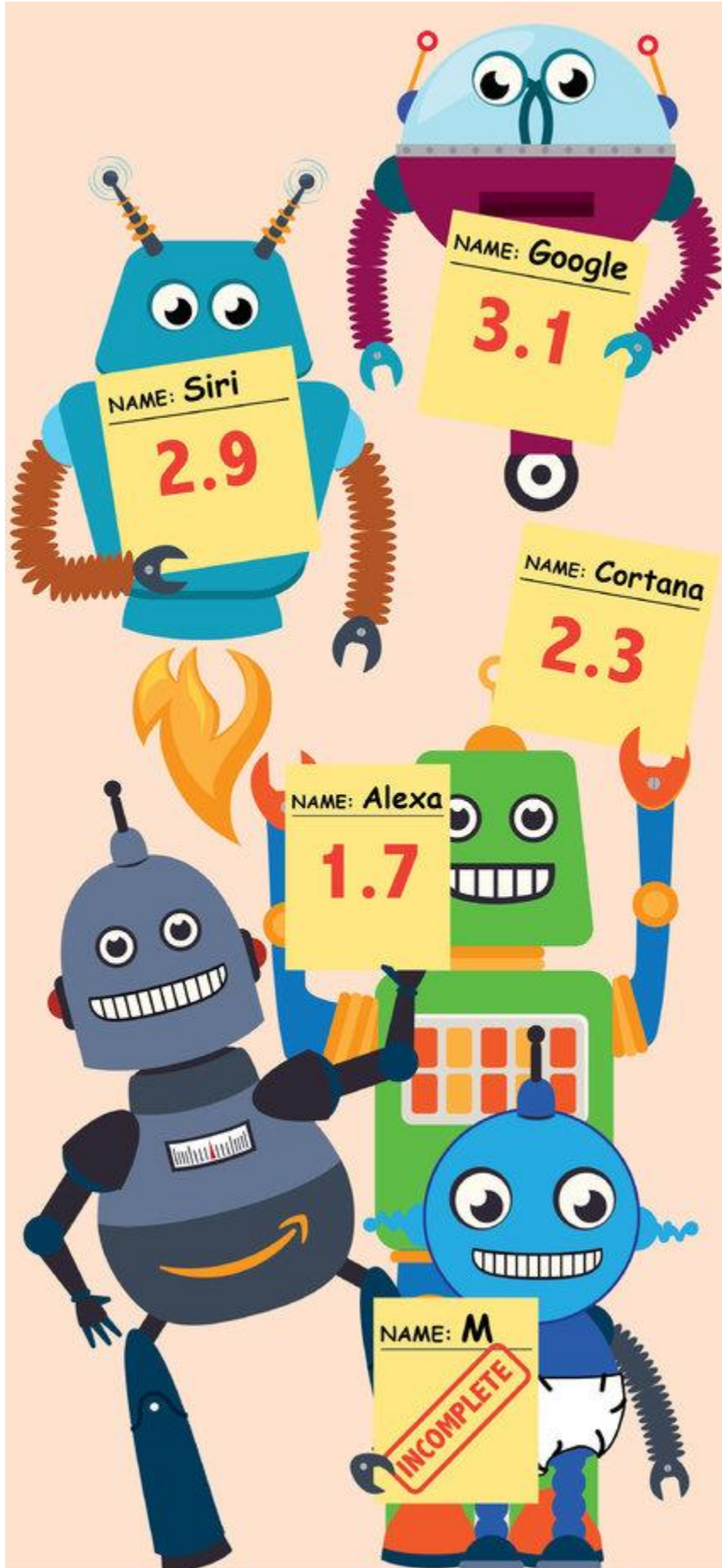


[Personal Tech](#)

Siri, Alexa and Other Virtual Assistants Put to the Test



Credit Minh Uong/The New York Times

WHEN I asked Alexa earlier this week who was playing in the [Super Bowl](#), she responded, somewhat monotonously, “[Super Bowl](#) 49’s winner is New England Patriots.”

“Come on, that’s last year’s Super Bowl,” I said. “Even I can do better than that.”

At the time, I was actually alone in my living room. I was talking to the virtual companion inside [Amazon](#)’s wireless speaker, Echo, which was released last June. Known as Alexa, she has gained raves from Silicon Valley’s tech-obsessed digerati and has become one of the newest members of the virtual assistants club.

All the so-called [Frightful Five](#) tech behemoths — [Apple](#), Microsoft, [Amazon](#), [Facebook](#) and [Google](#), now part of Alphabet — now offer virtual assistants, which handle tedious tasks in response to voice commands or keystrokes, on various devices. [Apple](#)’s Siri is the best known, having been available since 2011, but Microsoft now has Cortana, [Facebook](#) is testing one called M, and [Google](#) builds its voice assistant into its search apps.

These companies are presenting [scorecards of their progress](#) with quarterly earnings reports in the next few weeks, so what better time to hand out report cards to their artificially intelligent assistants? With that in mind, I set up tests for the assistants and graded their abilities to accomplish 16 tasks in categories that most consumers generally enjoy: music, productivity, travel and commuting, dining, entertainment and interests like sports.

In the end, none of the voice assistants earned a report card that would make a strict parent proud. Here’s how they stacked up in terms of grade-point averages out of 4.0.

- Google (Google) — 3.1
- Siri (Apple) — 2.9
- Cortana (Microsoft) — 2.3
- Alexa (Amazon) — 1.7

Apple was the strongest at productivity tasks like calendar appointments and email; Google was the best at travel and commute-related tasks. Alexa excelled at music, and Cortana was mediocre across the board. Facebook was left out of the grading system because the company denied access to M, though I did hang out with her for two hours on a friend’s account. More on that later.

[Tech Fix](#)

A research-driven feature aimed at solving everyday problems related to consumer technology.

Apple said that Siri had “become faster and smarter” and spoke more languages than other assistants. Microsoft said it was “just scratching the surface” on how Cortana could help people. Google said that it wanted smartphones to do more of the heavy lifting, and that users could do a host of things just by speaking to Google. Amazon did not respond to requests for comment.

PLAY VIDEO 4:09

[Continue reading the main story](#) Video

Siri and Google Now, Meet Cortana

Molly Wood tests out mobile virtual assistants, including Microsoft’s Cortana, to see if Siri has met her match.

By Vanessa Perez, Ricky Montalvo, Rebekah Fergusson and Melanie Ruiz on Publish Date April 23, 2014. [Watch in Times Video »](#)

On the productivity front, Siri, summoned by pressing the home button on the iPhone or by saying, “Hey, Siri,” was best able to schedule a meeting with a friend in Hawaii, check what was on my calendar for tomorrow, send an email and read my most recent email aloud. Others could complete only some of those tasks — Google could not read my last email out loud, and Alexa could not compose an email or create a calendar event.

Siri also fared well in music-related tasks, but was bested by Alexa. Both assistants could play the song “Hey” by the Pixies, put on the latest episode of the “Radiolab” podcast and play music in the instrumentals genre. But Alexa, summoned simply by saying “Alexa,” could play a specific music station on Pandora, whereas Siri could only open the Pandora app.

Advertisement

[Continue reading the main story](#)

Google, which builds its voice-controlled assistant into the Google mobile app, achieved the highest marks for completing travel and commuting-related tasks. It responded perfectly to the question “What is the traffic like to 221 Main Street?” by showing me how long it would take to drive there.

When I said, “Take me to the Dogpatch Boulders gym,” it showed me a map and gave voice directions. When I said, “Find me plane tickets to New York next week,” it offered an impressive response: Flights from San Francisco to New York next week start at \$435, and the shortest flight is five hours and 10 minutes long.

On travel and commuting, Cortana could offer solutions for the questions about traffic and directions, but not the one about flights. Siri earned a C-minus in the category: She could not give traffic estimates, and in response to the question about flights to New York, she spat out an unhelpful list of web search results related to traveling to New York. And instead of taking me to a bouldering gym to hone my physique, she took me to a brewery.

Alexa got a D — she could offer traffic estimates for only one fixed location that was set up inside the app, like your office, and she added the task of finding a flight to New York to my to-do list. (I gave Alexa a pass on failing to map me to the gym. It seemed too much to ask from a home audio speaker.)

For food-related tasks, Google and Apple were even. Each of the assistants was able to find a list of nearby Indian restaurants. Only Google's voice assistant could order delivery food, but with an unintuitive process that required naming a specific restaurant that delivers food through one of the apps that Google has teamed up with. Siri was the only one capable of booking a restaurant table.

As for special interests, I asked each assistant two fairly obvious questions: Who won this past Sunday's football games, and who will be [playing at the Super Bowl](#)?

Google, Cortana and Siri loaded scores for Sunday's National Football League games. But only Google and Cortana could say the Carolina Panthers would face the Denver Broncos in the Super Bowl, whereas Siri could only say that the big game would take place on Feb. 7 at [Levi's Stadium in Santa Clara](#), Calif. Alexa, on the other hand, was as clueless about sports as I am: She couldn't answer either question.

That brings me to Facebook's elusive M assistant. The social network denied my request to meet her — it has granted [access](#) to only a small number of testers — so I used a privileged friend's Facebook Messenger account to meet M. According to the company, M is [controlled partly by artificial intelligence](#) and partly by humans; you talk to the assistant by sending messages to M through Facebook's Messenger service, just as you would send messages to a friend.

In my limited time with M, I asked her to handle some of the most mundane tasks: Call the water company to ask about my utility bill, find out what meats are on sale at the local Whole Foods store and research when would be the cheapest time for my editor to fly to Hong Kong (not that I was trying to get rid of my editor).

Advertisement

[Continue reading the main story](#)

Advertisement

[Continue reading the main story](#)

M pondered for a few minutes before answering each question, which made me suspect that a person handled most of the tasks. So I asked M to schedule a photo shoot with a studio owned by a friend of mine. Within minutes, the photo studio's phone rang, and my friend picked up.

"Hi, I'm calling on behalf of my boss," said M, who sounded like a young woman. "He wanted to find out if you guys have the ability for a photo shoot at 2 p.m. tomorrow."

M left a contact phone number with a 650 area code, which includes Menlo Park, Calif., where Facebook has its headquarters.

"I didn't catch your name," my friend at the photo studio said.

"First name is M," the not-so-virtual assistant replied. "Last name is Messenger."

"Is that Greek?" my friend at the photo studio asked. M laughed nervously.

In other words, M probably is more capable than all the virtual assistants, but largely because humans are on the other end of the puppet strings handling tasks that artificial intelligence cannot. That makes me doubt that many consumers will get to meet M, at least in its current state, in which it is free.

"M is still in its very, very early stages," Facebook said in a statement. "We're not yet focused on scaling M to a large number of people."

A version of this article appears in print on January 28, 2016, on page B8 of the New York edition with the headline: Putting Virtual Assistants to the Test. [Order Reprints](#) | [Today's Paper](#) | [Subscribe](#)