Naive 0.31 0.34 0.38 0.41 0.43 0.6 0.3 0.4 0.43 0.44 0 0.31 0.34 0.38 0.41 0.43 0.6 0.3 0.4 0.43 0.44 \vdash 0.34 0.38 0.41 0.43 0.3 0.31 0.4 0.43 0.44 0.6 \sim 0.31 0.57 0.73 0.36 0.38 0.69 0.42 0.43 0.73 0.73 $_{\odot}$ Training task 0.31 0.57 0.73 0.36 0.38 0.69 0.42 0.43 0.73 0.73 0.66 0.5 0.49 0.99 0.7 0.79 0.5 0.5 0.51 0.5 0.66 0.5 0.49 0.99 0.7 0.79 0.5 0.5 0.51 0.5 9 0.5 0.99 0.7 0.5 0.51 0.66 0.49 0.79 0.5 0.5 / 0.99 0.66 0.5 0.49 0.7 0.79 0.5 0.5 0.51 0.5 ∞ 0.95 0.52 0.65 0.02 0.02 0.71 0.24 0.5 0.6 0.5 0 6 0 1 2 3 5 7 8 9 **Evaluation task**

- 0.8 0.6 - 0.4

- 0.2