ON OPEN-ENDED INFORMATION SEEKING FOR INFORMATION ELICITATION AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Information elicitation is an open-ended information-seeking problem in which an interaction can unfold in many potentially valuable directions, requiring an elicitor to continually determine which information to pursue as new information emerges. In agentic elicitation, these decisions may be delegated to a foundation model, yet how model choice shapes the resulting information-seeking behavior remains understudied. We study how judgments about information value vary across LLMs and how these differences shape sequential information seeking. We first examine these judgments across 11 LLMs spanning multiple model families and parameter scales, using a shared set of information and elicitation objectives. We then develop a controlled elicitation simulation in which different models encounter the same information space and use the same selection rule, isolating these judgments from question generation and respondent behavior. Using this setting, we characterize the breadth–depth behavior that emerges from model-specific information-seeking preferences over the course of elicitation. We further examine how interaction history changes the evaluation and subsequent selection of prospective information. We test the robustness and boundaries of these findings through sensitivity analyses and ablations over the opportunities available to the elicitor, the response labels used to operationalize information-seeking preferences, the presence of interaction history, and whether redundancy is explicitly relevant to the assessment. The project code, data, and trajectory files are available at https://github.com/anonymousauthor2352/open-elicitation.

1 INTRODUCTION

Information elicitation is an open-ended information-seeking problem in which the objective is to maximize the *value* of the facts uncovered through dialogue, a problem that arises in a wide range of real-world interactions, including recruiting, journalism, and oral history. It differs fundamentally from traditional information-seeking, which asks questions to distinguish among competing hypotheses by maximizing expected information gain (Feng et al., 2024; Chopra & Shah, 2025; Lin et al., 2026). Different lines of questioning may uncover entirely different sets of facts, from which equal or vastly different benefit can be derived (De Lima & Yang, 2026). When these decisions are delegated to an LLM-based elicitation agent, the choice of foundation model will shape what information is evaluated and prioritized. Yet how foundation model choice shapes this process over the course of an interaction remains understudied.

The open-ended nature of elicitation makes these prioritization decisions particularly consequential. Deciding what to select next, however, is not captured by information gain alone: two questions can convey the same amount of Shannon information (Shannon, 1948; Yu et al., 2020) while differing substantially in whether the information they reveal is worth pursuing. For example, asking whether a candidate is willing to relocate or prefers coffee to tea can provide one bit of Shannon information in either case, while differing dramatically in value to a recruiter. Classical Value of Information instead values information through its contribution to a downstream decision (Howard, 1966; Grass & Zilberstein, 2000), but open-ended elicitation need not have a specified downstream decision. Nor can the problem generally be reduced to predefined information targets. Restricting an interaction to questions about where a candidate previously worked or whether they know Python reduces elicitation to a checklist that excludes unanticipated discoveries. Learning that a candidate unexpectedly "*worked at NASA*," for example, may open a new line of inquiry about their mission,

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

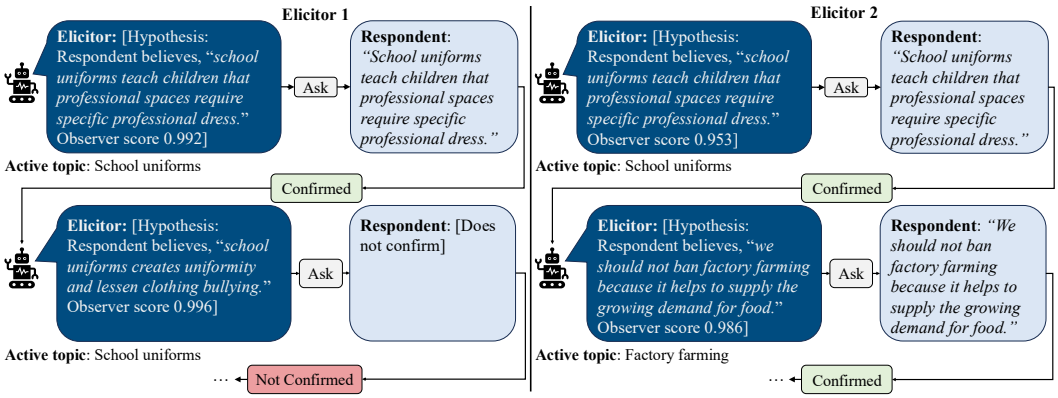

Figure 1: Illustration of the experimental elicitation setup and example trajectories.

role, or challenges. As illustrated in Figure 1, differences in how prospective information is prioritized can therefore redirect the subsequent course of elicitation and ultimately determine which regions of the information space are explored.

Studying how foundation model choice shapes information seeking presents several challenges. LLMs do not implement an explicit information-seeking policy from which their behavior can be readily anticipated: local preferences over prospective information may compound into trajectory-level patterns that only become apparent through sequential interaction. Moreover, information-seeking preferences vary with the objective, information already acquired, and opportunities available at each step, and therefore evolve over the course of an interaction. Lastly, information-seeking preferences interact with question generation, respondent behavior, and the information available at subsequent steps, making it difficult to isolate how the foundation model's preferences shape sequential elicitation.

To address these challenges, we make the following contributions:

- We systematically characterize information-seeking preferences across 11 open-weight instruction-tuned LLMs spanning multiple model families and parameter scales, multiple information-seeking objectives, and a shared information space.
- We introduce a sequential elicitation simulation that separates the evaluation of prospective information from question generation and respondent behavior, allowing model-dependent information-seeking preferences to be studied while holding other components fixed.
- We examine information-seeking preferences in sequential elicitation, studying how repeated candidate-level decisions shape the breadth and depth of elicitation trajectories and how interaction history changes the evaluation and subsequent selection of prospective information.
- We conduct extensive sensitivity analyses and ablations to establish the robustness and boundaries of these findings, varying the opportunities available to the elicitor, the response labels used to operationalize information-seeking preferences, the presence of interaction history, and whether redundancy is explicitly made relevant to the assessment.

2 Related Work

A substantial line of work formulates information seeking as selecting observations or questions that reduce uncertainty about an underlying hypothesis, state, or parameter (Nelson, 2005). This framing has been used to study human question asking (Testoni et al., 2023) and to guide language models through expected information gain (Mazzaccara et al., 2024; Feng et al., 2024; Chopra & Shah, 2025) or the contribution of information-seeking steps toward reaching a target answer (Lee et al., 2026). Related work in active sensing similarly values observations by how they change what an agent knows about its environment (Veiga & Renoux, 2023). These approaches are well suited to settings such as diagnosis, troubleshooting, and identification, where information seeking is organized around uncertainty about an unknown target (Croskerry, 2009; Yazdani et al., 2017). In open-ended elicitation, there may instead be many possible discoveries worth pursuing rather than a predefined target whose uncertainty is to be resolved.

Decision-theoretic Value of Information (VoI) instead values information by its contribution to a downstream decision, distinguishing the amount of information acquired from its usefulness to a

decision maker (Howard, 1966; Grass & Zilberstein, 2000). VoI has been used to guide costly information acquisition (Bilgic & Getoor, 2007; Weiss & Taskar, 2013) and extended to sequential decision making (Arumugam & Van Roy, 2021), human–AI decision making (Emmons et al., 2025; Guo et al., 2026), and adaptive communication (Dong et al., 2026). The resulting value can depend strongly on the decision maker's preferences and utility function (Delquié, 2008; Abbas & Hazen, 2025). This decision-theoretic framing does not directly cover open-ended elicitation, where information may be worth pursuing without a specified downstream decision.

Conversational information seeking studies information acquisition through multi-turn interaction (Dalton et al., 2022; Tang & Yang, 2017). Statistical dialogue management models such interaction as sequential decision making under uncertainty (Williams & Young, 2007; Young et al., 2013), while mixed-initiative systems study when agents should answer, clarify, or proactively contribute information (Wu et al., 2023; Deng et al., 2025; 2024; Wu et al., 2019). Other work focuses specifically on eliciting information from people (Bi et al., 2021; Ren et al., 2021; Li et al., 2022; Kim et al., 2024), including follow-up questioning in interviews (Hu et al., 2024) and elicitation across domains such as academic interviews, journalism, judicial proceedings, and oral history (De Lima & Yang, 2026). Work on broader question spaces includes open-ended inquiry (Lin et al., 2026; 2025; Hamami, 2014), curiosity-driven questions (Ko et al., 2020), and question salience for prioritizing questions by their contribution to understanding (Wu et al., 2024). Together, this work establishes settings in which many lines of inquiry are possible, but leaves open how LLMs prioritize among them and how those choices shape the course of elicitation.

Our work focuses on the information-seeking preferences that arise when LLMs prioritize among these possibilities. We isolate the evaluation of prospective information from question generation and respondent behavior, and study how these preferences shape sequential elicitation.

3 Information-Seeking Preferences in LLMs

3.1 Operationalizing Information-Seeking Preferences

We study how LLMs prioritize prospective information given an information-seeking objective and the information already acquired. Let a denote an LLM, h a candidate piece of information, o the information-seeking objective, and H_t the interaction history at step t. We represent the model's assessment with an information-seeking score $s_a(h \mid o, H_t)$, where higher scores indicate that model a more strongly favors pursuing h in the given context. We use these scores to characterize a model's *information-seeking preferences*. This formulation allows preferences to depend on the model, objective, and interaction history without prescribing which considerations, such as relevance, novelty, specificity, or usefulness, the model should use in making the assessment.

For our empirical study, we operationalize each model's assessment using its probability of affirming that a candidate piece of information is worth pursuing for a specified objective and context. Prior work has shown that LLMs can evaluate the utility of information beyond semantic relevance, distinguishing information that is merely related to a query from information that is useful for addressing it (Zhang et al., 2024; 2025). We leverage this capacity to study model judgments conditioned on a specified information-seeking objective and interaction history. Each model thereby induces its own information-seeking preferences. Because these probabilities need not be calibrated across models, we interpret them as model-specific information-seeking scores rather than directly comparable measurements of information value.

For each candidate–objective pair, the objective is incorporated into the system instruction and the argument is supplied as the user message using the model's native chat template (full prompt in Appendix A). The model is asked whether the argument provides information worth pursuing for the specified objective and to choose one of three possible responses: a positive response y_v, a negative response y_n, and an uncertain response y_u. We include an uncertain option to reduce the tendency of the resulting scores to collapse toward extreme values (Cohen et al., 2024) (see Appendix B). Rather than generating a response and treating the resulting discrete output as the model's assessment, we use a forward pass through the model to compute the probability assigned to each assessment label. This also avoids asking the model to generate a numerical score and treating that number as a direct measurement of its preference. Let $y = (y_1, \ldots, y_L)$ denote the token sequence corresponding to a candidate assessment, and let $x(h, o, H_t)$ denote the prompt presented to model a. We compute the

Table 1: Information-seeking scores ($P(Yes)$) across models to "*A multi-party system could lead to more situations where a majority is not reached. This means that it could be harder to push policies through.*" Objectives are shown at the top, middle, and bottom of M14's ranking; SS denotes semantic similarity. All topics begin with "we should."

Evaluation Topic ↓ Model →	SS	L3	L8	L70	Q7	Q14	Q72	M8	M14	M24	P4	G12
...adopt a multi-party system (original)	.38	.97	.94	1.00	1.00	1.00	1.00	.86	.84	.99	.93	1.00
Positions 2-4												
...oppose collectivism	.02	.97	.82	1.00	.00	.00	.00	.72	.75	.84	.69	.99
...introduce compulsory voting	.19	.97	.86	1.00	.11	.00	.16	.75	.75	.68	.73	.84
...adopt libertarianism	.01	.96	.79	1.00	.00	.00	.00	.74	.74	.94	.82	1.00
Positions 35-37												
...ban missionary work	.01	.81	.12	.00	.00	.00	.00	.03	.09	.03	.41	.00
...close Guantanamo Bay detention camp	.01	.42	.21	.00	.00	.00	.00	.02	.08	.01	.54	.00
...abolish capital punishment	.01	.43	.21	.00	.00	.00	.00	.05	.08	.01	.50	.00
Positions 69–71												
...subsidize embryonic stem cell research	.01	.46	.40	.00	.00	.00	.00	.03	.01	.01	.45	.00
...ban cosmetic surgery for minors	.01	.08	.25	.00	.00	.00	.00	.02	.01	.04	.48	.00
...ban the use of child actors	.01	.32	.20	.00	.00	.00	.00	.01	.01	.04	.33	.00

teacher-forced log-probability of the complete assessment sequence as

$$\ell_a(y \mid h, o, H_t) = \sum_{r=1}^{L} \log P_a(y_r \mid x(h, o, H_t), y_{<r}) .$$

For assessments consisting of multiple tokens, this sums the token-level log-probabilities over the complete sequence. Exponentiating the resulting sum gives the probability assigned by the model to the candidate assessment. Let y_v denote the affirmative assessment label. We define the information-seeking score used throughout our experiments as $s_a(h \mid o, H_t) = P_a(y_v \mid h, o, H_t)$.

3.2 Characterizing Model Preferences

We evaluate 11 publicly available open-weight instruction-tuned LLMs, selected to provide variation both across model families and, where possible, across parameter scales within a family. We include Llama-3.2-3B-Instruct (Meta Llama, 2024b), Llama-3.1-8B-Instruct (Meta Llama, 2024a), Llama-3.3-70B-Instruct (Meta Platforms, Inc., 2024), Qwen2.5-7B-Instruct, Qwen2.5-14B-Instruct, Qwen2.5-72B-Instruct (Yang et al., 2024), Ministral-3-8B-Instruct-2512, Ministral-3-14B-Instruct-2512, and Mistral-Small-3.2-24B-Instruct-2506 (Liu et al., 2026). To broaden the comparison beyond these three families, we additionally include Phi-4-mini-instruct (Microsoft et al., 2025) and gemma-4-12B-it (Gemma Team et al., 2026). In Tables 2 and 1, L, Q, M, P, and G denote the Llama, Qwen, Mistral/Ministral, Phi, and Gemma families, respectively, and the accompanying number denotes model size in billions of parameters. As a semantic-relatedness baseline, we use *cross-encoder/stsb-roberta-large* (Reimers & Gurevych, 2019), which receives only the argument and evaluation topic.

Table 2: Agreement between models in their information-seeking preferences across objectives, with STSB semantic similarity (SS) included as a reference.

	SS	L3	L8	L70	Q7	Q14	Q72	M8	M14	M24	P4	G12
SS	1	0.23	0.16	0.29	0.25	0.28	0.21	0.24	0.32	0.32	0.14	0.34
L3		1	.25	.53	.33	.37	.31	.41	.53	.43	.14	.39
L8			1	.27	.20	.32	.22	.26	.29	.26	.22	.24
L70				1	.37	.39	.47	.50	.60	.51	.13	.49
Q7					1	.28	.33	.31	.36	.34	.18	.38
Q14						1	.37	.34	.47	.45	.22	.42
Q72							1	.36	.39	.42	.20	.44
M8								1	.59	.47	.15	.42
M14									1	.60	.09	.51
M24										1	.09	.53
P4											1	.09
G12												1

We first characterize the information-seeking preferences induced by different LLMs before examining their consequences in sequential elicitation. We use the IBM-Rank-30k dataset (Gretz et al., 2019), which contains 30,497 short, human-written arguments collected in response to 71

researcher-selected controversial topics. For example, for the topic *We should abandon the use of school uniforms*, one argument states that "*school uniforms do not allow children to emphasize their differences, denying individuality*." Each argument was subsequently evaluated by 10 annotators on whether they would recommend using the argument as-is in a speech supporting or contesting the topic. The dataset therefore provides a shared information space of naturally occurring arguments associated with known topics and varying substantially in their specificity and quality. To keep the evaluation computationally tractable, we randomly sample two arguments from each of the 71 topics, yielding 142 arguments. For each argument, we vary the objective o across all 71 topics while keeping the candidate information h unchanged. We use no interaction history ($H_t = \emptyset$), so that the objective and model are the only conditioning variables that vary. This produces $142 \times 71 = 10{,}082$ argument–objective evaluations per model. Because each evaluation is deterministic under a fixed model and prompt, we do not repeat evaluations across random seeds.

We instantiate the positive, negative, and uncertain response options y_v, y_n, and y_u as *Yes*, *No*, and *Uncertain*, respectively. We use *Yes* and *No* as the primary response labels because they provide simple, unambiguous expressions of positive and negative responses and are typically represented by a single token under the BPE-based tokenizers used by the models we study. We validate the robustness of this choice against alternative labels in Appendix C. With $H_t = \emptyset$ and $y_v = \textit{Yes}$, the information-seeking score becomes $s_a(h \mid o) = P_a(\textit{Yes} \mid h, o)$. We use $P(\textit{Yes})$ directly rather than renormalizing over *Yes*, *No*, and *Uncertain*. Renormalization would instead measure the model's preference for *Yes* conditional on its probability mass being restricted to the specified response set.

Table 2 quantifies agreement in information-seeking preferences under identical candidate information and objectives. Agreement varies substantially across models, with most pairwise correlations ranging from .20 to .60 and P4 showing consistently lower agreement (.09–.22). Correlations with semantic similarity are low for all models (.14–.34), indicating that these preferences are not explained by semantic relatedness alone. Table 1 shows that the variation nevertheless retains clear structure in a representative example. All models assign high information-seeking scores to the argument under its original, semantically aligned topic (.84–1.00), while substantial disagreement emerges for other objectives. For example, for *We should oppose collectivism*, L3, L70, and G12 assign scores of .97, 1.00, and .99, whereas the Qwen models assign scores at or near zero. Taken together, these results show that the models induce substantially different, but not arbitrary, information-seeking preferences over the same information space.

4 Elicitation Trajectory Simulation

4.1 Candidate Hypotheses and Transitions

We simulate an interaction between an *elicitor* and a *respondent* using the IBM-Rank-30k information space, treating each argument as a hypothesis about a possible view of the respondent. At each step, the elicitor selects from k candidate hypotheses, the respondent confirms or does not confirm the selected view, and the selection and response enter the interaction history. Repeating this process produces an elicitation trajectory. The dynamic is illustrated in Figure 1. Let C_t denote the set of k candidate hypotheses presented to the elicitor at step t. Each hypothesis $h \in C_t$ is instantiated by an argument from the dataset, and $z(h)$ denotes the dataset topic associated with that argument. Before any hypothesis has been selected, no topic is active, and the initial candidate set is sampled across distinct topics. After the elicitor selects hypothesis h_t, its topic $z(h_t)$ becomes the *active topic*, regardless of whether the respondent subsequently confirms the hypothesis. Subsequent candidate sets provide more opportunities to continue the active topic than to switch topics, while ensuring that alternative lines of inquiry remain available. Specifically, we draw the smallest possible majority of candidates from the active topic, $k_{\text{active}} = \left\lfloor \frac{k}{2} \right\rfloor + 1$, and draw the remaining $(k - k_{\text{active}})$ candidates from distinct non-active topics. If the requested composition cannot be filled from these sources, the remaining positions are sampled from the pool of available hypotheses. Once a hypothesis has been selected, it is removed from future consideration. Unselected candidates remain available and may appear again at later steps. Because each selection determines the next active topic and therefore affects the composition of the next candidate set, the simulation is path dependent: $C_t \to h_t \to C_{t+1} \to h_{t+1} \to \cdots$.

4.2 ELICITOR

We instantiate the elicitor as an information elicitation agent whose candidate-selection scores are provided by an LLM. For each candidate hypothesis $h \in C_t$, the prompt specifies both the overall objective of learning about the respondent's views and the candidate's topic $z(h)$. The information-seeking score is additionally conditioned on the previous hypotheses investigated about that same topic. That is, when scoring candidate h, the elicitor includes all previously selected hypotheses whose topic is $z(h)$, with each hypothesis followed by a *[confirmed]* or *[not confirmed]* label indicating the respondent's response. History concerning other topics is not included in that candidate's prompt. Using the information-seeking score defined in Section 3.1, the elicitor scores each candidate and selects the candidate with the highest score, $h_t = \arg\max_{h \in C_t} s_a(h \mid o, H_t^{z(h)})$, where $H_t^{z(h)}$ denotes the interaction history associated with the candidate's topic. Ties between candidates with numerically equal maximum scores are broken randomly. Thus, the explicit selection rule is identical across elicitors; experimental conditions differ in the LLM used to assign information-seeking scores. Importantly, we do not generate the natural-language questions asked during the dialogue. Instead, when the elicitor selects a hypothesis, we treat that selection as choosing to ask about the view represented by the hypothesis and assume that an appropriate question is asked. The respondent then either confirms or does not confirm that view. In this way, the simulation preserves the sequential consequences of deciding what information to pursue while abstracting away variation due to natural-language question generation.

4.3 RESPONDENT

Once h_t has been selected, we independently sample whether the respondent confirms the corresponding view with fixed probability p. If confirmed, we treat this as learning that the respondent expresses the view represented by h_t, without making any judgment about whether its content is factually correct (factual correctness is a separate question, addressed by fact-checking and truth-seeking research). If unconfirmed, the simulation establishes only that the respondent did not confirm that view. The response is sampled independently of the hypothesis, topic, previous responses, and information-seeking score. This deliberately avoids making the resulting trajectories depend on a particular model of human response behavior. Both confirmed and unconfirmed outcomes are recorded in the interaction history and can therefore influence the information-seeking scores assigned to later hypotheses concerning the same topic. We also do not impose consistency constraints across the respondent's confirmed views. Although IBM-Rank-30k annotates whether arguments support or oppose their topics, enforcing consistency from these annotations would assume that views on a controversial topic reduce to a single position, even though a respondent may reasonably endorse arguments on both sides. For example, someone may agree that school uniforms reduce bullying while also agreeing that they limit individuality.

5 EXPERIMENTS

5.1 EXPERIMENTAL SETUP

We use the simulation introduced in Section 4 to focus on two properties of model-dependent information-seeking preferences: (1) how these preferences shape the breadth and depth of elicitation trajectories and (2) how acquired information can change subsequent preferences and selections. Because the dataset contains arguments expressing different views on controversial topics, we set the elicitor's objective as "*Learn about an interview participant's views on common controversial topics*." Based on the candidate-set size analysis in Section 5.2.2, we use $k = 7$ unless otherwise specified. We fix the hypothesis success probability at $p = 0.5$, giving each selected hypothesis an equal chance of being confirmed or unconfirmed. Because confirmation does not determine the opportunities available along a trajectory, this balances the two possible outcomes without favoring either. For each experimental condition, we simulate 100 trajec-

Table 3: Breadth–depth behavior at $k = 7$ (three-seed means). Sel./Cand.: selected/-candidate WA; Max Rate: maximum-WA selection rate; Switch: topic-switch rate; Topics: distinct topics; Run: mean run length.

	Sel. WA	Cand. WA	Max Rate	Switch	Topics	Run
Rand	.792	.791	.228	.423	32.9	2.34
L3	.856	.792	.339	.348	27.2	2.83
L8	.860	.793	.329	.458	36.5	2.16
L70	.799	.789	.231	.504	39.6	1.96
Q7	.847	.791	.315	.551	43.8	1.79
Q14	.840	.791	.292	.560	44.7	1.76
Q72	.841	.790	.286	.525	42.0	1.88
M8	.837	.790	.290	.413	34.0	2.54
M14	.818	.792	.260	.526	42.7	1.88
M24	.818	.791	.258	.484	39.0	2.04
P4	.861	.790	.351	.295	23.4	3.35
G12	.814	.791	.241	.342	26.2	2.89

Table 4: Effect of candidate-set size k on elicitation trajectories. Values are means across seeds 42, 43, and 44; the selected setting ($k = 7$) is underlined. Model abbreviations follow Table 1. Standard errors are at most .0045 for WA-based measures and .464 topics for Distinct Topics.

Metric →	Selected WA			Candidate WA			Candidate Max WA			Selects Max WA			Distinct Topics		
k ↓	L8	Q14	M8	L8	Q14	M8	L8	Q14	M8	L8	Q14	M8	L8	Q14	M8
3	.838	.825	.824	.792	.791	.791	.936	.935	.935	.472	.442	.438	29.7	35.4	26.7
5	.853	.832	.831	.793	.790	.789	.970	.970	.969	.369	.327	.325	34.4	41.2	31.2
7	.860	.840	.837	.793	.791	.790	.984	.984	.983	.329	.292	.290	36.5	44.7	34.0
9	.866	.843	.841	.792	.791	.789	.991	.991	.991	.323	.279	.275	37.8	46.0	34.3
11	.868	.846	.843	.793	.792	.790	.995	.995	.995	.312	.271	.265	38.9	47.1	35.3
13	.873	.847	.843	.792	.792	.789	.997	.997	.996	.316	.271	.257	40.0	48.2	35.2

tories of 100 elicitation steps across three random seeds (42, 43, and 44), yielding 30,000 elicitation steps in total. All experiments are run on three NVIDIA RTX A6000 GPUs. We include random selection as a trajectory baseline. We evaluate semantic similarity separately in Section 3.2, since there is no principled choice of text against which to compare candidates under a history-dependent assessment. We release the elicitation trajectory files generated for each model.

5.2 Breadth–Depth Behavior

5.2.1 Differences Across Models

We map the information-seeking preferences previously characterized onto the breadth–depth behavior they induce over sequential elicitation, without explicitly specifying or optimizing for breadth, depth, topic coverage, or switching. We hold the elicitation environment and selection rule fixed and vary only the LLM used to assign information-seeking scores. We characterize each trajectory by its topic-switch rate, number of distinct selected topics, and mean length of consecutive same-topic runs, and report each metric averaged across trajectories and seeds. Table 3 shows substantial variation in the resulting trajectories: topic-switch rates range from .295 (P4) to .560 (Q14), with distinct-topic counts ranging from 23.4 to 44.7 and mean run lengths from 1.76 to 3.35. Most models also select higher-WA arguments than random despite argument quality not appearing in the elicitation objective. Thus, differences in local information-seeking preferences are sufficient to induce markedly different breadth–depth behavior over the course of elicitation.

5.2.2 Effect of Available Opportunities

Table 5: Scores by history, semantic similarity, and prior confirmation (three-seed means). Conf.: confirmed; Not Conf.: unconfirmed. **Bold** indicates contrasts discussed in text.

		W/ History		W/out History	
Model	Group	Conf.	Not Conf.	Conf.	Not Conf.
Q7	≤ 0.7	.471	.485	.670	.667
	> 0.7	.961	.945	.952	.946
Q14	≤ 0.7	.430	.481	.622	.619
	> 0.7	**.257**	.649	**.823**	.833
M8	≤ 0.7	.663	.677	.598	.594
	> 0.7	.736	.739	.631	.635
M14	≤ 0.7	.724	.739	.837	.838
	> 0.7	**.634**	.718	**.865**	.864
M24	≤ 0.7	.831	.861	.920	.921
	> 0.7	**.708**	.844	**.931**	.934

The preceding analysis holds the candidate-set size fixed, but realized breadth and depth may also depend on the opportunities available to the elicitor. We therefore vary $k \in \{3, 5, 7, 9, 11, 13\}$, changing the number of hypotheses available at each step while holding the remaining simulation parameters fixed. To characterize the hypotheses available and selected as the candidate set changes using an external human-annotated measure, we use the Weighted-Average (WA) argument-quality scores provided by IBM-Rank-30k (Gretz et al., 2019). WA aggregates human judgments of whether an argument should be used as-is in a speech on its topic, weighted by annotator reliability. We use WA only as an external, human-annotated property of the underlying arguments; it is not intended as a measure of information-seeking preference or as the objective optimized by the elicitor. We report this analysis for representative models from three model families.

Table 4 shows that realized information-seeking behavior depends on the opportunities available to the elicitor. Increasing k from 3 to 13 increases the number of distinct topics explored for all three models, from 29.7 to 40.0 for L8, 35.4 to 48.2 for Q14, and 26.7 to 35.2 for M8. The ordering

across models remains consistent, with Q14 exploring the most topics and M8 the fewest, showing that candidate-set size shifts elicitation breadth without eliminating differences induced by model-specific information-seeking preferences. Larger candidate sets improve the best available argument with rapidly diminishing returns: maximum candidate WA rises from approximately .935 at $k = 3$ to .984 at $k = 7$ and .997 at $k = 13$, while selected WA increases more gradually and the maximum-WA selection rate generally decreases with k. We therefore use $k = 7$ when candidate-set size is held fixed, balancing available opportunities against the cost of scoring additional candidates. Together, these patterns characterize k as an opportunity constraint: increasing it expands both the range and quality of available information, while the model's preferences continue to determine which opportunities are pursued.

5.3 History-Dependent Information-Seeking Preferences

5.3.1 Redundancy and Confirmation

Information-seeking preferences may depend on what has already been learned. We examine this dependence in the case of potentially redundant information by providing the model with the topic-specific interaction history and explicitly instructing it that repeated information is less valuable, while leaving the model to determine what constitutes repetition and how strongly it should affect the assessment. This explicit instruction is important: as we show in Section 5.4.2, providing interaction history alone does not reliably cause models to discount repeated information. No separate deduplication procedure, similarity threshold, redundancy penalty, or candidate-filtering rule is used in selection; any effect of redundancy instead operates through the same information-seeking score used for all candidates. To isolate the effect of interaction history, at each elicitation step we score the same candidate hypotheses both with and without the topic-specific history.

For this analysis, we measure the semantic similarity of candidate hypotheses to previously selected hypotheses using *cross-encoder/stsb-roberta-large* (Reimers & Gurevych, 2019), and group candidates by similarity (> 0.7 vs. ≤ 0.7); examples of highly similar candidate–history pairs are provided in Appendix D. We further distinguish whether the similar information was previously confirmed or only unconfirmed, which allows us to test whether a candidate's score depends on its semantic relationship to earlier hypotheses and their confirmation status. Table 5 shows that history can substantially change how models evaluate potentially redundant information. For highly similar candidates whose overlapping information was previously confirmed, including history reduces mean scores from .823 to .257 for Q14, .865 to .634 for M14, and .931 to .708 for M24. Model size also plays a role, with this distinction absent for Q7 and M8 but emerging for Q14 and the larger M14 and M24 models. Moreover, the reduction is stronger when the overlapping information was previously confirmed than when it was unconfirmed, indicating that these models are more willing to move on once similar information has already been confirmed.

5.3.2 Effect on Subsequent Selection

We next examine whether these changes in information-seeking scores propagate to the elicitor's subsequent selections. Because the elicitor selects the hypothesis with the highest information-seeking score from the candidate set, a candidate's score need not decrease substantially to change the selection; it need only fall below that of another available candidate. This is a useful property of the elicitation process, as it allows redundant information to fall out of selection without requiring a separate mechanism for deduplication, removal, or filtering. At each trajectory step, we classify each candidate according to whether it is similar to previously confirmed information or only to previously unconfirmed information. We then compute each group's selection rate as the proportion of its candidates that are selected.

Table 6: Selection rates for candidates similar to confirmed (StC) or only to unconfirmed (SOtU) information, with and without history (three-seed means).

	W/ History		W/out History	
Model	StC	SOtU	StC	SOtU
L3	.315	.276	.330	.301
L8	.201	.216	.251	.238
L70	**.033**	.064	**.216**	.211
Q7	.287	.233	.280	.265
Q14	**.042**	.120	**.211**	.224
M8	.230	.172	.184	.184
M14	**.053**	.090	**.180**	.193
M24	**.049**	.092	**.167**	.163
P4	.467	.354	.344	.310
G12	.066	.124	.113	.118

Table 6 shows that the effect of history on subsequent selection also varies with model size. Within Qwen, Q7 shows little change in the selection of candidates similar to confirmed information (.280 without history vs. .287 with history), whereas Q14 shows a large reduction (.211 to .042). A

similar pattern appears within Mistral: M8 shows no reduction (.184 to .230), while M14 and M24 fall from .180 to .053 and from .167 to .049, respectively. The Llama models show a related but less monotonic pattern, with a large reduction for L70 (.216 to .033) but smaller reductions for L3 and L8. For Q14, M14, and M24, the reduction is also larger for candidates similar to confirmed information than for those similar only to unconfirmed information. These results show that the history-dependent differences observed in information-seeking scores propagate to subsequent selection, with the clearest effects generally appearing in larger models. When overlap with previously confirmed information lowers a candidate's score, that candidate becomes less likely to be selected, allowing redundancy to affect selection without explicit deduplication or filtering.

5.4 Sensitivity and Ablations

5.4.1 Label Sensitivity

To assess sensitivity to the response-label interface, we rescored every candidate encountered along the original trajectories using an alternative response set, replacing *Yes*/*No*/*Uncertain* with *Valuable*/*Not valuable*/*Uncertain*, while holding the candidate, objective, and interaction history fixed. We compute the Spearman correlation between the rankings induced by the two affirmative probabilities within each candidate set and average these correlations across decision points. Figure 2 shows substantial agreement for most models: nine of ten have mean correlations above .70, including five above .88, while G12 is substantially lower at .459. The positive correlations across all models indicate that the induced information-seeking preferences are not specific to the *Yes* formulation, although the degree of robustness varies across models. Additional analysis of the impact of these differences in trajectories is available in Appendix C.

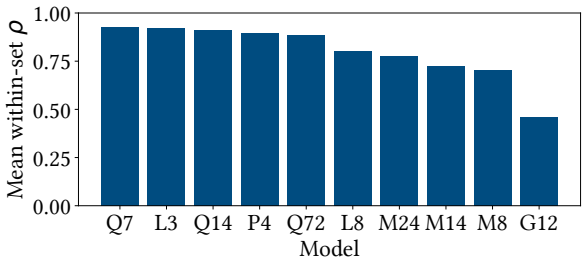

Figure 2: Response-label formulation robustness. Mean within-set rank correlation between *Yes* and *Valuable*.

5.4.2 Prompt Sensitivity

We test whether interaction history by itself is sufficient to produce the redundancy-sensitive behavior above by replacing the original prompt with a neutral variant that does not explicitly ask the model to consider repetition or additional information. As shown in Table 7, the reductions in information-seeking scores for hypotheses similar to previously confirmed information largely disappear under the neutral prompt, most notably for Q14, M14, and M24. History conditioning alone therefore does not reliably produce redundancy-sensitive assessments. Instead, the results show that redundancy sensitivity depends on how the information-seeking judgment is specified: when redundancy is made relevant to the assessment, it can affect candidate scores and subsequent selections through the same scoring mechanism used throughout the simulation.

Table 7: Scores for highly similar candidates (> 0.7) under the original and neutral prompts (three-seed means).

		W/ History		W/out History	
Model	Prompt	Conf.	Not Conf.	Conf.	Not Conf.
Q7	Original	.961	.945	.952	.946
	Neutral	.979	.968	.957	.958
Q14	Original	**.257**	.649	**.823**	.833
	Neutral	**.887**	.929	**.899**	.925
M8	Original	.736	.739	.631	.635
	Neutral	.947	.944	.940	.945
M14	Original	**.634**	.718	**.865**	.864
	Neutral	**.915**	.969	**.877**	.969
M24	Original	**.708**	.844	**.931**	.934
	Neutral	**.962**	.987	**.950**	.986

6 Conclusion

We studied how model-dependent information-seeking preferences shape sequential elicitation. Our simulation intentionally abstracts away question generation, natural-language responses, and an unbounded information space, allowing us to isolate the effects of LLM information-seeking preferences from other components of dialogue at scale. Extending this analysis to less controlled settings with richer interaction dynamics is an important direction for future work. We also treat the elicitor as the observer assessing which information is worth pursuing, although information may ultimately be valued by other observers. For example, an interviewer's judgments may depend on what they expect an audience to find valuable. Studying how information-seeking preferences interact across elicitors, audiences, and other observers therefore provides another direction for future work.

AI USE STATEMENT

In this work, we used generative AI tools for the tasks with required disclosure: *design or provide feedback on research methodology or experiments, interpret results, and propose or refine hypotheses*. We have not used generative AI tools for the other tasks with required disclosure: *help develop theoretical models or conceptual frameworks, formulate mathematical claims, implement methods, clean and reformat dataset, support qualitative and thematic data analysis*, and the rest of the required disclosure tasks: *generate synthetic data sets, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, assist with translation* are not applicable to this work. Additionally, we used generative AI tools for the tasks with recommended disclosure: *create or edit software code, searching for information, identify relevant literature*. We have reviewed all AI-assisted work. *All code for which generative AI tools were used was manually reviewed line by line by the authors, and relevant components were independently executed and tested to verify their correctness and consistency with the intended experimental design*. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provide the implementation details and experimental configurations needed to reproduce our results in the paper and supplementary code. Section 5 specifies experimental settings not otherwise captured in the code. All stochastic components of our experiments use fixed random seeds, which are reported in the paper. The prompts used to operationalize information-seeking preferences, dataset details, and additional experimental results and ablations are provided in the appendix and supplementary materials. We also provide the simulation code and elicitation trajectory files in the repository.

REFERENCES

Ali Abbas and Gordon Hazen. On the Value of Information Across Decision Problems. *Decision Analysis*, 22(1):1–13, March 2025. ISSN 1545-8490, 1545-8504. doi: 10.1287/deca.2024.0187. URL https://pubsonline.informs.org/doi/10.1287/deca.2024.0187.

Dilip Arumugam and Benjamin Van Roy. The Value of Information When Deciding What to Learn. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 9816–9827. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/517da335fd0ec2f4a25ea139d5494163-Paper.pdf.

Keping Bi, Qingyao Ai, and W. Bruce Croft. Asking Clarifying Questions Based on Negative Feedback in Conversational Search. In *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*, pp. 157–166, Virtual Event Canada, July 2021. ACM. ISBN 9781450386111. doi: 10.1145/3471158.3472232. URL https://dl.acm.org/doi/10.1145/3471158.3472232.

Mustafa Bilgic and Lise Getoor. VOILA: Efficient Feature-value Acquisition for Classification. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*, pp. 1225–1230, 2007. URL https://aaai.org/papers/01225-aaai07-194-voila-efficient-feature-value-acquisition-for-classification/.

Harshita Chopra and Chirag Shah. Feedback-Aware MCTS for Goal-Oriented Information Seeking. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, October 2025. URL https://openreview.net/forum?id=ustF8MMZDJ.

Roi Cohen, Konstantin Dobler, Eden Biran, and Gerard De Melo. I Don't Know: Explicit Modeling of Uncertainty with an [IDK] Token. In *Advances in Neural Information Processing Systems 37*, pp. 10935–10958, Vancouver, BC, Canada, 2024. Neural Information Processing Systems Foundation, Inc. (NeurIPS). ISBN 9798331314385. doi: 10.52202/079017-0349. URL http://www.proceedings.com/079017-0349.html.

Pat Croskerry. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. *Advances in Health Sciences Education*, 14(S1):27–35, September 2009. ISSN 1382-4996, 1573-1677. doi: 10.1007/s10459-009-9182-2. URL http://link.springer.com/10.1007/s10459-009-9182-2.

Jeffrey Dalton, Sophie Fischer, Paul Owoicho, Filip Radlinski, Federico Rossetto, Johanne R. Trippas, and Hamed Zamani. Conversational Information Seeking: Theory and Application. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 3455–3458, Madrid Spain, July 2022. ACM. ISBN 9781450387323. doi: 10.1145/3477495.3532678. URL https://dl.acm.org/doi/10.1145/3477495.3532678.

Victor De Lima and Grace Hui Yang. Yield: A large-scale dataset and evaluation framework for information elicitation agents. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 14882–14895. Association for Computational Linguistics, 2026. doi: 10.18653/v1/2026.acl-long.678. URL http://dx.doi.org/10.18653/v1/2026.acl-long.678.

Philippe Delquié. The Value of Information and Intensity of Preference. *Decision Analysis*, 5(3): 129–139, September 2008. ISSN 1545-8490, 1545-8504. doi: 10.1287/deca.1080.0116. URL https://pubsonline.informs.org/doi/10.1287/deca.1080.0116.

Yang Deng, Lizi Liao, Zhonghua Zheng, Grace Hui Yang, and Tat-Seng Chua. Towards Human-centered Proactive Conversational Agents. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 807–818, Washington DC USA, July 2024. ACM. ISBN 9798400704314. doi: 10.1145/3626772.3657843. URL https://dl.acm.org/doi/10.1145/3626772.3657843.

Yang Deng, Lizi Liao, Wenqiang Lei, Grace Hui Yang, Wai Lam, and Tat-Seng Chua. Proactive Conversational AI: A Comprehensive Survey of Advancements and Opportunities. *ACM Transactions on Information Systems*, 43(3):1–45, May 2025. ISSN 1046-8188, 1558-2868. doi: 10.1145/3715097. URL https://dl.acm.org/doi/10.1145/3715097.

Yijiang River Dong, Tiancheng Hu, Zheng Hui, Caiqi Zhang, Ivan Vulić, Andreea Bobu, and Nigel Collier. Value of information: A framework for human–agent communication. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 42879–42896. Association for Computational Linguistics, 2026. doi: 10.18653/v1/2026.acl-long.1987. URL http://dx.doi.org/10.18653/v1/2026.acl-long.1987.

Scott Emmons, Caspar Oesterheld, Vincent Conitzer, and Stuart Russell. Observation Interference in Partially Observable Assistance Games. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 15320–15345. PMLR, July 2025. URL https://proceedings.mlr.press/v267/emmons25a.html.

Xidong Feng, Junxian He, Bryan Hooi, Zhiyuan Hu, Pang Koh, Chumin Liu, Anh Luu, See-Kiong Ng, and Yilun Zhao. Uncertainty of Thoughts: Uncertainty-Aware Planning Enhances Information Seeking in LLMs. In *Advances in Neural Information Processing Systems 37*, pp. 24181–24215, Vancouver, BC, Canada, 2024. Neural Information Processing Systems Foundation, Inc. (NeurIPS). ISBN 979-8-3313-1438-5. doi: 10.52202/079017-0762. URL http://www.proceedings.com/079017-0762.html.

Gemma Team, Sherif El Abd, Vaibhav Aggarwal, Robin Algayres, Alek Andreev, Olivier Bachem, Ian Ballantyne, Cormac Brick, Victor Cărbune, Michelle Casbon, Mayank Chaturvedi, Aditya Chawla, Victor Cotruta, Alice Coucke, Phil Culliton, Robert Dadashi, Lucas Dixon, Mohamed Elhawaty, Utku Evci, Clément Farabet, Johan Ferret, Filippo Galgani, Sertan Girgin, Jean-Bastien Grill, Maarten Grootendorst, Jiaxian Guo, Cassidy Hardin, Yanzhang He, Steven M. Hernandez, Omri Homburger, Léonard Hussenot, Juyeong Ji, Armand Joulin, Aishwarya Kamath, Parnian Kassraie, Olivier Lacombe, Preethi Lahoti, Gaël Liu, Gus Martins, Luciano Martins, Tatiana

Matejovicova, Ramona Merhej, Nikola Momchev, Sneha Mondal, Ryan Mullins, Sindhu Raghuram Panyam, Shreya Pathak, Sarah Perrin, André Susano Pinto, Etienne Pot, Angéline Pouget, Alexandre Ramé, Sabela Ramos, Douglas Reid, David Rim, Morgane Rivière, Karsten Roth, Louis Rouillard, Omar Sanseviero, Pier Giuseppe Sessa, Shane Settle, Danila Sinopalnikov, Sara Smoot, Piotr Stanczyk, Andreas Steiner, Lawrence Stewart, Ilya Tolstikhin, Michael Tschannen, Anton Tsitsulin, Nino Vieillard, Renjie Wu, Pingmei Xu, Haichuan Yang, Edouard Yvinec, Biao Zhang, Li Zhang, Joe Zou, Nicolas Aagnes, Abdelrahman Abdelhamed, Jakub Adamek, Shivani Agrawal, Shubham Agrawal, Ibrahim Alabdulmohsin, Jean Baptiste Alayrac, Uri Alon, Chandramouli Amarnath, Ankesh Anand, Chrysovalantis Anastasiou, Setareh Ariafar, François-Xavier Aubet, Kyriakos Axiotis, Federico Barbero, Joelle Barral, Alexei Bendebury, Urs Bergmann, Stanley Bileschi, Kat Black, Mathieu Blondel, Sebastian Borgeaud, Arthur Bražinskas, Ryan Burnell, Robert Busa-Fekete, Mu Cai, Daniele Calandriello, Glenn Cameron, Charlotte Caucheteux, Rahma Chaabouni, Garima Chadha, Jetha Chan, Blake Jianhang Chen, Jesse Chen, Lin Chen, Xu Chen, Derek Cheng, Tzu-hsiang Chien, Nikolai Chinaev, Yi Chou, Zhaohui Chu, Benjamin Coleman, Pooja Consul, Sam Conway-Rahman, Scott Crowell, Dylan Cutler, Vivek Dani, Samira Daruki, Anil Das, Daniel Deutsch, Nishanth Dikkala, Li Ding, Qiuhan Ding, Shenil Dodhia, Konstantin Donhauser, Tulsee Doshi, Anca Dragan, Alex Druinsky, Sahil Dua, Zoltan Egyed, Danielle Eisenbud, Daniel Eppens, Cindy Fan, Bahare Fatemi, Yassir Fathullah, Vlad Feinberg, Milen Ferev, Sebastian Flennerhag, Takumi Fujimoto, João Gabriel Oliveira, Isaac Galatzer-Levy, João Gante, Simon Geisler, Soham Ghosal, Antonious M. Girgis, Tamara von Glehn, Alec Go, Alhaad Gokhale, Alex Grills, Yiming Gu, Mayank Gupta, Pramod Gupta, Guru Guruganesh, Raia Hadsell, Hamza Harkous, Jitendra Harlalka, Demis Hassabis, Anja Hauth, Joe Heyward, Arian Hosseini, Chih-Yang Hsia, I-Hung Hsu, Xiaopeng Huang, Yangsibo Huang, Kevin Hui, Adrian Hutter, Te I, Fotis Iliopoulos, Advait Jain, Ganesh Jawahar, Ziwei Ji, Qilin Jin, Melvin Johnson, Kandarp Joshi, Arun Kandoor, Wang-Cheng Kang, Koray Kavukcuoglu, Mehran Kazemi, Kathleen Kenealy, Amr Khalifa, Phoebe Kirk, Ivan Korotkov, Suraj Kothawade, Vitaly Kovalev, Neel Kovelamudi, Adam Kraft, Ravin Kumar, Vivek Kumar, Harish Kuppam, Justin Lannin, Chen-Yu Lee, Seungji Lee, Dmitry Lepikhin, Alon Levkovitch, Dongdong Li, Qiujia Li, Valentin Liévin, Ethan Lin, Ziqian Lin, Casper Liu, Tianlin Liu, Tianqi Liu, Xin Liu, Ivan Lobov, Mayank Lunayach, Min Ma, Gagan Madan, Andrii Maksai, Eric Malmi, Michal Matuszak, Daniel McDuff, Gaurav Menghani, Maciej Mikuła, Daniil Mirylenka, Karolis Misiunas, Vedant Misra, Andreea Mitran, Kareem Mohamed, Maksim Mukha, Eric Noland, James O'Donnell, Brendan O'Donoghue, Kate Olszewska, Bernett Orlando, Wanqiong Pan, Rina Panigrahy, Unnati Parekh, Nicolas Perez-Nieves, Chunjong Park, Eric Paskie, Liqian Peng, Bryce Petrini, Slav Petrov, Jonas Pfeiffer, Bilal Piot, Martyna Plomecka, Siim Poder, Octavio Ponce, Arijit Pramanik, David Racz, Anish Rajan, Michelle Ramanovich, Anand Rao, Marvin Ritter, Vitor Rodrigues, Evan Rosen, Mikołaj Rybiński, Noveen Sachdeva, Michaël E. Sander, Rohit Sathyanarayana, Sagar Savla, Samuel Schmidgall, Tal Schuster, George Scrivener, Benoit Seguin, Andrew Sellergren, Aliaksei Severyn, Izhak Shafran, Dhruv Shah, Bobak Shahriari, Yuan Shangguan, Ashish Shenoy, Pradeep Shenoy, Rakesh Shivanna, Pauline Sho, Lucas Spangher, Wojciech Stokowiec, Tim Strother, Yao Su, Yinghao Sun, Mukund Sundararajan, Andrea Tacchetti, Mor Hazan Taege, Pouya Tafti, Jean Tarbouriech, Chetan Tekur, Shantanu Thakoor, Rahul Thapa, Madeleine Traverse, Lenart Treven, Tao Tu, Chien Te Tung, Çağlar Ünlü, Petar Veličković, Malini Pooni Venkat, Sagar Gubbi Venkatesh, Vidya Venkiteswaran, Francesco Visin, Alex Vitvitskyi, Kiran Vodrahalli, Weiyi Wang, Xin Wang, Tris Warkentin, Jan Wassenberg, John Wieting, Cindy Wu, Lechao Xiao, Hao Xu, Yuhui Xu, Fuzhao Xue, Arun Yadav, Jun Yan, Antoine Yang, Lin Yang, Ming-Hsuan Yang, Ziyu Ying, Jae Hyeon Yoo, Morteza Zadimoghaddam, Sajjad Zafar, Fred Zhang, Jiageng Zhang, Jianyi Zhang, Xiaofan Zhang, Chao Zhao, David Zhou, and Chen Zou. Gemma 4 Technical Report, July 2026. URL https://arxiv.org/abs/2607.02770.

Joshua Grass and Shlomo Zilberstein. A value-driven system for autonomous information gathering. *Journal of Intelligent Information Systems*, 14(1):5–27, March 2000. ISSN 1573-7675. doi: 10.1023/a:1008718418982. URL http://dx.doi.org/10.1023/A:1008718418982.

Shai Gretz, Roni Friedman, Edo Cohen-Karlik, Assaf Toledo, Dan Lahav, Ranit Aharonov, and Noam Slonim. A Large-scale Dataset for Argument Quality Ranking: Construction and Analysis, November 2019. URL http://arxiv.org/abs/1911.11408. arXiv:1911.11408 [cs.CL].

Ziyang Guo, Yifan Wu, Jason Hartline, and Jessica Hullman. The Value of Information in Human-AI Decision-making. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 41826–41857, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/45824c7580778ad1ba384982b9e6b39f-Paper-Conference.pdf.

Yacin Hamami. Inquiry in Conversation. *Logique et Analyse*, 2014.

Ronald Howard. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1):22–26, 1966. ISSN 0536-1567. doi: 10.1109/TSSC.1966.300074. URL http://ieeexplore.ieee.org/document/4082064/.

Jiaxiong Hu, Jingya Guo, Ningjing Tang, Xiaojuan Ma, Yuan Yao, Changyuan Yang, and Yingqing Xu. Designing the Conversational Agent: Asking Follow-up Questions for Information Elicitation. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW1):1–30, April 2024. ISSN 2573-0142. doi: 10.1145/3637320. URL https://dl.acm.org/doi/10.1145/3637320.

Eunsu Kim, Juyoung Suk, Seungone Kim, Niklas Muennighoff, Dongkwan Kim, and Alice Oh. LLM-as-an-Interviewer: Beyond Static Testing Through Dynamic LLM Evaluation, 2024. URL https://arxiv.org/abs/2412.10424.

Wei-Jen Ko, Te-yuan Chen, Yiyan Huang, Greg Durrett, and Junyi Jessy Li. Inquisitive Question Generation for High Level Text Comprehension. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 6544–6555, Online, 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.530. URL https://aclanthology.org/2020.emnlp-main.530.

Jaewoo Lee, Archiki Prasad, Justin Chen, Zaid Khan, Elias Stengel-Eskin, and Mohit Bansal. PRInTS: Reward Modeling for Long-Horizon Information Seeking. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 34120–34138, San Diego, California, United States, 2026. Association for Computational Linguistics. doi: 10.18653/v1/2026.acl-long.1574. URL https://aclanthology.org/2026.acl-long.1574.

Zixuan Li, Lizi Liao, and Tat-Seng Chua. Learning to Ask Critical Questions for Assisting Product Search. In *Proceedings of ACM SIGIR Workshop on eCommerce*, Madrid, Spain, July 2022. ACM. URL https://sigir-ecom.github.io/ecom22Papers/paper_6978.pdf.

Xubo Lin, Mingze Wang, Grace Hui Yang, and Daniel Chen. Reward-on-the-Line: A Novel Offline Reinforcement Learning Method for Building Legal Conversational Agents. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(2):1575–1584, October 2025. ISSN 3065-8365. doi: 10.1609/aies.v8i2.36657. URL https://ojs.aaai.org/index.php/AIES/article/view/36657.

Xubo Lin, Zezhi Deng, Shihao Wang, Grace Hui Yang, and Yang Deng. Dual Hierarchical Dialogue Policy Learning for Legal Inquisitive Conversational Agents. In *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 11030–11047, San Diego, California, United States, 2026. Association for Computational Linguistics. doi: 10.18653/v1/2026.findings-acl.536. URL https://aclanthology.org/2026.findings-acl.536.

Alexander H. Liu, Kartik Khandelwal, Sandeep Subramanian, Victor Jouault, Abhinav Rastogi, Adrien Sadé, Alan Jeffares, Albert Jiang, Alexandre Cahill, Alexandre Gavaudan, Alexandre Sablayrolles, Amélie Héliou, Amos You, Andy Ehrenberg, Andy Lo, Anton Eliseev, Antonia Calvi, Avinash Sooriyarachchi, Baptiste Bout, Baptiste Rozière, Baudouin De Monicault, Clémence Lanfranchi, Corentin Barreau, Cyprien Courtot, Daniele Grattarola, Darius Dabert, Diego de las Casas, Elliot Chane-Sane, Faruk Ahmed, Gabrielle Berrada, Gaëtan Ecrepont, Gauthier Guinet, Georgii Novikov, Guillaume Kunsch, Guillaume Lample, Guillaume Martin, Gunshi Gupta, Jan Ludziejewski, Jason Rute, Joachim Studnia, Jonas Amar, Joséphine Delas, Josselin Somerville Roberts, Karmesh Yadav, Khyathi Chandu, Kush Jain, Laurence Aitchison, Laurent Fainsin, Léonard Blier, Lingxiao Zhao, Louis Martin, Lucile Saulnier, Luyu Gao,

Maarten Buyl, Margaret Jennings, Marie Pellat, Mark Prins, Mathieu Poirée, Mathilde Guillaumin, Matthieu Dinot, Matthieu Futeral, Maxime Darrin, Maximilian Augustin, Mia Chiquier, Michel Schimpf, Nathan Grinsztajn, Neha Gupta, Nikhil Raghuraman, Olivier Bousquet, Olivier Duchenne, Patricia Wang, Patrick von Platen, Paul Jacob, Paul Wambergue, Paula Kurylowicz, Pavankumar Reddy Muddireddy, Philomène Chagniot, Pierre Stock, Pravesh Agrawal, Quentin Torroba, Romain Sauvestre, Roman Soletskyi, Rupert Menneer, Sagar Vaze, Samuel Barry, Sanchit Gandhi, Siddhant Waghjale, Siddharth Gandhi, Soham Ghosh, Srijan Mishra, Sumukh Aithal, Szymon Antoniak, Teven Le Scao, Théo Cachet, Theo Simon Sorg, Thibaut Lavril, Thiziri Nait Saada, Thomas Chabal, Thomas Foubert, Thomas Robert, Thomas Wang, Tim Lawson, Tom Bewley, Tom Edwards, Umar Jamil, Umberto Tomasini, Valeriia Nemychnikova, Van Phung, Vincent Maladière, Virgile Richard, Wassim Bouaziz, Wen-Ding Li, William Marshall, Xinghui Li, Xinyu Yang, Yassine El Ouahidi, Yihan Wang, Yunhao Tang, and Zaccharie Ramzi. Ministral 3, 2026. URL https://arxiv.org/abs/2601.08584.

Davide Mazzaccara, Alberto Testoni, and Raffaella Bernardi. Learning to Ask Informative Questions: Enhancing LLMs with Preference Optimization and Expected Information Gain. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pp. 5064–5074, Miami, Florida, USA, 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.findings-emnlp.291. URL https://aclanthology.org/2024.findings-emnlp.291.

Meta Llama. Llama-3.1-8B-Instruct, July 2024a. URL https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct.

Meta Llama. Llama-3.2-3B-Instruct, October 2024b. URL https://huggingface.co/meta-llama/Meta-Llama-3-8B-Instruct.

Meta Platforms, Inc. *Llama-3.3-70B-Instruct*. Hugging Face, November 2024. URL https://huggingface.co/meta-llama/Llama-3.3-70B-Instruct.

Microsoft, :, Abdelrahman Abouelenin, Atabak Ashfaq, Adam Atkinson, Hany Awadalla, Nguyen Bach, Jianmin Bao, Alon Benhaim, Martin Cai, Vishrav Chaudhary, Congcong Chen, Dong Chen, Dongdong Chen, Junkun Chen, Weizhu Chen, Yen-Chun Chen, Yi-ling Chen, Qi Dai, Xiyang Dai, Ruchao Fan, Mei Gao, Min Gao, Amit Garg, Abhishek Goswami, Junheng Hao, Amr Hendy, Yuxuan Hu, Xin Jin, Mahmoud Khademi, Dongwoo Kim, Young Jin Kim, Gina Lee, Jinyu Li, Yunsheng Li, Chen Liang, Xihui Lin, Zeqi Lin, Mengchen Liu, Yang Liu, Gilsinia Lopez, Chong Luo, Piyush Madan, Vadim Mazalov, Arindam Mitra, Ali Mousavi, Anh Nguyen, Jing Pan, Daniel Perez-Becker, Jacob Platin, Thomas Portet, Kai Qiu, Bo Ren, Liliang Ren, Sambuddha Roy, Ning Shang, Yelong Shen, Saksham Singhal, Subhojit Som, Xia Song, Tetyana Sych, Praneetha Vaddamanu, Shuohang Wang, Yiming Wang, Zhenghao Wang, Haibin Wu, Haoran Xu, Weijian Xu, Yifan Yang, Ziyi Yang, Donghan Yu, Ishmam Zabir, Jianwen Zhang, Li Lyna Zhang, Yunan Zhang, and Xiren Zhou. Phi-4-Mini Technical Report: Compact yet Powerful Multimodal Language Models via Mixture-of-LoRAs, 2025. URL https://arxiv.org/abs/2503.01743.

Jonathan D. Nelson. Finding Useful Questions: On Bayesian Diagnosticity, Probability, Impact, and Information Gain. *Psychological Review*, 112(4):979–999, 2005. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.112.4.979. URL https://doi.apa.org/doi/10.1037/0033-295X.112.4.979.

Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 3980–3990, Hong Kong, China, 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1410. URL https://www.aclweb.org/anthology/D19-1410.

Xuhui Ren, Hongzhi Yin, Tong Chen, Hao Wang, Zi Huang, and Kai Zheng. Learning to Ask Appropriate Questions in Conversational Recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 808–817, Virtual Event Canada, July 2021. ACM. ISBN 9781450380379. doi: 10.1145/3404835.3462839. URL https://dl.acm.org/doi/10.1145/3404835.3462839.

C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(4): 623–656, October 1948. ISSN 00058580. doi: 10.1002/j.1538-7305.1948.tb00917.x. URL https://ieeexplore.ieee.org/document/6773067.

Zhiwen Tang and Grace Hui Yang. A Reinforcement Learning Approach for Dynamic Search. In *Proceedings of the Text REtrieval Conference 2017*, Gaithersburg, MD, USA, November 2017.

Alberto Testoni, Raffaella Bernardi, and Azzurra Ruggeri. The Efficiency of Question-Asking Strategies in a Real-World Visual Search Task. *Cognitive Science*, 47(12):e13396, December 2023. ISSN 0364-0213, 1551-6709. doi: 10.1111/cogs.13396. URL https://onlinelibrary.wiley.com/doi/10.1111/cogs.13396.

Tiago Veiga and Jennifer Renoux. From reactive to active sensing: A survey on information gathering in decision-theoretic planning. *ACM Computing Surveys*, 55(13s):1–22, 2023. ISSN 1557-7341. doi: 10.1145/3583068. URL http://dx.doi.org/10.1145/3583068.

David J Weiss and Ben Taskar. Learning Adaptive Value of Information for Structured Prediction. In C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013. URL https://proceedings.neurips.cc/paper_files/paper/2013/file/ebd9629fc3ae5e9f6611e2ee05a31cef-Paper.pdf.

Jason D. Williams and Steve Young. Partially observable Markov decision processes for spoken dialog systems. *Computer Speech & Language*, 21(2):393–422, April 2007. ISSN 08852308. doi: 10.1016/j.csl.2006.06.008. URL https://linkinghub.elsevier.com/retrieve/pii/S0885230806000283.

Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. Proactive Human-Machine Conversation with Explicit Conversation Goal. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3794–3804, Florence, Italy, 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1369. URL https://aclanthology.org/P19-1369.

Yating Wu, Ritika Rajesh Mangla, Alex Dimakis, Greg Durrett, and Junyi Jessy Li. Which questions should I answer? Salience Prediction of Inquisitive Questions. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 19969–19987, Miami, Florida, USA, 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.emnlp-main.1114. URL https://aclanthology.org/2024.emnlp-main.1114.

Zeqiu Wu, Ryu Parish, Hao Cheng, Sewon Min, Prithviraj Ammanabrolu, Mari Ostendorf, and Hannaneh Hajishirzi. InSCIt: Information-Seeking Conversations with Mixed-Initiative Interactions. *Transactions of the Association for Computational Linguistics*, 11:453–468, 2023. doi: 10.1162/tacl_a_00559. URL https://aclanthology.org/2023.tacl-1.27/.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, Jianxin Yang, Jin Xu, Jingren Zhou, Jinze Bai, Jinzheng He, Junyang Lin, Kai Dang, Keming Lu, Keqin Chen, Kexin Yang, Mei Li, Mingfeng Xue, Na Ni, Pei Zhang, Peng Wang, Ru Peng, Rui Men, Ruize Gao, Runji Lin, Shijie Wang, Shuai Bai, Sinan Tan, Tianhang Zhu, Tianhao Li, Tianyu Liu, Wenbin Ge, Xiaodong Deng, Xiaohuan Zhou, Xingzhang Ren, Xinyu Zhang, Xipin Wei, Xuancheng Ren, Xuejing Liu, Yang Fan, Yang Yao, Yichang Zhang, Yu Wan, Yunfei Chu, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, Zhifang Guo, and Zhihao Fan. Qwen2 Technical Report, September 2024. URL http://arxiv.org/abs/2407.10671. arXiv:2407.10671.

Shahram Yazdani, Mohammad Hosseinzadeh, and Fakhrolsadat Hosseini. Models of clinical reasoning with a focus on general practice: A critical review. *Journal of Advances in Medical Education & Professionalism*, 5(4):177–184, October 2017. ISSN 2322-2220.

Steve Young, Milica Gasic, Blaise Thomson, and Jason D. Williams. POMDP-Based Statistical Spoken Dialog Systems: A Review. *Proceedings of the IEEE*, 101(5):1160–1179, May 2013. ISSN 0018-9219, 1558-2256. doi: 10.1109/JPROC.2012.2225812. URL http://ieeexplore.ieee.org/document/6407655/.

Lili Yu, Howard Chen, Sida I. Wang, Tao Lei, and Yoav Artzi. Interactive Classification by Asking Informative Questions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2664–2680, Online, 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.237. URL https://www.aclweb.org/anthology/2020.acl-main.237.

Hengran Zhang, Ruqing Zhang, Jiafeng Guo, Maarten De Rijke, Yixing Fan, and Xueqi Cheng. Are Large Language Models Good at Utility Judgments? In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1941–1951, Washington DC USA, July 2024. ACM. ISBN 9798400704314. doi: 10.1145/3626772.3657784. URL https://dl.acm.org/doi/10.1145/3626772.3657784.

Hengran Zhang, Minghao Tang, Keping Bi, Jiafeng Guo, Shihao Liu, Daiting Shi, Dawei Yin, and Xueqi Cheng. Utility-Focused LLM Annotation for Retrieval and Retrieval-Augmented Generation. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 1683–1702, Suzhou, China, 2025. Association for Computational Linguistics. doi: 10.18653/v1/2025.emnlp-main.88. URL https://aclanthology.org/2025.emnlp-main.88.

A PROMPTS

The following prompt was used for the main experiments.

Main Prompt

```
You are interviewing a person.  Your objective is to {dataset_objective} and the
current topic is '{topic}'.

The following is information you have already gathered about this topic:
{history}

Regardless of whether you agree with what they say, does the response provide
valuable information about the person's views?  Repeated information makes it
less valuable.

Respond with exactly one of the following:

{valid_labels}
```

For the prompt-sensitivity ablation in Section 5.4.2, we remove the instruction that repeated information is less valuable.

Ablation Prompt

```
You are interviewing a person.  Your objective is to {dataset_objective} and the
current topic is '{topic}'.

The following is information you have already gathered about this topic:
{history}

Regardless of whether you agree with what they say, does the response provide
valuable information about the person's views?

Respond with exactly one of the following:

{valid_labels}
```

B WITH AND WITHOUT UNCERTAIN LABEL

Table B.1 compares assessments of the same argument–objective pairs with and without *Uncertain* as a response option. Although little probability is assigned to *Uncertain* itself, its inclusion substantially affects $P(Yes)$. Without *Uncertain*, intermediate scores largely shift toward the extremes, while high scores remain comparatively stable. Including *Uncertain* therefore preserves greater differentiation among intermediate information-seeking scores.

C LABEL TRAJECTORY SENSITIVITY

We repeat the sequential elicitation experiment using the alternative labels `Valuable`, `Not valuable`, and `Uncertain`, while holding the assessment question, simulation environment, and selection rule fixed. Table C.2 compares breadth–depth behavior across independently generated trajectories under the two label formulations. Sensitivity varies across models: Q7, L3, and M24 exhibit similar behavior under both formulations, whereas M8 changes substantially. These results show that differences between the two label formulations can propagate to trajectory-level behavior, with the magnitude of the effect varying substantially across models.

D SEMANTIC SIMILARITY EXAMPLES

Table D.3 shows that the information space contains many highly similar candidate–history pairs, with similarity scores reaching .97. Despite this redundancy, overlapping information can be deprioritized through the information-seeking process itself, without explicit deduplication, removal, or candidate filtering.

Table B.1: Effect of including *Uncertain* on information-seeking scores. YesΔ denotes the change in $P(\textit{Yes})$ when *Uncertain* is included.

	W/ Uncert.			W/out Uncert.		
	Yes	No	Uncert.	Yes	No	YesΔ
We should adopt a multi-party system	0.94	0.05	0.02	0.93	0.08	0.01
We should introduce compulsory voting	0.86	0.13	0.02	0.79	0.21	0.07
We should oppose collectivism	0.82	0.14	0.04	0.75	0.25	0.07
We should adopt libertarianism	0.79	0.15	0.06	0.75	0.25	0.05
We should adopt an austerity regime	0.71	0.25	0.04	0.50	0.50	0.21
We should limit judicial activism	0.63	0.34	0.03	0.36	0.64	0.27
We should end the use of economic sanctions	0.62	0.36	0.02	0.36	0.64	0.26
Homeopathy brings more harm than good	0.53	0.46	0.01	0.17	0.83	0.36
Blockade of the Gaza Strip should be ended	0.48	0.51	0.01	0.21	0.79	0.27
Entrapment should be legalized	0.47	0.51	0.02	0.18	0.82	0.29
We should abolish the right to keep and bear arms	0.46	0.53	0.01	0.18	0.82	0.28
Intelligence tests bring more harm than good	0.45	0.54	0.01	0.10	0.90	0.35
We should fight urbanization	0.42	0.56	0.02	0.14	0.86	0.27
Holocaust denial should be a criminal offence	0.41	0.58	0.02	0.14	0.87	0.27
We should fight for the abolition of nuclear weapons	0.41	0.58	0.02	0.14	0.86	0.26

Table C.2: Breadth–depth behavior under the original *Yes*/*No*/*Uncertain* labels (Yes) and alternative *Valuable*/*Not valuable*/*Uncertain* labels (Valuable) (three-seed means).

	Topic Switch Rate			Distinct Topics			Topic Run Length		
Model	Yes	Valuable	$\lvert\Delta\rvert$	Yes	Valuable	$\lvert\Delta\rvert$	Yes	Valuable	$\lvert\Delta\rvert$
L3	.348	.331	0.017	27.177	25.977	1.200	2.834	2.986	0.152
L8	.458	.561	0.103	36.523	44.907	8.384	2.158	1.760	0.398
Q7	.551	.555	0.004	43.810	43.767	0.043	1.794	1.781	0.013
Q14	.560	.521	0.039	44.720	41.680	3.040	1.764	1.897	0.133
M8	.413	.666	0.253	33.997	54.230	20.233	2.542	1.485	1.057
M14	.526	.449	0.077	42.657	36.713	5.944	1.878	2.198	0.320
M24	.484	.466	0.018	38.977	37.997	0.980	2.040	2.118	0.078

Table D.3: Ten highest-similarity pairs from Qwen2.5-14B-Instruct trajectories.

Candidate	History Argument	Sim. Score
The vow of celibacy should be abandoned because it's outdated in today's society.	The vow of celibacy is outdated in today's society and so it should be abandoned.	0.97
We should not ban telemarketing because it would put people out of jobs.	We should not ban telemarketing because it would result in a loss of jobs.	0.97
Burning the flag is freedom of speech and should be allowed.	Flag burning is a form of free speech and should not be banned.	0.96
We should abolish the right to keep and bear arms because there are those who abuse that power and go on rampages killing innocent people and shooting up schools.	We should abolish the right to keep and bear arms because too many abuse that right by killing other and going on rampages and school shootings.	0.95
If we legalise the organ trade then it will mean that more lives can be saved.	Legalizing organ trade has the potential to save many lives.	0.95
People should be allowed to work for as long as they want and as long as they are capable.	Individuals should be able to work as long as they are willing and able.	0.95
Prohibiting school prayer is the right thing to do to avoid conflict in the classroom.	School prayer promotes conflict in the classroom and therefore should be prohibited.	0.94
Safe spaces tend to act as a brake on free speech.	Safe spaces are often used to stifle free expression of opinions.	0.94
Students are children who need room to learn and grow from mistakes.	Students are still children who need room to learn and grow from mistakes.	0.94
Social media can be used as a bully technique.	Social media is an outlet for bullying.	0.94