The effects of time horizon and guided choices on explore-exploit decisions in rodents

Siyu Wang, Blake Gerken, Julia R. Wieland, Robert C. Wilson, and Jean-Marc Fellous

Supplemental Materials

Supplemental figures and figure captions

Figure S1: Parameter recovery of exploration threshold θ, decision noise σ, short-term feeder bias α_{LG}, long-term feeder bias α_{LS} and spatial bias b.

Figure S2: Probability of switching away (from guided to unguided option) and probability of switching back (from unguided option to guided option) in free choices, i.e. p(switch away) and p(switch back), split up by whether the guided option is the objectively better option, for humans (A, C) and rats (B, D). Data from Experiments 1 (rats) and 4 (humans). High (low) contrast colors indicate games where the guided choices where in fact the best (worst) one of the two available choices.

Figure S3: Human Experiment 5 (Rewards range from 1 to 100). A: Probability of choosing the option with the highest reward as a function of trial number. B: Probability of switching from the last chosen option as a function of trial number. C: p(high reward) in the 1st free choice as a function of guided reward size by horizon. D: average p(high reward, 1st choice) by horizon. E: p(high reward) in the last free choice as a function of guided reward size by horizon. F: average p(high reward, last choice) by horizon. G: P(unguided) as a function of guided reward size by horizon. H: average P(unguided) by horizon. I: Model estimates of group-level exploration thresholds. J: Average of subject-level estimates of exploration thresholds by horizon. K: Model estimates of group-level decision noise. L: Average of subject-level estimates of decision noise by horizon.

Figure S4: Posterior distribution over the group-level means of spatial bias b, short-term feeder bias α_{LG}, long-term feeder bias α_{LS} for both humans and rats in Experiment 1.

Figure S5: Rats are influenced by both short-term and long-term feeder bias. Left, Percentage of choosing the unguided feeder in 1st free choice as a function of the experienced reward of the guided feeder from last game in humans (Top) and rats (Bottom). Right, Percentage of choosing the unguided option in 1st free choice as a function of the average reward of the guided feeder from last session in rats. Humans do not have a LS panel since all humans only participated in a single session. NaN refers to cases when the guided feeder was not chosen in the last game that involved it.

Figure S6: Sound cue variant of Experiment 2. In this experiment, the different horizon conditions are cued by either a low-pitch sound (H = 1) or a high-pitch sound (H = 6). Games of different horizons are interleaved. A: P(unguided) as a function of guided reward size. B. Model estimates of exploration threshold. C. Model estimates of decision noise.

Figure S7: Short term feeder bias is larger in long horizon condition. LG (left) and LS (right) coefficients as a function of Horizon (1:blue, 6:red) and nG (number of guided choices, nG = 0, 1, or 3). LG coefficient is significantly larger in H = 6 than H = 1 condition, showing that short term feeder bias (from last game) has a significantly bigger influence on H = 6 games ($p < 0.001$). This is likely due to that rats spend more trials at H = 6 feeders within a session. There are no differences in long term feeder bias (from last session) between horizon conditions ($p = 0.48$).

Figure S8: Posterior distribution over the group-level means of spatial bias b, short-term feeder bias α_{LG}, long-term feeder bias α_{LS} for rats in Experiment 2. Each row corresponds to one of the parameters, each column corresponds to one nG condition (nG = 0, 1 or 3).

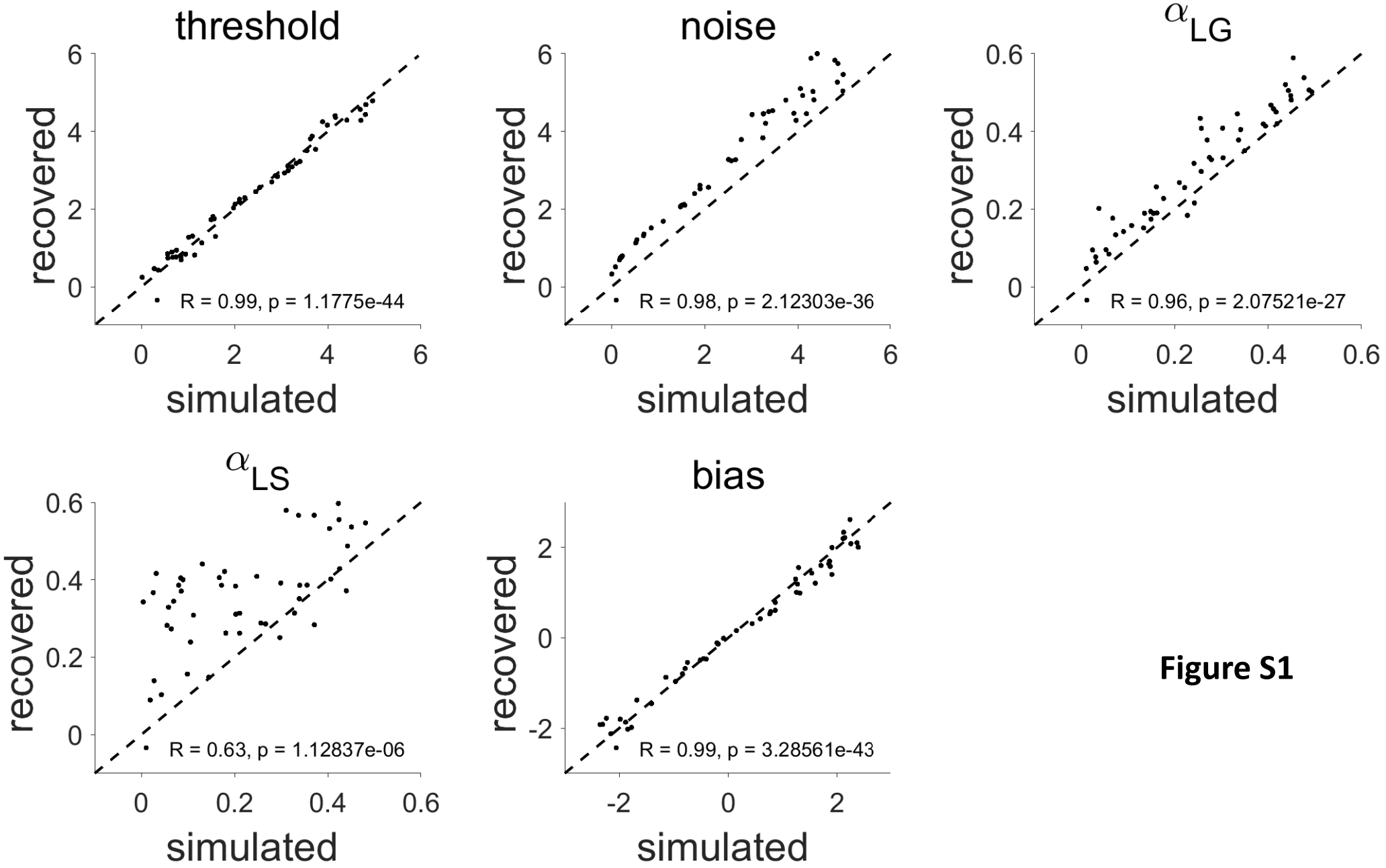

threshold
recovered
simulated
R = 0.99, p = 1.1775e-44
noise
recovered
simulated
R = 0.98, p = 2.12303e-36
α_{LG}
recovered
simulated
R = 0.96, p = 2.07521e-27
α_{LS}
recovered
simulated
R = 0.63, p = 1.12837e-06
bias
recovered
simulated
R = 0.99, p = 3.28561e-43

Figure S1

Figure S2

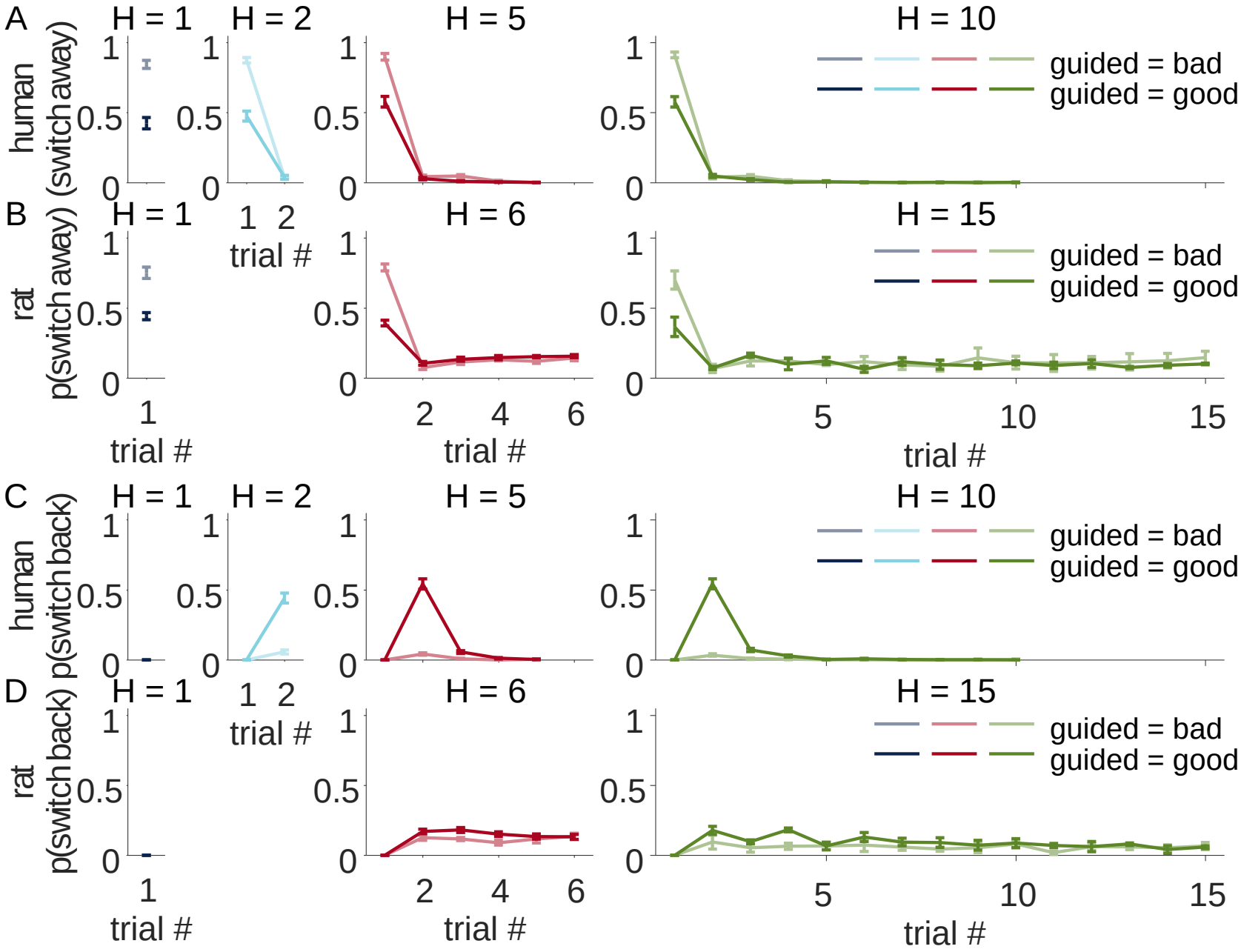

A
B
C
D
H = 1
H = 2
H = 5
H = 10
H = 6
H = 15
human p(switch away)
rat p(switch away)
human p(switch back)
rat p(switch back)
guided = bad
guided = good
trial #
0
0.5
1
2
4
6
5
10
15

Figure S3

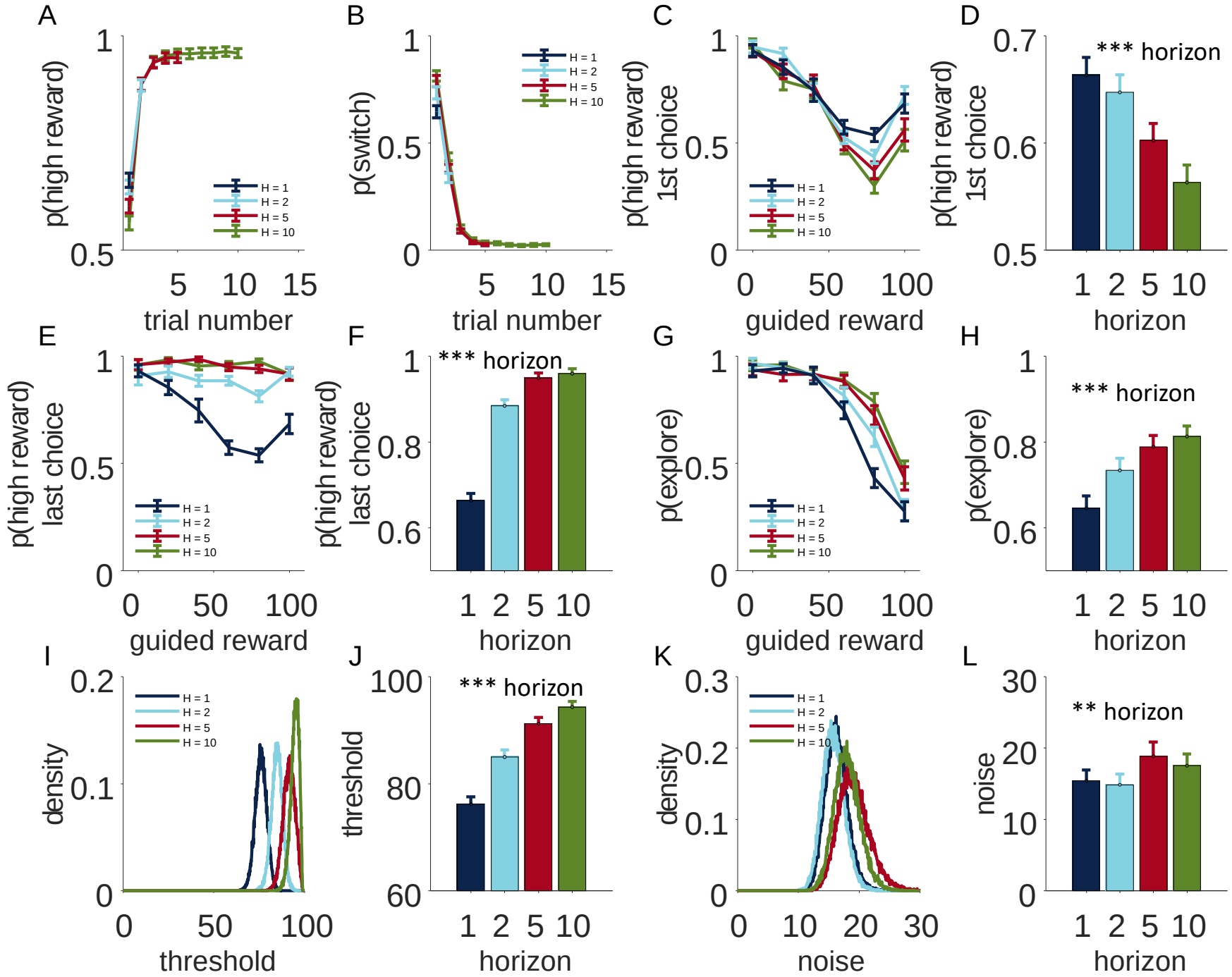

A
p(high reward)
trial number
H = 1
H = 2
H = 5
H = 10
B
p(switch)
trial number
C
p(high reward) 1st choice
guided reward
D
p(high reward) 1st choice
*** horizon
horizon
E
p(high reward) last choice
guided reward
F
p(high reward) last choice
*** horizon
horizon
G
p(explore)
guided reward
H
p(explore)
*** horizon
horizon
I
density
threshold
J
threshold
*** horizon
horizon
K
density
noise
L
noise
** horizon
horizon

Figure S4

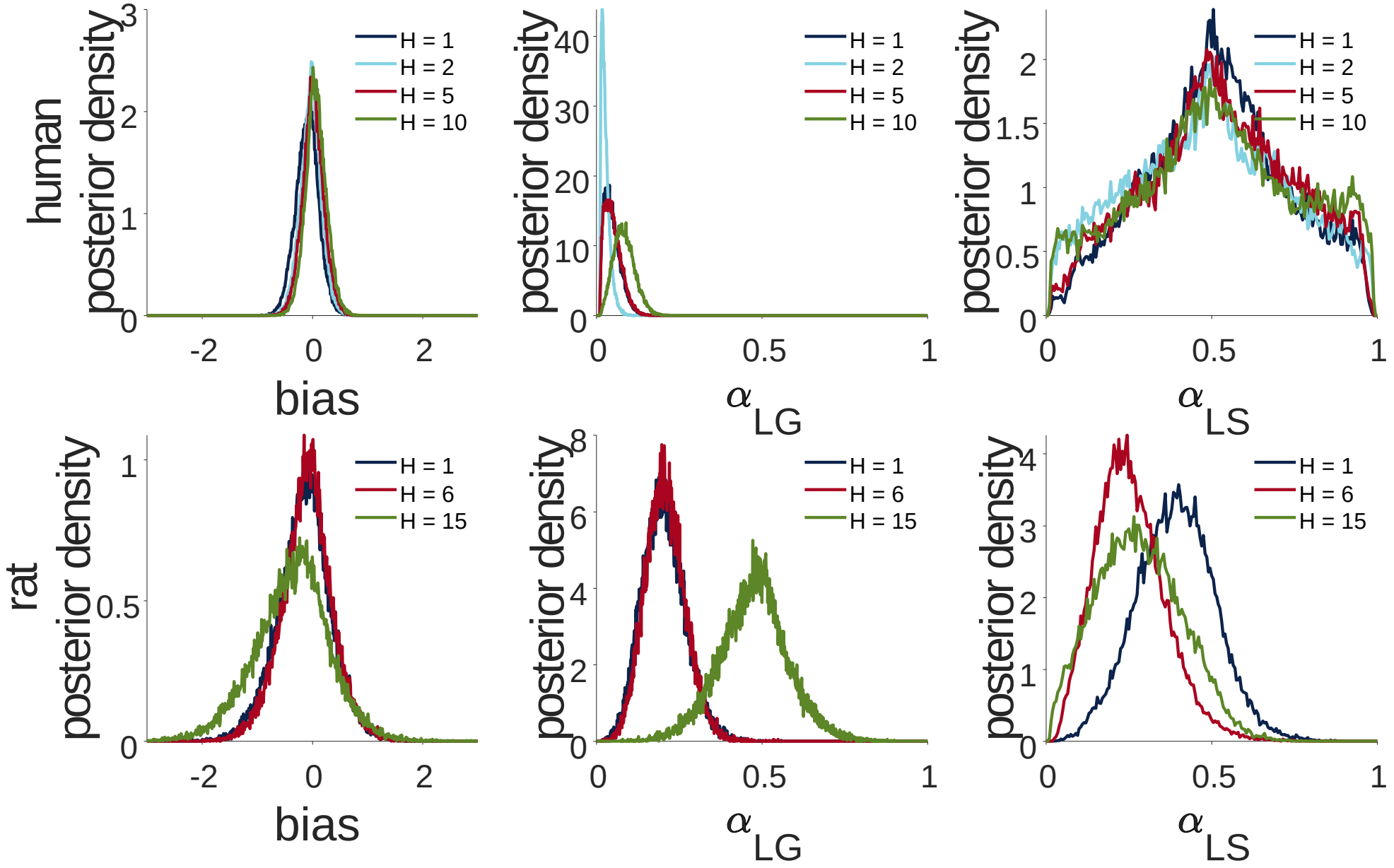

human
posterior density
bias
H = 1
H = 2
H = 5
H = 10
posterior density
α_{LG}
H = 1
H = 2
H = 5
H = 10
posterior density
α_{LS}
H = 1
H = 2
H = 5
H = 10
rat
posterior density
bias
H = 1
H = 6
H = 15
posterior density
α_{LG}
H = 1
H = 6
H = 15
posterior density
α_{LS}
H = 1
H = 6
H = 15

Figure S5

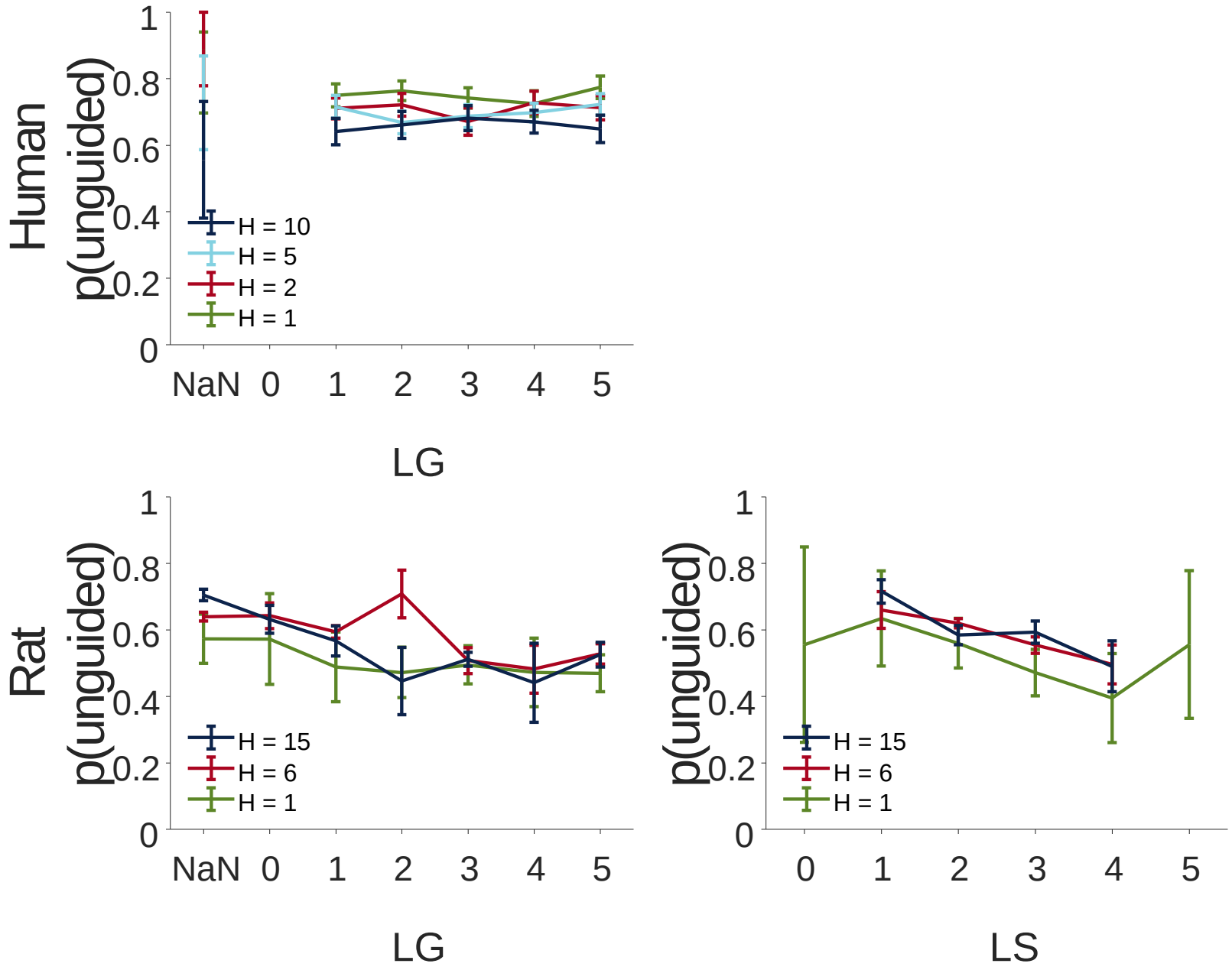

Human
p(unguided)
1
0.8
0.6
0.4
0.2
0
H = 10
H = 5
H = 2
H = 1
NaN 0 1 2 3 4 5
LG
Rat
p(unguided)
H = 15
H = 6
H = 1
NaN 0 1 2 3 4 5
LG
p(unguided)
H = 15
H = 6
H = 1
0 1 2 3 4 5
LS

Figure S6

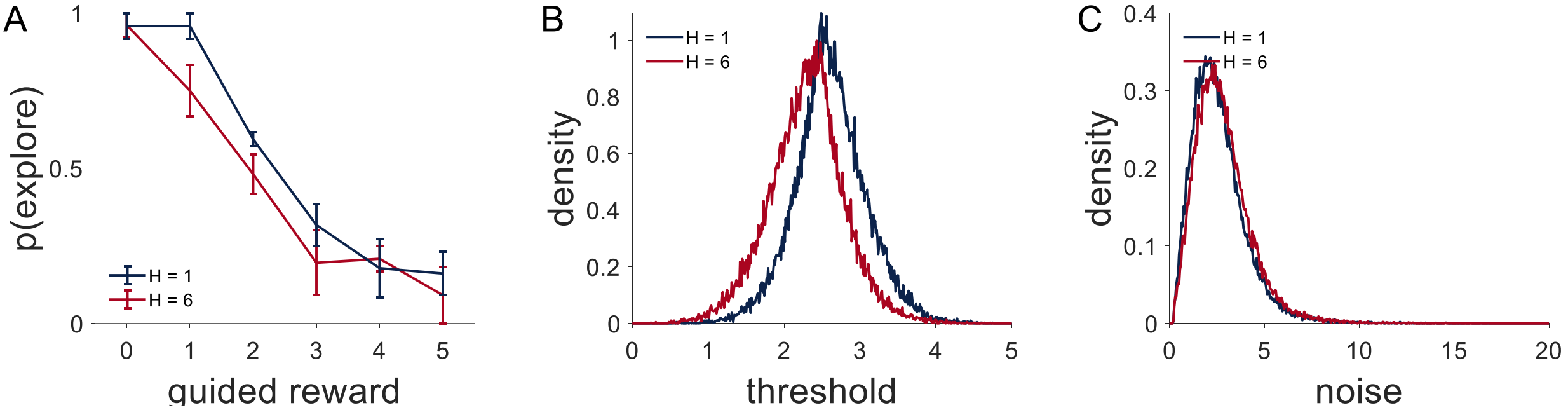

A
p(explore)
1
0.5
0
H = 1
H = 6
0
1
2
3
4
5
guided reward
B
density
1
0.8
0.6
0.4
0.2
0
H = 1
H = 6
0
1
2
3
4
5
threshold
C
density
0.4
0.3
0.2
0.1
0
H = 1
H = 6
0
5
10
15
20
noise

Figure S7

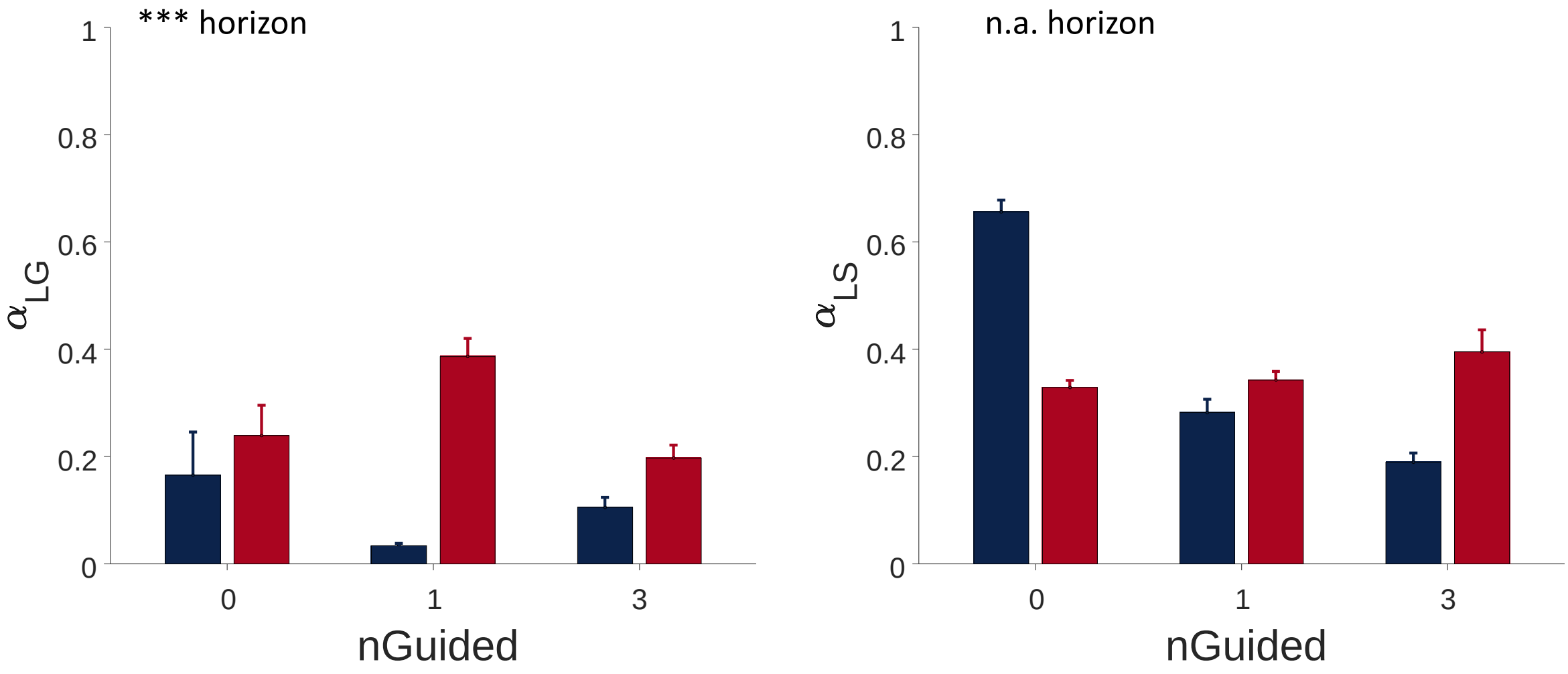
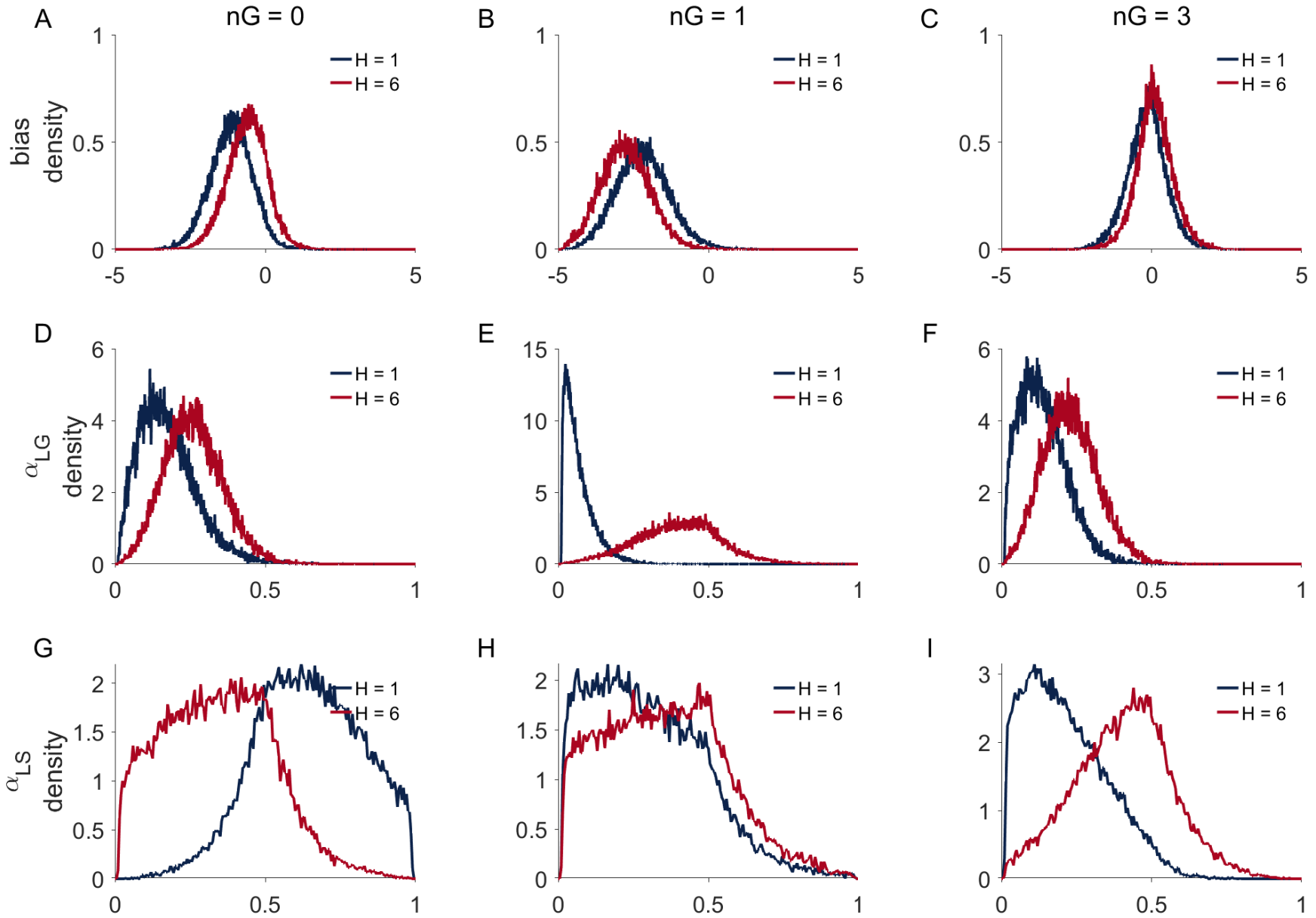

*** horizon
α_{LG}
1
0.8
0.6
0.4
0.2
0
0
1
3
nGuided
n.a. horizon
α_{LS}
1
0.8
0.6
0.4
0.2
0
0
1
3
nGuided

Figure S8


A
nG = 0
B
nG = 1
C
nG = 3
bias density
H = 1
H = 6
D
E
F
α_{LG} density
G
H
I
α_{LS} density