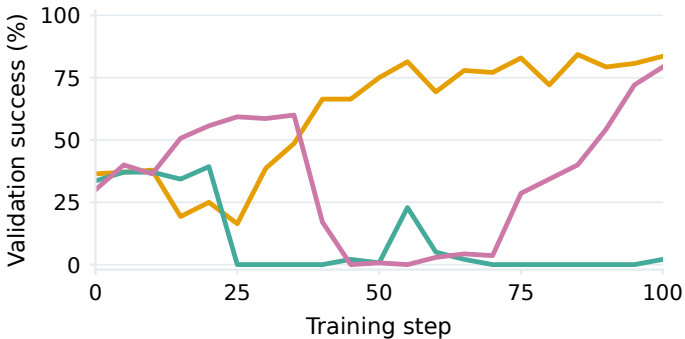


GRPO GRPO+OPSD AgentOPSD OCSD

(a) Intermediate instability



(b) Late regression

