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  $\vdash$ 0.28 0.01 0.43 0.42 0.43 0.4 0.58 0.57 0.54 0.43 2 0.28 0.58 0.57 0.01 0.43 0.42 0.43 0.54 0.43 0.4  $_{\odot}$ Training task 0.51 0.34 0.42 0.88 0.03 0.49 0.54 0.45 0.58 0.4 0.56 0.55 0.47 0.98 0.08 0.75 0.41 0.51 0.76 0.36 0.42 0.58 0.44 0.97 0.11 0.62 0.54 0.51 0.66 0.31 9 0.58 0.42 0.97 0.11 0.62 0.54 0.51 0.44 0.66 0.31 / 0.97 0.11 0.65 0.59 0.44 0.43 0.63 0.47 0.67 0.31  $\infty$ 0.17 0.47 0.02 0.41 0.35 0.41 0.54 0.93 0.5 0.47 6 2 0 1 3 4 5 6 7 8 9 **Evaluation task** 

- 0.8

- 0.6

- 0.4

- 0.2