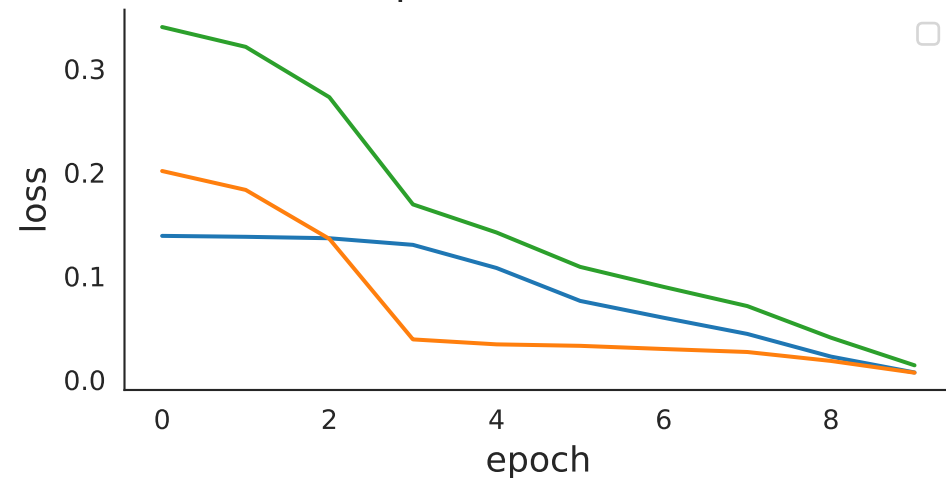
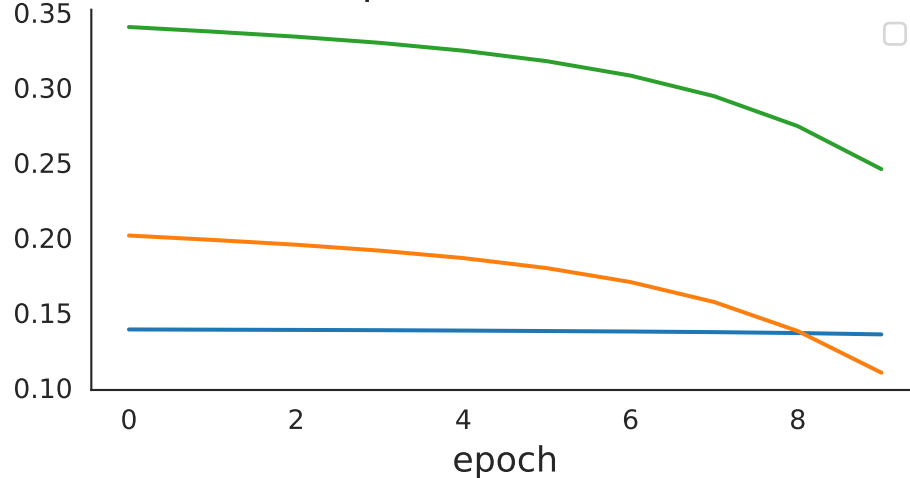


# Pretraining loss when reducing learning rate

10 pretr. + lr 0.0005



10 pretr. + lr 0.0001



100 pretr. + lr 0.0001

