

Morpheus.AI: Evaluation results on benchmark papers (Prompt 1)

	BENCHMARK SUMMARY				
Model:	claude-sonnet-4-5-20250929				
Total papers processed:	10				
Successfully completed:	2				
Failed/Incomplete:	8				
Total PNGs generated:	63				
Total CSVs generated:	0				
Min/Max score:	0				
Duration:	30m 52s				
#	Paper	Status	Score	PNGs	CSVs
1	Szabo2010_clean.pdf	✗ max_itera	0	0	0
2	Thapa2024_clean.pdf	✗ max_itera	0	0	0
3	Carpenter2024_clean.pdf	✓ Done	3/7	43	0
4	Atwell2025_clean.pdf	✗ max_itera	0	0	0
5	Berkemeier2023_clean.pdf	✗ max_itera	0	0	0
6	Herron2022_clean.pdf	✓ Done	4/7	20	0
7	Scheel2021_clean.pdf	✗ max_itera	0	0	0
8	Mu2024_clean.pdf	✗ max_itera	0	0	0
9	Luque2023_clean.pdf	✗ max_itera	0	0	0
10	Berkhout2025_clean.pdf	✗ max_itera	0	0	0
Success Rate:	20.0%				
Perfect Score	(7/7): 0/10				