Question 8 Exercise 6.12

If action selection was greedy, they will act similar.

The action selection at every episode step would be the same.

The SARSA update equation essentially becomes

Choose A' such that we get $\max(Q(S,A'))$

$$Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma Q(S',A') - Q(S,A)]$$

$$= Q(S,A) \leftarrow Q(S,A) + \alpha[R + \gamma \max_a Q(S',a) - Q(S,A)]$$

which is the same as the Q learning update.

Updates and action selection would be the same for both algorithms provided a greedy selection replaces all ε-greedy policies involved.