



Comparison Table

	OpenAI - gpt-4o-mini	OpenAI - gpt-4.1-mini	OpenAI - gpt-5-mini	Google - gemini-2.5-flash	Google - gemini-2.5-pro	Notes
Prompt #1	Quiz on Logic Gates	Quiz on Logic Gates	Quiz on Logic Gates			Pydantic model was used
Prompt Technique #1	Zero-shot	Zero-shot	Zero-shot			Same technique used for accurate comparison
Output #1	An MCQ based quiz on the said topic	An MCQ based quiz on the said topic	An MCQ based quiz on the said topic			Pydantic based response model was received
Count of output tokens #1	206	185	1280			GPT-5-mini is designed to give more logical and descriptive responses.
Latency of output #1	4.9 sec	5.62 sec	27.5 sec			as the response is longer, hence the longer latency
Total Cost #1	0.000159	0.000390	0.002619			the costlier the model, the longer the response
Prompt #2	Same Quiz	Same Quiz	Same Quiz			Pydantic model was used
Prompt Technique #2	RAG	RAG	RAG			Same technique used for accurate comparison
Output #2	An MCQ based quiz on the said topic	An MCQ based quiz on the said topic	An MCQ based quiz on the said topic			Pydantic based response model was received
Count of output tokens #2	270	115	1684			GPT-5-mini is designed to give more logical and

						descriptive responses.
Latency of output #2	6.35 sec	6.04 sec	27.81 sec			as the response is longer, hence the longer latency
Total Cost #2	0.000183	0.000239	0.003402			the costlier the model, the longer the response

Conclusion: