Not addressed
Abulhair Saparov et al., 2022	Deterministic dataset generation from symbolic ontologies	Probabilistic model performance evaluation	Code and outputs made available for reproducibility
Michael J.Q. Zhang et al., 2023	Not explicitly discussed	Not explicitly discussed	Not addressed

The only study explicitly implementing deterministic generation is Saparov et al., where the PRONTOQA dataset is generated from symbolic ontologies and proofs , though this concerns reasoning evaluation rather than interrogative structure. Chan et al.'s conformal prediction approach offers distribution-free uncertainty measurement , which provides robustness against model probability variations but does not ensure deterministic outputs. Most systems rely fundamentally on probabilistic language model outputs , making the separation of deterministic interrogative structure from stochastic response behavior—a central goal of the proposed instrument—theoretically relevant but not empirically demonstrated in current literature.

Evaluation Metrics and Performance

The studies employ diverse evaluation approaches, with information gain being the most common metric related to the research question's focus on interrogative effectiveness.

Study	Primary Metrics	Question Effectiveness Measure	Baselines
Sachithra Hemachandra et al., 2014	Entropy reduction, map accuracy	Entropy reduction	Not mentioned
Jimmy Wang et al., 2025	Accuracy, perplexity, expected calibration error, Brier score	Reliability diagrams	Base LLM, In-Context Tuning
Robin Deits et al., 2013	Shannon entropy reduction, command understanding accuracy	Entropy reduction maximization	Baseline question-selection strategies
Dylan Hutson et al., 2025	Information gain metrics (Bayesian, entropy-based)	IG correlation with game efficiency	Not mentioned
Hongxin Ding et al., 2025	6.29% improvement over state-of-the-art, 54.45% gain over reactive paradigm	Shapley Information Gain	Partial-information medical benchmarks
Yu Feng et al., 2024	AUROC, accuracy, abstention rate, truthfulness score	Accuracy and truthfulness	Four uncertainty baselines, seven hallucination baselines
Stefanie Tellex et al., 2012	Correctly grounded noun phrases (%), questions asked (%)	Disambiguation performance	No questions, all questions
Kwan Ho Ryan Chan et al., 2025	Predictive performance, query chain length	Information gain	Previous IP approaches, chain-of-thought methods
Wasu Top Piriyakulkij et al., 2023	Task performance, user interactions, information gain per round	Entropy minimization, model change maximization	Vanilla LLM, ReAct LLM
Abulhair Saparov et al., 2022	Strict/skip/broad/valid proof accuracy	Proof step correctness	Label accuracy comparison
Michael J.Q. Zhang et al., 2023	AUROC, contrastive accuracy, answer recall, 3-way classification accuracy	Performance improvement under fixed interaction budget	Various LLMs

No study employs metrics specifically designed to measure interrogative structure independent of content. The GuessingGame protocol represents the most direct attempt to quantify question quality through information gain, demonstrating that a one-standard-deviation increase in information gain reduces expected game length by 43% . This finding supports the premise that question form affects downstream efficiency, though it does not isolate the contribution of interrogative category (Who, What, etc.) from content.

Key Findings Related to Interrogative Structure and Uncertainty

Several studies provide findings relevant to understanding how question form structures uncertainty, though none directly validate the six-dimensional geometric model.

Question Form and Uncertainty Reduction: Multiple studies demonstrate that strategic question selection significantly reduces uncertainty. Hemachandra et al. show that targeted questions reduce entropy in spatial-semantic representations, improving map accuracy . Deits et al. report improved command understanding accuracy while asking fewer questions compared to baselines . Hutson et al. find that higher information gain strongly predicts questioning efficiency, with constrained prompting enabling weaker models to improve performance .

Separating Structure from Variability: The literature provides limited evidence for separating question structure from answer variability. Ding et al.'s Shapley Information Gain focuses on question value rather than answer variability . Zhang et al. discuss disentangling uncertainty estimation from ambiguity , which aligns conceptually with separating interrogative structure from response stochasticity. Feng et al. demonstrate that models often fail to consistently retrieve correct answers under diverse questions even when knowing the answer , suggesting that question form variations affect response behavior independently of underlying knowledge.

Performance Improvements: Quantitative improvements from optimized questioning include: ProMed outperforming state-of-the-art methods by 6.29% and achieving 54.45% gains over reactive paradigms ; DiverseAgentEntropy improving accuracy by 2.5% compared to self-consistency approaches ; and Piriyakulkij et al.'s entropy reduction algorithm outperforming baselines while using fewer questions .

Limitations Identified: Several studies note limitations relevant to the proposed instrument. Tellex et al. identify failures due to unhelpful answers and complex environments . Feng et al. note that models struggle with parametric knowledge retrieval under varied questioning . Piriyakulkij et al. acknowledge that their binary question constraint limits generalizability .

Synthesis

The systematic review reveals a fundamental gap between current approaches and the proposed geometric measurement instrument. While the literature demonstrates substantial interest in information-theoretic approaches to question selection and uncertainty quantification, no existing work implements the specific architectural features described in the research question: a six-dimensional cubic field representing classical interrogatives, deterministic spatial injections, or interrogative entropy computed over field states rather than response distributions.

The absence of the six-dimensional interrogative framework across all 12 studies may reflect either the novelty of this approach or a fundamental methodological divergence. Current systems optimize question selection based on expected information gain from responses , treating questions as instruments for reducing uncertainty about external states rather than as geometric objects with intrinsic structural properties.

The determinism requirement presents particular challenges. Most systems rely on probabilistic language models , making the separation of deterministic interrogative structure from stochastic response behavior technically difficult. Chan et al.'s conformal prediction approach offers distribution-free uncertainty measurement that could provide robustness, while Saparov et al.'s symbolic ontology generation demonstrates deterministic structure in evaluation contexts.

Several findings indirectly support the proposed instrument's premises. The consistent finding that question selection affects downstream performance suggests that interrogative form carries measurable information. Feng et al.'s observation that response consistency varies with question form supports the conceptual separation of interrogative structure from response variability. Rothe et al.'s treatment of questions as formal programs provides a precedent for deterministic modeling of interrogative structure.

The practical applications tested—robotics , medical diagnosis , adaptive assessment , and LLM evaluation —suggest domains where a geometric interrogative instrument could provide diagnostic value. The emphasis on interactive settings across studies indicates sustained interest in understanding question-answer dynamics that the proposed instrument could formalize.

References

Abulhair Saparov, and He He. "Language Models Are Greedy Reasoners: A Systematic Formal Analysis of Chain-of-Thought." *International Conference on Learning Representations*, 2022.

Anselm Rothe, B. Lake, and T. Gureckis. "Question Asking as Program Generation." *Neural Information Processing Systems*, 2017.

Dylan Hutson, Daniel Vennemeyer, Aneesh Deshmukh, Justin Zhan, and Tianyu Jiang. "GuessingGame: Measuring the Informativeness of Open-Ended Questions in Large Language Models." *arXiv.org*, 2025.

Hongxin Ding, Baixiang Huang, Yue Fang, Weibin Liao, Xinke Jiang, Zheng Li, Junfeng Zhao, and Yasha Wang. "ProMed: Shapley Information Gain Guided Reinforcement Learning for Proactive Medical LLMs." *arXiv.org*, 2025.

Jimmy Wang, Thomas P. Zollo, Richard Zemel, and Hongseok Namkoong. "Adaptive Elicitation of Latent Information Using Natural Language." *arXiv.org*, 2025.

Kwan Ho Ryan Chan, Yuyan Ge, Edgar Dobriban, Hamed Hassani, and Ren'e Vidal. "Conformal Information Pursuit for Interactively Guiding Large Language Models." *arXiv.org*, 2025.

Michael J.Q. Zhang, and Eunsol Choi. "Clarify When Necessary: Resolving Ambiguity Through Interaction with LMs." *arXiv.org*, 2023.

Robin Deits, Stefanie Tellex, Pratiksha Thaker, D. Simeonov, T. Kollar, and N. Roy. "Clarifying Commands with Information-Theoretic Human-Robot Dialog." *J. Hum. Robot Interact.*, 2013.

Sachithra Hemachandra, Matthew R. Walter, and S. Teller. "Information Theoretic Question Asking to Improve Spatial Semantic Representations." *AAAI Fall Symposia*, 2014.

Stefanie Tellex, Pratiksha Thaker, Robin Deits, T. Kollar, and N. Roy. "Toward Information Theoretic Human-Robot Dialog." *Robotics: Science and Systems*, 2012.

Wasu Top Piriyakulkij, Volodymyr Kuleshov, and Kevin Ellis. "Active Preference Inference Using Language Models and Probabilistic Reasoning." *arXiv.org*, 2023.

Yu Feng, Phu Mon Htut, Zheng Qi, Wei Xiao, Manuel Mager, Nikolaos Pappas, Kishaloy Halder, Yang Li, Yassine Benajiba, and Dan Roth. "DiverseAgentEntropy: Quantifying Black-Box LLM Uncertainty Through Diverse Perspectives and Multi-Agent Interaction." *arXiv.org*, 2024.