In 1967, Marvin Minksy, a founder of the field of artificial intelligence (AI), made a bold prediction: “Within a generation…the problem of creating ‘artificial intelligence’ will be substantially solved.” Assuming that a generation is about 30 years, Minsky was clearly overoptimistic. But now, nearly two generations later, how close are we to the original goal of human-level (or greater) intelligence in machines?

Some leading AI researchers would answer that we are quite close. Earlier this year, deep-learning pioneer and Turing Award winner Geoffrey Hinton told. *Technology Review*, “I have suddenly switched my views on whether these things are going to be more intelligent than us. I think they’re very close to it now and they will be much more intelligent than us in the future.” His fellow Turing Award winner Yoshua Bengio voiced a similar opinion in a recent blog post: “The recent advances suggest that even the future where we know how to build superintelligent AIs (smarter than humans across the board) is closer than most people expected just a year ago.”

These are extraordinary claims that, as the saying goes, require extraordinary evidence. However, it turns out that assessing the intelligence—or more concretely, the general capabilities—of AI systems is fraught with pitfalls. Anyone who has interacted with ChatGPT or other large language models knows that these systems can appear quite intelligent. They converse with us in fluent natural language, and in many cases seem to reason, to make analogies, and to grasp the motivations behind our questions. Despite their well-known unhumanlike failings, it’s hard to escape the impression that behind all that confident and articulate language there must be genuine understanding.