RL Training Dynamics: Train Reward, Val Reward, Test GSM8K