1. *The nonplanning method looks particularly poor in Figure 8.3 because it is a one-step method; a method using multi-step bootstrapping would do better. Do you think one of the multi-step bootstrapping methods from Chapter 7 could do as well as the Dyna method? Explain why or why not.*

   I think so. An n-step bootstrpping algorithm with big enough n (or a Monte Carlo method) would update at the end of the first episode all the action-values we encountered during the episode. I expect the policy based on these updated action-values to be quite good.