

Table 1: Infinite Loops, Average Token Count, and Thinking Format Adherence

Model (LR)	AIME24			MATH500			GPQA		
	Loops	Tokens	Format	Loops	Tokens	Format	Loops	Tokens	Format
SFT_2e4	24	7821	8	228	4418	220	133	6039	367
SFT_1e4	23	7685	7	253	6130	195	143	4623	357
SFT_1e5	14	5338	0	78	2357	0	76	2598	0
SFT_5e6	6	2840	0	51	1544	0	34	1552	0
SFT_1e6	1	1311	0	22	974	0	4	796	0

Note: 'Loops' refers to the Number of Responses that went into infinite repeating Loop. 'Tokens' refers to the Model Average Response Token Count. 'Format' refers to the Number of Model Responses following the Thinking Output Format.