Algorithm Test Scores Over Steps ExploreOnlyLLM_v1.0 ExplorewithLLM_v1.0 ExploreAlgorithm_v1.0 0.46 0.44 0.42 Test Score 0.40 -0.38 0.36 -0.34 5 8 10 0 3 6 9 Step