

Table 1: RQ2: Abalation study on SCCG

SCCG	HumanEval ET			MBPP ET			CodeContest		
	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek
Our Method	30	11	15	40	21	34	6	13	18
w/o Code Check	27	10	13	35	18	27	6	11	16
w/o Plan Interpretation	25	8	10	32	16	24	5	10	14
w/o Multi-prompt	21	6	11	27	15	22	4	9	14
Only Plan Interpretation	20	5	10	27	14	20	5	7	12
Only Multi-prompt	19	4	9	24	12	20	3	8	10
Only Code Check	14	4	5	20	12	17	2	5	8

Table 2: RQ2: Abalation study on MetaGPT

SCCG	HumanEval ET			MBPP ET			CodeContest		
	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek
Our Method	33	13	17	23	16	20	4	13	14
w/o Code Check	29	10	14	20	12	16	4	12	13
w/o Plan Interpretation	27	9	13	17	12	15	4	10	12
w/o Multi-prompt	24	7	11	17	10	15	4	10	10
Only Plan Interpretation	24	9	11	16	11	13	3	9	10
Only Multi-prompt	20	5	10	17	8	10	3	7	8
Only Code Check	15	5	7	10	5	9	1	5	6

Table 3: RQ2: Abalation study on PairCoder

SCCG	HumanEval ET			MBPP ET			CodeContest		
	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek
Our Method	8	7	7	14	16	14	9	7	8
w/o Code Check	7	7	5	12	12	12	8	6	6
w/o Plan Interpretation	6	6	5	11	13	12	7	5	6
w/o Multi-prompt	6	4	5	11	10	10	7	4	5
Only Plan Interpretation	6	4	5	10	10	11	7	4	6
Only Multi-prompt	4	2	4	7	7	10	7	4	4
Only Code Check	3	1	2	6	8	6	3	3	4