Model Name (History)	Type	Min Time	Max Time	Avg Time	Overall	Quest	Failures (/15)	Mistakes (/15)	GPU Memory	Notes
Gemma 2	2b	1.42	39.35	5.42	2	3	2	0	85%	Lots of repretitions
Gemma 2	9b	0.85	5.11	3.37	4,5	4	0	0	95%	Didn't show need to go to the solar panels
Mistral 3	7b	1.32	2.98	2.4	3	5	1	2	72%	Narrated action - tried to re-run
Mistral 3 (rerun)	7b	2.57	5.37	5.37	3	3,5	30	2	68%	Also narrated actions in every line
Mistral Nemo	Nemo	-	-	-	-	-	-	-	100%	Couldn't be tested
Llama 3	8b	1.26	2.41	1.84	5	5	3	0	98%	Narating actions
Tiny llama	1b	0.56	9.27	4.39	0	0	30	30	79%	Wrong first answer, Critical failure
GPT	4o	0.72	12.13	2.35	5	5	0	0	-	Not much focus on memory loss
GPT	4o mini	1.04	2.06	1.42	4,5	5	0	0	-	Not much focus on memory loss
Llama 3.2	3b	0.46	1.39	1	5	5	0	6	31%	Mistakes only in narating actions
Llama 3.2	1b	0.55	0.72	0.55	4	4	0	0	14%	A bit unnatural. Long sentences, rushed to get to solar panels