$$J_\theta = E_{\tau\sim(\tau;\theta)}[r(\tau)] = \int_\tau r(\tau)\,p(\tau;\theta)\,d\tau$$

$$\nabla_\theta J_\theta = \int_\tau r(\tau)\nabla_\theta p(\tau;\theta)\,d\tau = \int_\tau r(\tau)\,p(\tau;\theta)\nabla_\theta \log p(\tau;\theta)\,d\tau$$

$$= E_{\tau\sim p(\tau;\theta)} r(\tau)\nabla_\theta \log p(\tau;\theta) = E_{\tau\sim p(\tau;\theta)} r(\tau)\sum_{t\geq 0}\nabla_\theta \log \pi_\theta(a_t \mid s_t)$$

$$loss \approx \frac{1}{|batch|}\sum_{batch} r(\tau)\sum_{t\geq 0}\log \pi_\theta(a_t \mid s_t)$$

$$\approx \frac{1}{|batch|}\sum_{batch}\sum_{t\geq 0}[r(\tau)-b_t]\log \pi_\theta(a_t \mid s_t)$$