Reward as a function of time. R_tot_naive = -1.07 and R_tot_smart = -0.96