TABLE I
RQ1 Full Results (PAT Dataset).

Pipeline	Stage	CSR	FPR	APR
<o3, Claude>	Direct Generation		0.8077	0.8378
	Refine Round 1		0.8077	0.8784
	Refine Round 2	1.0000	0.8462	0.9324
	Refine Round 3		0.9231	0.9595
	Refine Round 4		1.0000	1.0000
	Refine Round 5		1.0000	1.0000
<R1, Claude>	Direct Generation		0.6538	0.7432
	Refine Round 1		0.7308	0.7432
	Refine Round 2	0.8846	0.7692	0.8243
	Refine Round 3		0.8077	0.8378
	Refine Round 4		0.8077	0.8378
	Refine Round 5		0.8462	0.8514
<R1, R1>	Direct Generation		0.4615	0.5000
	Refine Round 1		0.5000	0.5270
	Refine Round 2	0.6923	0.5385	0.5676
	Refine Round 3		0.5769	0.5811
	Refine Round 4		0.5769	0.5811
	Refine Round 5		0.6538	0.6081
<R1, o3>	Direct Generation		0.5385	0.5405
	Refine Round 1		0.5385	0.5541
	Refine Round 2	0.6923	0.5769	0.5811
	Refine Round 3		0.6154	0.6081
	Refine Round 4		0.6923	0.6351
	Refine Round 5		0.6923	0.6351
<o3, R1>	Direct Generation		0.7308	0.6486
	Refine Round 1		0.7308	0.6757
	Refine Round 2	0.8462	0.7692	0.7297
	Refine Round 3		0.7692	0.7297
	Refine Round 4		0.8077	0.7297
	Refine Round 5		0.8077	0.7297
<o3, o3>	Direct Generation		0.7308	0.7297
	Refine Round 1		0.7308	0.8108
	Refine Round 2	0.8846	0.7308	0.8243
	Refine Round 3		0.8077	0.8514
	Refine Round 4		0.8077	0.8649
	Refine Round 5		0.8077	0.8649

TABLE II
RQ1 Full Results (UCS Dataset).

Pipeline	Stage	CSR	FPR	APR
<o3, Claude>	Direct Generation		0.8333	0.9474
	Refine Round 1		0.8333	0.9474
	Refine Round 2	1.0000	1.0000	1.0000
	Refine Round 3		1.0000	1.0000
	Refine Round 4		1.0000	1.0000
	Refine Round 5		1.0000	1.0000
<R1, Claude>	Direct Generation		0.6667	0.8421
	Refine Round 1		0.8333	0.9474
	Refine Round 2	1.0000	1.0000	1.0000
	Refine Round 3		1.0000	1.0000
	Refine Round 4		1.0000	1.0000
	Refine Round 5		1.0000	1.0000
<R1, R1>	Direct Generation		0.5000	0.6316
	Refine Round 1		0.6667	0.6842
	Refine Round 2	0.6667	0.6667	0.6842
	Refine Round 3		0.6667	0.6842
	Refine Round 4		0.6667	0.6842
	Refine Round 5		0.6667	0.6842
<R1, o3>	Direct Generation		0.3333	0.5789
	Refine Round 1		0.5000	0.6316
	Refine Round 2	0.6667	0.5000	0.6316
	Refine Round 3		0.5000	0.6316
	Refine Round 4		0.5000	0.6316
	Refine Round 5		0.5000	0.6316
<o3, R1>	Direct Generation		0.8333	0.9474
	Refine Round 1		0.8333	0.9474
	Refine Round 2	1.0000	0.8333	0.9474
	Refine Round 3		0.8333	0.9474
	Refine Round 4		0.8333	0.9474
	Refine Round 5		0.8333	0.9474
<o3, o3>	Direct Generation		0.6667	0.6842
	Refine Round 1		0.6667	0.6842
	Refine Round 2	0.6667	0.6667	0.6842
	Refine Round 3		0.6667	0.6842
	Refine Round 4		0.6667	0.6842
	Refine Round 5		0.6667	0.6842

TABLE III
RQ1 FULL RESULTS (A4F DATASET).

Pipeline	Stage	CSR	FPR	APR
<o3, Claude>	Direct Generation		0.5000	0.6750
	Refine Round 1		0.6250	0.8250
	Refine Round 2	1.0000	0.8750	0.9750
	Refine Round 3		1.0000	1.0000
	Refine Round 4		1.0000	1.0000
	Refine Round 5		1.0000	1.0000
<R1, Claude>	Direct Generation		0.5000	0.6750
	Refine Round 1		0.5000	0.7500
	Refine Round 2	0.8750	0.7500	0.8000
	Refine Round 3		0.7500	0.8000
	Refine Round 4		0.7500	0.8000
	Refine Round 5		0.7500	0.8000
<R1, R1>	Direct Generation		0.2500	0.5000
	Refine Round 1		0.3750	0.6750
	Refine Round 2	0.8750	0.7500	0.8000
	Refine Round 3		0.7500	0.8000
	Refine Round 4		0.7500	0.8000
	Refine Round 5		0.7500	0.8250
<R1, o3>	Direct Generation		0.3750	0.6250
	Refine Round 1		0.7500	0.8000
	Refine Round 2	0.8750	0.7500	0.8000
	Refine Round 3		0.7500	0.8000
	Refine Round 4		0.7500	0.8000
	Refine Round 5		0.7500	0.8000
<o3, R1>	Direct Generation		0.3750	0.5250
	Refine Round 1		0.5000	0.6500
	Refine Round 2	0.8750	0.6250	0.7500
	Refine Round 3		0.6250	0.7500
	Refine Round 4		0.6250	0.7750
	Refine Round 5		0.7500	0.8250
<o3, o3>	Direct Generation		0.5000	0.7250
	Refine Round 1		0.7500	0.9000
	Refine Round 2	1.0000	0.8750	0.9250
	Refine Round 3		0.8750	0.9250
	Refine Round 4		0.8750	0.9250
	Refine Round 5		0.8750	0.9250

TABLE IV
RQ1 FULL RESULTS (OVERALL).

Pipeline	Stage	CSR	FPR	APR
<o3, Claude>	Direct Generation		0.7500	0.8045
	Refine Round 1		0.7750	0.8722
	Refine Round 2	1.0000	0.8750	0.9549
	Refine Round 3		0.9500	0.9774
	Refine Round 4		1.0000	1.0000
	Refine Round 5		1.0000	1.0000
<R1, Claude>	Direct Generation		0.6250	0.7368
	Refine Round 1		0.7000	0.7744
	Refine Round 2	0.9000	0.8000	0.8421
	Refine Round 3		0.8250	0.8496
	Refine Round 4		0.8250	0.8496
	Refine Round 5		0.8500	0.8571
<R1, R1>	Direct Generation		0.4250	0.5188
	Refine Round 1		0.5000	0.5940
	Refine Round 2	0.7250	0.6000	0.6541
	Refine Round 3		0.6250	0.6617
	Refine Round 4		0.6250	0.6617
	Refine Round 5		0.6750	0.6842
<R1, o3>	Direct Generation		0.4750	0.5714
	Refine Round 1		0.5750	0.6391
	Refine Round 2	0.7250	0.6000	0.6541
	Refine Round 3		0.6250	0.6692
	Refine Round 4		0.6750	0.6842
	Refine Round 5		0.6750	0.6842
<o3, R1>	Direct Generation		0.6750	0.6541
	Refine Round 1		0.7000	0.7068
	Refine Round 2	0.8750	0.7500	0.7669
	Refine Round 3		0.7500	0.7669
	Refine Round 4		0.7750	0.7744
	Refine Round 5		0.8000	0.7895
<o3, o3>	Direct Generation		0.6750	0.7218
	Refine Round 1		0.7250	0.8195
	Refine Round 2	0.8750	0.7500	0.8346
	Refine Round 3		0.8000	0.8496
	Refine Round 4		0.8000	0.8571
	Refine Round 5		0.8000	0.8571