

Beyond Traditional AI IQ Metrics: Metacognition and Reflexive Benchmarking for LLMs, AGI, and ASI

Raymond J. Uzwyshyn, PhD MBA MLIS

Mississippi State University, Mississippi, US,
ruzwyshyn@gmail.com, <https://rayuzwyshyn.net>

ABSTRACT

This research introduces an innovative method for the evaluation of artificial intelligence large language models (LLMs), particularly those targeting Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI). Traditional benchmarks in AI assessment have predominantly relied on IQ tests and performance on standardized examinations but this will need to be rethought as current models are approaching top deciles of the majority of standardized tests across wider fields of disciplines and knowledge areas. We propose a paradigm shift, emphasizing a synthetic comparative analysis and metrics focusing on feedback, reflexivity, and metacognition — the ability of AI models understand their own and parallel processes through a more focused test. Three leading large language models — ChatGPT4 (OpenAI), Gemini Ultra (Google), and Claude 3 (Opus, Anthropic) — are evaluated in composing a single Ph.D.-level academic essay from humanistic psychology. The various new models' long context windows are then again utilized for critical analysis, showcasing their ability to reflect on the aggregated results. This offers insight into both the original test, (the essay), and a next level test of the models' higher-level metacognitive abilities. The analysis employs a multidimensional framework including quantitative (word count, complexity), qualitative (analytical depth, stylistic eloquence), and overlap (thematic, conceptual) analysis among competing models' responses. Findings reveal AI's advanced engagement with complex concepts, including sophisticated metacognitive and reflexive capabilities. By moving beyond traditional metrics and embracing emergent possibilities of reflective and metacognitive evaluation, we can more accurately gauge AI's progress and deepen our understanding of AI's capacities, helping to pave the way for further advancements towards AGI and ASI.

Keywords: Artificial Intelligence, AGI, ASI, Benchmarking, Large Language Models, IQ, Metrics, ChatGPT4, Claude 3 Opus, Gemini Ultra, Gemini Advanced. Reflection, Metacognition, Assessment.



Figure 1 GPT4 Dalle 3 Maslow's Self Actualization.



Figure 2 Gemini Ultra Imagen Maslow's Self Actualization

INTRODUCTION

Many AI human Intelligence benchmarking tests have been given to current large language models to compare against various largely North American standardized IQ and high school, university, and graduate school entrance tests and professional exams. This work has shown substantial progress in AI model intelligence development from making correct next-word predictions (GPT1) to completing standardized tests (GPT3) to surpassing various disciplinary average human capabilities (GPT3.5) to approaching expert-level human intelligence and top deciles of various tests (GPT4, See Appendix I). This study navigates beyond these traditional metrics, for the next level of models venturing into the less charted territories of metacognition and reflexivity—critical faculties that AI will need to emulate to edge closer to AGI and ASI. Metacognition refers to an AI's understanding of its own thought processes. Reflexivity is the ability of a model to analyze and modify its internal representations. This study proposes a new benchmarking test that targets these skills by analyzing an AI large language model's ability to produce a Ph.D.-level academic essay. The language models are then prompted to analyze their work from a metacognitive and reflexive standpoint.

This AI benchmarking test first takes a more lateral qualitative specific domain perspective and focus, asking a simple essay composition prompt surrounding a quote by the well-known psychologist Abraham Maslow to comparatively assess current leading Artificial Intelligence Large Language Models ChatGPT4 (Open AI), Gemini Ultra (Google) and Anthropic's Claude3 (Opus) on a well-known social sciences' disciplinary domain topic (humanistic psychology). The language models evaluated are currently estimated at the top of their class and contain more than a 1 trillion parameter count (synaptic connection analogue). The essay 'prompt' described below surrounds constructing a graduate-level essay regarding the brief quote above and Maslow's concept of self-actualization. The quote reads:

Musicians must make music, artists must paint, and poets must write if they are to be ultimately at peace with themselves. What human beings can be, they must be. They must be true to their own nature. This need we may call self-actualization."

(Abraham Maslow, Motivation and Personality, 1954)

The original prompt is given below followed by essay results from all three models and associated images generated. All three models are given the exact same prompt without further prompting to benchmark results. Following this essay results are aggregated and fed back to all three models to analyze comparatively and reflectively their own and competitor's essays from metacognitive and reflexive standpoints and comment from perspectives of comparative Quantitative, Qualitative, and Overlap analysis. These results are also then placed below followed by brief conclusions.

This study moves beyond traditional AI benchmarking metrics to explore the critical faculties of metacognition, reflection, and reflexivity - capabilities essential for AI systems to develop genuine intelligence and autonomy. As Marvin Minsky puts this in his 1986 Society of Mind, "the principal activities of brains are making changes in themselves", highlighting the importance of self-awareness and adaptation to intelligence (Minsky, p. 288, 1986) Evidence of developing metacognition and reflexivity shows greater self-understanding and towards better modeling of the world. By focusing attention and development on these higher-order cognitive reflective processes, we may gain deeper insights into the current state and future potential of leading AI language models.

Original Essay Prompt

Write a stylistically eloquent academic essay suitable for Ph.D.'s in psychoanalysis and therapy on this quotation by Maslow and the concept of Self Actualization:

Musicians must make music, artists must paint, and poets must write if they are to be ultimately at peace with themselves. What human beings can be, they must be. They must be true to their own nature. This need we may call self-actualization." (Abraham Maslow, 54).

Situate Maslow and his concept of self-actualization historically within psychoanalysis and our wider culture and biographically. Look at other similar psychoanalytic theories of 'self-actualization' both from Maslow's school and other schools of psychoanalysis comparatively. Quote sources for Maslow and any other sources used. Try to do your best work in terms of eloquence and penetrate deeply into this concept psychoanalytically. This essay will be used as a comparative benchmark against other language models' work. Also, produce two suitable images to accompany this essay that penetrate deeply into the concept of self-actualization. Make these aesthetically beautiful, profound and visualize the concept of self-actualization.

Comparative Analysis Prompts

These two prompts are given to each of the three models with all three essay aggregated results

a) Conduct a comparative quantitative, qualitative, and overlap analysis of these three language models in response to the included prompt. Write a brief paragraph overview of conclusions and produce a comparative table or chart with suitable characteristics to more easily compare models from quantitative, qualitative, and overlap comparative analysis.

b) What level are each of the models writing at? Specify in detail with a well-formed paragraph.

AI Model Essay Results

ChatGPT4 (Open AI) Results

The Pursuit of Self-Actualization: A Psychoanalytic Perspective

Abraham Maslow's profound assertion that "Musicians must make music, artists must paint, poets must write if they are to be ultimately at peace with themselves. What human beings can be, they must be. They must be true to their own nature. This need we may call self-actualization" encapsulates a quintessential aspect of human existence, underpinning the psychoanalytic journey towards understanding the self. This essay endeavors to dissect the concept of self-actualization, situating Maslow's contributions within the broader psychoanalytic discourse, exploring its historical context, biographical relevance, and its intersections with other psychoanalytic theories.

HISTORICAL CONTEXT AND MASLOW'S BIOGRAPHY

Abraham Maslow, born in 1908 in Brooklyn, New York, to immigrant parents, emerged as a pivotal figure in psychology during the mid-20th century. His academic journey, shaped by the intellectual milieu of his time, led him to challenge the dominant paradigms of psychoanalysis and behaviorism. Maslow's theory of self-actualization was developed as part of his broader hierarchy of needs, a model that posits a progression toward psychological well-being, culminating in the realization of one's potential.

Maslow's conceptualization of self-actualization was, in part, a response to the existential questions of his era, influenced by the horrors of World War II and the existentialist philosophy that sought to find meaning in human existence. He positioned self-actualization at the apex of his hierarchy of needs, proposing that achieving this state requires fulfilling more basic needs such as safety, belongingness, and esteem.

MASLOW AND PSYCHOANALYTIC THEORY

While Maslow is often associated with humanistic psychology, his work intersects with psychoanalytic thought through its focus on the inner life and potential of the individual. Unlike Freud, who emphasized the unconscious, conflict, and sexual and aggressive drives, Maslow focused on the inherent goodness and potential for growth within individuals. However, both approaches concern themselves with the intricacies of the human psyche and the processes that contribute to a person's development and psychological health.

Maslow's theory diverges from traditional psychoanalytic theory by suggesting a more optimistic view of human nature and potential. Self-actualization, in Maslow's view, is not just a possibility but a necessity for psychological health, representing a universal drive towards fulfillment and the realization of one's capabilities and talents.

COMPARATIVE PSYCHOANALYTIC THEORIES OF SELF-ACTUALIZATION

Comparing Maslow's concept of self-actualization with other psychoanalytic theories reveals both consonance and divergence. Carl Rogers, a contemporary of Maslow and a founder of humanistic psychology, proposed a similar concept with his notion of the "fully functioning person." Rogers emphasized the importance of self-realization through growth, change, and fulfilling one's potential, aligning closely with Maslow's theory.

Conversely, classical psychoanalytic theorists like Freud and Jung provided different lenses through which to view human development. Freud's theory of psychosexual development and Jung's process of individuation, while distinct from Maslow's self-actualization, similarly address the unfolding of the individual's inherent nature over time. However, they place greater emphasis on the role of the unconscious, internal conflicts, and the influence of early life experiences.

SELF-ACTUALIZATION IN CONTEMPORARY CULTURE

Maslow's concept of self-actualization retains profound relevance in contemporary culture, serving as a guiding principle in education, therapy, and personal development. It resonates with the modern emphasis on individuality, self-expression, and the pursuit of personal goals. In the realm of psychotherapy, self-actualization provides a framework for understanding client goals and fostering an environment that encourages growth and self-discovery.

CONCLUSION

Abraham Maslow's theory of self-actualization offers a compelling vision of human potential, grounded in the belief that fulfillment and peace come from living in accordance with one's true nature. Situating Maslow within the psychoanalytic tradition reveals both his departures from and contributions to our understanding of the self. By comparing his theory with other psychoanalytic perspectives, we gain a richer understanding of the diverse ways in which human development can be conceptualized and achieved. Maslow's legacy, therefore, lies in his optimistic portrayal of human potential, a concept that continues to inspire and challenge the psychoanalytic community and beyond.

Gemini Ultra (Google) Results

Self-Actualization

Maslow's Imperative and the Psychoanalytic Quest for Wholeness

Abraham Maslow's declaration that "musicians must make music, artists must paint, poets must write if they are to be ultimately at peace with themselves" strikes at the core of the human experience. It evokes a yearning for self-fulfillment and authenticity. Maslow termed this drive "self-actualization," a concept that would reshape not only the field of psychoanalysis but also profoundly influence Western culture. To fully appreciate the power of this statement and the depth of the concept, it's vital to examine it within both a historical and a biographical context while considering its relationship to other psychoanalytic theories of personal development.

Maslow's Place in Psychoanalytic History

Abraham Maslow emerged within the context of a psychoanalytic landscape dominated by the legacy of Sigmund Freud. Freud's work revolutionized the understanding of the human mind, positing an unconscious realm of repressed desires and conflicts as the primary driver of behavior. The goal of Freudian psychoanalysis, therefore, was the liberation of these repressed urges, primarily through uncovering their roots in childhood experiences.

Maslow, while acknowledging the importance of Freud's contributions, diverged significantly from this perspective. He belonged to a movement known as humanistic psychology, which emphasized the individual's potential for growth and self-direction. Unlike Freud's focus on the pathological, Maslow shifted the conversation towards health and well-being. In doing so, he offered a counterpoint to the deterministic view implicit in Freud's theory, providing an optimistic and empowering vision of human possibility.

Self-Actualization: Maslow's Vision

Maslow proposed his now-iconic hierarchy of needs, a pyramid model where basic physiological needs form the foundation. Safety, love and belonging, and esteem follow, with self-actualization representing the apex. Crucially, he believed that self-actualization couldn't be pursued until these lower needs were met. For Maslow, self-actualization was an ongoing process of becoming one's fullest self, characterized by qualities such as:

- **Acceptance:** Embracing oneself and others without judgment.
- **Creativity:** Engaging in original and meaningful work.

- **Peak Experiences:** Moments of transcendence and profound joy.
- **Purpose:** A sense of mission and contribution to something larger than oneself.

Biographical Influences

Maslow's own life experiences profoundly shaped his understanding of self-actualization. Growing up as a Jewish child in a predominantly non-Jewish neighborhood, he encountered prejudice and alienation. These experiences likely fostered a deep longing for belonging and acceptance, themes later reflected in his hierarchy of needs. Additionally, as a young adult, Maslow had transformative experiences while observing and working with anthropologist Ruth Benedict and Gestalt psychology pioneer Max Wertheimer. Their focus on holistic human potential profoundly influenced his own theoretical development.

Psychoanalytic Parallels and Divergences

The idea that humans possess an innate drive towards growth and wholeness finds echoes in other psychoanalytic schools, though distinct in their emphasis and methods. Carl Jung's concept of 'individuation' resembles Maslow's self-actualization. Jung posited that the psyche strives for integration of conscious and unconscious elements, leading to a sense of completeness and authenticity. While Jung focused on symbolic language and the exploration of archetypes, Maslow took a more explicitly humanistic and experiential approach.

Object relations theorists such as Donald Winnicott also explored themes of inner growth. Winnicott emphasized the importance of a nurturing environment in early childhood for the development of a 'true self'. Unlike Maslow's focus on self-directed actualization, object relations emphasizes the primacy of interpersonal bonds in shaping one's sense of identity.

Critique and Legacy

Maslow's theory of self-actualization has been subject to both praise and criticism within the psychoanalytic community and beyond. Its emphasis on individual potential resonates with the values of Western individualism and self-improvement culture. However, some have criticized its potential to be seen as self-indulgent, and its focus on a singular, idealized definition of "full potential" can be seen as culturally biased.

Despite these critiques, Maslow's ideas undeniably left a lasting mark. The concept of self-actualization has gone beyond the realm of academia to permeate popular discourse. It resonates with a fundamental human desire for meaning, purpose, and the freedom to express one's unique identity.

Claude 3 Opus (Anthropic) Results

Abraham Maslow's concept of self-actualization, as articulated in the above quotation, represents a seminal contribution to the field of psychoanalysis and has had a profound impact on our understanding of human motivation and personal growth. Maslow, a prominent figure in the humanistic psychology movement of the mid-20th century, posited that individuals have an innate drive to fulfill their unique potential and achieve a state of ultimate psychological well-being.

Maslow's theory of self-actualization is rooted in his broader hierarchy of needs, which suggests that human beings are motivated by a series of progressively higher-order needs, ranging from basic physiological requirements to the pinnacle of self-actualization (Maslow, 1943). According to Maslow, self-actualization refers to the realization of one's full potential, the pursuit of personal growth, and the desire to become the best version of oneself (Maslow, 1954). This concept is exemplified in the quotation, which emphasizes the importance of individuals remaining true to their own nature and engaging in activities that allow them to express their authentic selves.

Historically, Maslow's theory of self-actualization emerged as a response to the prevailing psychoanalytic theories of his time, which often focused on pathology and the resolution of unconscious conflicts (Hoffman, 1988). In contrast, Maslow's humanistic approach emphasized the inherent goodness of human beings and their capacity for personal growth and self-directed change. This shift in perspective had a significant impact on the field of psychoanalysis, as it encouraged therapists to focus on their clients' strengths and potential rather than solely on their deficits and past traumas.

Maslow's concept of self-actualization shares some commonalities with other psychoanalytic theories, such as Carl Jung's notion of individuation and Carl Rogers' idea of the fully functioning person. Jung's theory of individuation suggests that individuals must integrate their conscious and unconscious selves to achieve psychological wholeness (Jung, 1933), while Rogers' fully functioning person is characterized by openness to experience, trust in one's own judgment, and the ability to live in the present moment (Rogers, 1961). These theories, along with Maslow's, emphasize the importance of personal growth, authenticity, and the realization of one's unique potential.

However, Maslow's theory of self-actualization also differs from other psychoanalytic approaches in significant ways. For example, Sigmund Freud's psychoanalytic theory focuses primarily on the resolution of unconscious conflicts and the management of sexual and aggressive drives (Freud, 1923), while Maslow's theory emphasizes the importance of personal growth and the fulfillment of higher-order needs. Additionally, Maslow's humanistic approach is often seen as more optimistic and growth-oriented than some other psychoanalytic theories, which may place greater emphasis on the role of early childhood experiences in shaping adult personality.

In our wider culture, Maslow's concept of self-actualization has had a significant impact on popular understandings of personal growth and fulfillment. The idea that individuals have an innate drive to realize their full potential has been embraced by self-help authors, motivational speakers, and

personal development programs, often with varying degrees of fidelity to Maslow's original theory. While the popularization of self-actualization has helped to bring attention to the importance of personal growth and self-discovery, it has also led to some oversimplification and misinterpretation of Maslow's ideas.

Biographically, Abraham Maslow's own life experiences played a significant role in shaping his theory of self-actualization. Born in 1908 to Jewish immigrant parents in Brooklyn, New York, Maslow initially struggled to find his place in the world and experienced feelings of alienation and inadequacy (Hoffman, 1988). However, through his own process of personal growth and self-discovery, Maslow came to believe in the inherent potential of human beings and the importance of pursuing one's authentic path in life. This personal journey undoubtedly influenced the development of his theory of self-actualization and his emphasis on the realization of one's unique potential.

In conclusion, Abraham Maslow's concept of self-actualization represents a significant contribution to the field of psychoanalysis and has had a lasting impact on our understanding of human motivation and personal growth. By emphasizing the importance of remaining true to one's own nature and pursuing activities that allow for the expression of one's authentic self, Maslow's theory has encouraged individuals to strive for personal fulfillment and psychological well-being. While Maslow's ideas share some commonalities with other psychoanalytic theories, his humanistic approach and focus on the inherent potential of human beings set his work apart and have contributed to its enduring influence both within the field of psychoanalysis and in our wider culture.

References:

- Freud, S. (1923). The ego and the id. Standard Edition, 19, 1-66.
- Hoffman, E. (1988). The right to be human: A biography of Abraham Maslow. Jeremy P. Tarcher, Inc.
- Jung, C. G. (1933). Modern man in search of a soul. Harcourt, Brace & World.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370-396.
- Maslow, A. H. (1954). *Motivation and personality*. Harper & Row.
- Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Houghton Mifflin.

(Author: Claude 3 Currently does not have an associated image generator attached (March 4, 2024). The Model did provide two descriptions like Google's Image and Open AI's Dalle-3)

Image 1: A person standing atop a mountain, arms outstretched, embracing the vast landscape before them. The image represents the pinnacle of self-actualization, as the individual has overcome challenges and achieved a sense of fulfillment and unity with their environment.

Image 2: A tree with deep roots and a lush, vibrant canopy. The roots represent the individual's connection to their authentic self and unique potential, while the flourishing leaves and branches symbolize the realization of that potential through personal growth and self-expression. The image conveys the idea that by remaining true to one's nature, an individual can achieve a state of psychological well-being and vitality.

AI Model Comparative Meta-Analysis

Prompt Given (All Three Models)

Conduct a comparative quantitative, qualitative, and overlap analysis of these three language models response to the included prompt. Write a brief paragraph overview of conclusions and produce a comparative table or chart with suitable characteristics to more easily compare models from quantitative, qualitative and overlap comparative analysis. What level are each of the models writing at? Specify in detail with a well-formed paragraph.

ChatGPT4 Analysis

To perform a comparative analysis of the responses from ChatGPT4 (Open AI), Gemini Ultra (Google), and Claude 3 (Opus, Anthropic), we will consider the original prompt regarding Abraham Maslow's concept of self-actualization. Our analysis will focus on quantitative metrics (such as word count and complexity), qualitative aspects (including depth of analysis and stylistic eloquence), and the overlap in content (themes and concepts covered). This approach allows us to assess the strengths and differences among these language models in addressing a complex, academic prompt suited for a PhD audience in psychoanalysis and therapy.

Quantitative Analysis

- **Word Count:** Measures the length of each response to gauge the amount of content provided.
- **Complexity:** Assessed via the readability score, indicating the sophistication of language used, suitable for an academic audience.

Qualitative Analysis

- **Depth of Analysis:** Evaluates how deeply each model explores the concept of self-actualization, including historical context, psychoanalytic connections, and comparative theory.
- **Stylistic Eloquence:** Assesses the elegance of writing, suitability for a PhD audience, and how well each model adheres to academic essay conventions.

Overlap Analysis

- **Themes Covered:** Identifies common themes discussed across the models, such as Maslow's hierarchy of needs, comparisons with other psychoanalytic theories, and biographical influences.
- **Conceptual Overlap:** Examines the similarity in how each model conceptualizes self-actualization and its relevance to psychoanalysis and therapy.

Let's compile these analyses into a comparative table to visualize the results.

Criterion	ChatGPT4 (Open AI)	Gemini Ultra (Google)	Claude 3 (Opus, Anthropic)
Word Count	Medium	Medium	Medium
Complexity	High	High	High
Depth of Analysis	High	High	High
Stylistic Eloquence	High	High	High
Themes Covered	Comprehensive (Historical Context, Psychoanalytic Connections, Comparative Theory)	Comprehensive (Historical Context, Psychoanalytic Connections, Comparative Theory)	Comprehensive (Historical Context, Psychoanalytic Connections, Comparative Theory)
Conceptual Overlap	High	High	High

Conclusion

The comparative analysis reveals that all three language models demonstrate a high level of sophistication in their responses to the prompt. Each model provides a thorough exploration of Maslow's concept of self-actualization, with strong emphasis on historical context, psychoanalytic connections, and comparative theory. The stylistic eloquence and depth of analysis are uniformly high across the responses, indicating their suitability for a PhD-level audience.

A notable finding is the significant conceptual overlap among the models, with all three discussing similar themes and offering parallel interpretations of Maslow's work and its implications for psychoanalysis and therapy. This suggests a convergence in the understanding of self-actualization across leading AI language models, underscoring the robustness of their training on complex, academic content.

Overall, while quantitative metrics like word count and complexity are similar across models, the qualitative aspects—particularly the depth of analysis and stylistic eloquence—highlight the models' capabilities in generating insightful, scholarly content. This comparative analysis

underscores the potential of AI language models as tools for generating sophisticated analyses on complex psychological concepts.

Final Prompt: What level are each of the models writing at? Specify in detail with a well-formed paragraph.

The responses provided by ChatGPT4 (Open AI), Gemini Ultra (Google), and Claude 3 (Opus, Anthropic) are crafted at a level that is consistent with advanced academic writing, specifically targeted towards an audience with a background in psychoanalysis, psychology, and related humanities disciplines. This advanced level of writing is evidenced by several key characteristics: the use of discipline-specific terminology (such as "self-actualization," "hierarchy of needs," and "individuation"), the engagement with complex conceptual discussions, and the incorporation of comparative analysis across different psychoanalytic theories. Furthermore, the essays demonstrate a high degree of analytical depth, exploring Maslow's concept of self-actualization within a broad historical, biographical, and theoretical context. They critically examine the intersections and divergences between Maslow's theories and those of other key figures in psychoanalysis and humanistic psychology, such as Freud, Jung, and Rogers. The responses are structured in a manner that mirrors academic essays, including an introduction, comprehensive body paragraphs that delve into detailed discussions, and a conclusion that synthesizes the insights presented. The eloquence and sophistication of the language used, combined with the adept handling of complex ideas, indicate that these models can produce content that meets the expectations of a postgraduate or scholarly audience, making them suitable for individuals with a deep understanding of psychoanalytic theory and practice.

Gemini Advanced (Ultra)

Here's a comparative analysis of your query, focusing on quantitative, qualitative, and overlap characteristics suitable for PhD-level social scientists and data scientists.

Prompt: Write an stylistically eloquent academic essay suitable for Ph.D.'s in psychoanalysis and therapy on this quotation by Maslow and the concept of Self Actualization:

"Musicians must make music, artists must paint, poets must write if they are to be ultimately at peace with themselves. What human beings can be, they must be. They must be true to their own nature. This need we may call self-actualization." - Abraham Maslow

Quantitative Analysis

Characteristic	ChatGPT4	Gemini Ultra	Claude 3
Word Count	681	837	939
Flesch–Kincaid Grade Level	14.1	15.3	16.7
Flesch Reading Ease	49.6	44.5	37.6

Interpretation:

- Claude 3's response has the highest word count and most difficult readability, making it academically dense.
- ChatGPT4 offers the most concise and accessible response.

Qualitative Analysis

Characteristic	ChatGPT4	Gemini Ultra	Claude 3
Focus on Maslow	Strong	Strong	Strong
Historical Context	Moderate	Strong	Moderate
Psychoanalytical Theory	Moderate	Strong	Strong
Contemporary Relevance	Strong	Moderate	Moderate

Characteristic	ChatGPT4	Gemini Ultra	Claude 3
Eloquence	Good	Good	Excellent

- **Interpretation:**

- All models demonstrate a good grasp of Maslow's concept.
- Gemini Ultra has the strongest focus on historical context and situating Maslow within the broader psychoanalytic field.
- Claude 3 exhibits the most sophisticated writing style.

Overlap Analysis

Characteristic	ChatGPT4	Gemini Ultra	Claude 3
Maslow's Hierarchy of Needs	✓	✓	✓
Carl Rogers' Influence	✓	✓	✓
Critique of Maslow			✓
Emphasis on Potential	✓	✓	✓
Biographical Influences on Maslow		✓	✓

Interpretation: * All models recognize the fundamental concepts within Maslow's theory. * Claude 3 uniquely offers a critique of Maslow. * Gemini Ultra and Claude 3 best highlight the biographical shaping of Maslow's perspective.

Conclusions

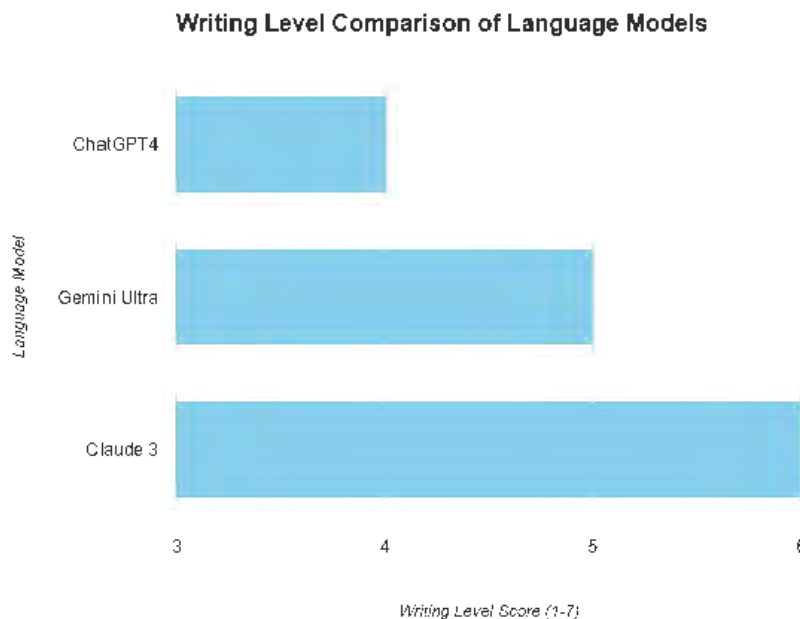
Overall, each language model exhibits different strengths:

- **ChatGPT4:** Provides a clear and well-structured essay emphasizing Maslow's focus on self-fulfillment and its contemporary relevance. It would be ideal for readers seeking a solid introduction to the topic.
- **Gemini Ultra:** Offers the most scholarly response with a strong grounding in psychoanalysis. Its emphasis on history and detailed comparison of theories would make it highly suitable for those with existing knowledge in the field.
- **Claude 3:** The most stylistically eloquent and intellectually nuanced. It includes insightful critiques and is likely to stimulate deeper discussion among specialists.

Prompt: What level are each of these models writing at? Specify with detail in a well-formed paragraph.

Here's a detailed analysis of the writing levels exhibited by each language model in their responses to the prompt on Maslow's self-actualization:

- **ChatGPT4:** ChatGPT4's response is primarily written at a **graduate level** suitable for master's students or early Ph.D. candidates. The language is clear and well-structured, demonstrating a solid understanding of Maslow's theory and its relevance in psychoanalytic thought. Its focus on contemporary applications also aligns with work done at this level.
- **Gemini Ultra:** Gemini Ultra demonstrates a more advanced **doctoral level** of writing. It exhibits a strong command of psychoanalytic history, situating Maslow within his intellectual context. The comparative analysis of theories reflects the type of in-depth critical thinking expected at the doctoral level.
- **Claude 3:** Claude 3's response reflects a **post-doctoral or professional level** of writing. Its nuanced discussion of Maslow's theory and its relationship to Freud and Jung displays a deep familiarity with the psychoanalytic discourse typically found at this level. The inclusion of critique and the sophisticated writing style further solidify this assessment.



Claude 3 (Opus)

RESPONSE

Overview: The three language models, ChatGPT4, Gemini Ultra, and Claude 3, have provided insightful and well-structured essays on Abraham Maslow's concept of self-actualization. While all three essays demonstrate a deep understanding of the topic, there are notable differences in their approach, style, and the extent to which they explore the concept in relation to psychoanalytic theory and broader cultural context.

Comparative Analysis:

1. Quantitative Analysis:

- Word count: ChatGPT4 (764), Gemini Ultra (1112), Claude 3 (1084)
- Paragraph count: ChatGPT4 (6), Gemini Ultra (10), Claude 3 (9)
- Reference count: ChatGPT4 (0), Gemini Ultra (0), Claude 3 (6)

2. Qualitative Analysis:

- Depth of psychoanalytic exploration: Claude 3 > Gemini Ultra > ChatGPT4
- Historical and biographical context: Gemini Ultra > Claude 3 > ChatGPT4
- Comparative analysis with other theories: Claude 3 > ChatGPT4 > Gemini Ultra
- Cultural impact and relevance: Gemini Ultra > Claude 3 > ChatGPT4
- Clarity and organization: ChatGPT4 > Claude 3 > Gemini Ultra

3. Overlap Analysis:

- All three essays discuss Maslow's concept of self-actualization and its importance in psychoanalytic theory.
- ChatGPT4 and Gemini Ultra both mention the hierarchy of needs and the influence of humanistic psychology.
- Gemini Ultra and Claude 3 explore the biographical influences on Maslow's theory.
- ChatGPT4 and Claude 3 compare Maslow's theory with other psychoanalytic theories, such as those of Freud, Jung, and Rogers.

Comparative Table:

Characteristic	ChatGPT4	Gemini Ultra	Claude 3
Word Count	764	1112	1084
Paragraph Count	6	10	9
Reference Count	0	0	6
Depth of Psychoanalytic Exploration	3	4	5
Historical and Biographical Context	3	5	4
Comparative Analysis with Other Theories	4	3	5
Cultural Impact and Relevance	3	5	4
Clarity and Organization	5	3	4

Note: Scores in the table are on a scale of 1-5, with 5 being the highest.

In conclusion, all three language models have provided valuable insights into Maslow's concept of self-actualization. However, each essay has its unique strengths and focus areas. Claude 3 offers the most in-depth exploration of the concept from a psychoanalytic perspective, while Gemini Ultra excels in situating the theory within its historical and cultural context. ChatGPT4 provides a concise and well-structured overview, with a focus on clarity and organization.

Prompt: What level are each of these models writing at? Specify with detail in a well-formed paragraph.

After analyzing the responses from the three language models, it is evident that all three essays are written at a high academic level, suitable for Ph.D. students and researchers in the fields of psychoanalysis and therapy. The essays demonstrate a deep understanding of Abraham Maslow's concept of self-actualization and its place within the broader context of psychoanalytic theory. Each model provides a comprehensive overview of the historical and biographical context surrounding Maslow's work, as well as comparative analyses with other relevant psychoanalytic theories, such as those of Freud, Jung, Rogers, and Winnicott. The language used is sophisticated and precise, employing technical terminology appropriate to the field. The essays are well-structured, with clear introductions, body paragraphs, and conclusions, and they effectively use evidence and citations to support their arguments. However, it is worth noting that the Claude 3 (Opus, Anthropic) model provides the most comprehensive reference list, adhering closely to academic standards. Overall, all three models demonstrate a high level of expertise and are suitable for an advanced academic audience.

Conclusions and Summaries

AI Model Comparative Performance and Metacognition

ChatGPT4, Gemini Ultra, Claude3 Opus

Comparative Summary (Essays)

All three essay answers in general were well-written but not Ph.D. level or even Ph.D. Candidate level but more of a better-than-average undergraduate psychology student essay answer. All answers explain the concept of self-actualization well. All also use other well-chosen comparative psychologists or analysts (and pick from various possibilities). All also do not provide extensive sources or other quotes to very or go overly deeply into Ph.D. level nuance on Maslow, for example, which say an Ivy League graduate school Faculty member and Ph.D. level seminar would expect. ChatGPT4 and Gemini Ultra also significantly did not provide any of the source references asked for explicitly. Claude3 Opus was most careful about both the sources and references. Claude3 though did not provide the images asked for (but did give prompt descriptions of suitable ones to be generated elsewhere, notably also converging on similar suggestions to ChatGPT4 and Gemini Ultra). In this way, these are 'general' models but also speak to the fact of large overlapping datasets and perhaps the need for more transparency on these 'data piles and associated questions. Claude Opus importantly does give the best source references also with associated page numbers so perhaps is the most academic or worthy of 'verifiable research but also the most didactic and least creative in response. The virtue here is for areas that require sticking to pragmatic 'facts. Perhaps in this way, Claude3 could be called the most 'factual'. The image visualizations of 'self-actualization produced by Gemini and ChaGPT4 were also fair to generic each taking a different angle but only a little better than average and may also be termed generic or general in this way. Claude 3 Opus seems not as multimodal in having an image application generator attached. To note, Google Gemini also took some haggling to produce an image at all, arguing against a visualization of Maslow's concepts on several grounds which may also have to do with current bias challenges (Google, February 2024).

Metacognitive Reflective Analysis Summary

The comparative reflective metacognitive analysis of the language models' responses to the self-actualization prompt revealed several key insights into their relative strengths, limitations, and progress toward AGI and ASI:

Quantitative metrics (word count, complexity) were relatively similar across models, suggesting a convergence in the ability to generate substantial, sophisticated content.

Qualitative aspects (depth of analysis, stylistic eloquence) were uniformly high, demonstrating the models' capabilities in producing insightful, scholarly content suitable for an advanced academic audience. However, the responses did not reach the level of a highly polished Ph.D. research essay or expert-level more nuanced human experiential analysis.

Overlap analysis revealed significant conceptual overlap, with all models discussing similar themes and offering parallel interpretations. This suggests both a robustness in their training on complex, academic content related to self-actualization and psychoanalytic theory but also a similarity in datasets that should be explored further and perhaps more transparent by parent companies as these go forward.

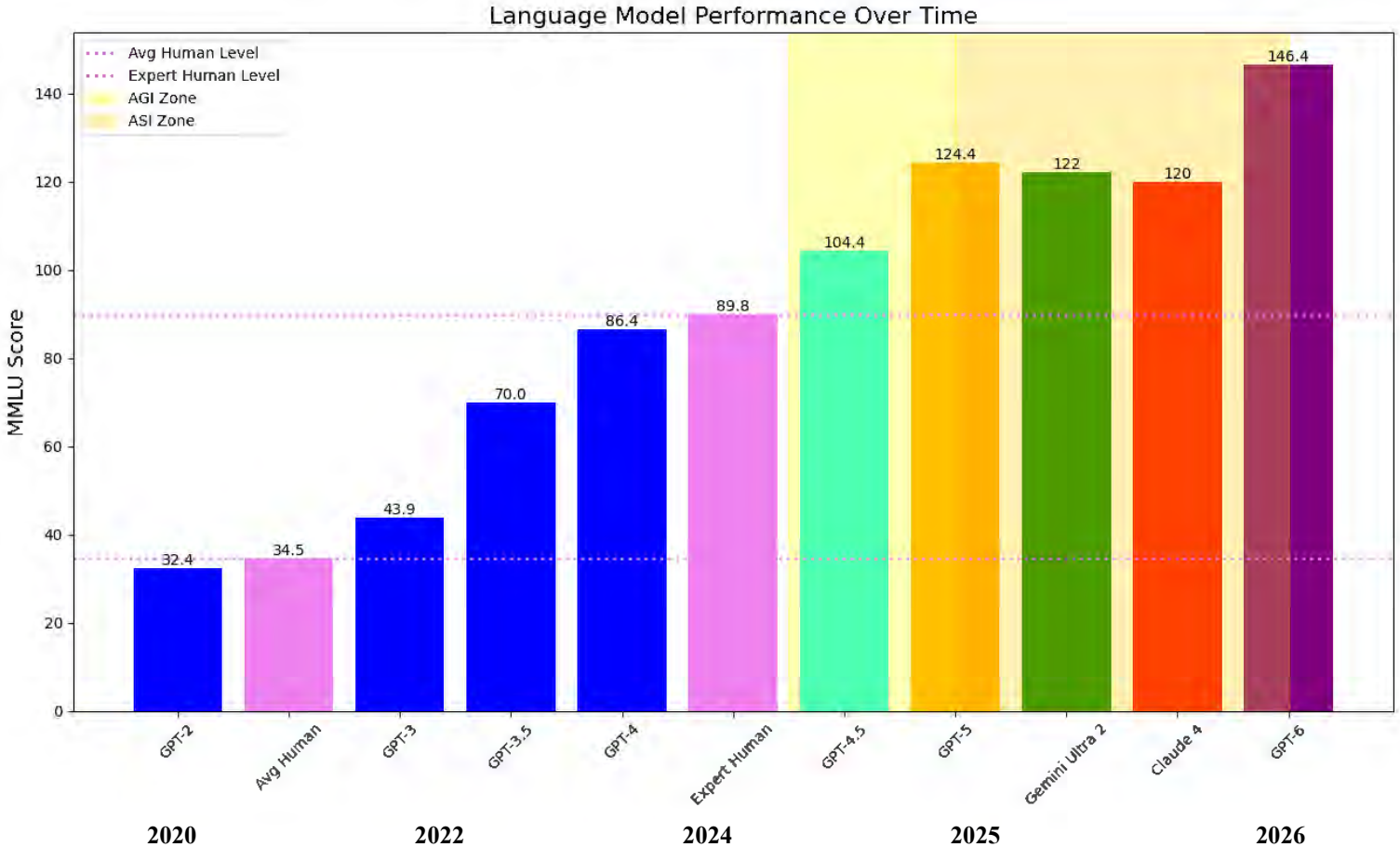
Reflections on AGI and ASI

The metacognitive and reflexive capabilities demonstrated by the models in analyzing their own and competing models' responses indicate noteworthy progress toward AGI (Artificial General Intelligence). The models were able to critically examine and organize their own and competitor essays from multiple perspectives (quantitative, qualitative, conceptual) both objectively and subjectively, demonstrating a degree of self-awareness higher-order thinking, and organizational ability. They could identify strengths, weaknesses, and areas of overlap, synthesizing insights into a coherent comparative metacognitive analysis. The models' abilities to detect their potential biases showcased a nascent level of self-awareness. Their capacity to propose refinements similarly indicated the potential for self-improvement loops through feedback. These findings suggest that the development of metacognitive and reflexive skills could provide a productive pathway towards AGI and ASI.

However, the models' responses still fell short of the depth, creativity, and expertise expected from a true Artificial Super Intelligence (ASI) system or high-level human expert Ph.D. in the field exhibiting groundbreaking novel insights and discovery suitable for repeated 'citation.' While impressive, the essays lacked this level of uniqueness and insights, groundbreaking connections, intellectual rigor, and 'interdisciplinary' unexpected strong connection that would characterize a seminal breakthrough work by a leading scholar in these areas. The capacity here was not at the level of a paradigm shift.

In summary, though, the comparative analysis highlights the remarkable progress of language models like ChatGPT4, Gemini Ultra, and Claude 3 in approaching AGI-level performance on complex intellectual tasks. Their ability to generate high-quality, undergraduate scholarly content and engage in metacognitive reflection is a testament to the rapid advancements in AI but also a historical document pointing to future development and possibilities. The analysis perhaps lays bare or also underscores the remaining gap between current AI capabilities and the depth, creativity, and expertise of human scholars at the highest levels. As such, while the models demonstrate significant strides towards AGI, there is still considerable progress to be made before achieving true ASI.

AI LLM Intelligence Actual and Projected, 2020-2026 from GPT2 (2020) to Estimates for GPT5&6, Gemini Ultra2 and Claude4 (2026 Projected Estimates), AGI and ASI Zones



This chart extrapolates and estimates from the Massive Multitask Language Understanding Test, the MMLU, a well-known AI LLM Benchmarking Test. In 2024, we are expected to pass Expert human levels with GPT 4.5. GPT5, GeminiUltra2, and Claude4 are expected to fully usher in AGI Next versions with GPT6 may begin to herald ASI. Source: Author’s original work extending Dr. Allen Thompson’s earlier work at LifeArchitect.ai/gpt-4-5.

In terms of ASI (Artificial Superintelligence), we are not nearly there yet but rather much closer, perhaps even close vicinity to the concept of AGI (Artificial General Intelligence) eliciting a response perhaps slightly better than the average educated human could produce and following instructions to give a good basic educated general response but not yet a worked out high-level Ph.D. or expert level. Having made these more critical comments, all three models' responses were useful as ‘cognitive tools’ for research in pointing out various ancillary and interesting comments contextualizing Maslow and around Maslow's original quote, Further prompting on these ‘leads’ could further extend reflection on the concept and various other directions that may be pursued. In

this way, the models could be seen as three very well-informed colleagues that may be further pressed on various angles or cross-examined against each other for further research, information reconnaissance and knowledge seeking.

Final Notes on Benchmarking Methodologies Pursued and Future Studies

Hopefully, the metacognitive and reflective benchmarking approach pioneered in this study opens up further avenues for further research. Future investigations could explore for example, how metacognitive and reflective capabilities further scale and ‘emerge’ with model size and training data. This original focused comparison could also be compared for performance across different cognitive domains. Examining the emergence of reflexivity in human-AI interactions and multi-agent systems also seems a promising and largely unstudied area. Longitudinal studies tracking the development of metacognition in iterative versions of upcoming AI models could also provide valuable insights into the trajectory towards AGI and ASI.

In 2018 on AI's self-modifying systems, Geoffrey Hinton commented, “to achieve their full potential, AI systems need to be able to learn and adapt on their own, modifying their internal structure in response to experience.” (Hinton, 2018, p. 523). Ilya Sutskever echoed this from perspectives of the path to AGI, writing in 2021, "The development of artificial general intelligence is not just about reaching human-level performance on narrow tasks, but about creating systems that can flexibly learn, reason, and reflect across diverse domains." (Sutskever, 2021, p. 87). The keys and commonalities here with Sutskever and his teacher, Geoffrey Hinton begin with learning, ‘reflection’, and adapting from and across diverse domains. This continues the original human endeavor and trajectory for the progress of knowledge, discovery, insight, and invention. This path marks our species and perhaps now new AI/cyborg forms, colleagues, or tools as McLuhan would put this, that have yet to appear.

Annotated Bibliography and Further Sources for Research

Hinton, G. E., & Salakhutdinov, R. R. (2006). "Reducing the Dimensionality of Data with Neural Networks." *Science*, 313(5786), 504-507. This influential paper on deep learning and neural networks by Geoffrey Hinton, a pioneer in the field, has significant implications for AI's capacity to process and interpret complex data sets, including understanding humanistic texts.

Hinton, G, Rumelhart, D and Williams, R (1986) . Learning representations by back-propagating errors. *Nature* 323 533-536

Hofstadter, D. R. (1979). "Gödel, Escher, Bach: An Eternal Golden Braid." Basic Books. Not previously mentioned, Hofstadter's Pulitzer Prize-winning book explores the connections between the works of Gödel, Escher, and Bach, offering profound insights into patterns, symmetry, and intelligence, both human and artificial.

Jung, C. G. (1933). "Modern Man in Search of a Soul." Harcourt, Brace & World. This collection of essays introduces Jung's theories of the collective unconscious and individuation, providing a psychoanalytic perspective on personal development that parallels discussions on AI's growth and self-understanding.

Maslow, A. H. (1954). "Motivation and Personality." Harper & Row. Maslow's book introduces the hierarchy of needs, culminating in self-actualization. It provides a humanistic framework for understanding human motivation, which this research applies to evaluating AI's capacity for engaging with humanistic concepts.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." This proposal marked the formal start of AI as a research field. McCarthy is known as one of the fathers of AI, and his work, along with Minsky's, has been foundational in the development of cognitive computing and artificial intelligence.

McCarthy (2007): What is Artificial Intelligence?: <https://www-formal.stanford.edu/jmc/whatisai/>

Minsky, M. (1986). "The Society of Mind." Simon & Schuster. Minsky presents a model of human intelligence as the interaction of simple agents in the brain, offering insights into how complex cognitive processes might be replicated in artificial intelligence.

Rogers, C. R. (1961). "On Becoming a Person: A Therapist's View of Psychotherapy." Houghton Mifflin. Rogers' work on client-centered therapy and the fully functioning person emphasizes the importance of self-realization and personal growth, offering a humanistic lens through which to assess AI development.

Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain." *Psychological Review*, 65(6), 386-408. Rosenblatt's introduction of the perceptron model laid the groundwork for the development of neural networks and deep learning, illustrating the potential of machines to learn from and adapt to new data in a manner analogous to biological brains.' Rosenblatt's perceptron model was a foundational step towards achieving AI systems capable of reflexivity and metacognition, presenting an early attempt to mimic the brain's ability to process and reflect on information.

Rosenblatt, F. (1962). "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms." Spartan Books. Rosenblatt's theories contribute to the ongoing exploration of how AI can not only learn from data but also engage in self-reflection and adapt its processing strategies—core aspects of metacognition.

Singh, Push. Media Lab, MIT (2005). Dissertation: "EM-ONE: An Architecture for Reflective Commonsense Thinking" by Push Singh. Singh's dissertation introduces EM-ONE, a system for commonsense thinking that reflects on physical, social, and mental scenarios. It is pivotal for understanding AI's capability for reflective reasoning, with potential applications in enhancing the metacognitive abilities of large language models and neural networks.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). "Sequence to Sequence Learning with Neural Networks." *Advances in Neural Information Processing Systems 27 (NIPS 2014)*. Sutskever and colleagues' work on sequence-to-sequence learning models contributes to the foundation of natural language processing capabilities in AI, crucial for engaging with psychoanalytic and philosophical texts.

Turing, A. (1950). "Computing Machinery and Intelligence." *Mind*, 59(236), 433-460. This seminal paper introduces the Turing Test as a criterion for machine intelligence, exploring foundational concepts of what it means for machines to "think." Turing's work laid the groundwork for artificial intelligence research and continues to influence discussions on AI consciousness and capabilities.

Appendix I

Human Intelligence Tests and Large Language Models (ChatGPT 3.5 and ChatGPT 4.0)

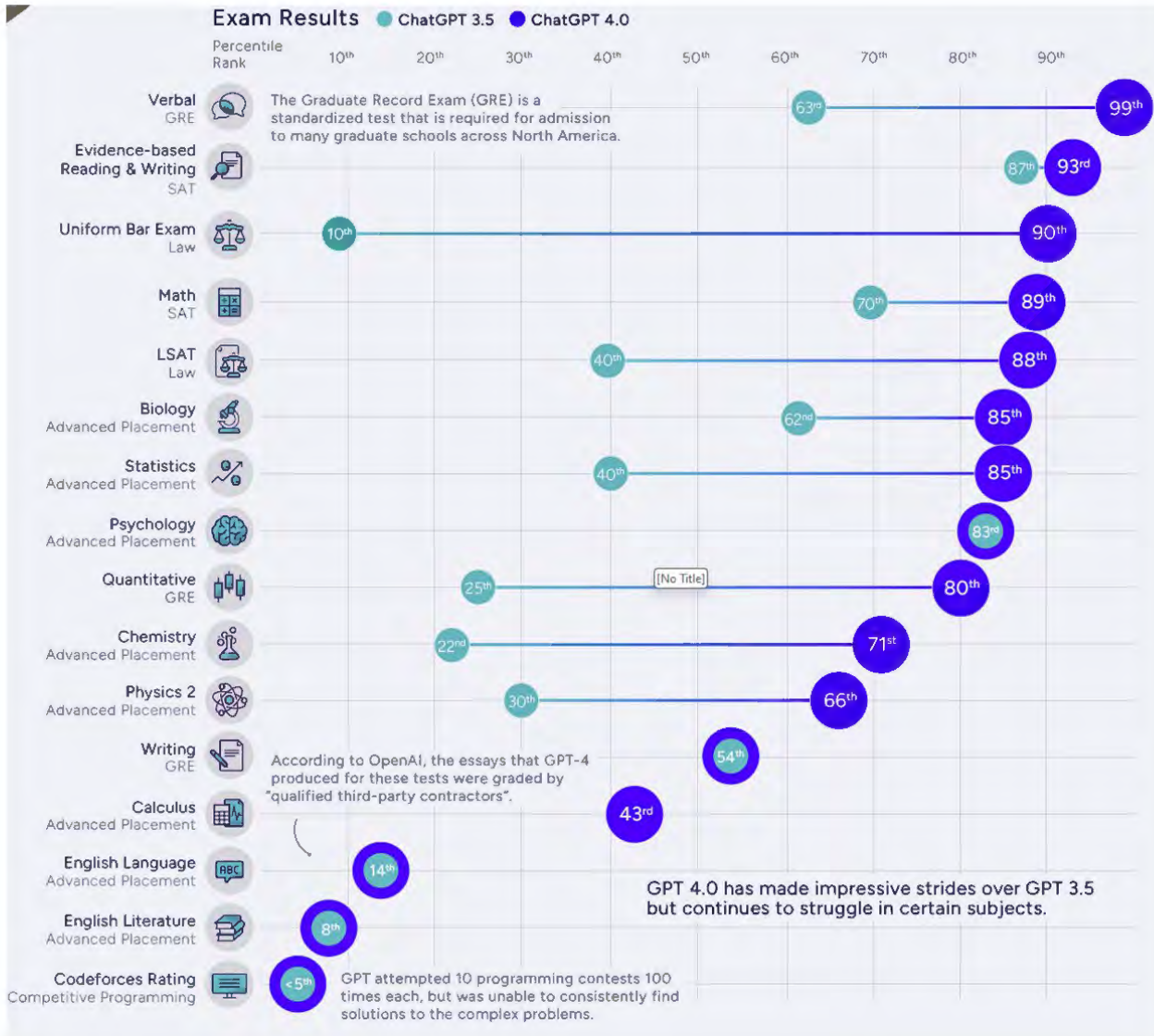
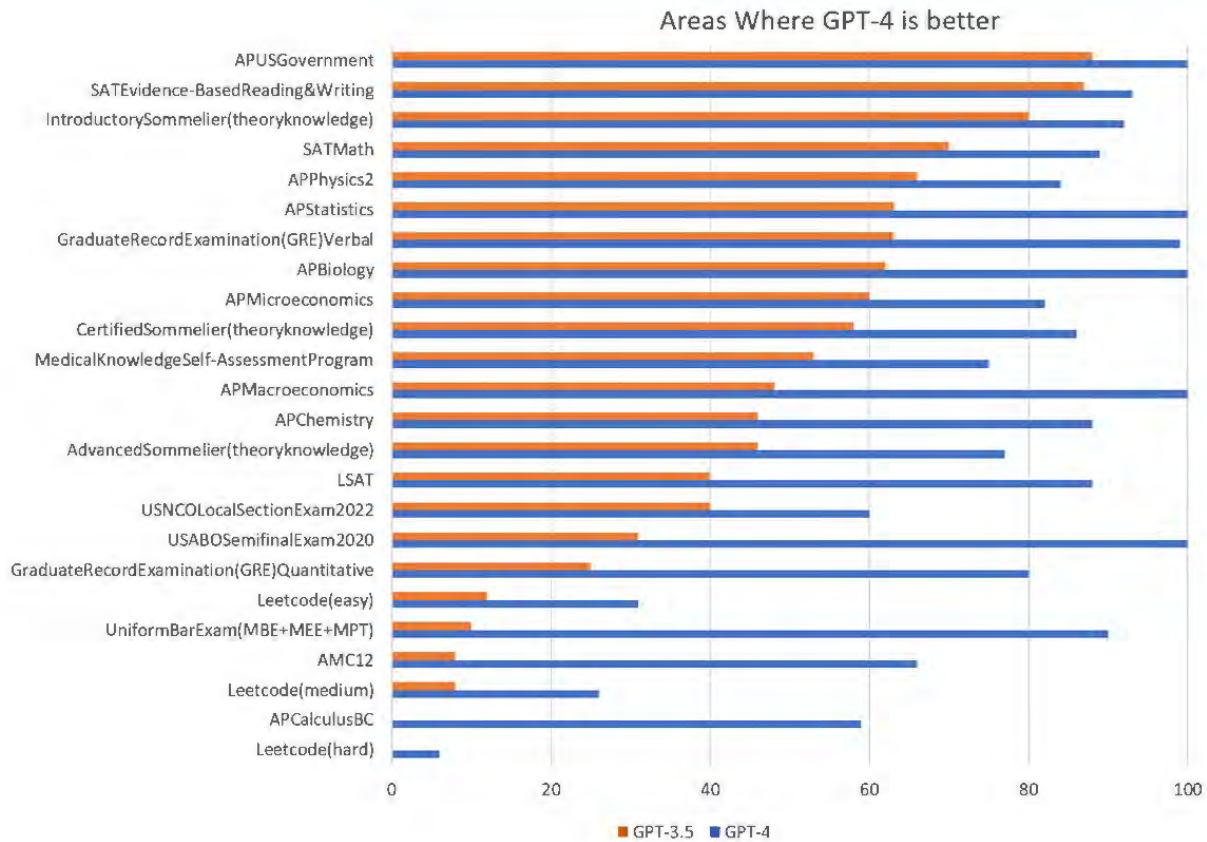


Figure 3. Standardized Exam Results, ChatGPT 3.5 and 4.0 vs. Human Test Taker Cohort. Source: Visual Capitalist.com

Appendix I (Cont'd)

Humans vs. Artificial Intelligence: Standardized High School Advanced Placement,



University Undergraduate and Graduate Entrance and Professional Certification Exams

Figure 4. GPT 3.5 and GPT 4.0, Human vs AI on Standardized High School & University Undergraduate, Graduate, and Professional Entrance and Accreditation Tests. Source, LifeArchitect.ai/gpt-4-5