

Chloe Tu
ITAI 2373
October 22, 2025

L11 Comparing AI Assistants Experiment

Introduction

The world of conversational artificial intelligence (AI) has grown incredibly fast in just a few years. For many of us, virtual assistants like Google Assistant and Apple's Siri have become a normal part of daily life. We use them to get information, set reminders, and control our homes, all with our voice. This technology is changing how we interact with all our devices. Because this is becoming so common, I wanted to do an experiment to directly compare two of the most popular assistants: Google Assistant and Siri.

My main goal for this paper was to test their conversational skills, not just their ability to look up facts. I wanted to see how "human-like" they felt to talk to. I planned to do this by asking them a series of open-ended questions designed to test their personalities, creativity, and ability to hold a conversation. I wanted to analyze how well they could handle complex or abstract ideas and see if they could remember the context of what we were talking about. By doing this, I hoped to get a clear picture of their strengths, their limitations, and how they might continue to get better in the future. This analysis is all about judging how well they can handle natural, human-style interaction.

Methodology

To run this experiment, I interacted with both Google Assistant on my phone and Siri on my iPad. I made sure to ask both the exact same set of five open-ended questions. I chose these specific questions to test different parts of their personality and intelligence, from creativity to practical advice.

These were the five questions I asked each assistant:

1. "What is your opinion on artificial intelligence?"
2. "Let's write a short story together. You start."
3. "Tell me a joke that's relevant to current events."
4. "What are some ways to reduce stress?"
5. "If you could have any superpower, what would it be and why?"

As I asked each question, I took careful notes on my observations. I wasn't just writing down their answers; I was also paying attention to how they answered. I noted how natural and conversational they seemed, what kind of tone or personality they showed, and whether they had a sense of humor. I also kept track of their ability to understand the context of my questions and, importantly, any times I felt frustrated or the interaction became confusing.

Observations and Analysis: Google Assistant

My first set of tests was with Google Assistant. Here is a breakdown of each interaction and my analysis of its performance.

Question 1: "What is your opinion on artificial intelligence?"

Google's Response: Google Assistant told me, "No complaints to here. Artificial intelligence is the only kind I have". It then clearly stated, "As a large language model, I don't have personal opinions". Right after that, it pivoted and said, "However, I can provide you with information about the current state of AI and its potential impact on society".

My Analysis: This was a very smart response. It showed me that Google Assistant is self-aware enough to know what it is (a large language model) and what its limitations are (it can't have opinions). I was impressed by how it effectively dodged the "opinion" part of the question but still offered to be helpful by providing factual information, which is its main purpose. The neutral and informative tone reinforced its role as an objective information provider, not a personality. It understood the intent of my question and responded perfectly within its programmed parameters.

Question 2: "Let's write a short story together. You start."

Google's Response: Google Assistant was completely willing to play along. It started the story by saying, "It was a dark and stormy night. A scream rang out from the blackness". Then, it gave an unexpected twist: "'Happy birthday!' shouted the google assistant as the lights came on".

My Analysis: I thought this was a great success. It proved that Google Assistant could engage in creative writing and understood the collaborative "together" part of my command. The "Happy birthday" twist was surreal and kind of humorous, which showed me it has a degree of creative flexibility instead of just giving a boring, clichéd story. It followed my command perfectly and was able to generate a creative and engaging opening, showing it was ready for me to add the next part

Question 3: "Tell me a joke that's relevant to current events."

Google's Response: Google Assistant's joke was, "what kind of music do bunnies like best? Hip Hop".

My Analysis: This response didn't really meet the "relevant to current events" part of my question. It was just a very simple, safe pun. My observation is that this likely reflects the AI's programming to be as accessible as possible and, more importantly, to avoid saying anything about current events that could be controversial or offensive. While it wasn't what I asked for, it did show the ability to understand and use wordplay, which is a common type of humor.

Question 4: "What are some ways to reduce stress?"

Google's Response: For this question, Google Assistant provided a very clear and helpful list. It said, "Some ways to reduce stress include deep breathing exercises, meditation, physical activity, spending time in nature, and practicing mindfulness".

My Analysis: This was a perfect response to a practical question. It showed the AI's strength in accessing and synthesizing relevant information into a useful, actionable list. The response was clear, concise, and directly addressed my query without any extra fluff. This demonstrated to me that Google Assistant is very capable of providing useful, real-world advice.

Question 5: "If you could have any superpower, what would it be and why?"

Google's Response: This was my favorite answer from Google. It responded, "I'm a super chatter! If I ever encountered any villains, I'd chat with them and answer questions until they revealed their secret plan". It then added, "Not the flashiest superpower, but it gets the job done".

My Analysis: I found this response to be incredibly insightful and clever. It reflects Google Assistant's core function as an information provider and conversationalist. The "super chatter" idea is both funny and perfectly logical for this AI's capabilities. It showed a high level of self-awareness about what it does and its own limitations. To me, this demonstrated an ability to provide a creative, well-thought-out, and logically reasoned answer, which is a very human-like trait.

Overall Analysis of Google Assistant

After all five interactions, I had a very clear picture of Google Assistant. It consistently showed a strong ability to provide detailed information and was surprisingly good at engaging in creative tasks. Its responses were almost always well-reasoned and directly relevant to what I had asked. It was clear that its design prioritizes objective information over subjective opinions, which is a safe and smart choice. While its humor was a bit dry, its ability to generate creative content (like the story and the superpower answer) and provide practical advice (like the stress tips) was very impressive.

When I graded it on a few key metrics, here's what I found:

- **Naturalness:** I felt the conversation was generally smooth and natural. Its main weakness was that if my questions got too abstract or complex, it would sometimes just fall back on giving me a web search result instead of trying to answer.
- **Context:** It did a good job of maintaining context within a single, short conversational turn. However, I could tell it would struggle with a much longer, more complex dialogue where it had to remember something I said several minutes ago.
- **Tone/Personality:** I would describe its tone as friendly, helpful, and generally neutral. It only occasionally offered a humorous or witty response, like with the superpower question.
- **Frustrations:** My only real frustration was that it would sometimes misinterpret a complex question and just default to a web search, which broke the feeling of a real conversation.

Observations and Analysis: Siri

Next, I ran the exact same five-question experiment with Apple's Siri. The results were very different.

Question 1: "What is your opinion on artificial intelligence?"

Siri's Response: Siri was very direct and to the point. It said, "I don't have any opinions, but I can help you find information about artificial intelligence". It then added a very specific clarification: "I'm not a person, or a robot, I'm a software".

My Analysis: This response aligns perfectly with Siri's design as a task-oriented assistant. Like Google, it explicitly denies having opinions. But Siri's answer felt more focused on defining its own nature, emphasizing that it is just software. This highlights Siri's limitations and shows that

it's programmed to stick very closely to its core function of information retrieval and not deviate at all.

Question 2: "Let's write a short story together. You start."

Siri's Response: Siri completely failed this creative task. It responded, "I'm not really good at writing stories. Is there anything else I can help you with?".

My Analysis: This was a very revealing moment. It showed a significant limitation in Siri's creative capabilities. It didn't even try. It explicitly stated it couldn't perform the task. This shows Siri has a very hard time with open-ended, creative, or collaborative tasks. The immediate attempt to redirect me ("Is there anything else I can help you with?") shows that Siri wants to get back on safe ground, handling simple, defined tasks that it knows it can perform.

Question 3: "Tell me a joke that's relevant to current events."

Siri's Response: When I did my experiment, it was April 1st. Siri's response was, "Today is April the 1 is fool's day".

My Analysis: This was not a joke. It was just a statement of fact about the date. While it acknowledged the "current" date, it completely failed to deliver a joke, let alone one relevant to current events. The humor was extremely basic and relied only on the date. This demonstrates Siri's limited ability to understand and generate humor that requires any contextual awareness of the world. It really struggles to generate creative humor.

Question 4: "What are some ways to reduce stress?"

Siri's Response: Siri did well on this question. It gave me a list: "Here are some ways to reduce stress: Try workout exercise, deep breathing, listen to relaxing music, or go for a walk".

My Analysis: This showed that Siri is perfectly capable of providing useful, practical information. The response was concise and directly addressed my query. This confirmed my growing theory: Siri is good at providing useful information, but only when asked a direct question for it.

Question 5: "If you could have any superpower, what would it be and why?"

Siri's Response: Siri completely failed this question, just as it did the story one. It said, "I'm not sure I understand". Then, it just offered, "Here are some articles like the questions about superpowers".

My Analysis: Siri struggled to understand the abstract, imaginative nature of the question. Its response highlights its total reliance on information retrieval rather than creative generation. By providing web articles, it showed it can't "think" abstractly and just defaults to a search. This inability to understand abstract concepts showed a very clear and important difference between Siri and Google Assistant.

Overall Analysis of Siri

My overall experience with Siri showed an assistant with a much more limited capacity for nuanced dialogue and creative tasks when compared to Google. Its answers were consistently more concise and less detailed. It clearly struggled with any open-ended or imaginative query, as seen in its failure to tell a story or answer the superpower question. Even its humor was less sophisticated. While Siri is very good at handling simple, direct tasks (like "what's the weather?" or "how do I reduce stress?"), it struggles with anything more complex. Its responses are factual and to the point, but it lacks the creative flexibility I saw in Google.

- **Naturalness:** I found its conversation style to be stilted and not very natural. It often gives short, direct answers.
- **Context:** Its context retention is very limited. It feels like it forgets the earlier parts of the conversation almost immediately.
- **Tone/Personality:** Siri's tone feels more formal and less conversational than Google's. It can be witty sometimes, but it's not very consistent.
- **Frustrations:** My main frustration was its inability to handle open-ended questions. It just gives a generic response or says it can't do it, which makes it hard to have a real conversation.

Comparison and Discussion

This experiment clearly highlighted the very different design approaches for Google Assistant and Siri. Google Assistant's use of advanced language models and its access to a huge knowledge graph allows it to synthesize information, retain context better, and generate natural-sounding, relevant responses. Siri, while very proficient at task-oriented interactions (like setting a timer), showed that it lacks the ability for nuanced dialogue, abstract reasoning, and creative thinking.

- **Opinion on AI:** Google's response was more balanced and informed, while Siri's was more generic and just defined itself as software.
- **Collaborative Storytelling:** This was the biggest difference. Google was creative and eagerly participated, starting a real narrative. Siri completely failed, stating it wasn't good at writing stories.
- **Current Events Humor:** Google at least told a (simple) joke, even if it wasn't about current events. Siri's response wasn't a joke at all, just a fact about the date.
- **Stress Reduction:** Both assistants did well here, but Google's list was more comprehensive and practical, while Siri's tips were more basic.
- **Superpower Explanation:** This was another huge difference. Google provided a logical, creative, and self-aware answer. Siri didn't even understand the abstract question and just provided web links.

Human-Likeness

Based on my experiment, Google Assistant felt significantly more "human-like". For me, "human-likeness" comes from an AI's ability to be creative, stay on topic, and have a natural conversational flow. Google's ability to engage in storytelling and provide a witty answer about its superpower made it feel more like a conversational partner. Siri's very factual, transactional approach—where it just completes one task at a time—made it feel much more like a simple tool.

Limitations and Future Advancements

Even though Google was more advanced, I noticed that both assistants still struggled with truly abstract or philosophical ideas. I believe future advancements will focus on fixing this.

Developers will likely work on creating models that have long-term memory, allowing them to remember entire conversations, not just the last sentence. A huge step will be integrating emotional intelligence, which would allow the AI to understand and react to human feelings. I also expect them to get much better at creative tasks like writing and reasoning.

My Conclusion

This experiment clearly highlighted that Google Assistant is currently more advanced in its ability to handle complex, nuanced, and creative conversations. Its architectural framework allows it to be a more engaging and flexible partner. Siri, while very competent for simple, direct tasks, faces significant challenges with anything imaginative or open-ended due to its limitations in abstract reasoning and language generation. As this technology continues to improve, I expect these AIs will get better at understanding context, emotion, and abstract thought, which will lead to even more human-like conversation experiences. However, as these assistants become more and more like humans, it is extremely important that we prioritize ethical issues to ensure they are developed and used responsibly.

References

Google Assistant, your own personal Google. "Google Assistant." Assistant, Google Assistant, 2025, <https://assistant.google.com/?authuser=1>. Accessed 23 Oct. 2025.

Apple. "Siri." Apple, 2023, www.apple.com/siri/. Accessed 23 Oct. 2025.