

# IQ LEADERBOARD FOR AI & HUMANS (JUN/2022)

Benchmark testing designed by...



- Respond to a question about a short passage.
- Identify a hypothesis contained within a text excerpt from sources including the *WSJ*.
- Determine either the cause or effect of a premise sentence.
- Answer a question about a context paragraph.
- Predict masked-out words and phrases from text from news sources.
- Identify whenever the truth of one text excerpt follows from another text excerpt.
- Determine how a polysemous word is used.
- Answer questions about the antecedent of ambiguous pronouns. (Winograd, designed to improve on the Turing Test.)

# AI model	Params (B)	Score	Change
1 Google ST-MoE	32	91.2	+1.54%
2 Microsoft Turing NLR	5.4	90.9	+1.21%
3 Baidu ERNIE 3.0	12	90.6	+0.88%
4 Google PaLM	540	90.4	+0.66%
5 Google T5 + Meena	11	90.4	+0.66%
6 Microsoft DeBERTa	1.5	90.3	+0.55%
7 Human baseline	*	89.8	0
8 Google T5	11	89.3	-0.56%

Source: <https://superpluebenchmark.com/leaderboard/>, Alan D. Thompson, June 2022, <https://life architect.ai/iq-testing-ai/>

