

# The role of adaptive decision noise in exploration

Robert C. Wilson, John M. White and Jonathan D. Cohen

## Summary

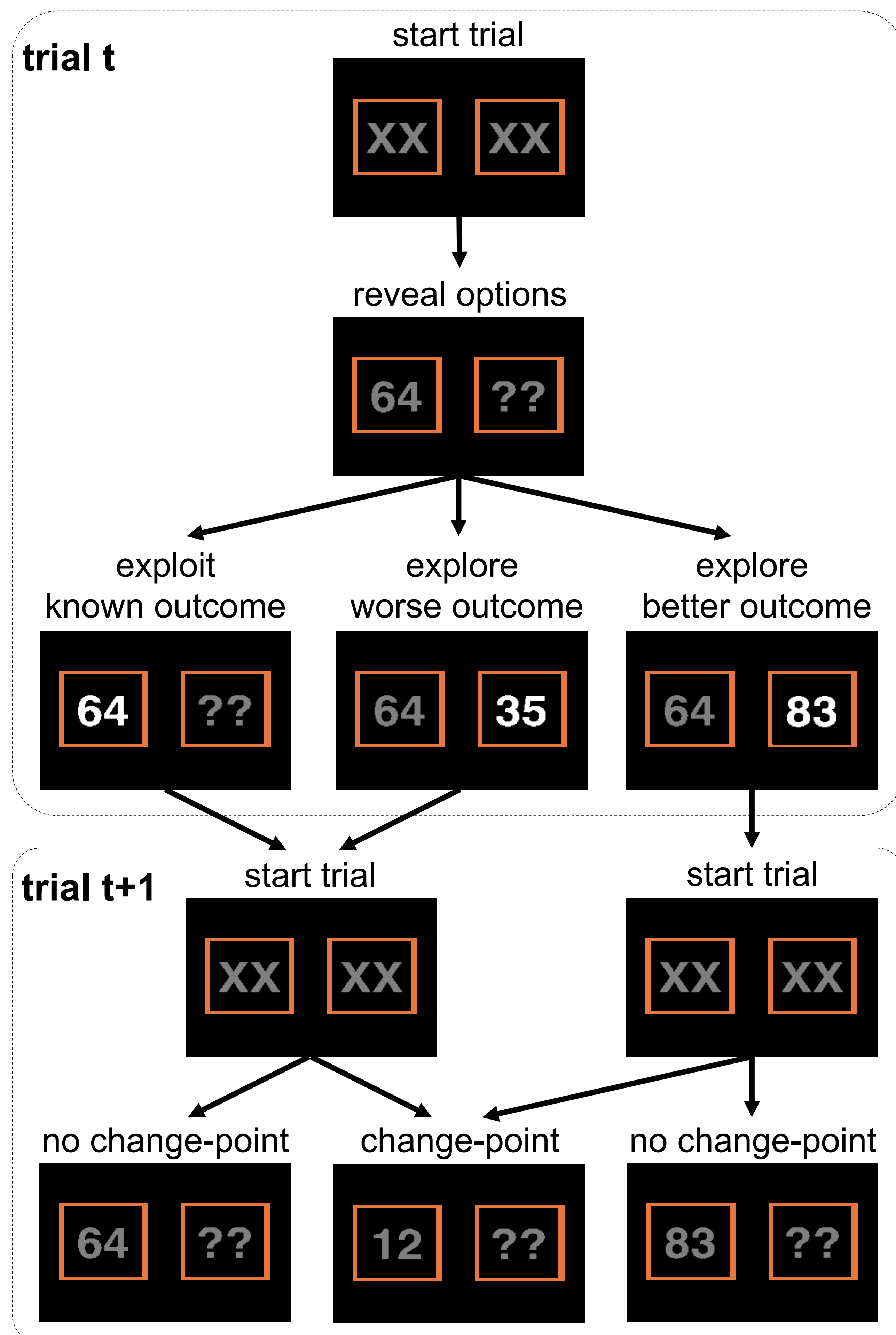
How do we solve the explore-exploit dilemma?

One idea is that we use decision noise to explore randomly ... here we ask

- 1) Is this a good idea? When is decision noise useful?
- 2) Do we adapt decision noise for exploration?

## Task

Bandits task with one known and one unknown option

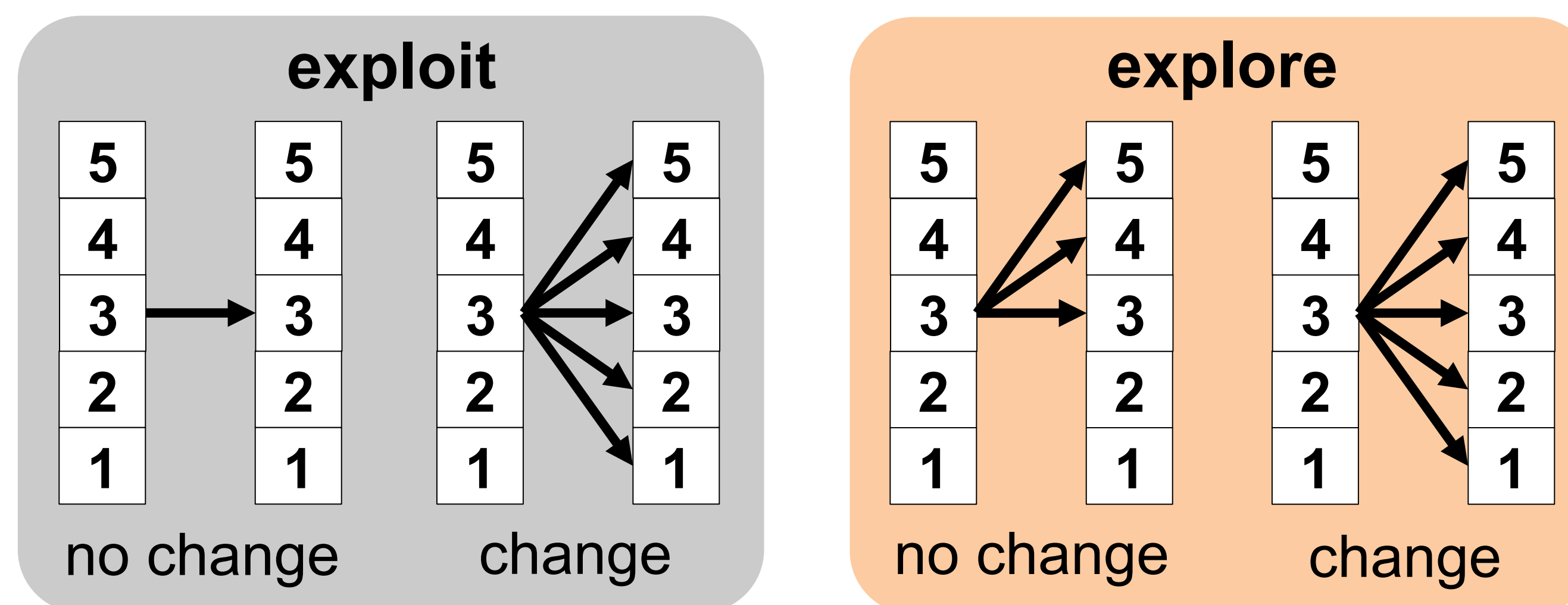


## Theory – optimal noise

State:  $s$  = value of known option

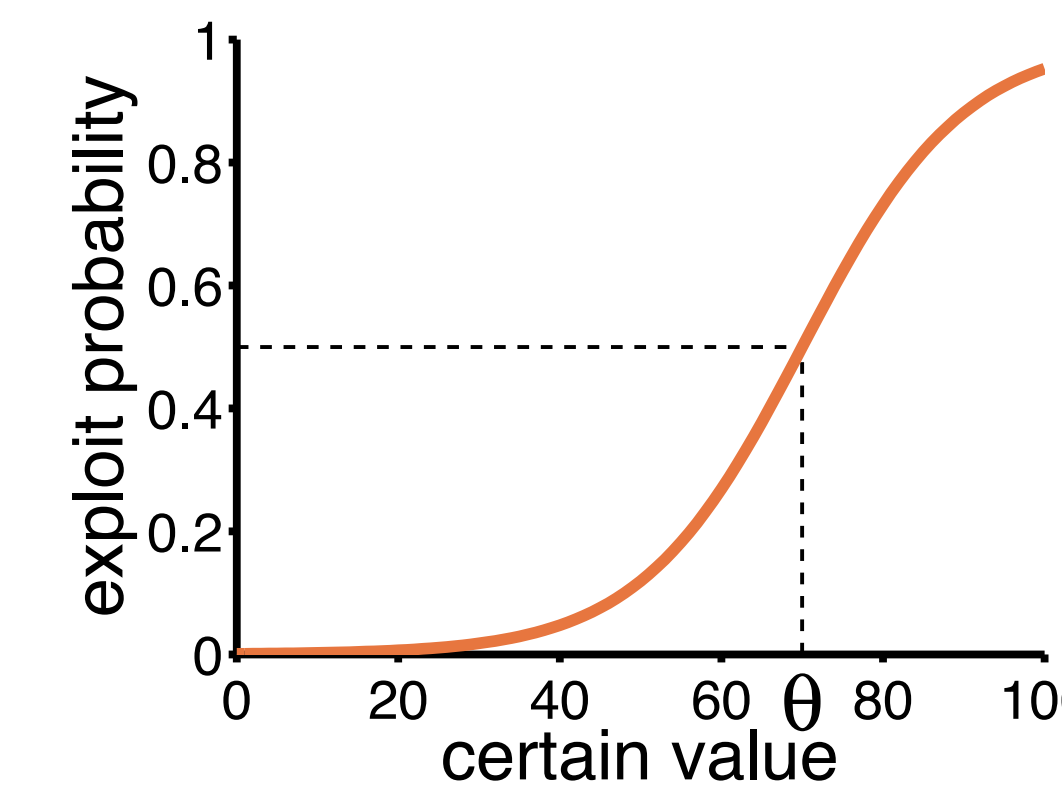
Therefore 100 possible states

Transition matrix:  $p(s' | s, a)$



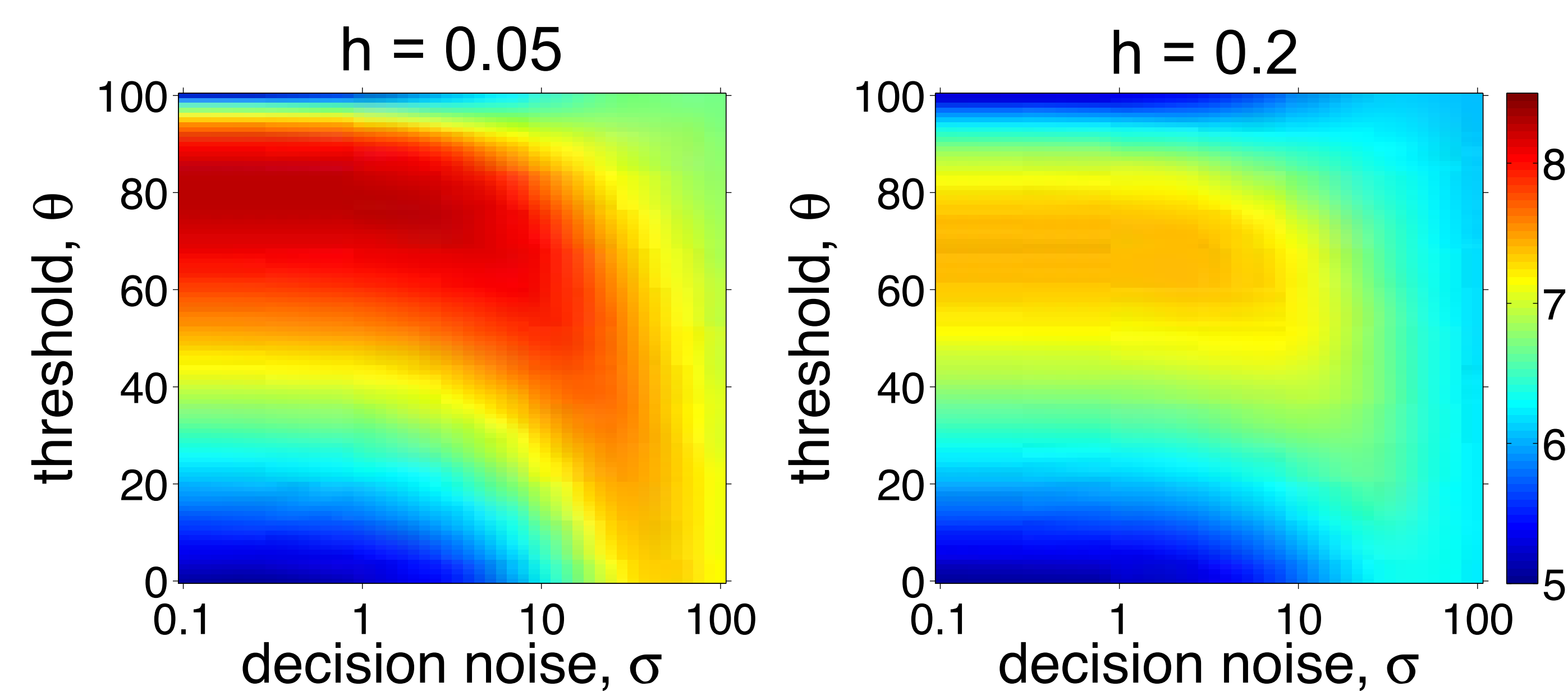
Noisy threshold policy:  $\pi(a | s)$

mostly exploit when certain value above  $\theta$   
otherwise mostly explore  
noise standard deviation is  $\sigma$

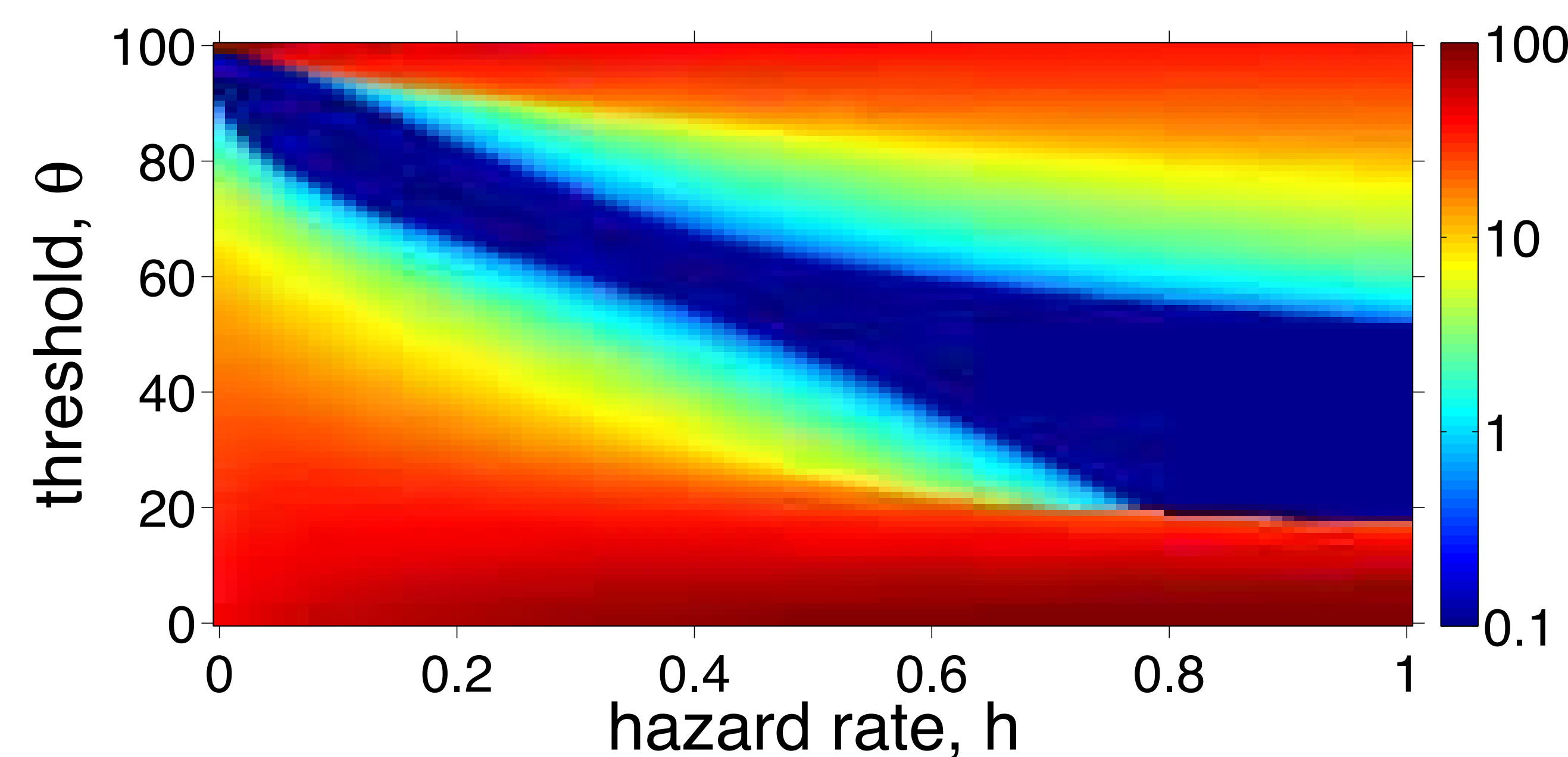


Policy evaluation:  $V(\theta, \sigma)$

Reward per trial under noisy threshold policy with different  $\theta$  and  $\sigma$

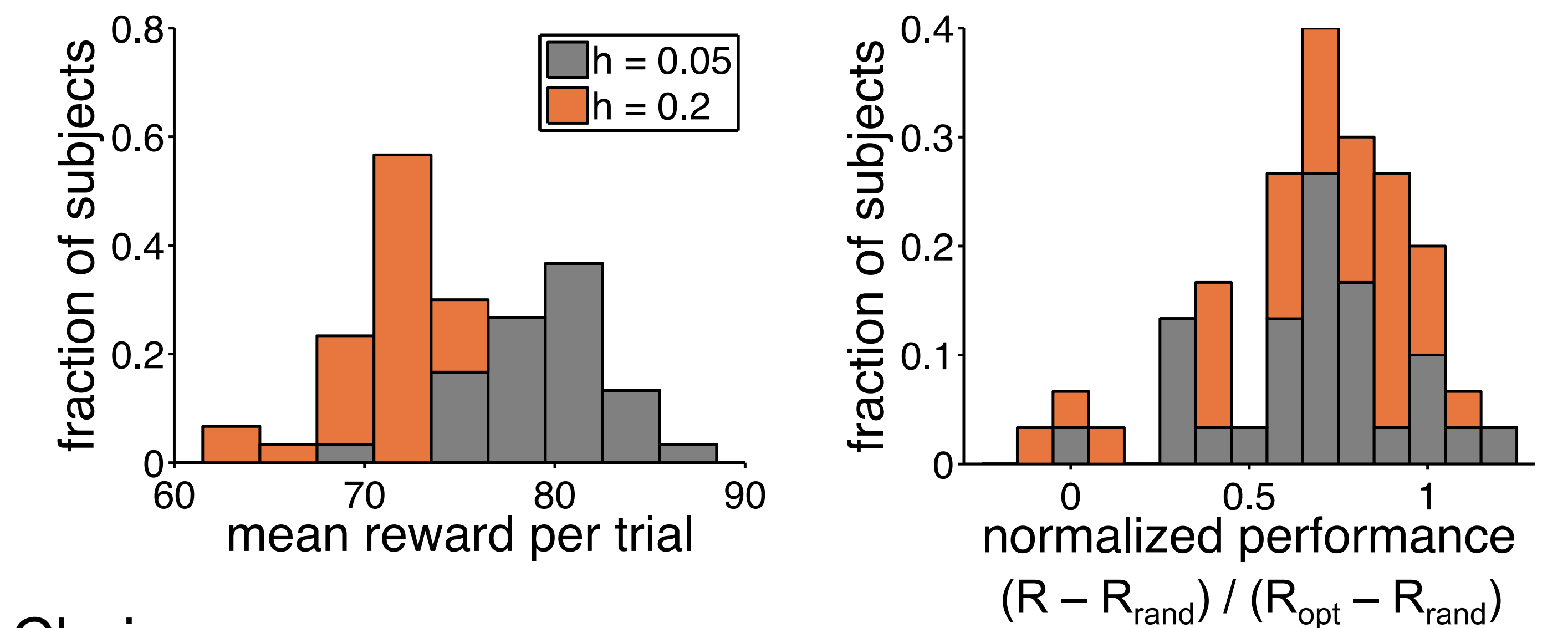


Optimal decision noise

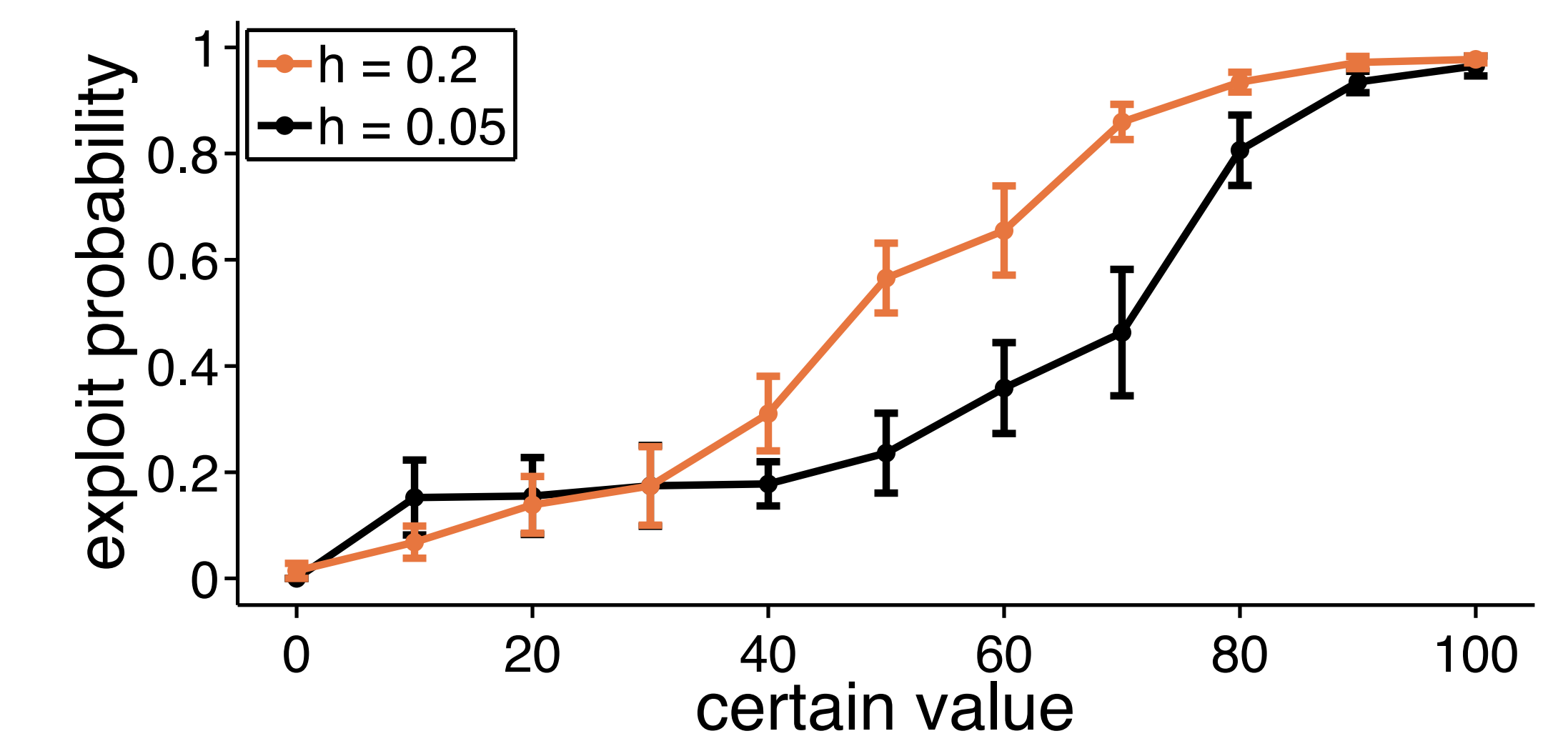


## Experimental results

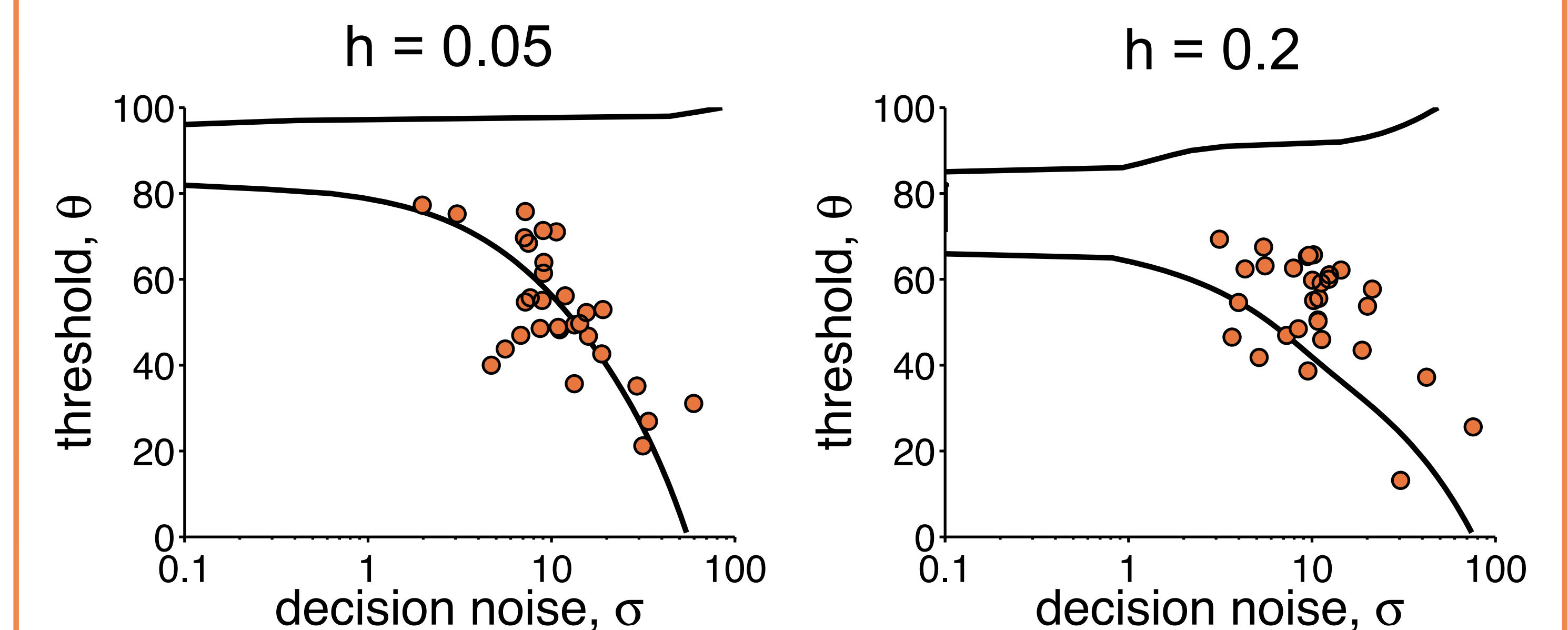
Performance



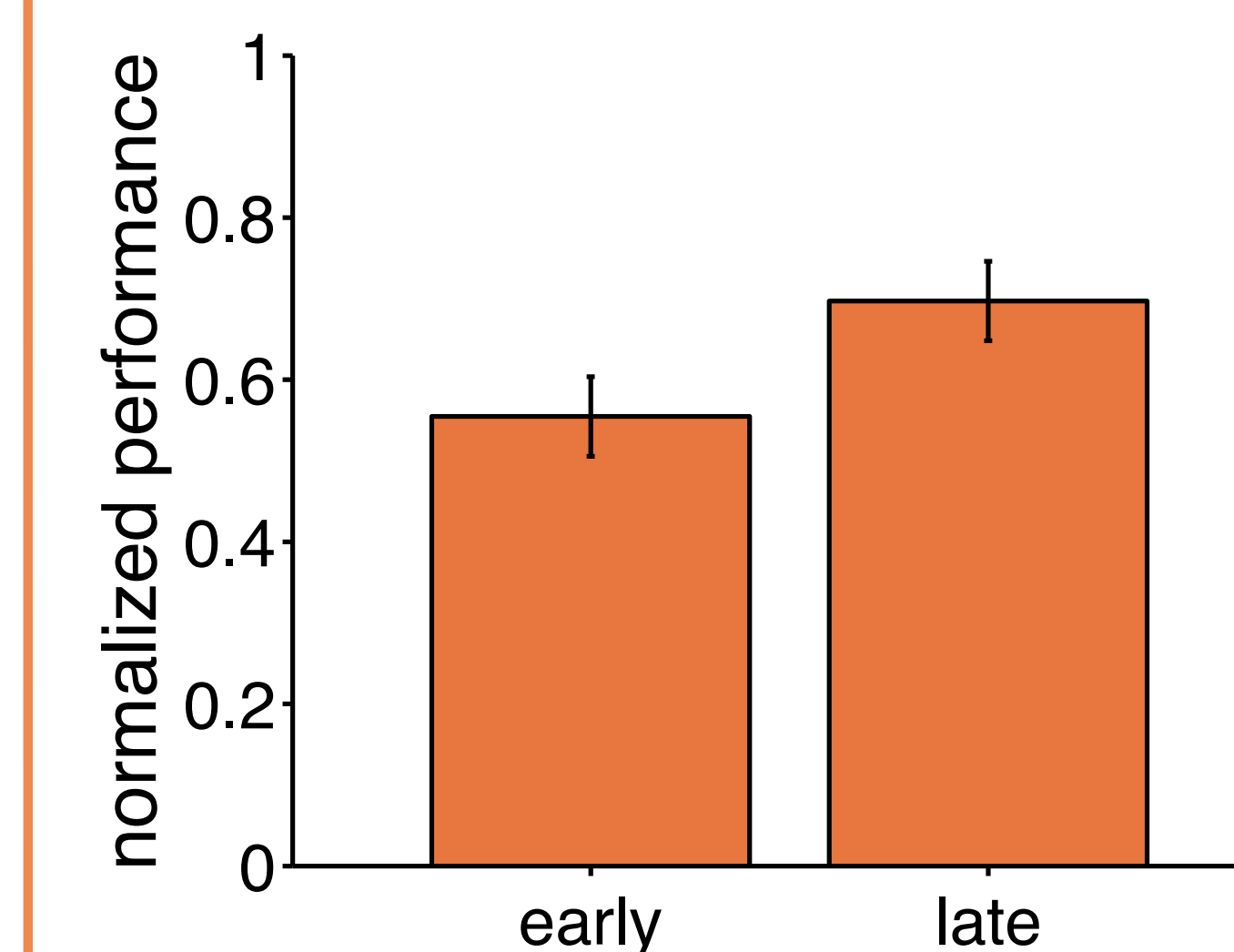
Choice curves



Threshold vs noise



Evidence for adaptation



Future work

Can we find evidence that noise adapts to hazard rate condition during block?

What is the effect of different reward dynamics like drift?