

reward_noisy_soft_seed0 — preference validation loss (BCE)

