

Table 1: RQ2: Abalation study on SCCG

SCCG	HumanEval ET			MBPP ET			CodeContest			CoderEval		
	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek
Our Method	30	11	15	48	21	34	9	13	8	9	5	7
w/o Code Check	27	10	13	40	18	27	8	11	7	8	4	6
w/o Plan Interpretation	25	8	10	28	16	24	8	10	7	8	4	5
w/o Multi-prompt	21	6	11	38	15	22	7	9	6	8	4	6
Only Plan Interpretation	20	5	10	39	14	20	6	7	5	7	4	4
Only Multi-prompt	19	4	9	29	12	20	6	8	5	7	3	4
Only Code Check	14	4	5	20	12	17	5	5	4	6	2	3

Table 2: RQ2: Abalation study on MetaGPT

SCCG	HumanEval ET			MBPP ET			CodeContest			CoderEval		
	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek	GPT-3.5	GPT-4o	Deepseek
Our Method	37	13	17	23	16	20	5	13	14	8	5	8
w/o Code Check	33	10	14	20	12	16	4	12	13	6	3	7
w/o Plan Interpretation	28	9	13	21	12	15	4	10	12	7	3	5
w/o Multi-prompt	30	7	11	17	10	15	4	10	10	7	4	6
Only Plan Interpretation	33	9	11	18	11	13	3	9	10	7	4	7
Only Multi-prompt	20	5	10	19	8	10	2	7	8	5	2	4
Only Code Check	13	5	7	14	5	9	2	5	6	4	2	4

Table 3: RQ2: Abalation study on PairCoder

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