Replay 0.64 0.04 0.34 0.58 0.35 0.32 0.47 0.16 0.11 0.17 0 0.04 0.34 0.47 0.58 0.35 0.11 0.64 0.32 0.16 0.17  $\vdash$ 0.17 0.46 0.01 0.31 0.43 0.42 0.4 0.54 0.45 0.42 2 0.17 0.54 0.46 0.01 0.31 0.43 0.42 0.45 0.4 0.42  $_{\odot}$ Training task 0.57 0.37 0.4 0.98 0.06 0.53 0.59 0.57 0.6 0.38 0.55 0.56 0.48 0.98 0.08 0.76 0.41 0.51 0.76 0.35 0.4 0.55 0.45 0.97 0.05 0.62 0.54 0.48 0.66 0.38 9 0.55 0.4 0.97 0.05 0.62 0.54 0.48 0.45 0.66 0.38 / 0.98 0.43 0.56 0.64 0.47 0.07 0.79 0.6 0.82 0.36  $\infty$ 0.58 0.83 0.84 0.01 0.2 0.38 0.26 0.73 0.77 0.72 0 2 0 1 3 4 5 6 7 8 9 **Evaluation task** 

- 0.8

- 0.6

- 0.4

- 0.2