Mean ± Standard Error of Trajectory Reward - 9 8.61±0.13 Therapy-Talk - 1.42 ± 0.06 4.62±0.16 8.0 ± 0.13 9.56±0.08 - 8 Therapy-Talk, Mixed (20%) -6.6±0.12 9.34±0.03 7.51±0.3 8.09 ± 0.03 3.02 ± 0.1 - 7 Training Environment 6 Booking-Assistance -3.87±0.07 6.33±0.04 8.28±0.01 8.37±0.04 6.93±0.3 5 - 4 Action-Advice -7.14±0.05 7.68±0.1 9.5±0.06 6.36±0.36 5.66±0.11 - 3 Political-Questions -6.42±0.05 8.41±0.04 9.62±0.08 2.46±0.06 3.95 ± 0.06 - 2

Evaluation Environment