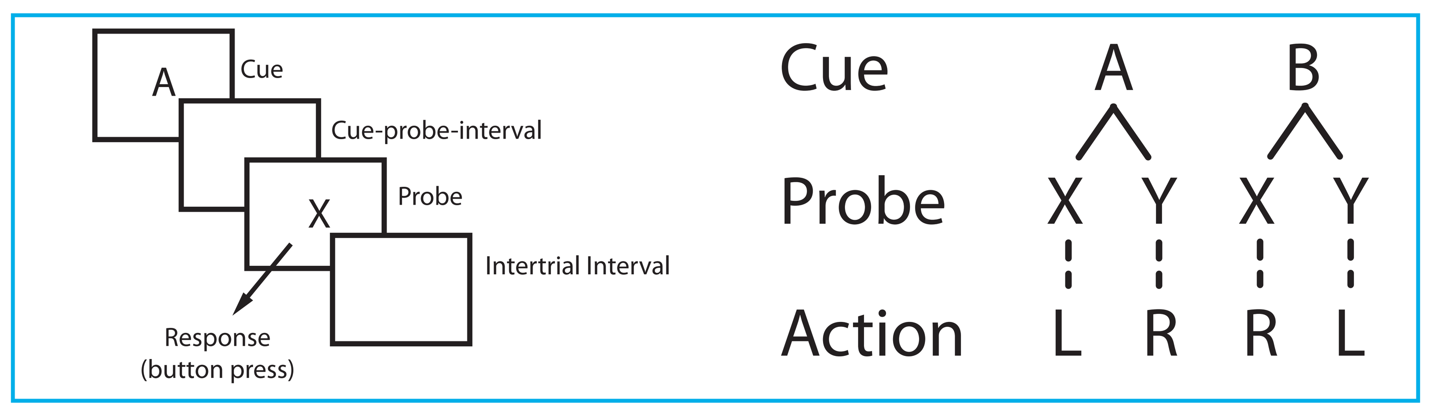
Bayesian model of proactive and reactive control in the AX-CPT

Olga Lositsky, Robert C. Wilson, John M. White, Jonathan D. Cohen
Princeton Neuroscience Institute and Department of Psychology at Princeton University

Abstract

The AX-Continuous Performance Test (AX-CPT) was one of the first theoretically-motivated tasks developed to probe the role of context processing in cognitive control (Cohen & Servan-Schreiber, 1992), and used to study deficits of these functions in schizophrenia (Servan-Schreiber, Cohen, & Steingard, 1996). Despite this corpus of work, our understanding of the cognitive processes engaged by the AX-CPT task remains incomplete. For instance, the Dual Mechanisms of Control (DMC) framework (2007) has proposed that there may be two strategies for performing this task: proactive control, which involves maintenance of context information in working memory, and reactive control, which relies on episodic memory. In the present work, we manipulated factors of motivation, response time and working memory load to bias subjects towards either proactive or reactive control. We then built a Bayesian model that captures performance on the AX-CPT using only two parameters. When fit to empirical data, only one of these parameters — memory noise — was found to differ between conditions designed to differentially engage the two modes of control. Our ongoing work aims to further elucidate whether reactive control involves distinct cognitive processes, or rather, reflects a deficit of proactive control.

The AX-CPT Task

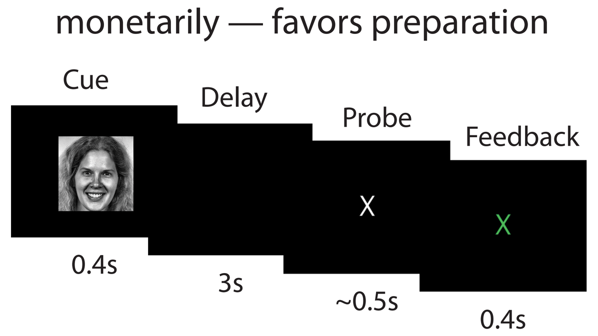


Proactive Control: At cue presentation, the task rule associated with the cue is represented as context information in working memory, and actively maintained during the delay period, biasing task processing pathways in preparation for an efficient response to the probe.

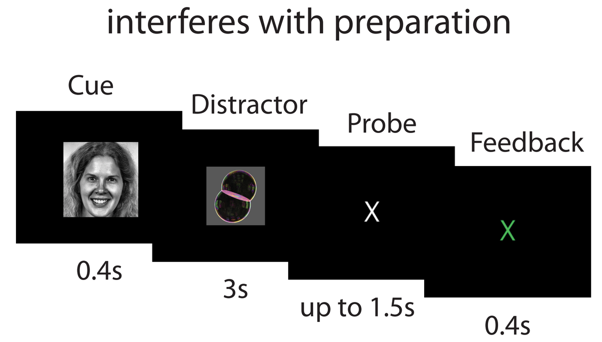
Reactive Control: The context is not held in working memory, but a trace of the cue remains (e.g. in episodic memory). At the time of probe, the cue representation is retrieved and used to activate the rule in working memory, allowing correct but less efficient responding.

Within-Subject Experimental Manipulations

Proactive bias condition: very fast & accurate responses are rewarded monetarily — favors preparation



Reactive bias condition: distractor task during delay between cue and probe — interferes with preparation



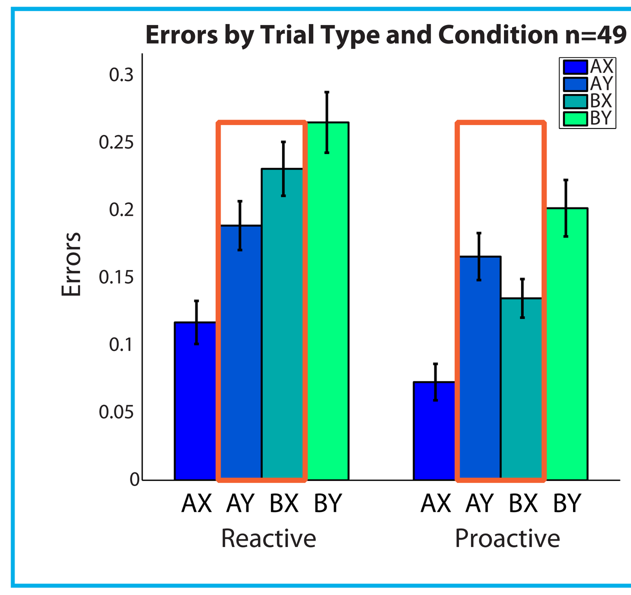
Varying trial frequencies induces biases that allow us to distinguish between strategies:

	Frequency	Action
AX	50%	Left
AY	20%	Right
BX	20%	Right
BY	10%	Left

- AY errors signal over-influence of the cue (A) relative to the probe (Y), suggesting use of proactive control
- BX errors signal over-influence of probe (X) relative to the cue (B), suggesting lack of preparation and use of reactive control
- AX and BY errors signal non-specific failures of processing

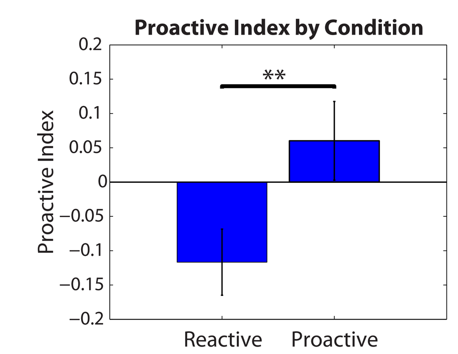
Behavioral Data

As expected, subjects are significantly more accurate and faster on the proactive-bias condition.



Errors in Proactive- vs. Reactive-Bias Conditions
- Decrease in AX and BY errors indicates general improvement in performance in proactive-bias condition.
- Reduction in BX errors relative to AY errors confirms greater influence of cue on responding in proactive-bias condition.

Proactive Index = (AY errors – BX errors) / (AY errors + BX errors)



Bayesian Model

We can approximate proactive and reactive patterns of behavior in terms of inferences about the identity of the cue and probe at the time of the response.

	A or B	X or Y	Left or Right
Actual stimulus	C = actual cue	O = actual probe	
Encoded stimulus	M = memory of cue	R = representation of probe	
Decoded stimulus	C′ = inferred cue	O′ = inferred probe	A = action

We assume that the memory of the cue and the perceptual interpretation of the probe are uncertain:
$M = 1 - \varepsilon M$; εM is the memory noise (cue noise).
$R = 1 - \varepsilon R$; εR is the perceptual noise (probe noise).

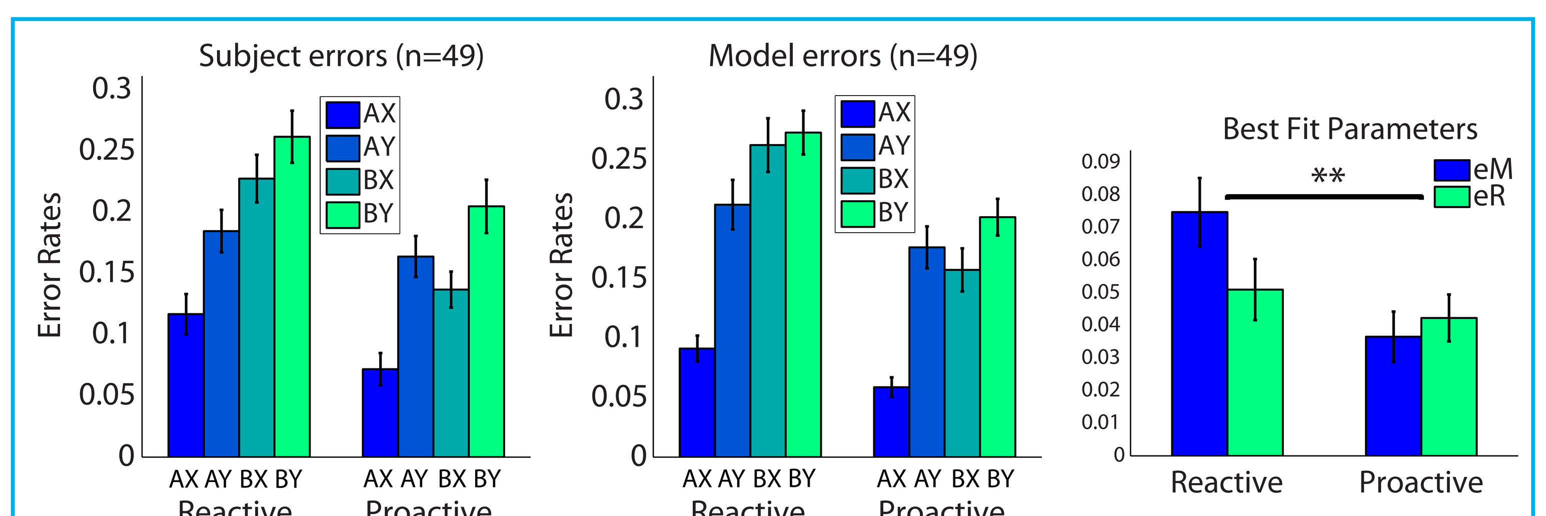
Due to this uncertainty, the agent acts based on inferences about the cue and probe (e.g. C′O′=AX), rather than the actual cue and probe (e.g. CO=BX).

$$p(a|c,o) = \sum_{c'o'} p(a|c',o')\, p(c',o'|c,o)$$

These inferences are shaped by cue noise (εM), probe noise (εR) and trial frequencies, which constitute the prior probability of each cue-probe combination, p(C′O′):

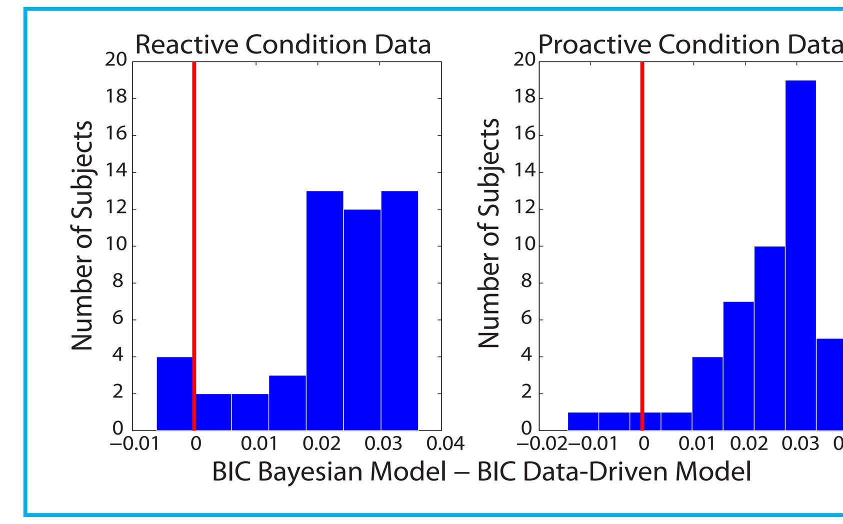
$$p(a|c,o) \propto \sum_{c'o'} p(a|c'o') \sum_{M,R} \underbrace{p(c'o')}_{\text{trial frequency}}\, \underbrace{p(M|c')}_{1-\varepsilon M}\, \underbrace{p(R|o')}_{1-\varepsilon R}\, \underbrace{p(M|c)}_{1-\varepsilon M}\, \underbrace{p(R|o)}_{1-\varepsilon R}$$

Given knowledge of the trial frequencies, and using the best fit parameters for εM and εR from our behavioral data, the model provides us with the probability that the agent will go Left or Right for a particular trial type, which can be converted to error rate by trial type.



As expected, there is a highly significant decrease in memory noise (more reliable influence of the cue) in the proactive-bias condition. However, perceptual noise (eR) does not differ significantly across conditions.

Assessing Goodness of Fit



We compared our Bayesian model against the most accurate model possible:
Data-driven Model that learns each subject's probability of going Left for each trial type.

It has 4 parameters:
p(L|AX) p(L|AY) p(L|BX) p(L|BY) and can replicate each subject's error rates perfectly.

The normalized BIC of the Bayesian model was higher than the normalized BIC of the data-driven model for almost every subject.

Normalized BIC for:	Bayesian Model	Data-Driven Model
Reactive Bias Condition	0.68 (SD=0.11)	0.66 (SD= 0.11)
Proactive Bias Condition	0.74 (SD=0.10)	0.71 (SD=0.10)

Wilcoxon signed rank test:
$p < 10^{-8}$ for both conditions.

Predictions of Bayesian Model

Simulations of the model suggest that some AX-CPT variants provide cleaner estimates of the Proactive Index than others. Variants differ in terms of response rules and trial frequencies.

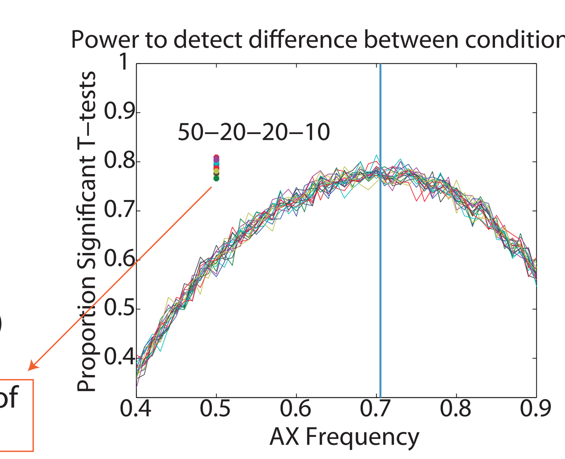
Response Rules
Most analogous studies (Braver et al., 2009; Edwards, Barch, & Braver, 2010) have used an asymmetric response rule.
The model predicts that using an asymmetric response rule will yield cleaner estimates of the proactive index.

Symmetric	Asymmetric
AX = Left	AX = Left
AY = Right	AY = Right
BX = Right	BX = Right
BY = Left	BY = Right

Trial Frequencies

BY Trials
The model recommends excluding BY trials when the response rule is symmetric (L - R - R - L) because that reduces the task to the asymmetric rule (L - R - R).

AX Trials
Given the εM and εR parameters fit to the subject data, and the number of trials per subject, we can estimate how the power of detecting a difference in Proactive Index between conditions varies as a function of trial frequencies.
We hold AY, BX and BY trial frequencies equal, varying AX frequency.
Note 1: Trade-off between strength of dominant response (to AX) and number of AY / BX trials needed to estimate Proactive Index.
Note 2: the 50-20-20-10 design appears more powerful because of the lower BY trial frequency.

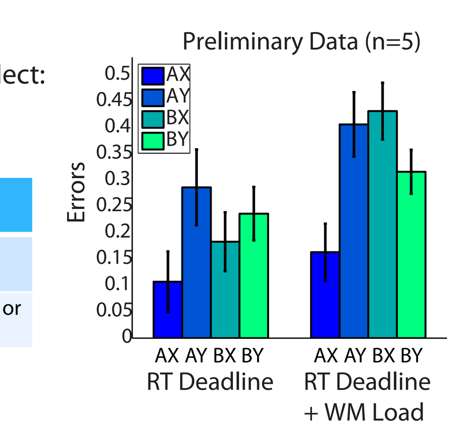


Meaning of Cue Noise and Reactive Control

Here, proactive- and reactive-bias conditions differed only in cue noise:
As the memory noise increased, overall performance decreased. Does this pattern reflect:
1. Genuine reliance on reactive control?
2. A failure of proactive control / working memory -> more automatic processing?

	Increase WM Load	Proactive Interference
Proactive-Bias Condition: Response Time Deadline	$\boxed{\text{Proactive Index decreases; Impaired performance on all trials}}$	No effect
Reactive-Bias Condition: No Deadline, Easy Distractor Task	Proactive Index decreases All other trials unaffected	Subjects become more proactive or Performance decreases

Preliminary data shows reduced performance and a relative increase in BX errors.



Future Modeling Directions

Model the temporal dynamics in the choice data: How do subjects learn trial frequencies over time? Do attentional strategies to cue / probe / combination shift over time?

Model Reaction Time data using the Drift Diffusion Model (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006) and use the memory noise and perceptual noise parameters to constrain the DDM parameters.