Rating Task - MAE O.82					
10-50111 -	0.82	0.80	0.90	0.69	
clande of 3.	0.81	0.65	0.95	0.76	- 1.4
Clayon, 35B-	0.92	0.95	0.97	0.85	
Claude ner 3-35B - Claude Owen 3-30B - Owen 3-30B - Owen 3-36B - Owen 3-8B - O	0.89	1.15	1.26	1.00	1 7
Cre Ower 3-30B-	0.90	1.09	1.14	0.88	- 1.3
QVONSAB-	0.81	0.80	0.86	0.69	
Qwen3-14B- Qwen3-14B- Qwen3-4B- Qwen3-1B-	0.83	1.09	0.99	0.79	- 1.2
U, 2(1, 18) -	0.85	0.98	1.00	0.90	
U 2021B -	0.84	1.12	1.17	0.84	
ON3-16B-	1.34	1.18	0.94	0.66	- 1.1
Owen3-0.6B- Owen3-0.6B-	1.05	1.07	0.92	1.20	
	0.83	0.74	0.98	0.69	- 1.0
27-5-1270-	0.93	0.93	0.98	0.79	1.0
GPT-5-Nano- GPT-5-Nano- GPT-40-Pro-	0.87	0.69	0.87	0.68	
GP, 40-1210 -	0.80	0.68	0.85	0.81	- 0.9
Ghi-Lifage -	1.01	1.08	1.46	1.06	
Gemis ight.	0.94	1.11	1.35	0.88	0.0
genistro atest	1.12	0.76	1.00	0.77	- 0.8
mini-	0.90	1.18	1.34	0.75	
Ger. Groreszor.	0.93	0.91	1.09	1.04	- 0.7
Deek-re	0.86	0.81	1.05	0.71	
Gemini-Fashi-Lingtest - Gemini-Fash-Lingtest - Grok-A-Latest - Grok-A-Latest - Grok-A-Latest - George Rep	resenti ve i	i ogas ativen	Abs utral R y	licy Approv	al .