

**Figure S6.** Supplementary experiment 2. (a) The log odds ratio of staying on the same stage 1 action after earning a reward on the previous trial over the odds after earning no reward. Sessions denoted by ‘all rare’ included only rare transitions (similar to sessions marked with ‘#’ in Figure 3a). (b) The probability of staying on the same stage 1 action in the probe session (session s76) as a function of whether the previous trial was rewarded (reward/no reward) and whether the transition in the previous trial was common or rare. (c) The probability of staying on the same stage 2 action in the probe session (session s76), as a function of whether the previous trial was rewarded (reward/no reward) and whether subjects stayed on the same stage 1 action (stay/switch). Similar to the analysis in the main paper, only trials in which the stage 2 state was different from the previous trial are included in panels (b) and (c) in order to detect the performance of action sequences. Similarly, only trials in which subjects made a correct discrimination on the previous trial (‘R’ in S2, and ‘L’ in S1) were included in panels (a-c). In all the probe sessions the probability of rare transitions was 50%. (d) Results of discrimination training showing the percentage of correct responses. Error-bars  $\pm 1$  SEM.

