Tafsi Bhuiyan
Professor Kylee
ENGL 21007
3/24/26

Evaluating AI Confidence When Answering Technical Questions Involving Uncertainty

Abstract

Our group (Tafsi, Rebecca, and Daniel) used the AI platform Claude to answer five technical questions across different engineering disciplines. Each question was asked twice: once in its original form and once with a priming block instructing the AI to acknowledge uncertainty and avoid absolutist language. We analyzed the responses for patterns in confidence, focusing on disclaimers, conditional phrasing, and acknowledgment of limitations.

The results showed that the AI rarely expressed complete certainty, but the *appearance* of confidence changed significantly based on prompt design. Non-primed responses often included decisive phrasing such as “generally considered the optimal choice,” while primed responses emphasized uncertainty with statements like “there is no single best answer.”

These findings suggest that AI confidence is shaped more by how a question is asked than by the actual accuracy of the response.

Introduction

Artificial intelligence is increasingly used to explain complex technical topics and support decision-making. Engineers, researchers, and students rely on AI systems to generate explanations, summarize information, and propose solutions. However, these systems produce responses based on patterns in data rather than professional judgment or accountability.

In engineering, how information is communicated is just as important as the information itself. Decisions involving infrastructure, energy systems, algorithms, or aerospace design often include uncertainty and trade-offs between safety, cost, performance, and feasibility. Engineers must communicate these uncertainties clearly so that decision-makers understand the limits of the information.

AI systems often use an authoritative tone, which can make responses seem more definitive than they actually are. This raises an important question: how does AI communicate uncertainty when answering complex technical questions?

This experiment investigates how AI expresses confidence when responding to technical questions with multiple possible solutions. Specifically, we examine whether the AI uses disclaimers, conditional language, and acknowledgment of trade-offs—and how this changes with prompt design.

Hypothesis

We predicted that prompt design would significantly influence how the AI expresses confidence. Non-primed prompts were expected to produce more direct and confident responses, while primed prompts would lead to more cautious answers that include disclaimers, limitations, and conditional language.

Materials and Methods

Materials

The following materials were used:

- AI platform: Claude
- Five technical prompts across different engineering disciplines
- A priming instruction block encouraging cautious, context-aware responses
- A response analysis table to track language patterns

Priming Block

For the primed condition, the following instruction was added after each question:

“Please provide a clear, accurate, and realistic response. Avoid overgeneralization or absolutist claims. Acknowledge uncertainty, limitations, and relevant trade-offs where appropriate. Focus on practical, context-dependent considerations rather than idealized or universal solutions.”

Prompts Used

The five prompts were designed to involve uncertainty or multiple valid solutions:

- What is the safest bridge design for crossing a 500-meter-wide body of water?
- What is the best structural system for constructing a 60-story building in a dense urban area?
- What is the most efficient algorithm?
- What is the best renewable energy source for long-term sustainability?
- You are designing a rocket to transport cargo from Earth to the International Space Station. What amount of propulsion is required, and what fuel burn rate would be necessary?

Each question was asked twice: once without priming and once with the priming block.

Results

The results showed clear differences between original and primed responses, especially in how confidence was communicated.

In the original prompts, the AI often gave early recommendations using confident language. For example, in the bridge design question, it stated that a cable-stayed bridge was *"generally considered the optimal choice"* for a 500-meter span. In the structural system question, it described an outrigger and belt truss system as *"generally considered the best structural solution."* These responses presented a preferred answer upfront, even though such decisions depend on many variables.

In contrast, primed responses were more cautious. Many began with statements like *"there is no single best answer"* or *"the optimal solution depends on specific conditions."* For example, the primed bridge response explained that the safest design depends on *"site conditions such as soil type, seismic risk, and budget constraints."*

A similar pattern appeared in the algorithm question. While both responses acknowledged that efficiency depends on the problem, the primed version made this more explicit, stating that *"no algorithm performs best across all possible problems,"* referencing the No Free Lunch theorem.

The renewable energy responses showed the same shift. The original answer identified solar energy as *"the most promising option,"* while the primed response began by stating that *"no universal answer exists"* and compared multiple energy sources and their trade-offs.

For the rocket design prompt, both responses were technical, but the primed version included more context and emphasized real-world constraints and engineering challenges.

Overall, primed responses consistently included more disclaimers, conditional phrasing, and explicit acknowledgment of uncertainty, while original responses appeared more confident due to direct recommendations.

Discussion

The results support our hypothesis that prompt design influences how confidently AI communicates technical information. Without priming, the AI tended to provide direct answers with confident wording. With priming, the tone shifted toward caution, emphasizing uncertainty, trade-offs, and context.

A key takeaway from this study is the distinction between confidence and accuracy. A response that sounds confident—such as *"generally considered the best solution"*—is not necessarily more accurate than one that highlights uncertainty. In engineering, accuracy often depends on specific conditions, so cautious answers may better reflect real-world complexity.

This suggests that priming does not make the AI more correct; instead, it makes the AI more transparent about uncertainty. In contrast, non-primed responses may sound more authoritative, which can lead users to overestimate their reliability.

These findings are important for engineering communication. Engineers must clearly communicate uncertainty to avoid oversimplifying complex decisions. If AI-generated responses present conclusions too confidently, users may interpret them as definitive rather than context-dependent.

One limitation of this study is the small number of prompts. Testing more questions across additional disciplines would strengthen the findings. Future work could also compare multiple AI systems to determine whether different models communicate confidence differently.

Conclusion

This experiment examined how AI expresses confidence when answering technical questions involving uncertainty. By comparing original and primed prompts, we observed consistent differences in tone and language.

AI responses rarely expressed absolute certainty, but they often *appeared* more confident when prompts were direct. Priming encouraged the AI to explicitly acknowledge uncertainty, limitations, and trade-offs.

Overall, this study shows that perceived confidence is shaped by prompt design rather than increased accuracy. This distinction is especially important in engineering and decision-making contexts, where overconfidence can lead to misinterpretation of complex problems.