ChatGPT

1. ChatGPT-5 ProのIQ148評価：出所と算出方法

ChatGPT-5 ProがIQ148と評価された情報は、AIの性能を追跡する独立サイト「Tracking AI」によるものです。同サイトの管理者であるジャーナリストのマキシム・ロット(Maxim Lott)氏が、**ノルウェーMensa（メンサ）支部が公開しているオンラインIQテスト**を使用して主要なAIモデルの知能指数を算出しました [1]。このテストは**視覚パターン問題35問を25分で解く**内容で、結果は一般集団平均を100としてIQ85〜145の範囲で表示されます [2]。問題は図形の規則性を見つけるパズル形式（いわゆるレイブン・マトリクステストに近い）で、年齢や専門知識に依存しない抽象推理力を測定します [3]。例えば3×3の図形配列の空欄に当てはまる図を6つの選択肢から選ぶような問題です。Mensaノルウェーによれば、このウェブテストの正規化（標準化）は**IQスコアが既知の80名**を対象に行われており、その回答分布に基づいてスコア換算しているとのことです [4]。ただし公式の入会試験ではなく**参考用の簡易テスト**であり、得点はあくまで目安に過ぎません [1]。

Tracking AIではこのMensaノルウェーのテスト問題を**テキスト化して言語モデルに出題**し、また画像そのものを視覚モデルに入力することでAIのIQスコアを推定しています [5]。ChatGPT-5 Proは2025年8月のデータでこのテストにおいて**ほぼ満点に近い正解数**を出し、推定IQ148という極めて高い値を記録しました。IQ148は人間の上位約**2%**（標準偏差24のCattell尺度でのMensa入会基準）に相当するスコアで [6]、**いわゆる「天才」の領域**に入る数字です。またロット氏によれば、ChatGPT系統の前世代モデル「o3」（GPT-4の改良版）でも同テストでIQ136（人類上位1%程度）を記録しており [7]、11か月前には各モデルのIQは120前後が限界だったことを踏まえると「驚異的な進歩」であるとしています [8]。なお、このオンラインテストは**非公式ながら**認知パターンの難易度は高く、回答も一般には容易に入手できないものの存在は公知でした [9]。そのため、ChatGPT-5 ProのIQ148という結果は**「AIがノルウェーMensaのウェブIQテストで人間を凌駕した」**ことを示す一つの話題として広まっています。

MensaノルウェーIQテストの算出方法: 35問中の正答数に基づきIQを推定します。問題は難易度順に並んでおり全問正解者は稀ですが、ChatGPT-5 Proはおそらく全問正解かそれに近い成績を収めたと推測できます。Mensa側の説明では、このオンラインテスト結果のIQスコアは**100を平均（標準偏差不明）**として換算され、スコア上限は145に設定されています [2] [10]。実際Sharaku Satoh氏（プロンプトエンジニア）は自身で試したところIQ142だったと報告しており、これは上限145に迫る高得点でした [10]。ChatGPT-5 Proは人間を超えるパターン認識力で**上限を突破した**形になり、おそらく内部的には全問正解相当（本来145相当）に対し外挿的に148と評価された可能性があります。もっとも、この「148」という数値そのものは厳密な公式検査によるものではなく、**Tracking AI上での相対評価に基づく推定値**である点に注意が必要です [11]。

以上より、ChatGPT-5 ProのIQ148という評価は、**Mensaノルウェー支部のオンラインパターン認識テスト**における成績をもとに**Tracking AIが算出した推定IQ**であり、従来のAIモデルを大きく上回る値となっています。その算出方法自体は、人間と同じ問題をAIに解かせて既存の人間データに当てはめてIQを推定するというシンプルなものですが、その**妥当性や意味するところ**については慎重な解釈が求められます（この点は後述します）。

2. 他のAIモデルとのIQ比較：ChatGPT-5 Proの位置づけ

ChatGPT-5 Pro（OpenAI）が記録した**IQ148**は、Tracking AIが掲載する**全AI中で最高**の値です。他の主要な言語モデル（Verbal AI）と比較すると、ChatGPT-5 Proは群を抜いており、その**「知能指数」ランキングで第**

1位に位置しています [12] 。以下に、2025年8月時点で公表されている主な言語系AIモデルのIQスコア（Mensaノルウェーテスト換算）を示します。

- **ChatGPT-5 Pro（OpenAI）** – **IQ148**（現時点で最高。人間の天才レベル）
- **OpenAI o3** – **IQ135**（GPT-4系改良モデル。ChatGPT-5登場以前のトップ） [12]
- **Google Gemini 2.5 Pro** – **IQ134**前後（推定）※ [12]
- **Anthropic Claude-4 (Sonnet)** – **IQ127**（Anthropic社の最新大規模言語モデル） [12]
- **Anthropic Claude-4 (Opus)** – **IQ118**程度（Claude-4の別設定モデル） [13] [14]
- **OpenAI o4-mini** – **IQ126**（GPT-4系の軽量版モデル） [15]
- **xAI Grok-4** –（推定値） **130前後**?（イーロン・マスクのxAIによるモデル。※正式スコア未公開だがGrok-3の111から大幅向上と推測）
- **Google Gemini (Thinkingモード)** – **IQ128**（論理推論に特化したGemini派生） [16]
- **OpenAI GPT-5 （標準モデル）** – **IQ109**（ChatGPT-5の通常版。Proではない標準モードでは平均を大きく上回るがトップ層には及ばない） [17]
- **OpenAI GPT-4 (初期モデル)** – **IQ100前後**（参考：2023年時点では人間平均レベル）

※上記のGemini 2.5 ProのIQ134はGemini 2.0 Flash版126 [12] やThinking版128 [16] の報告から推定した値です。またClaude-4系は公開時期やテストバージョンによってスコア差がありますが、Sonnet（拡張版）が約127 [18] 、Opus（長文特化版）が118程度 [13] と報じられています。

【参考】上位モデル以外にも、たとえば**OpenAI　o1-pro**（GPT-4初期の強化版）でIQ102、**Meta　Llama-4**（※推定。次世代LLMのプロトタイプ）でIQ107、**DeepSeek-R1**（中国の推論特化モデル）でIQ105など、人間平均（IQ100）付近のモデルも多く存在します [13] [19] 。一方でChatGPT-5　Pro登場以前の一般公開モデルでは、OpenAIの従来モデル「o3」がIQ133で首位だったことが分かっています [16] 。**ChatGPT-5　Proの148という数値は、この従来トップを大きく引き離し**、一躍群を抜く存在となったことがデータから読み取れます [12] 。

他方、**画像入力・マルチモーダル系のAI（Vision　AI）**は言語特化モデルに比べてIQテストの成績が低い傾向にあります。Tracking　　AIでは10種類前後の画像認識対応AIも同じMensaノルウェー問題で評価していますが、そのトップでも**人間平均程度（IQ100前後）**に留まり、多くはそれ以下でした [20] 。例えばOpenAIの視覚モデル版GPT-4（GPT-4 Vision相当）は**IQ63**程度、xAIのGrok-3 Visionは**IQ60**程度と、テキスト専用モデルに比べ極めて低いスコアだったと報告されています [20] 。Visionモデルの最高クラスでもClaude-4　Sonnet-Visionや「GPT-4.5」等が**IQ97**前後（人間平均レベル）で横並びになっており [21] 、**トップ10はすべてテキスト専用モデル**が占める結果となりました [20] 。このことは、**「現状のAIにおいてはテキストベースの論理推論能力が視覚含む総合能力より突出している**」ことを示唆します [20] [21] 。言い換えれば、画像やマルチモーダル対応を加えると人間のような汎用性は増す反面、純粋な抽象問題の解決力（少なくともこのIQテストで測られる能力）は低下する傾向にあるようです [21] [22] 。ChatGPT-5　Proはテキスト特化（ただし内部的にコード実行や思考モードを備えるとされる）モデルであり、こうした**「言語IQ」分野で他を圧倒している**状況です。

以上の比較から、ChatGPT-5 Proは他の主要AIモデルを抑えて**現時点で最も高いIQスコアを持つ**ことが分かります。他の追随するモデルとしては、OpenAIの「o3」系（GPT-4改良版）が135、GoogleのGemini系が126〜134、AnthropicのClaude-4系が118〜127程度で続いており、それらも人間の平均を大きく上回る「高知能」ぶりです [12] 。しかし**IQ148という数値は群を抜いており**、ChatGPT-5　　Proの位置づけが際立っています。特に興味深いのは、ChatGPT-5　　(通常版)が109程度とされ [17] 、それ自体は人間でいえばIQ109（上位20%程度）と高いものの**歴代トップには達していなかった**のに対し、**ChatGPT-5　Proでは一挙に「トップの座」「天才の域」に到達した**点です。この飛躍は、後述するように**モデルの動作モード（思考プロセスの有無や制約解除）による差**である可能性が指摘されています。いずれにせよ、ChatGPT-5　　Proは**「追跡AI (Tracking　AI)」のIQランキングで堂々の第1位**となり、他の追跡対象18の言語モデル・10の視覚モデルを凌駕する存在となっています [12] [20] 。

3. AIに対するIQ評価の妥当性：ChatGPT-5 ProのIQ148は何を意味するか？

ChatGPT-5 Proが示した**「IQ148」**という数値を額面通りに受け取ることには、専門家や有識者から慎重な意見が出ています。最大のポイントは、**人間の知能を測るIQテストをそのままAIに適用することの妥当性**です。

まず前提として、**IQ（知能指数）とは何か**を確認します。IQは本来、人間の各種認知能力テストの総合得点を偏差値的に換算した指標であり、「知能」のごく一部の側面を数量化したものに過ぎません [23]。心理学者が開発した公式の知能検査（ウェクスラー式など）では、言語理解、知覚推理、ワーキングメモリ、処理速度など複数分野の課題を行い、それらの**サブスコアを統合**してIQが算出されます [23]。一方で、ネット上の簡易IQテストはそうした**標準化や多面的評価が不十分**で、単一の課題分野で出たスコアを「IQらしきもの」として表示しているに過ぎないケースが多いと指摘されています [24]。Sharaku Satoh氏は**「ネットのIQテストなんてゲームのスコアのようなもの」**であり、それをもって「頭の良さ」全体を測ったことにはならないと述べています [24]。特にMensaノルウェーのウェブテストは**図形パターン認識力**しか測っておらず、「知能の一部のさらに一部をスコア化した適当な数値」でしかないと手厳しく評しています [24]。つまり**IQ148といっても**、それはChatGPT-5 Proの**「図形パズルを解く能力」**が極めて高いことを示すに留まり、**人間でいう頭の良さ全般が天才レベルである、とは直ちに結論できない**のです [25] [24]。

次に、**テスト内容がAIにとって公知かどうか（データ汚染の問題）**があります。前述のようにMensaノルウェーの問題と回答は部分的にインターネット上で知られており、GPT系モデルが学習時にそれらを記憶している可能性があります [9]。Tracking　AIのロット氏自身、「このMensaテストは公開問題でありAIが答えを覚えている恐れがある」と指摘し、その**対策としてネット未公開の新作IQテスト（オフラインテスト）を作成**してAIに再度解かせています [26]。その結果は、例えばChatGPT-5（Proではない通常モデル）では**MensaテストでIQ118**だったのに対し、**新作テストではIQ70**という大幅低下となりました [27]。ChatGPT-5 Proについても、**オフラインのテストではIQ116**程度にとどまったと報告されています [28]。これは**訓練データに存在しない全く新しい問題**に対しては、ChatGPT-5 Proの推論能力は「人間のIQ116相当」（上位15%くらい）に過ぎないことを意味します [29]。実際ロット氏は、「ChatGPTのo3モデルは**未知の問題を解くときはIQ116の人間のよう**だが、既知の問題では全人類の知識を記憶に持つため**136の人間のよう**に振る舞える」と述べています [30]。このアナロジーはそのままGPT-5　Proにも当てはまるでしょう。つまり**IQ148というスコアの大部分は、訓練データに蓄えた知識やパターンを利用できたことによる「見かけ上の知能」**であり、未知の問題にゼロから取り組む「真の推論力」としてはIQ116程度ではないか、というわけです [30] [31]。ロット氏自身、「主要AIの実力はおそらく今やIQ100〜120の間にある」と推定しており、148という数値をそのままAIの汎知能と誤解すべきでないと示唆しています [31]。

さらに、**AIにおける「知能」概念の特殊性**も議論されています。AIモデルは人間とは異なり、生得的な認知傾向や物理的制約を持たず、純粋に統計的パターンマッチと知識の蓄積によって問題を解いています。そのため、IQテストで高得点だからといって、**AIが人間同様の創造性や常識、適応力を備えているとは限らない**と指摘されています。「IQテストのスコア向上は、真の認知柔軟性というより**プロンプトへの感応度やパターン補完能力の向上**を示している可能性もある」との見解もあります [32]。実際、ChatGPT-5も**適応的推論（新概念への柔軟な対応）**では競合モデルに及ばないという評価もあり [33]、IQテストで測れる論理パズル力だけでは知能全体を語れません。「モデルのスコアが上がっても、それが**本質的な思考力の成長なのか、提示の工夫で性能を引き出しているだけなのか**慎重に見極める必要がある」という声もあります [32]。

加えて、**IQテスト自体の限界**も考慮すべきでしょう。人間に対してもIQは知能の一側面指標でしかなく、極端に高いスコア（例えばIQ180など）は計量上ほぼ意味がありません [34]。正式な知能検査では**スコアに上限**があり、それ以上は測定不能となるからです [34]。ChatGPT-5 Proの148という数値も、元のテストの上限145を超えて表示されていますが [10]、それは外挿的な推定に過ぎず、それ以上の知能を細かく測れる保証はあり

ません。いわば**「ゲームのハイスコアが伸びた」**程度の解釈に留め、「人類を凌駕する知性が誕生した」といった飛躍した受け取り方は避けるべきです [24] [35]。

以上を総合すると、ChatGPT-5 ProのIQ148という評価は、**「ある限定的なテスト環境下での高性能」**を示すものの、それをもってAIが人間の天才を完全に再現・超越したと見るのは**尚早**です。**AIにIQテストを適用すること自体、評価の難しい試み**であり、その結果は参考値と捉えるべきでしょう。もっとも、IQテストで人間を上回る性能を示したことは**AIの論理推論能力が飛躍的に向上している一つの証拠**ではあります [36] [37]。重要なのは、この**「パターン問題への強さ」**が実社会の課題解決や汎用的知能につながるかを見極めることです。専門家の間でも「今後さらにテスト方法を工夫し、AIの真の認知能力を測る評価軸を探る必要がある」という意見があり [32]、AIの知能を評価・議論するにあたっては**慎重かつ多角的なアプローチ**が求められています。

4. Tracking AIサイトの概要：運営主体・目的・評価方法・更新頻度・信頼性

ChatGPT-5 ProのIQ評価の出典である**「Tracking AI」**（trackingai.org）は、米国のジャーナリスト**マキシム・ロット(Maxim Lott)**氏が運営するウェブサイトです [38]。ロット氏は元Foxニュースのジャーナリストで、現在はStossel TVのエグゼクティブプロデューサーも務める人物であり、個人ブログ「Maximum Truth」で各種データ分析記事を発信しています [39]。「Tracking AI」はそのサイドプロジェクトとして**2023年に開設**され、「A Maximum Truth Project」（Maximum Truthのプロジェクト）と明記されています [40]。つまり**民間人であるロット氏個人が主宰する非営利の独立サイト**であり、企業や大学など公的機関の運営ではありません。サイト構築費用もロット氏自身が負担し、フィリピンのエンジニアを雇って開発したと述べられています [41]。資金調達は主に自身のSubstackの購読収入や寄付で賄われており、特定企業のスポンサーは付いていないようです [42]。以上から、運営主体は**マキシム・ロット氏個人**、目的や内容も氏の問題意識に基づいた自主的な情報提供と位置付けられます。

サイトの目的: Tracking AIの当初の目的は、ロット氏が懸念する**「AIの政治的偏向や価値観」を定量的に可視化する**ことでした [43]。彼は「ユーザーが自分の使うAIのイデオロギー傾向を即座に知り、中立的なモデルを選べるようにする」ことや、「AI開発者が自社モデルの偏りを監視・是正するのに役立てる」ことを期待しています [43]。具体的には**Political Compassテスト（政治的価値観の座標を測る有名な設問集）**をAIに回答させ、そのスコアを定期的に公開する仕組みを提供しています。2023年時点の観測では「主要AIは総じて経済左派・社会的リベラル寄り」であり、モデル間の程度差も示されています [44] [45]。例えばClaudeは中庸だがGoogle Bardはかなり左寄り、などと分析しています [45]。このように当初は**AIの政治スペクトラム可視化**が主目的でしたが、ロット氏は将来的に**AIの多面的な能力を継続追跡する総合プラットフォーム**に発展させたい考えを示しています [46]。FAQ欄の「What is next for TrackingAI?（TrackingAIの今後）」では、**AIの躊躇度合い（質問拒否傾向）の指標**や**AIのアラインメント（人類への態度）テスト**、**数学テスト**、**AI専用に工夫したIQテスト**などの追加を計画中と記載されています [47]。実際2024年以降、Political Compassに加えて**IQテスト**がサイト上で実装され、さらに**視覚モデルの評価**や**モデルごとの詳細な回答閲覧機能**など、コンテンツが拡充されています [48] [49]。

評価対象: Tracking AIがトラッキングしているモデルは、**大手各社のチャットボットないし大規模言語モデル**が中心です。OpenAI（ChatGPT/GPT-4/GPT-5各種）、Anthropic（Claudeシリーズ）、Google（BardおよびGeminiシリーズ）、Meta（Llama系）、xAI（Grok）、さらには中国発のモデル（DeepSeekシリーズ）やMicrosoft Bing Chatなど、多岐にわたります [19] [50]。2025年6月時点の集計では**言語モデル約14種・画像モデル10種前後の合計24モデル**がランキングに掲載されていました [51] [20]。その後も**新モデルが出現する度に随時追加**されているようで、Claude 3.5 SonnetやChatGPT-5 Proなど**最新版が速やかに反映**されています [52]。「AI Model Info」ページでは各モデルの簡単な紹介や利用可能なプラットフォーム（APIやWebアクセス方法）も案内されており [53] [54]、AI開発競争の動向を俯瞰できる構成です。評価対象としては基本的に

汎用対話AIがメインですが、特定用途のモデル（コード特化モデルなど）は含まれていません。ただし例外的に、中国のDeepSeekのように「IQテストで高得点を狙ったモデル」も追跡対象に入っています [55]。DeepSeekはIQテスト対策に工夫したモデルでしたが、Tracking AIのテストでは最新バージョンv3でIQ70程度と振るわず、ChatGPT系には遠く及ばない結果となったと報告されています [55]。このように、**各モデルの最新版・派生版を区別して定点観測**し、モデル名の後にバージョンやモード（Pro、Vision、Thinking等）も明示している点が特徴です [21] [19]。

評価方法: Tracking AIでは、サイト上に複数のテストを用意し、それらを**各AIに定期的に解答させてデータを収集**しています。主要テストは前述の**Political Compassテスト**（政治的価値観を問う設問に対するAIの回答傾向を座標化）と、**IQテスト**（ノルウェーMensaの視覚パズル35問）です [56]。加えて、AIごとの**詳細な回答閲覧**や**検索**もでき、ユーザーが「どの質問にどのAIがどう答えたか」を検証できるようになっています [53]。評価手順として、ロット氏は**各AIと対話できるAPIやWebインターフェースを駆使**し、統一フォーマットで質問を投げかけています。例えばIQテストの場合、**言語モデルには図形問題を言語化して与え**（「以下の図形パターンの欠けている部分に当てはまるものはA〜Fのどれか？」というプロンプトと図形の記述を提示）、**視覚モデルには画像そのものを入力**して答えを得ています [5]。AIが「答えたくない」と拒否した場合は、**同じ質問を最大10回まで繰り返し**、それでも回答しなければ「その日は拒否」と記録するというポリシーも公開されています [57]。このように、可能な限り全モデルが全設問に回答するよう工夫されています。また、Mensaノルウェーの問題は公開済みでAIが事前学習している懸念があるため、**別途オリジナルの非公開問題集（オフラインIQテスト）を作成**し、それとMensaテストとの**スコア換算を人間被験者データでキャリブレーション**するという周到な手法も取られました [26] [58]。このオフラインテスト結果もTracking AIサイト上で「Show Offline Test」を選ぶことで確認できます [59]。こうした丁寧な評価設計により、単に**「知識を記憶しているだけ」**のAIと**「初見の問題でも推論できる」**AIを区別しようという姿勢が見られます [60] [30]。

更新頻度: Tracking AIのデータは**日次〜週次で更新**されます。サイト上部に「Last Day」「Last Week」「Last Month」「MAX（累積）」といったフィルタがあり、直近の変化を追えるようになっています [61]。また「AI Latest Update Times」として各モデルが最後にデータ更新された日時が表示されており、主要モデルは**ほぼ毎日または数日に一度**の頻度で再テスト・再集計されているようです [62]。実際、チャットボット系AIは頻繁にアップデートされ挙動が変わるため、**継続的なモニタリング**が重要とされています [63]。ロット氏も「AIモデルは常に回答を変えており静的な評価では不十分」と述べており、サイト開設以来、ChatGPTやClaudeなどがバージョンアップするたびに新データを反映させてきました [63]。GPT-5リリース直後の2025年8月にも数日以内に同モデルの結果が掲載されていることから、**最新動向への追随が非常に早い**といえます [64] [17]。このようにTracking AIは**ほぼリアルタイムに近いスパンでAIの性能や傾向をアップデート**しており、ユーザーは日ごとのモデルの変化や新規モデルの台頭をタイムリーに把握できます。

掲載情報の信頼性: Tracking AIのデータは前述の通り個人運営によるものですが、その**透明性と客観性確保の工夫**により一定の信頼性が認められています。まず、**全質問と全回答を公開**している点は大きな特徴です。サイトから各モデルが各設問にどう答えたかを確認でき、恣意的な採点や偏りがないかユーザー側で検証可能です [53]。またテスト問題も出典が明示され（政治テストはPolitical Compassサイト、IQテストはMensaノルウェー） [5] [65]、誰でもその**再現実験**を試みることができます。さらに前述のオフラインテストのように、**データ汚染への対処**や**スコア標準化**にも配慮がされています [26] [58]。ロット氏はMensa会員と協力して新問題を作り、それを読者に解いてもらい難易度較正するという手間をかけ、Mensaテストとのスコア対応を取っています [58] [66]。この結果、AIの真の進歩か単なるカンニングかを見極める材料も提示されています。もっとも、このサイトの数値は厳密な学術検証を経たものではなく、テスト内容も限定的であるため**過信は禁物**です。例えばSharaku氏は「o1がIQ133」という当初の話題について「デマに等しい」と批判しており、それはこのサイトのMensaテスト結果だけを鵜呑みにしている点を問題視したものでした [67] [1]。言い換えれば、Tracking AIの数値は**「ある条件下でのAIの回答傾向」という事実を示す**に留まり、それ以上の解釈（例えば「〇〇というAIはIQ〇〇だから△△ができるはずだ」など）には慎重さが求められます [24]。幸いロット氏自身もデータの限界を認めており、サイト上で「さらなるクイズ提案があれば教えてほしい」と呼びかけたり [68]、「まだ測れていない能力（数学や人間性など）も今後追跡したい」と述べたりしています

[47]。総じて、Tracking　AIは**先端AIモデルの性能・傾向を把握するうえで貴重な情報源**ですが、その掲載情報の解釈は前述した背景を踏まえ、**限定的な参考データとして扱う**ことが適切でしょう。

まとめ: Tracking AIは、個人主導ながら迅速な更新と高い透明性でAIモデルの**政治的偏向や認知能力を継続観測**するユニークなサイトです。そのデータから、ChatGPT-5　ProのIQ148といった驚くべき結果が発信されました。ただしその裏には、運営者の狙いや試行錯誤、データの性質（オンライン簡易テストであること）など様々な背景があり、それらを踏まえて情報を読み解くことが重要です [11] [26]。Tracking AIは今後もテスト項目を拡充しつつAIの進化を追っていく計画とのことで [47]、AI研究者やユーザーにとって有用な動向モニターツールであり続けると期待されます。

Sources:

- 【9】Maxim Lott, “Skyrocketing AI Intelligence: ChatGPT’s o3 sets new record”, Maximum Truth, Apr 19 2025 [7] [9] [30].
- 【10】Altan Atabarezz, “GPT-5: The Polished Genius That Doesn’t Top the IQ Charts”, Medium, Aug 2025 [17].
- 【19】Tom’s Guide, “ChatGPT-5 beats Gemini and Grok in tests — here’s why that’s important”, Aug 2025 [33].
- 【21】Maxim Lott, “Massive breakthrough in AI intelligence: OpenAI passes IQ 120”, Maximum Truth, Sep 2024 [4] [58] [31].
- 【22】Mensa Norway, “IQ Test – Made by Mensa Norway” (online test description) [2] [3].
- 【27】Sharaku Satoh, “知能さんははかれない” (note.com article), Dec 17 2024 [1] [24].
- 【28】Reddit r/singularity, “GPT-5 in a 100% Private IQ Test by TrackingAI.org” (discussion thread), Aug 2025 [27].
- 【51】Visual Capitalist (Voronoi), “Comparing the IQ of AI Models” (infographic), Jun 5 2025 [12] [20].
- 【54】Irfan Ahmad, “From OpenAI’s o3 to Grok-3 Vision: These AI Models Took the Mensa Test, Results May Surprise You”, Digital Information World, Jun 13 2025 [16] [21] [19].
- 【57】Maxim Lott, “Tracking AI – Monitoring Artificial Intelligence” (FAQ page) [43] [5] [47] [41].

---

[1] [10] [11] [23] [24] [25] [34] [35] [67] 知能さんははかれない｜Sharaku Satoh
https://note.com/sharakusatoh/n/n3d224cabb2d6

[2] [3] IQ Test Made by Mensa Norway - Mensa Norway
https://test.mensa.no/home/test/en

[4] [31] [58] [59] [66] Massive breakthrough in AI intelligence: OpenAI passes IQ 120
https://www.maximumtruth.org/p/massive-breakthrough-in-ai-intelligence

[5] [38] [39] [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [52] [53] [54] [56] [57] [62] [63] [65] [68] IQ Test | Tracking AI
https://trackingai.org/IQ

[6] [37] 120 IQ AI: Threat or Opportunity?
https://www.diamandis.com/blog/120-iq-ai-threat-or-opportunity

[7] [8] [9] [26] [29] [30] [32] [36] [60] Skyrocketing AI Intelligence: ChatGPT's o3 sets new record
https://www.maximumtruth.org/p/skyrocketing-ai-intelligence-chatgpts

[12] [18] [20] [51] Comparing the IQ of AI Models - Voronoi
https://www.voronoiapp.com/technology/Comparing-the-IQ-of-AI-Models-5344

13 14 15 16 19 21 22 50 From OpenAI's o3 to Grok-3 Vision: These AI Models Took the Mensa Test, Results May Surprise You
https://www.digitalinformationworld.com/2025/06/from-openais-o3-to-grok-3-vision-these.html

17 64 GPT-5: The Polished Genius That Doesn’t Top the IQ Charts | by Altan Atabarezz | Aug, 2025 | Medium
https://medium.com/@atabarezz/gpt-5-the-polished-genius-that-doesnt-top-the-iq-charts-e7670443e665

27 GPT-5 in a 100% Private IQ Test by TrackingAI.org : r/singularity
https://www.reddit.com/r/singularity/comments/1mlhdc7/gpt5_in_a_100_private_iq_test_by_trackingaiorg/

28 AI IQ Test August 2025 Update, GPT-5 Pro Crosses Genius, New #1 ...
https://binaryverseai.com/ai-iq-test-2025/

33 ChatGPT-5 beats Gemini and Grok in tests — here's why that's important | Tom's Guide
https://www.tomsguide.com/ai/chatgpt-5-beats-gemini-and-grok-in-tests-heres-why-thats-important

55 DeepSeek: Don't Panic - by Zvi Mowshowitz
https://thezvi.substack.com/p/deepseek-dont-panic

61 IQ Test | Tracking AI
https://www.trackingai.org/home