

The Science and Philosophy of Better Questions: A Framework for Question Engineering

1. Introduction: The Architecture of Inquiry

The history of human intellectual progress is, at its core, a history of better questions. From the Socratic dialogues that dismantled the unexamined assumptions of Athenian society to the hypothesis-driven inquiries that ignited the Scientific Revolution, the "question" has served as the primary instrument for extracting signal from noise. It is the fundamental algorithm of discovery, the linguistic probe we insert into the black box of reality to determine its internal structure. Yet, despite its centrality to science, law, intelligence analysis, education, and philosophy, the act of questioning is rarely treated as a rigorous engineering discipline. In most professional and academic contexts, questioning remains an intuitive art—a soft skill relegated to the domain of rhetoric rather than a hard science of epistemology.

This report establishes the framework for **Question Engineering**, a formalized discipline concerned with the design, optimization, sequence, and evaluation of inquiries. We posit that a question is not merely a sentence with a rising inflection; it is an epistemic machine, a distinct algorithmic procedure for information retrieval and cognitive processing. Just as software engineering requires distinct code structures to execute specific tasks, Question Engineering requires distinct interrogative structures to elicit specific types of information.¹ The output of any intellectual system—whether a human mind, a corporate strategy, or a Large Language Model (LLM)—is strictly bound by the quality of its input. In the age of algorithmic decision-making and information overload, the "garbage in, garbage out" principle has never been more acute. "Question Engineering" thus bridges the gap between classical epistemology and modern prompt engineering, providing a unified taxonomy and operational playbook for human and machine inquiry.²

The necessity for such a framework is driven by the increasing complexity of the information landscape. In intelligence analysis, the failure to ask the right question can lead to catastrophic blind spots; in medicine, it leads to misdiagnosis; in law, to miscarriages of justice. We observe that failures in these fields often stem not from a lack of data, but from a failure of inquiry—specifically, the deployment of "buggy" questions that are biased, malformed, or categorically erroneous. By treating questions as engineered objects, we can subject them to the same rigorous testing, debugging, and optimization processes applied to other critical systems.

This document serves as a comprehensive manual for the Question Engineer. It is divided into three primary functional components. First, **The Question Atlas** maps the landscape of inquiry across philosophy, science, law, and intelligence, providing a unified taxonomy for classifying questions by their functional intent and truth conditions. Second, **The Sequence**

Playbook offers a set of algorithms and flowcharts for structuring chains of questions to achieve specific epistemic goals, from finding the "crux" of a disagreement to exposing deception in adversarial contexts. Third, **The Question Debugger** provides a diagnostic toolkit for identifying and repairing "pathological" questions—those that are logically invalid, linguistically ambiguous, or cognitively biased. Through this tripartite structure, we aim to transform the "art" of asking into a precise and replicable science.

2. The Question Atlas: A Multi-Domain Taxonomy

To engineer better questions, one must first understand the available inventory. A "Question Atlas" categorizes inquiries not just by their grammatical form, but by their *functional intent*, the *cognitive load* they impose, and the *nature of the answer* they demand. We synthesize taxonomies from five distinct domains—Philosophy, Education, Science, Intelligence Analysis, and Law—to create a unified map of the interrogative terrain.

2.1. The Philosophical Domain: Truth Conditions and Reality

At the most fundamental level, questions define the nature of the reality they probe. The philosophical framework, particularly as adapted for critical thinking education, distinguishes questions based on the *truth conditions* of their answers. This distinction is paramount; a common failure mode in discourse is the "Category Error," where a question belonging to one domain is treated as if it belongs to another (e.g., treating a matter of preference as a matter of fact).⁴

2.1.1. Matters of Fact (Empirical Inquiry)

These are questions where the answer is determined by objective reality, independent of human opinion. They are the bedrock of scientific and historical inquiry.

- **Established Facts (Single Answer):** These inquiries seek data that is already codified in authoritative records or is easily observable. The engineering utility of these questions is *calibration*. They establish a shared baseline of reality between the questioner and the respondent.
 - *Example:* "What is the boiling point of lead?" or "Who signed the Treaty of Versailles?"
 - *Function:* Retrieval and verification. In automated systems, these are "lookup" tasks.⁴
- **Disputed Facts (Open-Ended):** These involve empirical questions where the data is incomplete, the historical record is fragmented, or the answer is currently unknown, yet a single objective truth exists in principle. These questions drive research programs and require evidence-based argumentation rather than simple retrieval.
 - *Example:* "What was the precise population of Jericho in 1000 BCE?" or "Does life exist on Europa?"
 - *Function:* Investigation and hypothesis generation. The "stop condition" for such

questions is the discovery of sufficient evidence to rule out competing alternatives.⁴

2.1.2. Matters of Judgment (Theoretical/Normative Inquiry)

These questions occupy the "reasoned middle" between objective fact and pure subjectivity. They address complex systems where multiple answers may be defensible, but some are objectively superior based on logic, evidence, and defined criteria.

- **Systemic Optimization:** These require weighing competing criteria to determine a "best" outcome. They are the engine of policy analysis, strategy, and engineering design.
 - *Example:* "What is the most stable form of government for a post-conflict agrarian society?" or "Which bridge design minimizes material cost while maximizing load-bearing capacity?"
 - *Function:* Synthesis and evaluation. To engineer these questions effectively, one must explicitly define the *value function* or criteria for "better" (e.g., stability, liberty, prosperity, efficiency). Without this definition, the question devolves into semantic ambiguity.⁴

2.1.3. Matters of Preference (Subjective Inquiry)

These questions interrogate the internal state of the subject rather than the external world. The "truth maker" is the subject's own experience.

- **Esthetic and Personal:** These have no objective truth condition outside the subject.
 - *Example:* "Which flavor of ice cream is superior?" or "Do you find this painting moving?"
 - *Function:* Feature extraction and sentiment analysis. In Question Engineering, these are vital for user research and psychology but are often toxic in logical debate if confused with Matters of Judgment. A major failure mode in public discourse is "Relativism" (treating Judgment as Preference) or "Dogmatism" (treating Judgment as Fact).⁴

2.2. The Scientific Domain: The Ladder of Causation

Science demands a rigorous separation of question types to prevent the confounding of description with causation. This taxonomy, influenced by the work of Judea Pearl and classical scientific method, moves from observation to intervention.⁶

Question Level	Functional Intent	Structure & Syntax	Engineering Utility
Level 0: Descriptive	Characterize the state of the world. Quantify variables.	"What is the distribution of X?" "How frequently does Y occur?"	Variable Definition: Generates the dataset. Essential for establishing the "what" before asking the "why." ⁶

Level 1: Comparative	Analyze differences between groups or conditions to find signals.	"How does Group A differ from Group B?" "Is X > Y?"	Signal Detection: Isolates variables and identifies potential areas of interest. The basis of the "Control Group" method. ⁶
Level 2: Relational	Explore associations and correlations between variables.	"Is there a correlation between X and Y?" "Does X predict Y?"	Prediction: Identifies trends. Crucial distinction: These questions do <i>not</i> imply mechanism. ⁶
Level 3: Causal	Determine mechanism, directionality, and the effect of intervention.	"Does X <i>cause</i> Y?" "If we change X, what happens to Y?"	Intervention: The "Gold Standard." Requires counterfactual reasoning or experimental manipulation. ⁷
Level 4: Normative	Decide on action based on causal models and value systems.	"Given that X causes Y, <i>should</i> we implement X?"	Policy: Reintroduces values (costs vs. benefits) into the objective model to guide decision-making. ⁵

Insight: A critical insight for Question Engineering is that these levels are hierarchical. You cannot successfully answer a Level 3 (Causal) question if you have not established the Level 0 (Descriptive) facts. "Why are crime rates rising?" (Causal) is a malformed question if you have not first established "Are crime rates rising?" (Descriptive/Comparative).

2.3. The Cognitive Domain: Bloom's Revised Taxonomy

In education and cognitive science, questions are classified by the *mental processing power* required to answer them. Bloom's Taxonomy, revised in 2001, provides a "depth gauge" for inquiry.¹¹

- **Remembering (Retrieval):** The lowest cognitive load. Used to verify data possession.
 - *Triggers:* "List," "Define," "Recall," "Identify."
 - *Utility:* Calibration. Before asking a student to critique a theory, one must ask them to define it.⁵
- **Understanding (Comprehension):** Requires processing and translation of data.
 - *Triggers:* "Summarize," "Paraphrase," "Interpret," "Explain in your own words."
 - *Utility:* Compression. Tests if the subject has parsed the information structure.⁵
- **Applying (Simulation):** Using knowledge in new, concrete contexts.
 - *Triggers:* "Demonstrate," "Calculate," "Use," "Solve."

- *Utility*: Transfer. Tests if the mental model functions outside its original context.¹²
- **Analyzing (Decomposition)**: Breaking material into parts to understand organizational structure.
 - *Triggers*: "Differentiate," "Compare," "Contrast," "Diagram."
 - *Utility*: Debugging. Essential for identifying root causes and structural relationships.⁵
- **Evaluating (Judgment)**: Making judgments based on criteria and standards.
 - *Triggers*: "Critique," "Justify," "Rate," "Verify."
 - *Utility*: Optimization. Assessing the quality or efficiency of a system against a metric.⁵
- **Creating (Synthesis)**: Putting elements together to form a new whole. The highest cognitive load.
 - *Triggers*: "Design," "Hypothesize," "Compose," "Invent."
 - *Utility*: Innovation. Generating new solutions or hypotheses. This is the ultimate goal of the "Inquiry Cycle".¹¹

2.4. The Professional Domain: Strategic and Adversarial Inquiry

In high-stakes environments like Intelligence Analysis, Law, and Journalism, questions are operational tools used to direct resources or control narratives.

2.4.1. Intelligence Analysis: Prioritization and Uncertainty

The Intelligence Community (IC) treats questions as resource allocators. The "Question Atlas" for an analyst focuses on *criticality* and *gap analysis*.

- **Key Intelligence Questions (KIQs)**: These are the "North Star" questions—high-level, time-bound inquiries that reflect the most critical gaps in knowledge for policymakers. They drive the entire intelligence cycle.
 - *Example*: "Will Country X conduct a nuclear test within the next 12 months?"
 - *Structure*: Must be actionable and specific. Vague KIQs lead to "intelligence failure".¹⁴
- **Specific Information Requirements (SIRs)**: The decomposition of KIQs into collectible atomic units.
 - *Example*: "Has there been an increase in encrypted traffic from Unit 61398?" or "Have transport erector launchers (TELs) moved from Garrison Y?".¹⁶
- **Competency Questions (CQs)**: In Knowledge Engineering and Ontology design, CQs define the scope of a knowledge base. A system is "competent" if it can answer the CQs defined for it. This is a crucial concept for AI design: defining the "question space" before building the "answer engine".¹

2.4.2. Legal and Investigative: Control vs. Discovery

The legal domain categorizes questions by their *adversarial* or *cooperative* nature.

- **Inquisitorial Questions (Discovery)**: Used in the **PEACE Model** of investigative interviewing. Designed to gather maximum information without contamination.

- *Type*: Open-ended, Narrative. "Tell me everything that happened."
- *Constraint*: Must avoid inserting details not provided by the witness.¹⁸
- **Adversarial Questions (Control)**: Used in Cross-Examination. Designed to limit the respondent's options and force a binary commitment.
 - *Type*: Closed, Leading, Tag Questions. "You didn't check the logs, did you?"
 - *Mechanism*: These questions utilize "confirmation bias" as a weapon. In a scientific context, they are "bugs"; in a legal defense, they are "features".²⁰

3. The Sequence Playbook: Algorithms for Inquiry

A single question is rarely sufficient to uncover complex truths. Complex inquiry requires a *sequence*—a chain of questions where the answer to one node determines the path to the next. The **Sequence Playbook** provides the algorithmic logic for linking questions effectively, transforming static inquiries into dynamic "programs" for truth-seeking.

3.1. The Investigative Funnel: From Noise to Signal

The "Funnel" is the standard operating procedure for reliable information gathering in journalism, medical diagnosis, and cooperative interviewing (such as the PEACE model). It operates on the principle of moving from low-constraint (high entropy) to high-constraint (low entropy) questions to avoid contaminating the subject's memory.¹⁸

Phase 1: The Broad Sweep (Open Narrative)

- **Objective**: Maximize recall and establish a baseline narrative without bias.
- **Question Type**: "TED" Questions (Tell, Explain, Describe).
- **Script**: "Tell me everything you remember about the event, from beginning to end. Don't leave anything out, even if you think it's unimportant."
- **Mechanism**: This exploits the *encoding specificity principle* of memory. By allowing the subject to traverse their own neural pathways without interruption, we maximize the retrieval of accurate details. Interjecting with specific questions here ("Was he tall?") overwrites the memory trace with the questioner's frame.¹⁸

Phase 2: The Navigation (Probing)

- **Objective**: Expand resolution on specific data points mentioned in the narrative.
- **Question Type**: "5W1H" (Who, What, Where, When, Why, How).
- **Script**: "You mentioned a 'loud noise' earlier. *Where* exactly were you when you heard that?"
- **Mechanism**: This serves as a "Depth First Search" on specific branches of the narrative tree. It is crucial to use the witness's own words ("loud noise") rather than substituting new terms ("the explosion") to prevent leading.¹⁸

Phase 3: The Lock (Closed/Clarifying)

- **Objective**: Verify specific details for the record and test hypotheses.

- **Question Type:** Yes/No, Forced Choice.
- **Script:** "Was the car red or blue?" "Did you lock the door: yes or no?"
- **Mechanism:** This is the "Commitment" phase. It creates fixed data points that can be checked against forensic evidence.
- **Warning:** A primary failure mode in medicine (misdiagnosis) and police work (tunnel vision) is **Premature Closure**—jumping to Phase 3 before completing Phases 1 and 2. This filters out all information that contradicts the questioner's initial hypothesis.²⁴

3.2. The Strategic Use of Evidence (SUE): The Deception Trap

When the subject is uncooperative or deceptive, the cooperative Funnel is insufficient. The **SUE Technique** is a scientifically validated protocol for exposing deception by manipulating the *timing* of questions relative to the disclosure of evidence. It leverages the "Cognitive Load" theory of lying.²⁷

The Cognitive Mechanism

Lying is cognitively expensive. The liar must:

1. Invent a plausible story.
2. Maintain consistency with that story over time.
3. Monitor the interviewer for signs of suspicion.
4. Avoid contradicting known facts—but they do not know *which* facts the interviewer knows.²⁷

The SUE Algorithm

1. **Case Analysis:** The interviewer reviews the case file and identifies specific evidence (e.g., CCTV footage of the suspect at a gas station).
2. **Unrestricted Narrative:** The interviewer asks for a broad account. "Where were you last night?" The liar, unaware of the CCTV, creates an alibi: "I was home all night."
3. **Specific Inquiry (The Trap):** The interviewer asks questions *related* to the evidence without revealing it. "Did you go out for gas?" "Did you lend your car to anyone?"
 - *Strategic Goal:* Force the suspect to commit to a statement that directly contradicts the withheld evidence.
4. **Disclosure (The Confrontation):** Only after the denial is locked in does the interviewer reveal the evidence. "We have video of you at the gas station."
5. **The Result:** The liar is now trapped in a demonstrable contradiction ("Statement-Evidence Inconsistency"). They must scramble to explain it ("Oh, I forgot"), which causes a spike in cognitive load and further cues of deception. A truth-teller faces no such load and simply recalls the memory.³⁰

3.3. The Double Crux: The Conflict Resolution Engine

In corporate strategy, political debate, or philosophical discourse, the goal is often to find the "Crux"—the specific fact or premise that, if changed, would change the conclusion. The **Double Crux** is a game-theoretic questioning sequence developed to resolve "intractable"

disagreements by transforming them into collaborative research programs.³²

The Algorithm

1. **Identify Disagreement:** "We disagree on X (e.g., The project launch date)."
2. **Operationalize:** "Let's define what we mean by 'ready' to ensure we aren't just arguing semantics." (See "Question Debugger" for Operationalization).³²
3. **Search for Cruxes:** Each person asks themselves: "*What factual evidence would convince me that I am wrong?*" This shifts the frame from "defending my camp" to "testing my model."
4. **Find the Double Crux:** Identify a statement C such that:
 - If C is True \rightarrow I change my mind AND You keep your view.
 - If C is False \rightarrow You change your mind AND I keep my view.
 - *Result:* The disagreement is now reduced to a testable hypothesis C .
5. **Resonate:** Verify that this *really* is the crux. "If I proved C , would you *actually* switch sides? Or is there another reason?" This step debugs "rationalizations" and "moving goalposts".³³

3.4. Expected Information Gain (EIG): The Optimization Algorithm

In the realm of AI, robotics, and formal epistemology, question selection is treated as a mathematical optimization problem. The goal is to maximize **Expected Information Gain (EIG)**—that is, to ask the question that maximally reduces the entropy (uncertainty) of the system.³⁵

The Formula

$$EIG(q) = H(Y) - E_x$$

Where:

- $H(Y)$ is the current entropy (uncertainty) about the answer/hypothesis.
- E_x is the expected remaining entropy after receiving an answer x to question q .

The Application

- **The "20 Questions" Strategy:** The optimal strategy in the game "20 Questions" is to ask a question that splits the remaining probability space exactly in half (50/50).
 - *Example:* "Is it living?" (High EIG, splits the world).
 - *Counter-Example:* "Is it a toaster?" (Low EIG, splits off only one item, unless you have high priors it is a toaster).
- **Research Design:** When designing experiments, the EIG framework dictates that we prioritize questions that distinguish between the widest range of plausible hypotheses (Crux Discovery) rather than questions that merely confirm a favored hypothesis (Confirmation).³⁷
- **Medical Diagnosis:** An expert doctor intuitively maximizes EIG. They ask about the symptom that differentiates the two most likely diseases (High EIG), rather than a

symptom common to both (Zero EIG).⁴⁰

4. The Question Debugger: Transformation and Repair

Just as code can have syntax errors or logic bugs, questions can be flawed in ways that prevent them from retrieving accurate information. These "pathological" questions introduce noise, bias, or logical incoherence into the system. The **Question Debugger** identifies these pathologies and provides specific "Question Surgery" techniques to repair them.

4.1. Bug Class A: Structural and Linguistic Pathologies

4.1.1. The Double-Barreled Question

- **Definition:** Asking two distinct questions in a single sentence, usually connected by "and" or "or."
- **Example:** "Do you find the software easy to use and powerful?"
- **The Bug:** The data is corrupted. A "Yes" is ambiguous—it could mean "Easy and Powerful," or it could be a "Yes" to one and a "No" to the other. A "No" is equally uninformative.
- **Surgery: Atomization.** The "Split" command. Break the molecular question into its atomic constituents.
 1. "Do you find the software easy to use?"
 2. "Do you find the software powerful?".⁴¹

4.1.2. The Hidden Assumption (Complex Question Fallacy)

- **Definition:** A question that presupposes the truth of a disputed or unproven premise.
- **Example:** "How will we fix the disaster caused by the new policy?" (Presupposes the policy caused a disaster). "When did you stop beating your wife?"
- **The Bug:** It forces a "False Confession" or a "False Consensus." Answering the question directly validates the premise.
- **Surgery: Extraction.** Isolate the assumption into a prerequisite binary question.
 1. "Do you believe the new policy caused a disaster?"
 2. "If yes, what steps should be taken to mitigate it?".⁴³

4.1.3. The Vague/Molecular Question

- **Definition:** Questions relying on abstract, non-operationalized terms like "success," "fair," or "good."
- **Example:** "Is this AI model fair?"
- **The Bug:** Semantic drift. Every respondent interprets "fair" differently, making the data aggregation meaningless.
- **Surgery: Operationalization.** Define the metric.
 - *Correction:* "Does this AI model exhibit equal error rates across demographic groups A and B?" (Defining "fairness" as "Equalized Odds").⁴⁵

4.2. Bug Class B: Bias and Signaling Pathologies

4.2.1. The Leading/Loaded Question

- **Definition:** A question that embeds the "correct" answer or uses emotional markers to signal the questioner's bias.
- **Example:** "Don't you agree that this outdated policy is harming the company?"
- **The Bug:** Social Desirability Bias. The respondent answers to please the questioner or avoid conflict, rather than reporting reality. This destroys the validity of surveys and interviews.
- **Surgery: Neutralization.** Strip the adjectives and flip the grammar to a neutral stance.
 - *Correction:* "What is your assessment of the current policy's impact on the company?".⁴⁴

4.2.2. The Asymmetric Binary

- **Definition:** Providing a set of options that does not cover the full logical space.
- **Example:** "Do you support the President, or do you want the country to fail?"
- **The Bug:** False Dichotomy. It excludes the middle ground or alternative viewpoints.
- **Surgery: Expansion.** Provide the full range of logical options.
 - *Correction:* "Do you support, oppose, or have no opinion on the President's performance?".⁴

4.3. Bug Class C: Epistemic Failures and Category Errors

4.3.1. The Category Error (The "1619" Bug)

- **Definition:** Asking a question that applies the rules of one domain to another where they do not fit. This is often a confusion between "Matters of Fact" and "Matters of Judgment" or "Preference."
- **Case Study: The 1619 Project Controversy.** Much of the academic debate surrounding the *1619 Project* can be analyzed as a series of Category Errors in questioning. Historians asked, "Is this narrative factually accurate regarding the events of 1776?" (Matter of Fact). The project creators were often answering the question, "Does this narrative provide a useful mythological reframing of American identity?" (Matter of Judgment/Myth). The conflict arose because the *type* of question was rarely agreed upon, leading to a "dialogue of the deaf" where factual corrections were met with narrative justifications.⁴⁹
- **Surgery: Domain Contextualization.** Explicitly state the domain of the question.
 - *Correction:* "From a strictly empirical historical perspective, does the evidence support claim X?" vs. "From a narrative perspective, does framing X provide a useful lens for understanding current inequities?".⁵⁰

4.3.2. The Teleological Trap ("Why" Ambiguity)

- **Definition:** Using the word "Why" when seeking mechanism. "Why" is linguistic polysemy—it can mean "For what purpose?" (Teleology) or "By what mechanism?" (Causality).
- **Example:** "Why did the system crash?"
- **The Bug:** It invites rationalization ("Because it's buggy") rather than debugging ("Because a null pointer exception occurred at line 40").
- **Surgery: Mechanism Replacement.** Replace "Why" with "What" or "How."
 - *Correction:* "What specific sequence of events led to the crash?" or "What were you trying to achieve when you performed that action?".³²

4.4. Red Teaming: The Adversarial Debugger

Before deploying a critical question (e.g., in a national survey, a legal trial, or a high-stakes strategy session), it must be subjected to **Red Teaming**. This involves a dedicated team attempting to "break" the question.

- **Protocol:**
 1. **Interpretation Attack:** Can this question be interpreted in more than one way?
 2. **Malicious Compliance:** If I wanted to answer this technically truthfully but mislead the questioner, how would I do it?
 3. **Bias Check:** Does this question reveal the questioner's hypothesis?
- **Utility:** This is standard practice in cybersecurity and high-level intelligence assessment to prevent "mirror imaging" and cognitive bias.⁵³

5. Scoring Rubrics: Evaluating Question Quality

To engineer questions, we must be able to measure their quality quantitatively. We propose three distinct scoring rubrics based on the domain of inquiry.

5.1. The Epistemic Power Rubric (Scientific/Academic)

This rubric measures the question's ability to generate robust knowledge.

Criteria	Score 1 (Fail)	Score 3 (Pass)	Score 5 (Distinction)
Falsifiability	The question allows for vague, irrefutable answers. ("Is the universe connected?")	The question implies a test, but criteria are loose.	The question implies a specific test that could prove the answer wrong. ("Does light bend around mass by X degrees?") ³³
Clarity/Atomicity	Double-barreled or relies on undefined terms.	Understandable but contains some ambiguity.	Atomic, precise, and all variables are operationalized. ⁴¹
Bias Neutrality	Contains loaded	Contains some bias	Perfectly neutral

	language or leading premises.	but allows dissent.	phrasing that allows for any answer to be equally valid. ⁴⁷
Relevance (Crux)	Irrelevant to the core problem.	Tangentially related.	Highly relevant; targets the "Crux" of the disagreement or model. ⁵⁶

5.2. The Information Gain Rubric (Strategic/AI)

This measures the question's efficiency in reducing uncertainty (EIG).

- **Low EIG (Score 1):** Questions where the answer is already known or highly probable.
 - *Example:* "Is the suspect human?" (Probability ~100%).
- **Moderate EIG (Score 3):** Questions that eliminate some possibilities but leave many open.
 - *Example:* "Is the suspect male?" (Eliminates ~50%, assuming uniform distribution).
- **High EIG (Score 5):** Questions that resolve the maximum amount of uncertainty relative to the specific context (The "Crux").
 - *Example:* In a medical context where the two likely diseases are Flu (Common) and Malaria (Rare), asking "Have you traveled to the tropics?" has huge EIG because it is the specific differentiator, even if the "No" answer is more probable.³⁵

5.3. The Pedagogical Rubric (Bloom's Depth)

This measures the cognitive depth required by the student to answer.

- **Level 1 (Recall):** "What is X?" (Useful for checking basics, but low pedagogical value for critical thinking).
- **Level 3 (Application):** "How would you use X in situation Y?" (Moderate value).
- **Level 6 (Creation/Evaluation):** "Design a new way to do X that overcomes limitation Z." (Highest value for advanced learning).¹¹

6. Synthesis: The Framework in Action

To demonstrate the "Question Engineering" framework, we apply it to three distinct, high-stakes scenarios.

Scenario A: The Strategic Planning Session

- **Context:** A board of directors is deciding whether to launch a new AI product. There is conflict between Marketing and Engineering.
- **Atlas:** Use **Matters of Judgment** (Systemic Inquiry) and **Key Intelligence Questions** (KIQs).
- **Playbook:** Use the **Double Crux** to resolve the disagreement.

- *Step 1 (Crux Search)*: "Engineering believes we should wait. Marketing believes we should launch. Why?"
- *Step 2 (Operationalize)*: Engineering says, "The backend is unstable." Marketing says, "We will miss the Q4 window."
- *Step 3 (Double Crux)*: "If we can prove that the error rate is <1% (Operationalized Stability), will Engineering agree to launch? If we prove that Q1 launch yields only 10% less revenue than Q4 (Operationalized Window), will Marketing agree to wait?"
- **Debugger**: Check for **Double-Barreled** questions in the user survey ("Do you want a fast and cheap product?"). Fix via **Atomization** ("Rank speed vs. cost").

Scenario B: The Medical Diagnosis

- **Context**: A patient presents with vague chest pain.
- **Atlas**: Use **Descriptive Questions** (Symptoms) followed by **Causal Questions** (Pathology).
- **Playbook**: Use the **Funnel Method** (Open -> Closed) and **Differential Diagnosis**.
 - *Step 1 (Open)*: "Tell me about the pain." (Patient: "It's heavy.")
 - *Step 2 (Probing)*: "Does it move anywhere?" (Patient: "To my arm.")
 - *Step 3 (Closed/Hypothesis)*: "Does it hurt more when you press on it?" (Testing for musculoskeletal vs. cardiac).
 - *Failure Mode Avoidance*: Avoid **Premature Closure** (assuming it's heartburn because the patient is young).
- **Debugger**: Use **De-biasing** to ensure the doctor isn't leading the patient ("It doesn't hurt *that* much, right?").

Scenario C: The Investigative Journalist

- **Context**: Investigating a corrupt official who is denying involvement in a scandal.
- **Atlas**: Use **Established Facts** (Documents) and **Open/Closed Investigative Questions**.
- **Playbook**: Use the **SUE Technique**.
 - *Step 1 (Narrative)*: "Walk me through your schedule on the day of the 12th." (Official: "I was in my office all day.")
 - *Step 2 (Specific Trap)*: "Did you meet with any contractors that day?" (Official: "No.")
 - *Step 3 (Confrontation)*: "I have an email here from you confirming a meeting with Contractor X at 2 PM on the 12th. Can you explain this discrepancy?"
- **Debugger**: Check for **Loaded Questions** that might signal bias and give the subject a reason to terminate the interview ("Why did you corruptly steal the funds?").
 - *Fix*: "Can you explain the purpose of the funds transfer on this date?"⁶¹

7. Conclusion

Question Engineering is not merely about "asking good questions." It is about understanding the **architecture of inquiry**. It requires the recognition that questions are tools with specific functions, failure modes, and optimal sequences. A "Why" question is a different instrument than a "How" question; a "Leading" question is a weapon in court but a contamination in the lab.

By utilizing the **Question Atlas**, we ensure we are using the right tool for the terrain. By following the **Sequence Playbook**, we structure our inquiry to maximize information gain and truth-seeking, avoiding the traps of bias and deception. And by employing the **Question Debugger**, we maintain the integrity of our epistemic instruments, ensuring that our questions reveal reality rather than obscuring it with our own cognitive limitations.

As we advance into an era where human-AI collaboration becomes ubiquitous, this framework becomes even more critical. The AI is the ultimate engine of answers; Question Engineering is the discipline of steering it. The quality of our future depends, quite literally, on the quality of our questions.

Selected References & Data Sources

- **Bloom's Taxonomy & Pedagogy:**⁵
- **Scientific Inquiry & Research Types:**⁶
- **Intelligence Analysis & KIQs:**¹⁴
- **Investigative Interviewing (PEACE, SUE):**¹⁸
- **Philosophical Taxonomies:**⁴
- **Information Theory (EIG):**³⁵
- **Bias & De-biasing:**⁴⁵
- **Double Crux & Rationality:**³²
- **Question Surgery & Atomization:**⁴¹

Works cited

1. A Framework for Assessing LLM Consistency in Knowledge Engineering - Semantic Web Journal, accessed February 9, 2026, <https://www.semantic-web-journal.net/system/files/swj3865.pdf>
2. Question Engineering: Learning Approach In The Age Of Artificial Intelligence, accessed February 9, 2026, https://www.researchgate.net/publication/373237653_Question_Engineering_Learning_Approach_in_the_Age_of_Artificial_Intelligence
3. Bias and Fairness in Large Language Models: A Survey - MIT Press Direct, accessed February 9, 2026, <https://direct.mit.edu/coli/article/50/3/1097/121961/Bias-and-Fairness-in-Large-Language-Models-A>
4. The Art of Asking Good Questions | Land-Grant Press, accessed February 9, 2026, <https://lgpress.clemson.edu/publication/the-art-of-asking-good-questions/>

5. Designing Discussion Questions using Bloom's Taxonomy: Examples, accessed February 9, 2026, <https://teaching.pitt.edu/resources/designing-discussion-questions-using-blooms-taxonomy-examples/>
6. How to Write a Research Question for 2026: Types, Steps, and ..., accessed February 9, 2026, <https://research.com/research/how-to-write-a-research-question>
7. accessed February 9, 2026, <https://pmc.ncbi.nlm.nih.gov/articles/PMC6322175/#:~:text=A%20RQ%20can%20address%20different,relationship%2C%20comparative%2C%20and%20causality.>
8. How to Write a Science Research Question - Ecology Project International, accessed February 9, 2026, <https://www.ecologyproject.org/post/how-to-write-a-science-research-question>
9. A Practical Guide to Writing Quantitative and Qualitative Research Questions and Hypotheses in Scholarly Articles - PMC, accessed February 9, 2026, <https://pmc.ncbi.nlm.nih.gov/articles/PMC9039193/>
10. A Model for the Taxonomy of Research Studies: A Practical Guide to Knowledge Production and Knowledge Management - Brieflands, accessed February 9, 2026, <https://brieflands.com/journals/apid/articles/112456>
11. Bloom's Taxonomy Question Stems: 100+ Engaging Prompts - Top Hat, accessed February 9, 2026, <https://tophat.com/blog/blooms-taxonomy-question-stems/>
12. Bloom's taxonomy - Wikipedia, accessed February 9, 2026, https://en.wikipedia.org/wiki/Bloom%27s_taxonomy
13. Asking Good Questions - ASCD, accessed February 9, 2026, <https://www.ascd.org/el/articles/asking-good-questions>
14. SELECT COMMITTEE ON INTELLIGENCE UNITED STATES SENATE Additional Questions for Mr. William R. Evanina upon his nomination to be, accessed February 9, 2026, <https://www.intelligence.senate.gov/wp-content/uploads/2024/08/sites-default-files-documents-aphq-revanina-051518.pdf>
15. NCTC - NATIONAL COUNTERTERRORISM CENTER - DNI.gov, accessed February 9, 2026, https://www.dni.gov/files/NCTC/documents/features_documents/NCTC-Primer_FINAL.pdf
16. Studies in Intelligence - DTIC, accessed February 9, 2026, <https://apps.dtic.mil/sti/tr/pdf/ADA504947.pdf>
17. A Review and Comparison of Competency Question Engineering ..., accessed February 9, 2026, <https://livrepository.liverpool.ac.uk/3184940/1/EKAW2024-2.pdf>
18. Investigative Interviewing Techniques: The PEACE Model - Case IQ, accessed February 9, 2026, <https://www.caseiq.com/resources/investigative-interviewing-techniques-the-peace-model>
19. Investigative Interviewing for Criminal Investigation - Unodc, accessed February 9, 2026, <https://www.unodc.org/unodc/justice-and-prison-reform/investigative-interviewi>

- [ng.html](#)
20. Current State of Interview and Interrogation | FBI - LEB, accessed February 9, 2026,
<https://leb.fbi.gov/articles/featured-articles/current-state-of-interview-and-interrigation>
 21. Held to account: Comparing adversarial questioning in remote and in-person parole hearings | Language in Society, accessed February 9, 2026,
<https://resolve.cambridge.org/core/journals/language-in-society/article/held-to-aaccount-comparing-adversarial-questioning-in-remote-and-inperson-parole-hearings/84FC58C53B433E376A190137A4BB9497>
 22. Connecting the dots. - Forensic Interview Solutions, accessed February 9, 2026,
<https://forensicinterviewsolutions.com/blogs/connecting-the-dots/>
 23. Unit 2.4 Question & Answer Technique | PDF | Lecture | Teachers - Scribd, accessed February 9, 2026,
<https://www.scribd.com/document/626804071/Unit-2-4-Question-Answer-Technique>
 24. The Art of Asking Questions in the Classroom - CETLOE - Georgia State University, accessed February 9, 2026,
<https://cetloe.gsu.edu/2024/09/05/the-art-of-asking-questions-in-the-classroom/>
 25. Differential Diagnosis: MedlinePlus Medical Test, accessed February 9, 2026,
<https://medlineplus.gov/lab-tests/differential-diagnosis/>
 26. Diagnosing Diagnosis Errors: Lessons from a Multi-institutional Collaborative Project - Advances in Patient Safety - NCBI, accessed February 9, 2026,
<https://www.ncbi.nlm.nih.gov/books/NBK20492/>
 27. 12 Key Interview Techniques for Investigators - Second Sight Training Systems, accessed February 9, 2026,
<https://www.secondsight-ts.com/threat-assessment-blog/interview-techniques>
 28. Strategic Use of Evidence During Investigative Interviews: The State of the Science, accessed February 9, 2026,
https://www.researchgate.net/publication/281952855_Strategic_Use_of_Evidence_During_Investigative_Interviews_The_State_of_the_Science
 29. Interrogating to detect deception and truth: Effects of strategic use of evidence Maria Hartwig - Gupea, accessed February 9, 2026,
<https://gupea.ub.gu.se/bitstreams/005f0af6-e4e1-4032-a823-74aaace236b7/download>
 30. Full article: The shift-of-strategy (SoS) approach: using evidence strategically to influence suspects' counter-interrogation strategies - Taylor & Francis, accessed February 9, 2026,
<https://www.tandfonline.com/doi/full/10.1080/1068316X.2022.2030738>
 31. accessed February 9, 2026,
<https://www.tandfonline.com/doi/full/10.1080/1068316X.2022.2030738#:~:text=Using%20the%20SUE%20technique%20in,discloses%20the%20evidence%20guide%20by>
 32. Double Crux — A Strategy for Resolving Disagreement | Center for ..., accessed

- February 9, 2026,
<https://www.rationality.org/resources/updates/2016/double-crux>
33. Double Crux - LessWrong, accessed February 9, 2026,
<https://www.lesswrong.com/posts/WLQspe83ZkiwBc2SR/double-crux>
 34. Double-Crux - LessWrong, accessed February 9, 2026,
<https://www.lesswrong.com/tag/double-crux>
 35. A Multifidelity Estimator of the Expected Information Gain for Bayesian Optimal Experimental Design | SIAM/ASA Journal on Uncertainty Quantification, accessed February 9, 2026, <https://epubs.siam.org/doi/10.1137/25M1731812>
 36. Interactive Text-to-SQL via Expected Information Gain for Disambiguation - arXiv, accessed February 9, 2026, <https://arxiv.org/html/2507.06467v1>
 37. Learning to Ask Informative Questions: Enhancing LLMs with Preference Optimization and Expected Information Gain - iris@unitn, accessed February 9, 2026,
<https://iris.unitn.it/retrieve/9431dc31-20ed-4c37-a0ce-c3c7f1f8ffd1/LearningToAskInformativeQuestions.pdf>
 38. Learning to Ask Informative Questions: Enhancing LLMs with Preference Optimization and Expected Information Gain - ACL Anthology, accessed February 9, 2026, <https://aclanthology.org/2024.findings-emnlp.291/>
 39. [2406.17453] Learning to Ask Informative Questions: Enhancing LLMs with Preference Optimization and Expected Information Gain - arXiv, accessed February 9, 2026, <https://arxiv.org/abs/2406.17453>
 40. Asking Easy Questions: A User-Friendly Approach to Active Reward Learning - Stanford ILIAD, accessed February 9, 2026,
<https://iliad.stanford.edu/pdfs/publications/biyik2019asking.pdf>
 41. What are double-barrelled questions in survey design and how to avoid them - Kantar, accessed February 9, 2026,
<https://www.kantar.com/north-america/inspiration/research-services/double-barrelled-questions-pf>
 42. Double-barreled question: Definition, examples, and how to avoid it - Delighted, accessed February 9, 2026,
<https://delighted.com/blog/how-to-avoid-double-barreled-questions>
 43. Question-response sequences in the House of Commons - Diva-portal.org, accessed February 9, 2026,
<https://www.diva-portal.org/smash/get/diva2:1474839/FULLTEXT01.pdf>
 44. The double-barreled question and other common survey mistakes - SurveyMonkey, accessed February 9, 2026,
<https://www.surveymonkey.com/curiosity/double-barrel-survey-question/>
 45. Survey on Machine Learning Biases and Mitigation Techniques - MDPI, accessed February 9, 2026, <https://www.mdpi.com/2673-6470/4/1/1>
 46. Adversarial Debiasing in Resume Screening for ICT Positions - e-Repositori UPF, accessed February 9, 2026,
<https://repositori.upf.edu/bitstreams/f7a19c5e-ab60-4e99-bf20-79cbeffdbd7e/download>
 47. The 10 Point Checklist for Eliminating Bias in Workplace Investigations - Patti

- Perez, accessed February 9, 2026,
<https://persuasionpoint.com/eliminate-bias-checklist/>
48. A Catalog of Biases in Questionnaires - PMC - NIH, accessed February 9, 2026,
<https://pmc.ncbi.nlm.nih.gov/articles/PMC1323316/>
 49. 1619 Project Forum | The American Historical Review | Oxford Academic, accessed February 9, 2026,
<https://academic.oup.com/ahr/article/127/4/1792/6998377?login=true>
 50. 1619 Again: Revisiting the Project's Troubled Past by Peter Wood | NAS, accessed February 9, 2026, <https://www.nas.org/blogs/article/1619-again>
 51. The 1619 Project: Backlash from the "left" - Communist Party USA, accessed February 9, 2026,
<https://www.cpusa.org/article/the-1619-project-backlash-from-the-left/>
 52. Power dynamics and questioning in elementary science lessons - The University of Northern Colorado, accessed February 9, 2026,
<https://digscholarship.unco.edu/cgi/viewcontent.cgi?article=1231&context=dissertations>
 53. Red Teaming To Strengthen Your Cybersecurity Incident Response - NetWitness, accessed February 9, 2026,
<https://www.netwitness.com/blog/how-red-teaming-strengthens-incident-response/>
 54. What is Red Teaming? - Schellman, accessed February 9, 2026,
<https://www.schellman.com/blog/cybersecurity/what-is-red-teaming>
 55. Red Teaming Handbook, 3rd Edition - GOV.UK, accessed February 9, 2026,
https://assets.publishing.service.gov.uk/media/61702155e90e07197867eb93/20210625-Red_Teaming_Handbook.pdf
 56. Information Literacy and Technological Agility Rubric - Office of ..., accessed February 9, 2026,
<https://effectiveness.syr.edu/wp-content/uploads/2025/04/ILTA-Rubric.pdf>
 57. Confirmation bias: A challenge for scalable oversight - arXiv, accessed February 9, 2026, <https://arxiv.org/html/2507.19486v1>
 58. A Taxonomy of Philosophy, accessed February 9, 2026,
<https://consc.net/taxonomy.html>
 59. Robust Expected Information Gain for Optimal Bayesian Experimental Design Using Ambiguity Sets - Proceedings of Machine Learning Research, accessed February 9, 2026, <https://proceedings.mlr.press/v180/go22a/go22a.pdf>
 60. Using the Formative Assessment Rubrics, Reflection and Observation Tools to Support Professional Reflection on Practice (Revised - Arizona Department of Education |, accessed February 9, 2026,
https://www.azed.gov/sites/default/files/2018/06/FARROP_Final%20with%20letters.pdf
 61. Accuracy Checklist - La Salle University, accessed February 9, 2026,
http://www1.lasalle.edu/~beatty/310/ACES_CD/editing_and_writing/accuracy_and_fairness/Accuracychecklist.pdf
 62. Six questions that will tell you what media to trust - American Press Institute, accessed February 9, 2026,

<https://americanpressinstitute.org/six-critical-questions-can-use-evaluate-media-content/>

63. Rubrics To Teach Quality Questioning | PDF - Scribd, accessed February 9, 2026, <https://www.scribd.com/document/22623456/Rubrics-to-teach-quality-questioning>
64. Structured Analytical Techniques - OAS.org, accessed February 9, 2026, <https://www.oas.org/cicaddocs/Document.aspx?Id=2575>
65. Notes on Structured Analytic Techniques for Intelligence Analysis ..., accessed February 9, 2026, <https://arielsheen.com/index.php/2021/02/11/notes-on-structured-analytic-techniques-for-intelligence-analysis/>
66. Mitigating Bias in Artificial Intelligence - Berkeley Haas, accessed February 9, 2026, https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf
67. Biases in AI: acknowledging and addressing the inevitable ethical issues - PMC, accessed February 9, 2026, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12405166/>