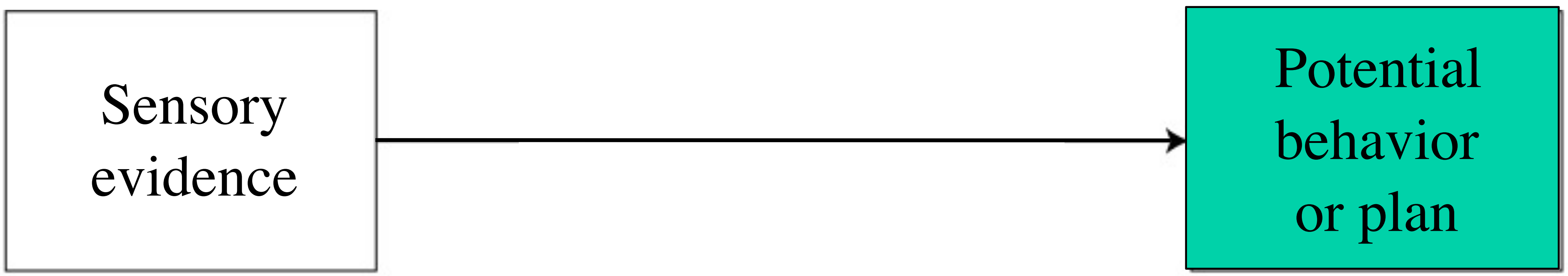
The Neurobiology of Decision Making

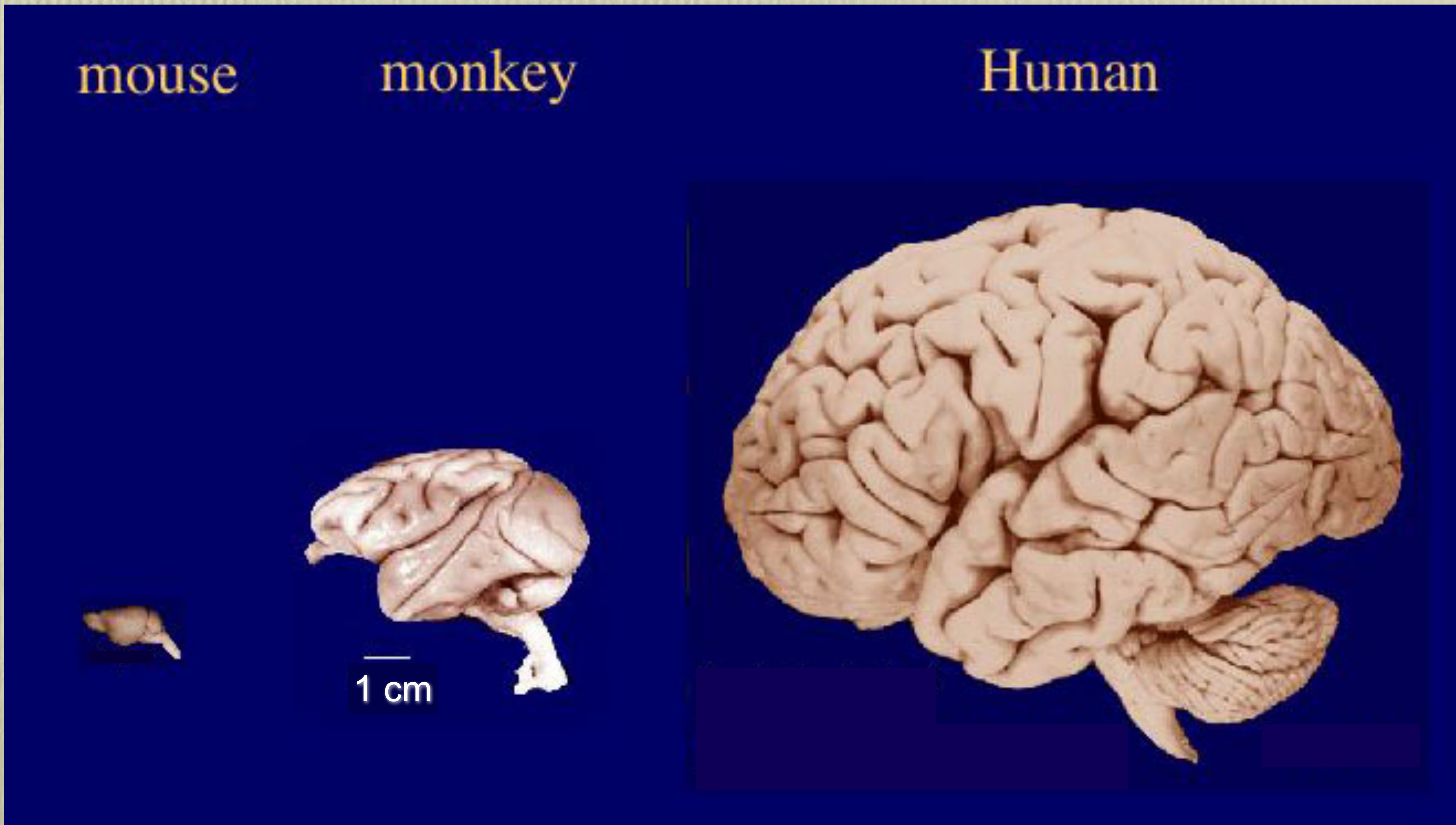
CSE 528
May 31, 2011

Michael Shadlen, MD PhD
Howard Hughes Medical Institute
Department of Physiology & Biophysics
National Primate Research Institute
University of Washington
Seattle, WA

www.shadlen.org

From sensorimotor integration to cognition




mouse
monkey
Human
1 cm

Outline

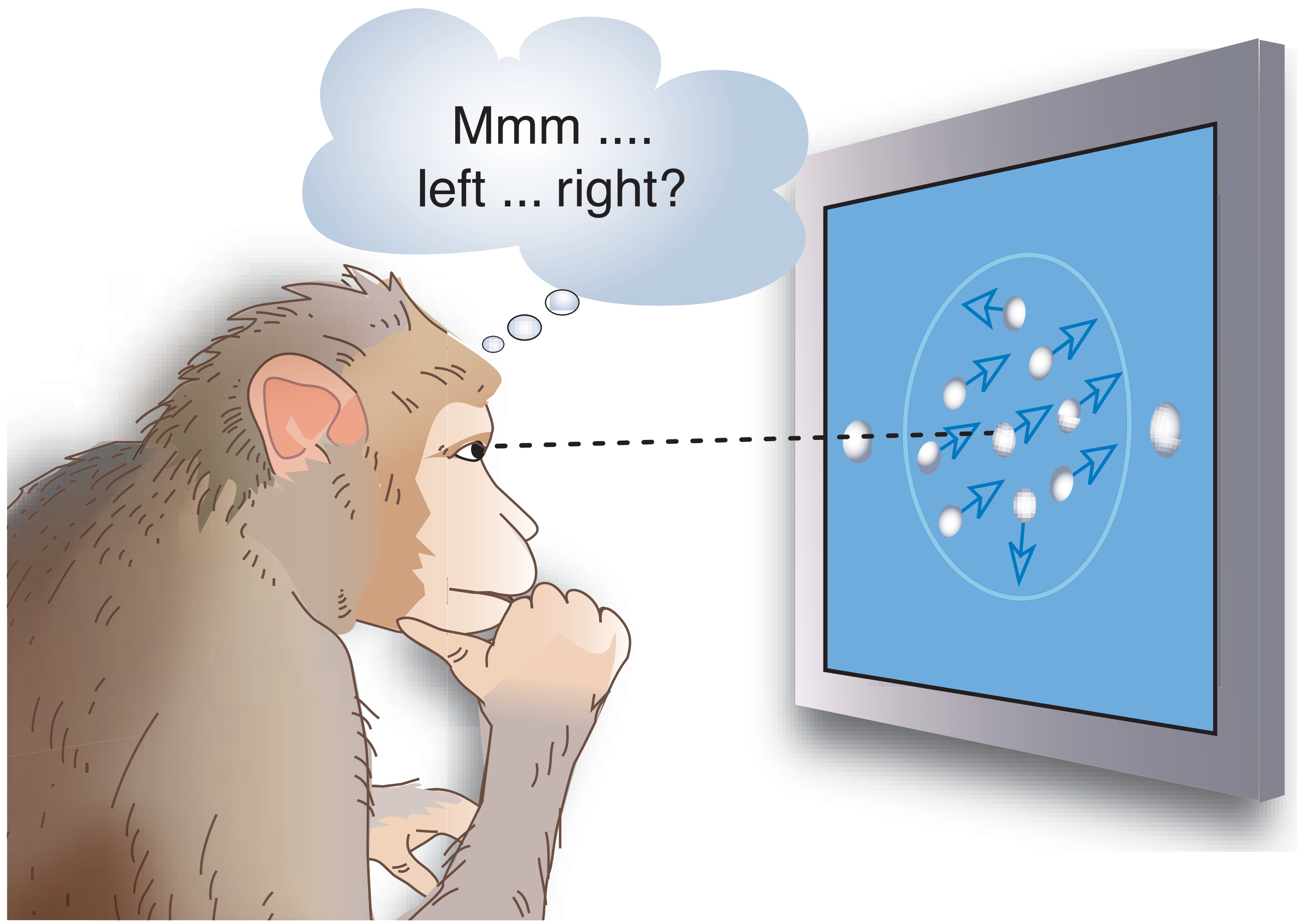
1. Probabilistic reasoning
2. Perceptual decisions: speed and accuracy
3. *Optional*: Sequential analysis, Wald's martingale, logistic choice function
4. Variance and covariance as signatures of neural computation
5. Confidence
6. *Optional*: Integration of prior probability & evidence

What is a decision?

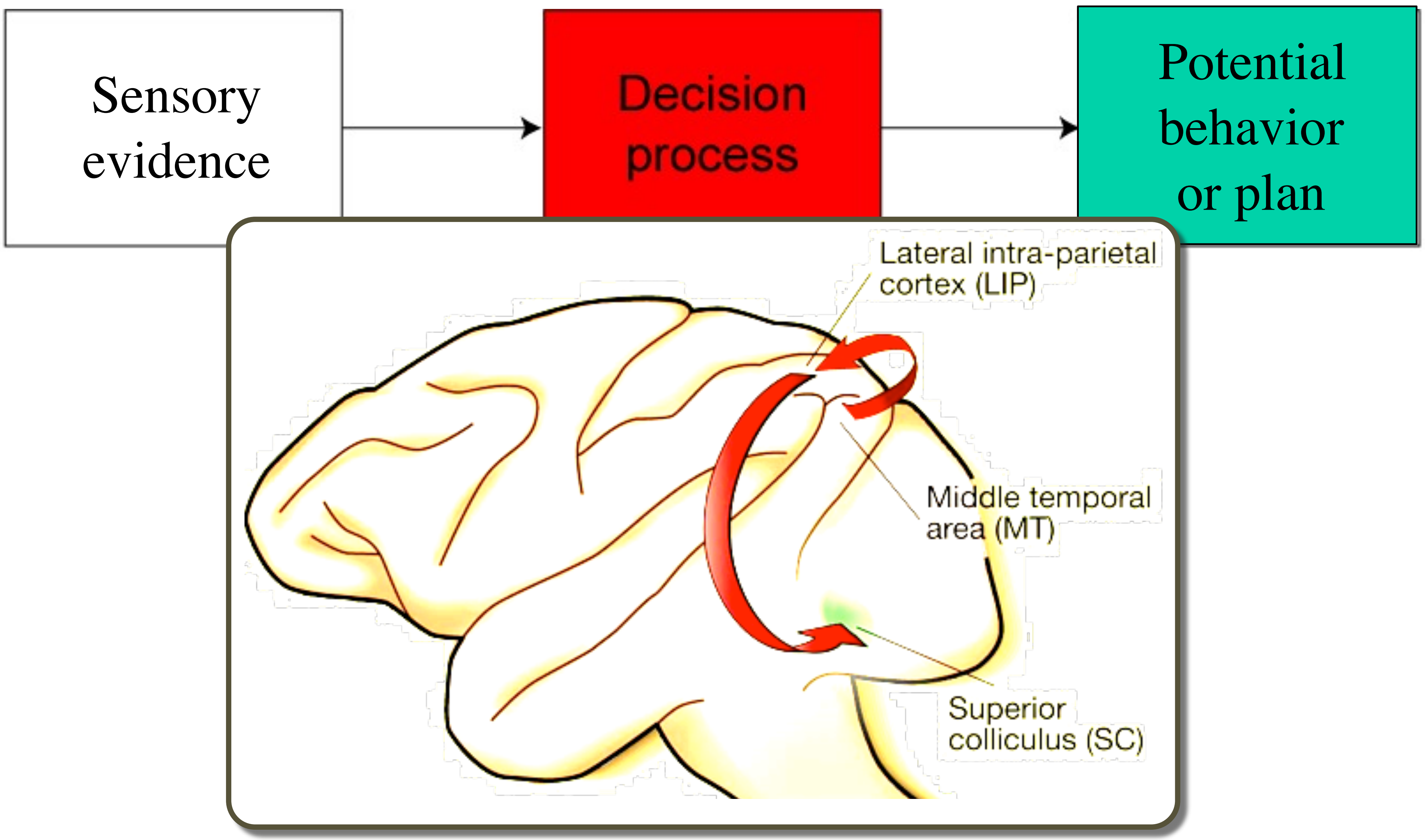
- A commitment to a proposition or plan of action, ...
  - based on evidence, prior knowledge, payoff, urgency
  - often requiring flexibility, contingency, interpretation

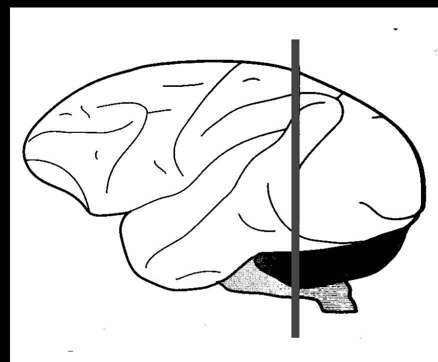
Some complex decisions

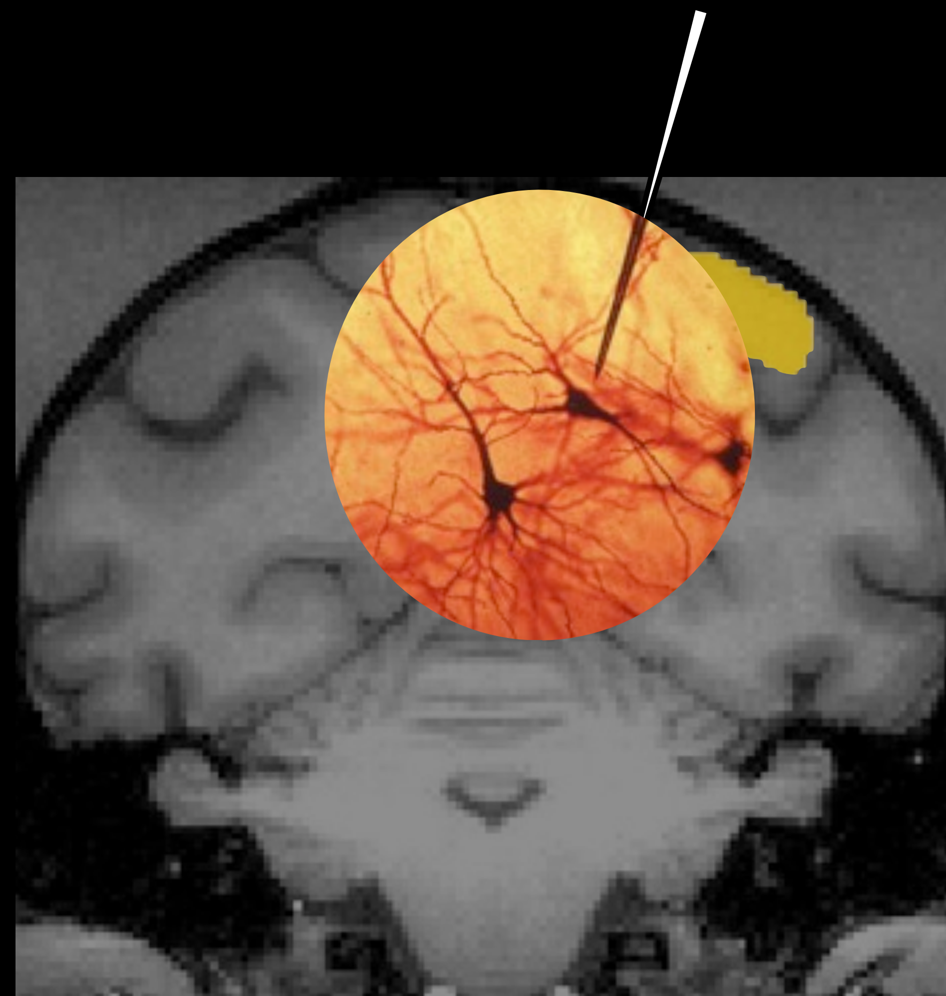
- Choosing a life partner
- Choosing a president
- Whether to invade Iraq


Mmm ....
left ... right?

From sensorimotor integration to cognition



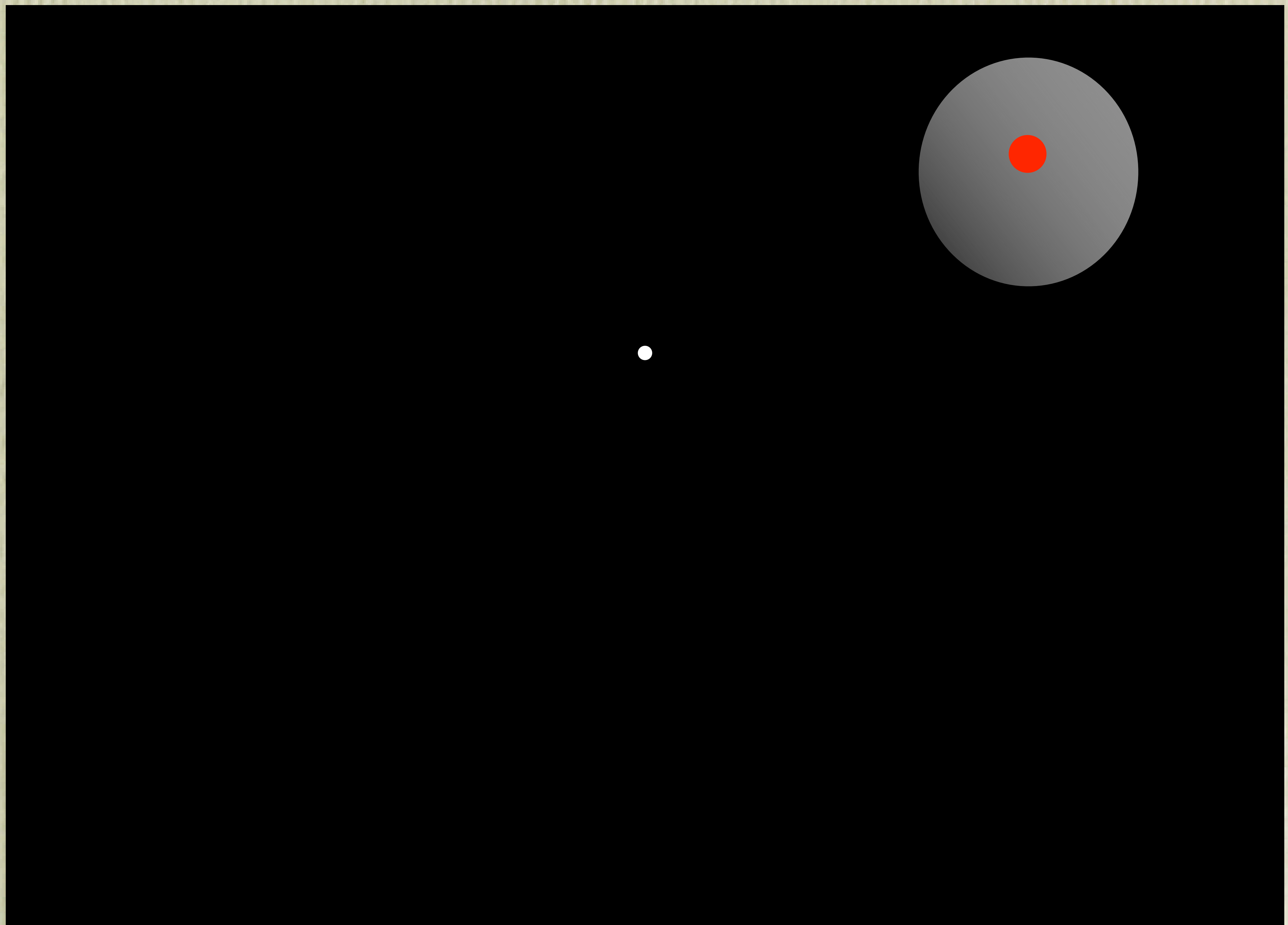




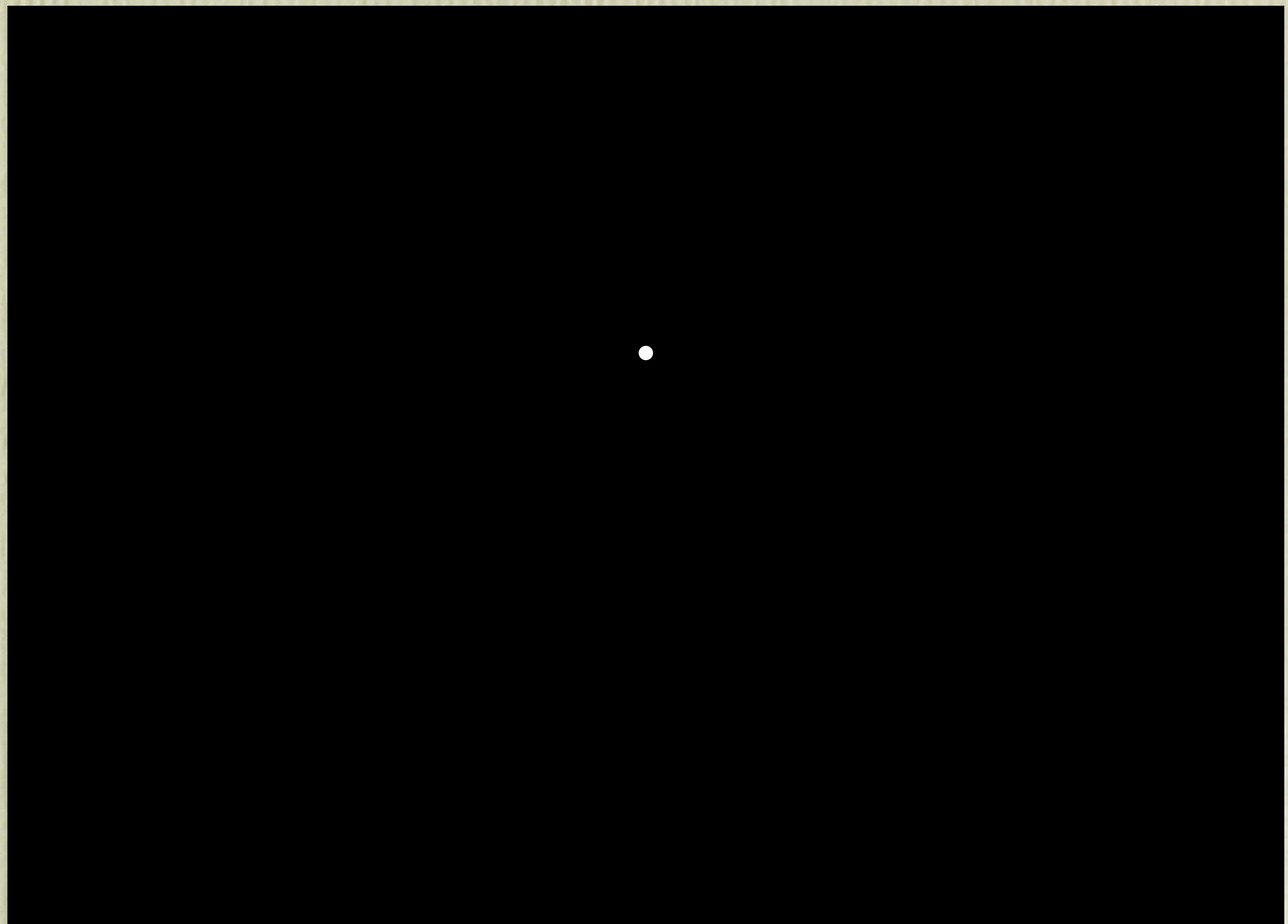
LIPv

Lewis and Van Essen, 2000

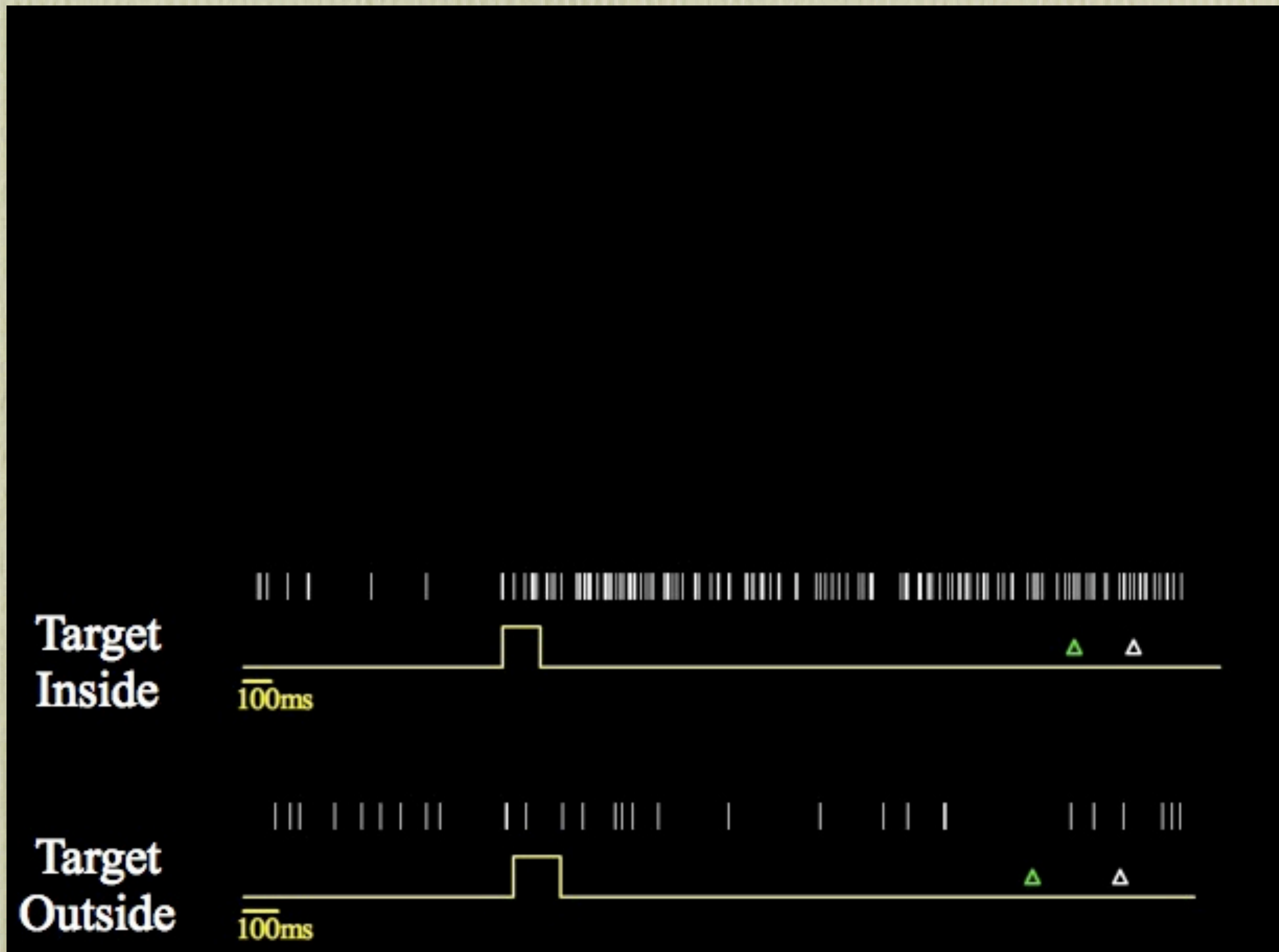
Spatially selective, persistent activity



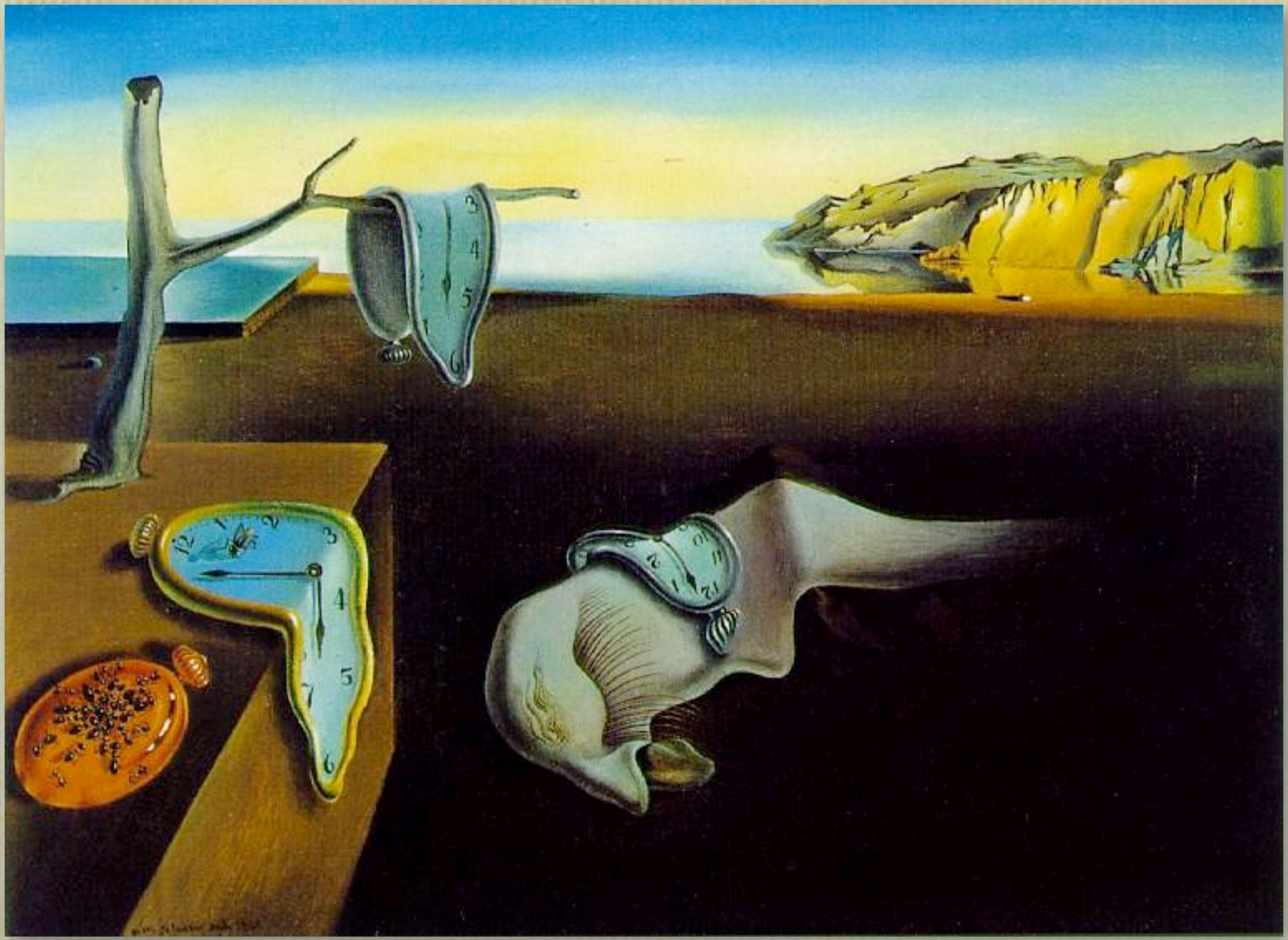
Spatially selective, persistent activity



Spatially selective, persistent activity



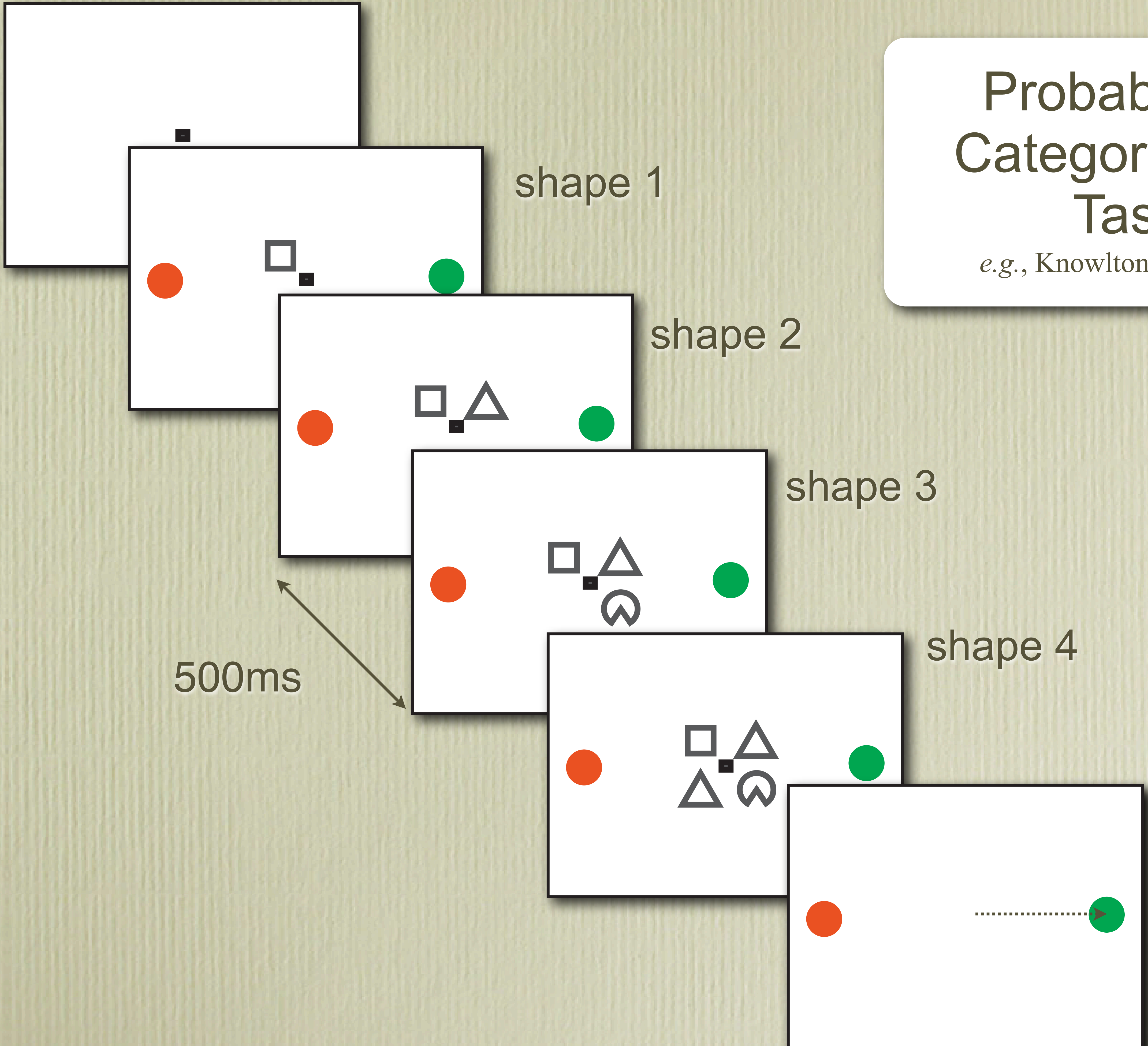
Freedom From Immediacy



S. Dali

Probabilistic Categorization Task

e.g., Knowlton *et al*, 1996



Tianming Yang

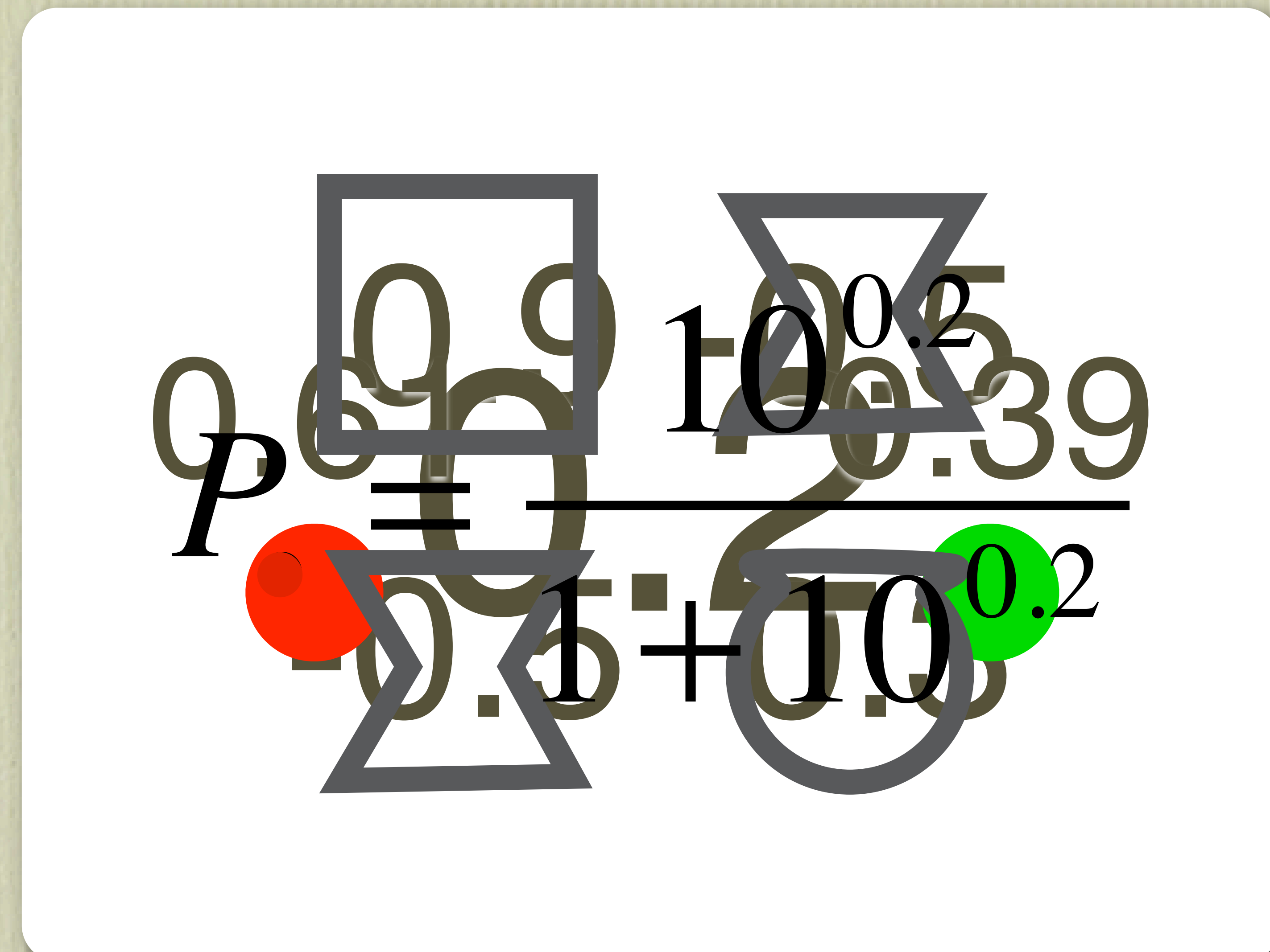
Yang & Shadlen (2007)
Nature 447: 1075-1080

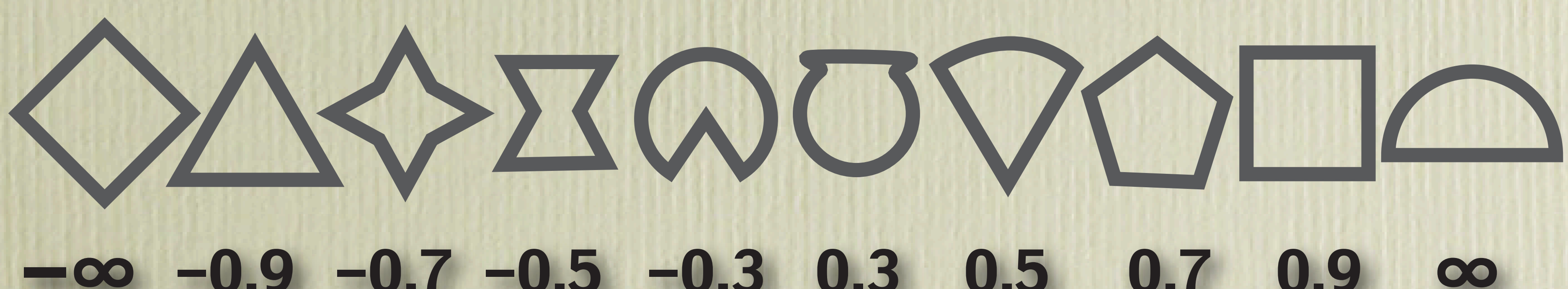
- 10 different shapes w/ different weights
- 4 shapes in a trial,drawn randomly with replacement
- each shape appears with equal probability
- sum of the weights is log odds in favor of red:

$$\log_{10} \frac{P(\text{red} \mid \text{shapes})}{P(\text{green} \mid \text{shapes})} = \text{logLR}$$

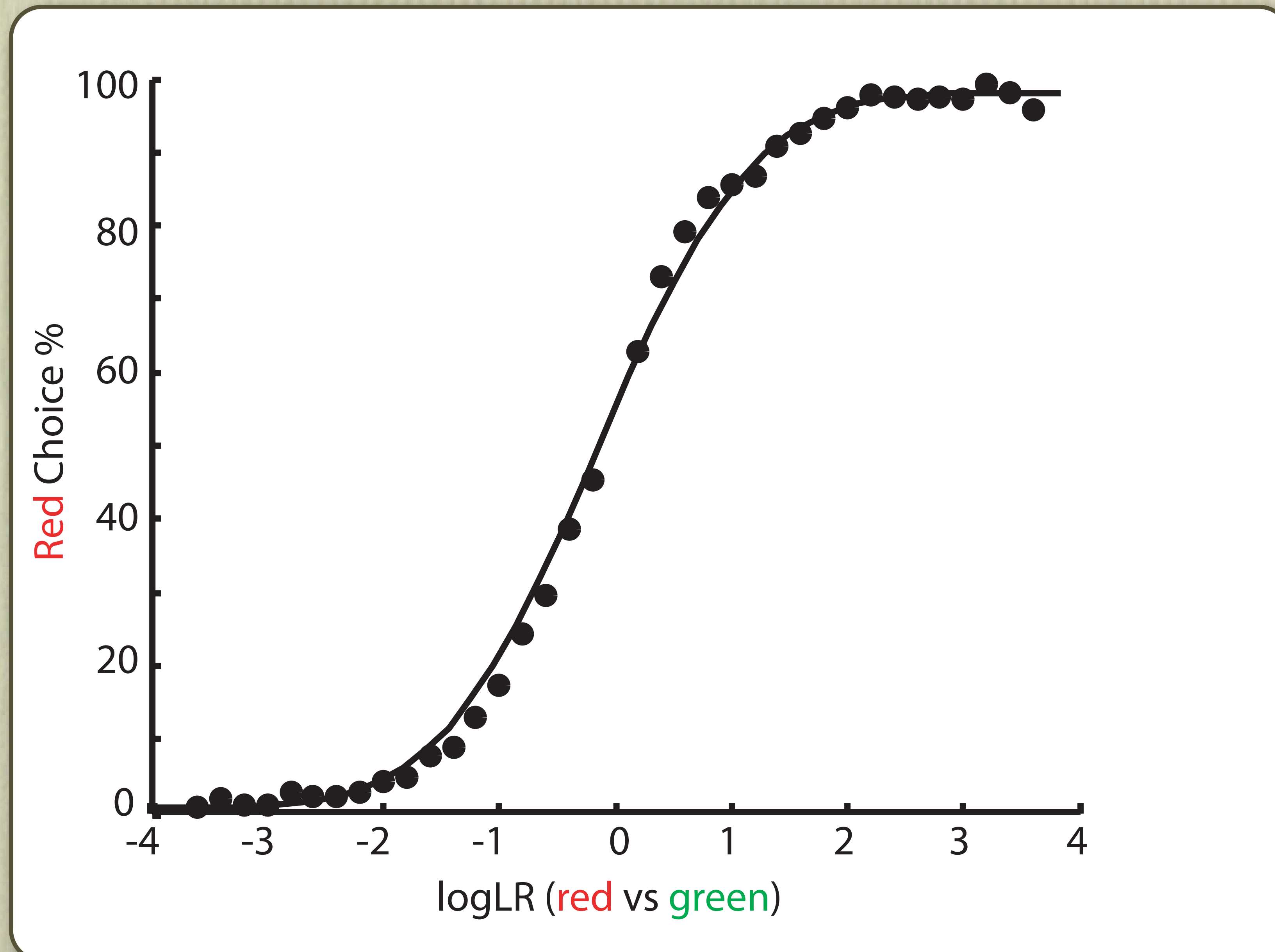
Yang & Shadlen (2007)
Nature 447: 1075-1080

Probabilistic Categorization



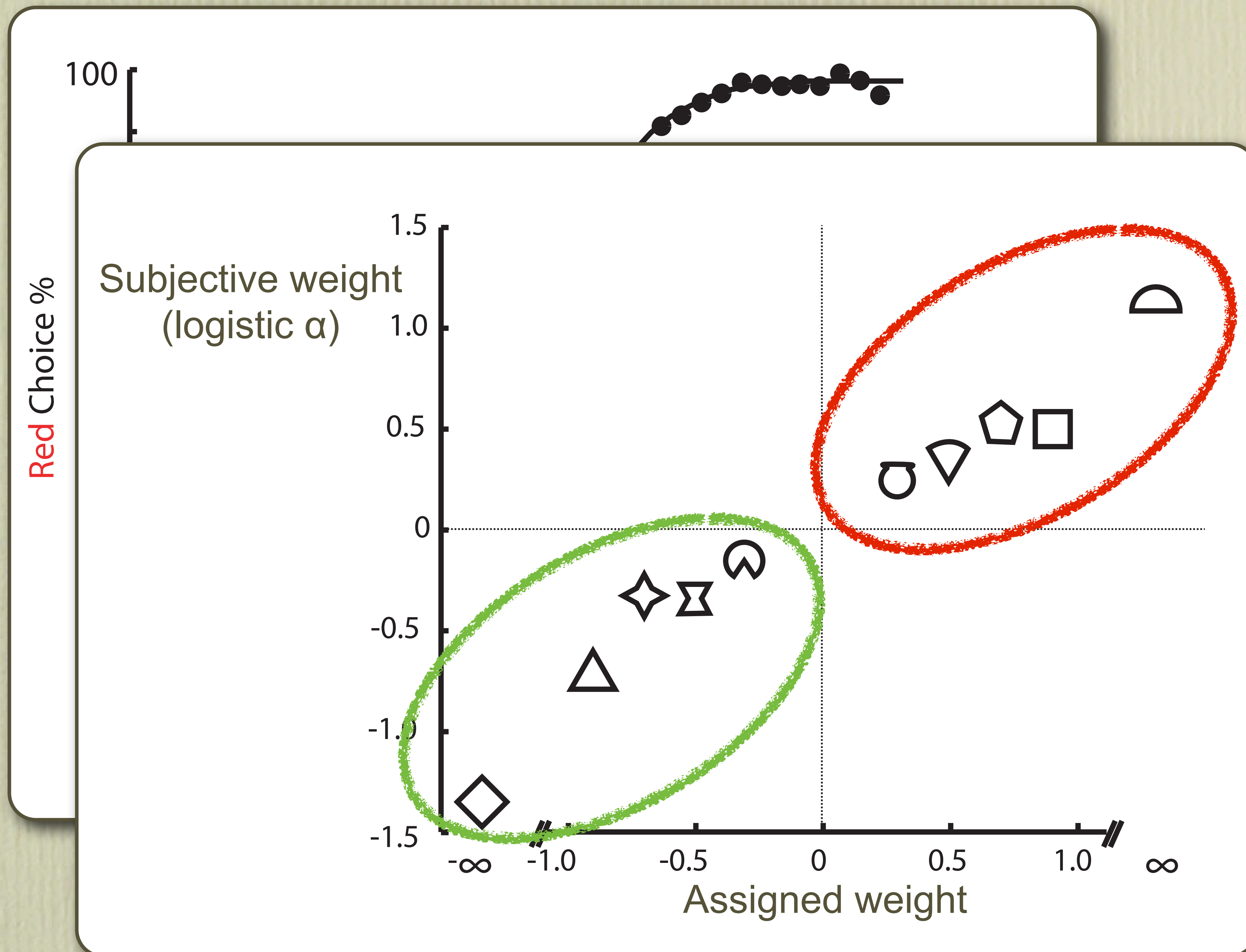


Yang & Shadlen (2007)
Nature 447: 1075-1080



$$P_{\bullet} = \frac{1}{1+10^{-\left(\alpha_{\Diamond} n_{\Diamond} + \alpha_{\triangle} n_{\triangle} + \alpha_{\diamond} n_{\diamond} + \ldots + \alpha_{\frown} n_{\frown}\right)}}$$

Yang & Shadlen (2007)
Nature 447: 1075-1080



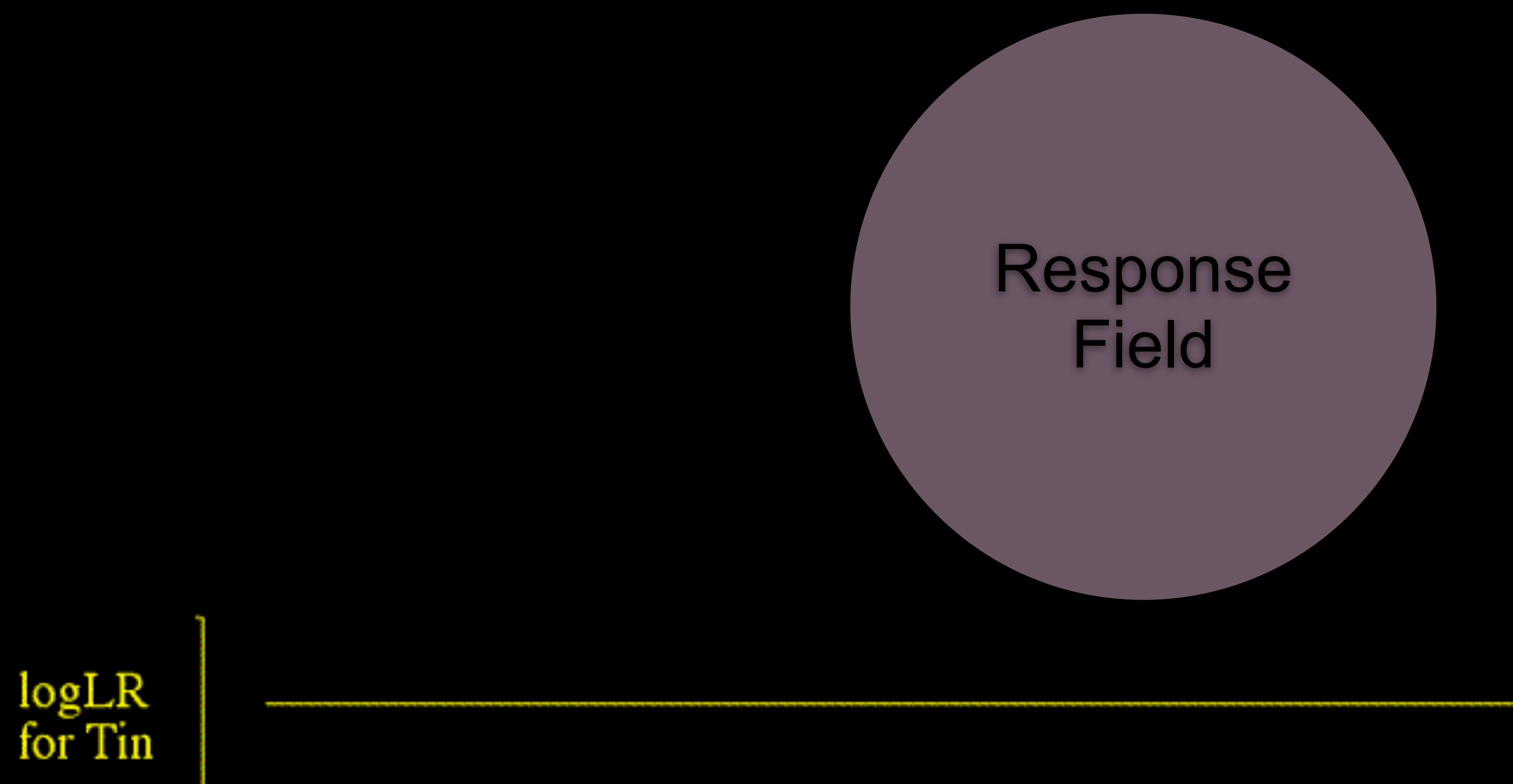
Fit with logistic function

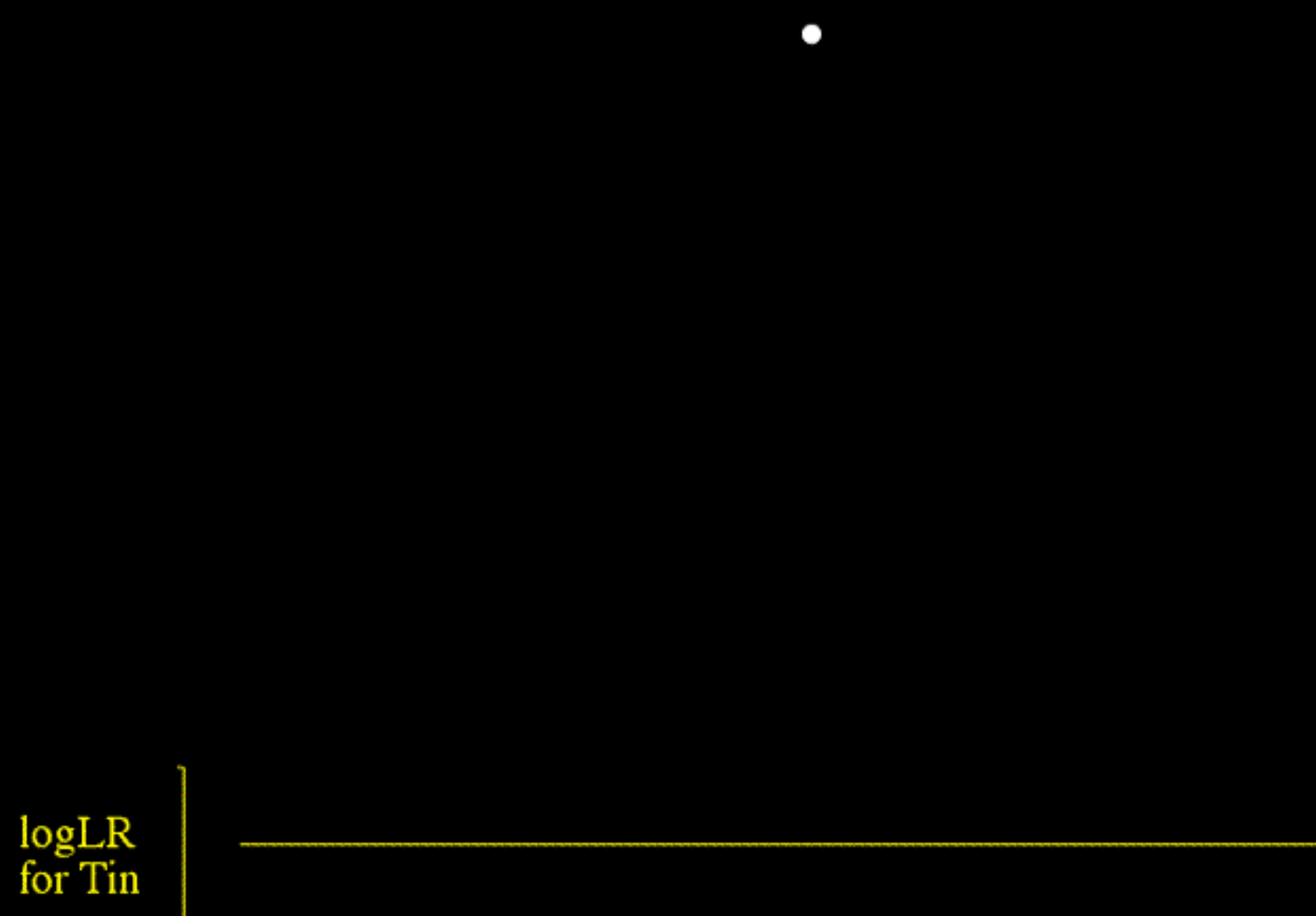
Allows us to ascertain...

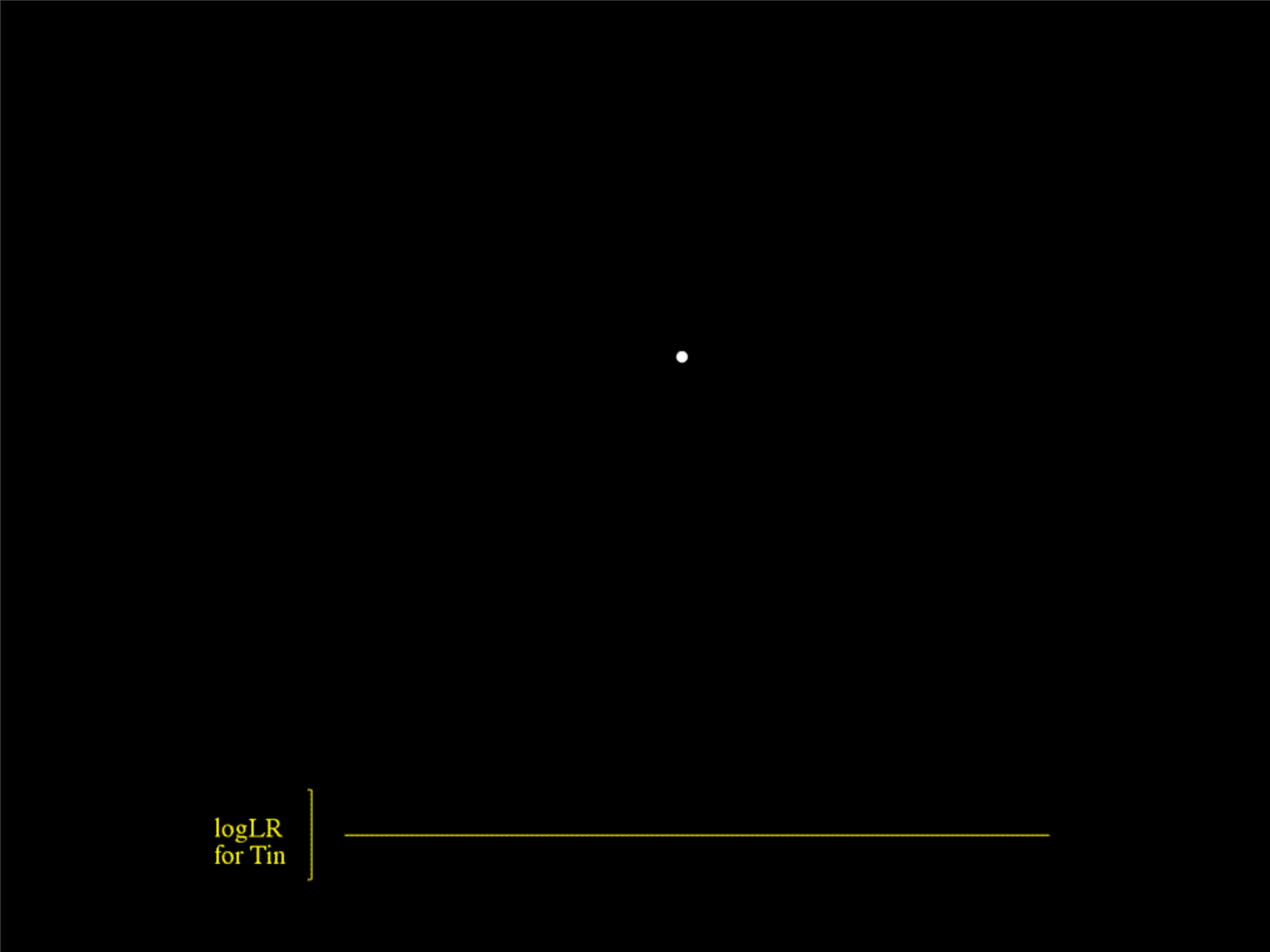
Yang & Shadlen (2007)
Nature 447: 1075-1080

eye position in yellow
One target in the RF of the LIP neuron; other outside

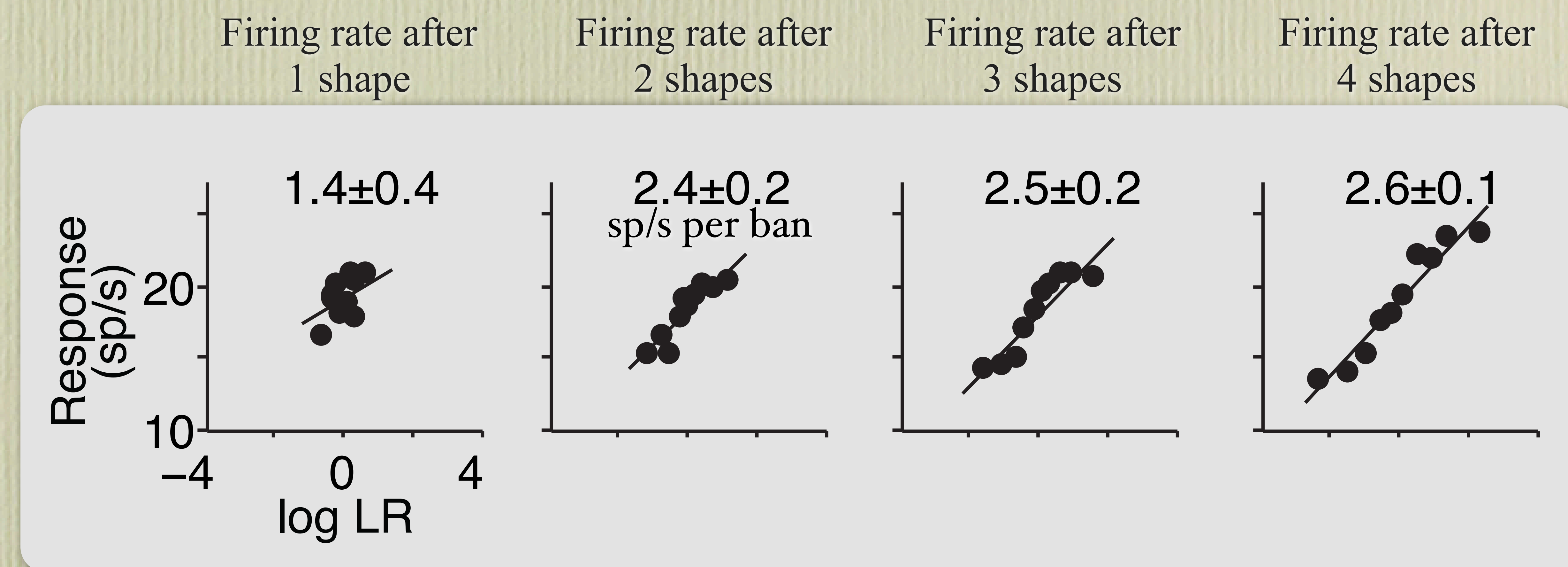
Spikes as shapes are added to the display




logLR
for Tin


logLR
for Tin

LIP represents accumulating evidence in units proportional to logLR



ban is unit of logLR

Yang & Shadlen (2007)
Nature 447: 1075-1080

Conclusions from probabilistic reasoning experiment

- Persistent activity represents accumulation of evidence:
  - a quantitative mapping between neural response and probability
- This permits "optimal" combination of cues with diverse reliability

I'm pleased we can teach monkeys to do this.

Convert observation, x, to Weight of Evidence or Degree of Belief

$$\log_{10} \frac{P(x|H_1)}{P(x|H_0)} = \text{logLR}$$

Alan Turing

X L U N N A R W S F Y T

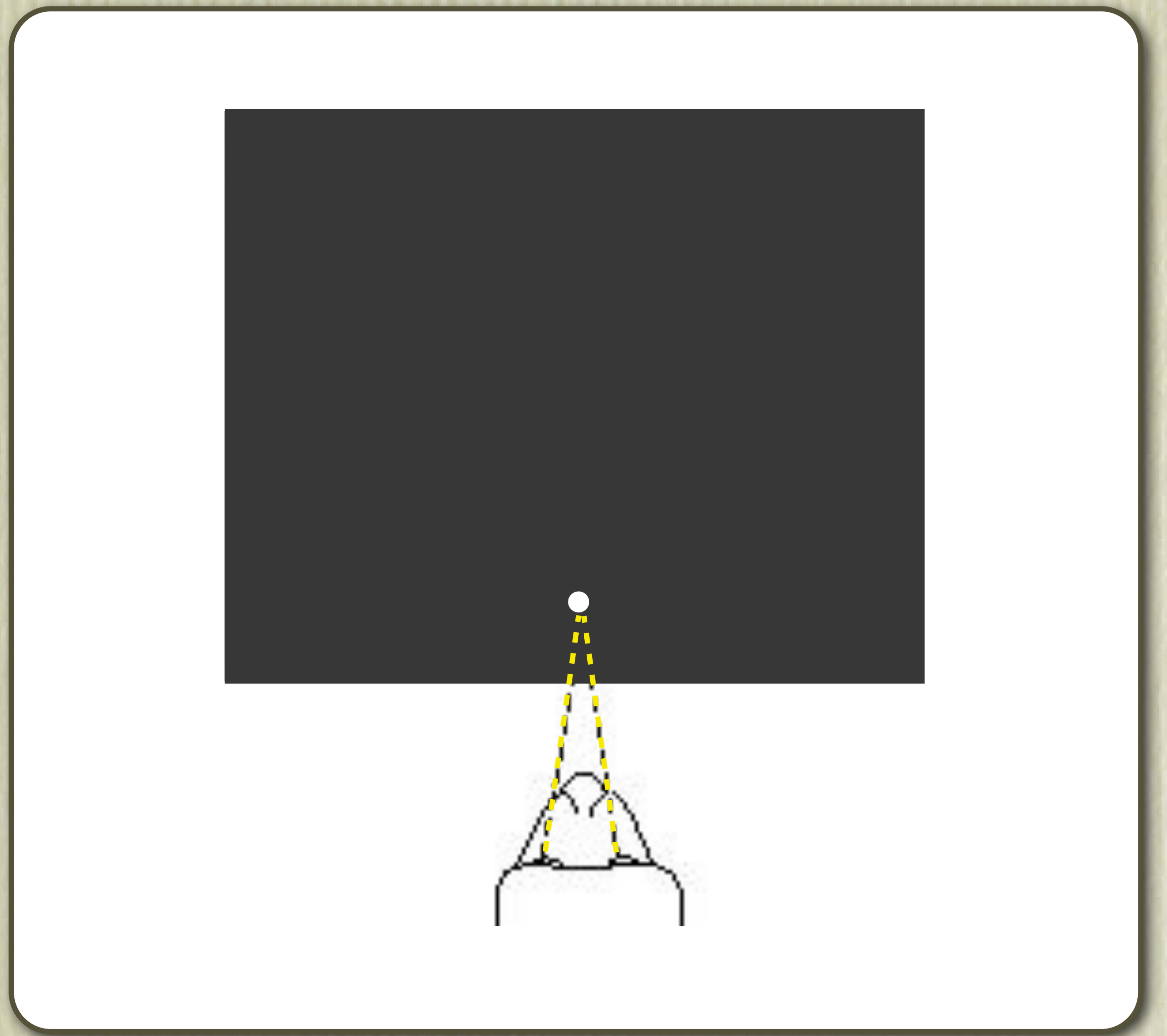
M Y U N X T S B S R P C

Abraham Wald

Outline

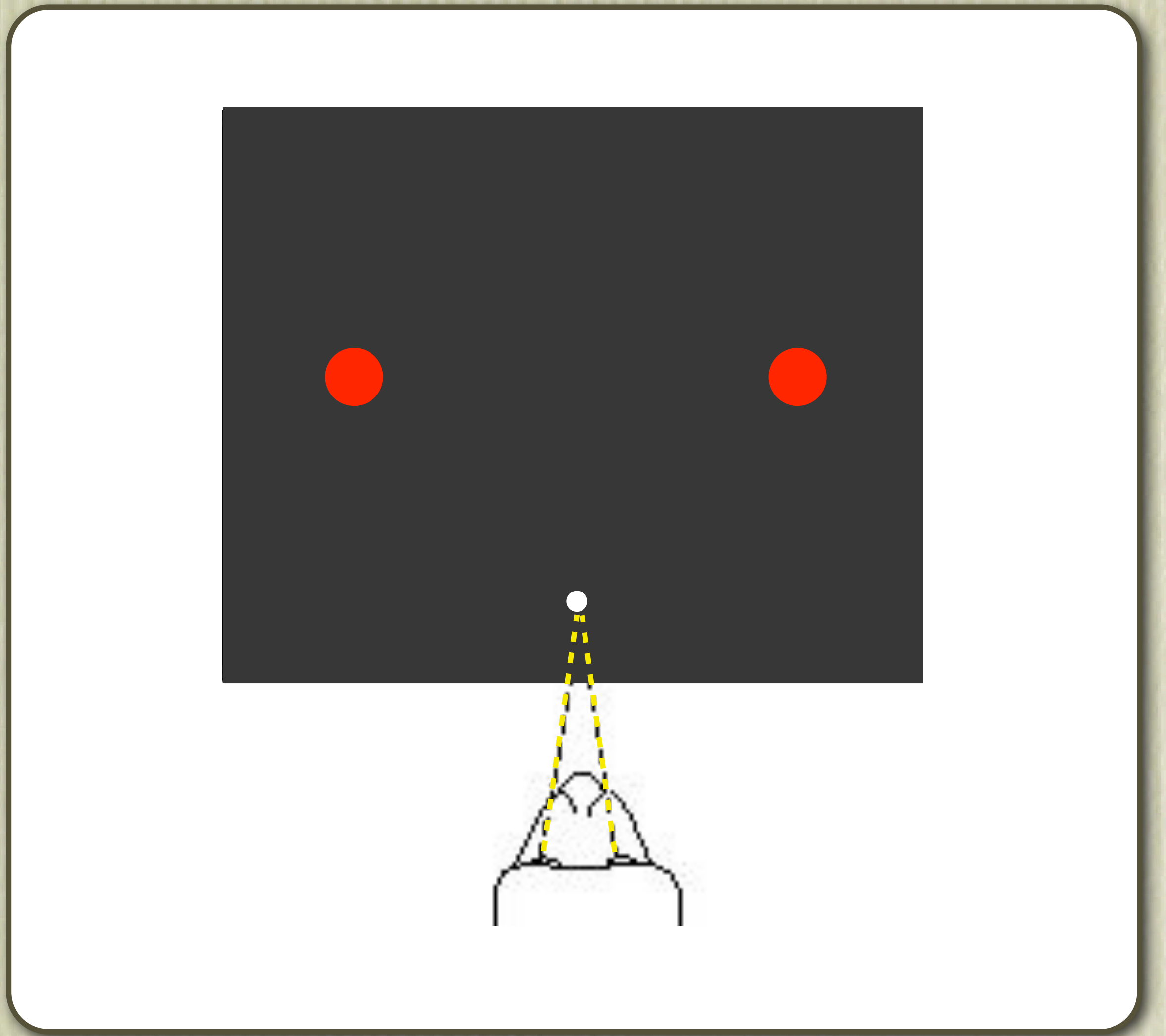
1. Probabilistic reasoning
2. Perceptual decisions: speed and accuracy
3. *Optional*: Sequential analysis, Wald's martingale, logistic choice function
4. Variance and covariance as signatures of neural computation
5. Confidence
6. *Optional*: Integration of prior probability & evidence

Direction-Discrimination Task



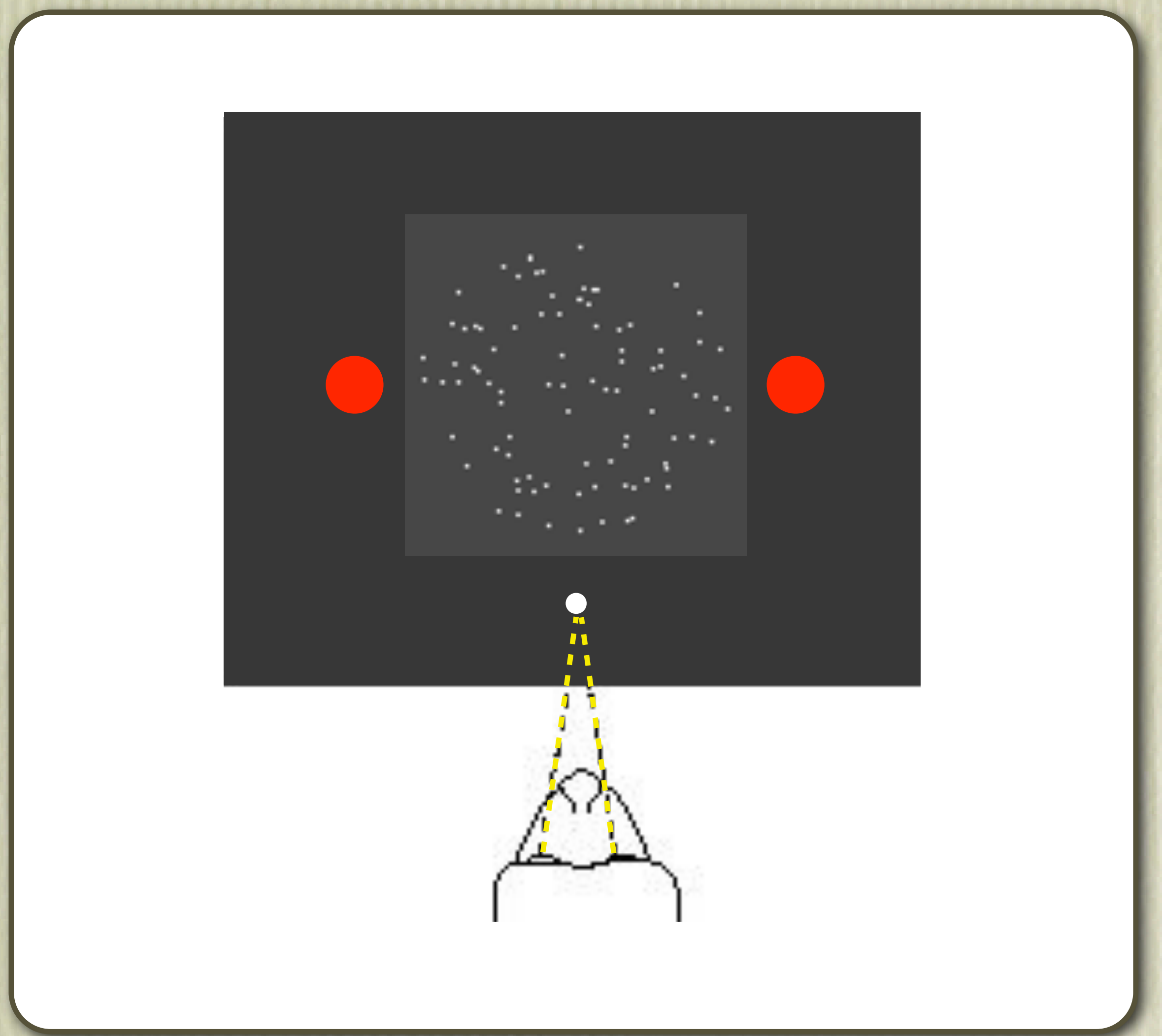
Newsome, Britten & Movshon, 1989

Direction-Discrimination Task



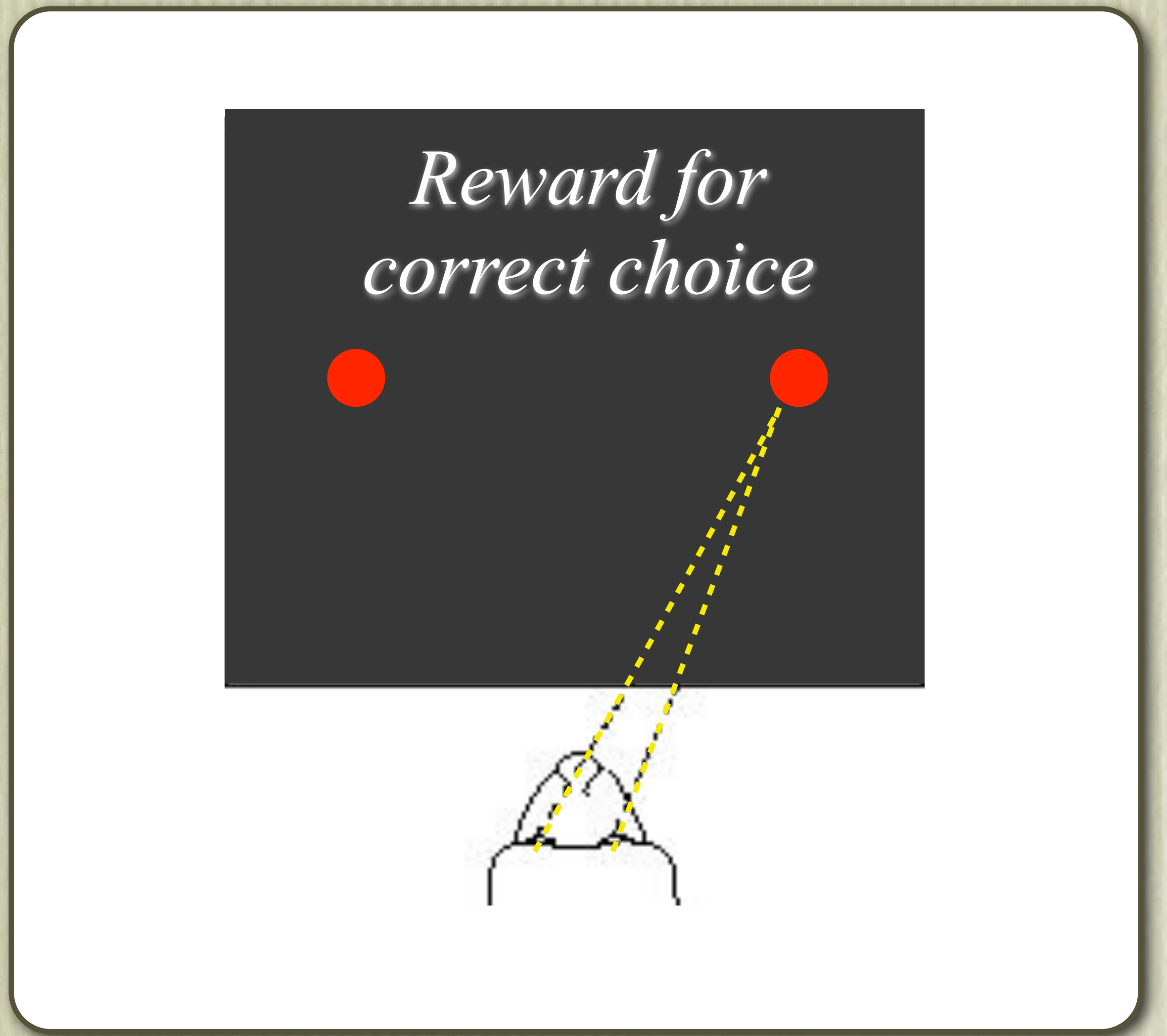
Newsome, Britten & Movshon, 1989

Direction-Discrimination Task

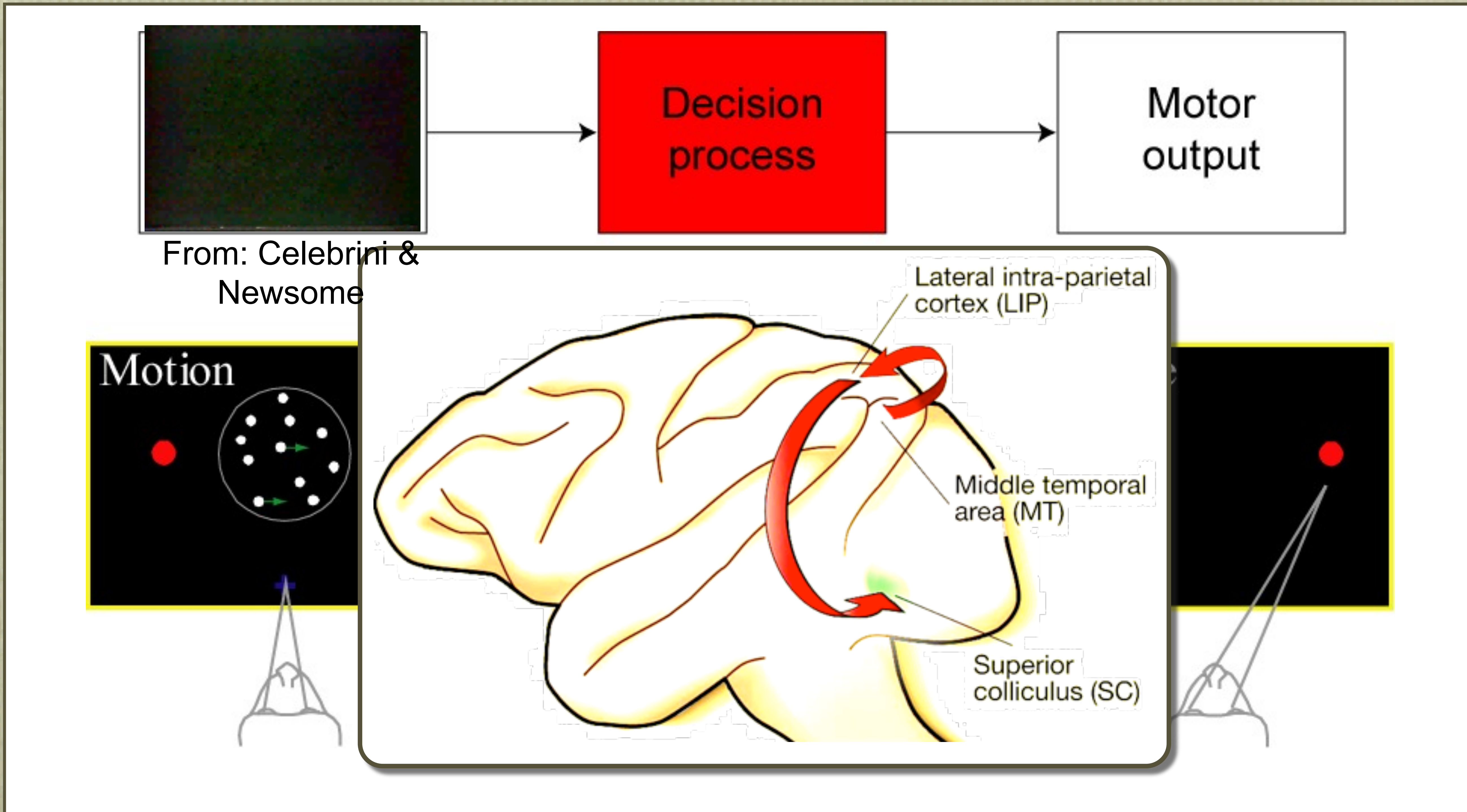


Newsome, Britten & Movshon, 1989

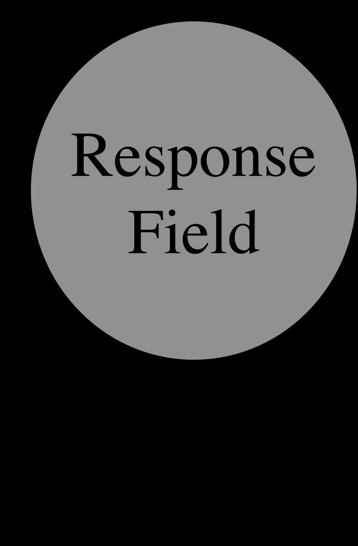
Direction-Discrimination Task
Reaction-time version



Direction selective neurons
Areas MT/V5 and MST

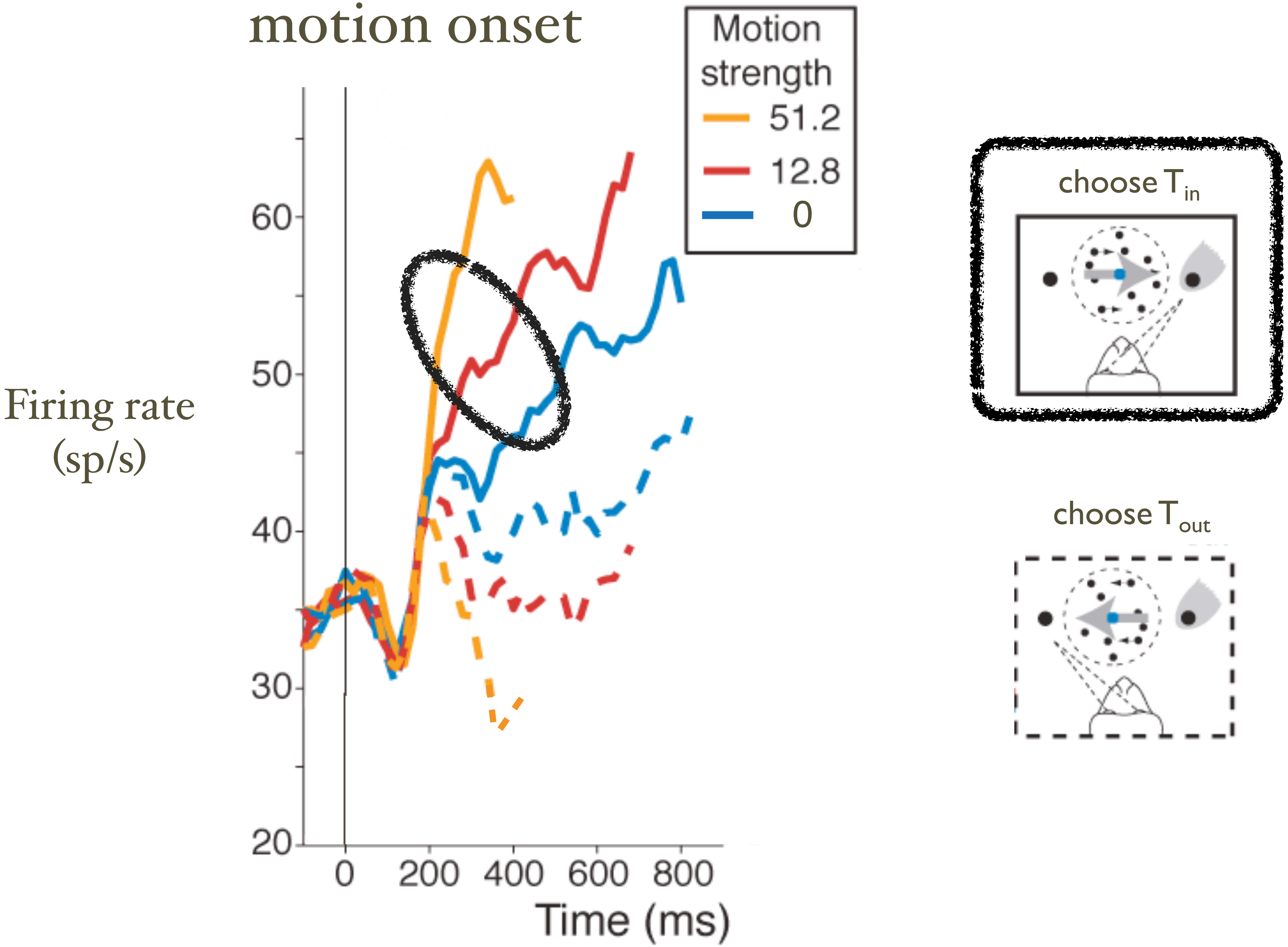


LIP activity during direction discrimination task



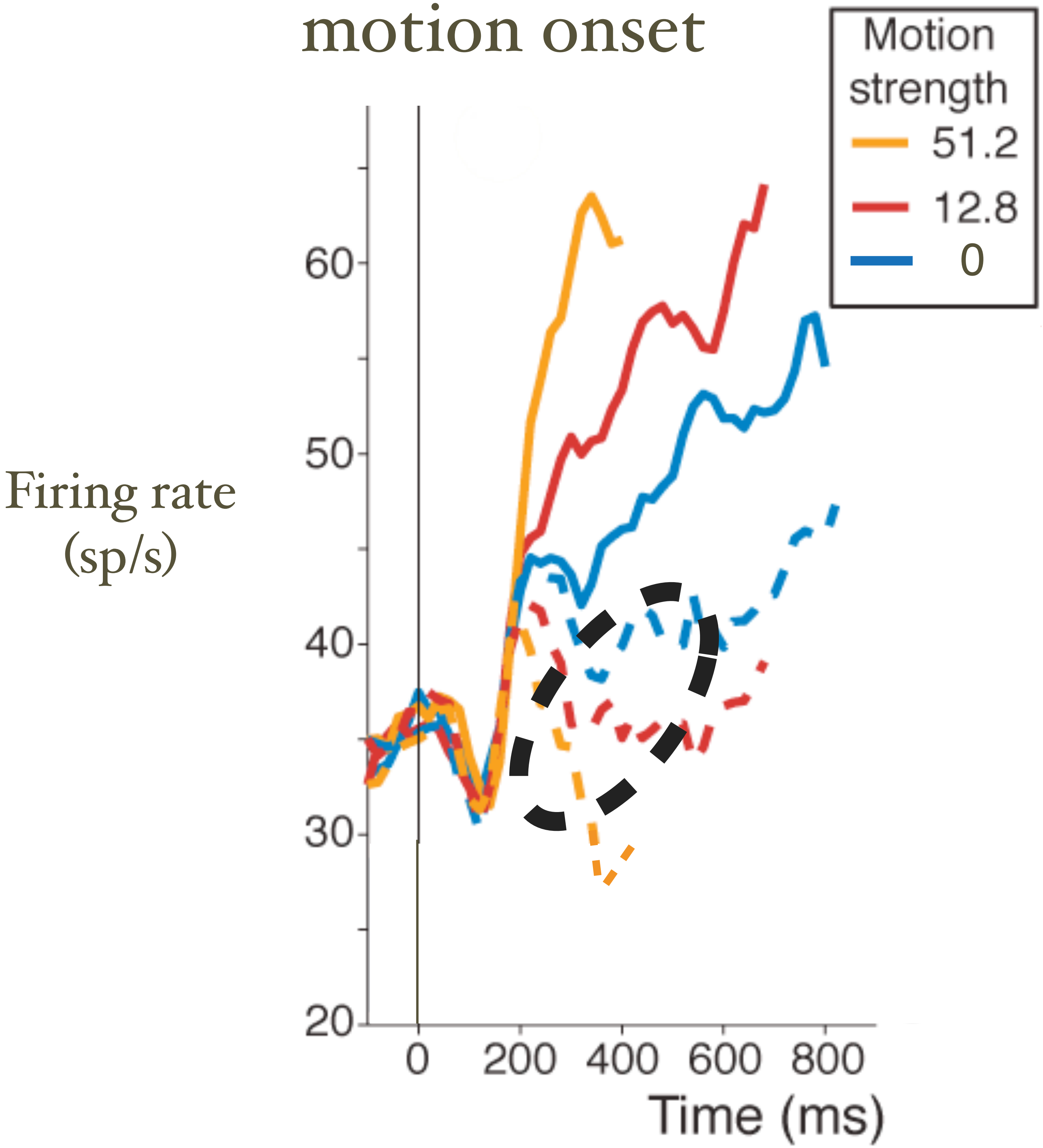
LIP activity during direction discrimination task

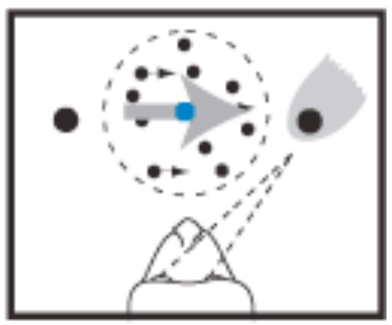
motion onset

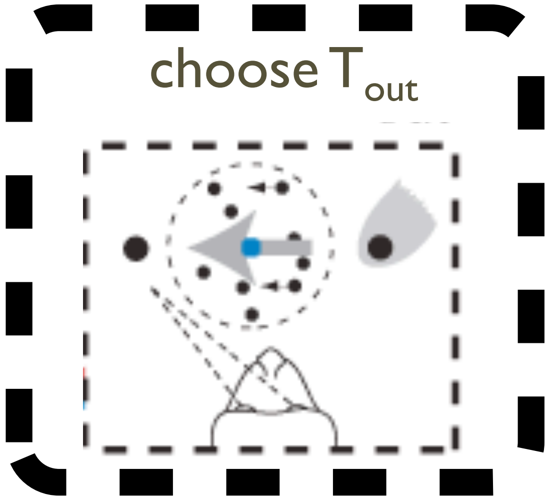


Roitman & Shadlen, 2002 J. Neurosci.

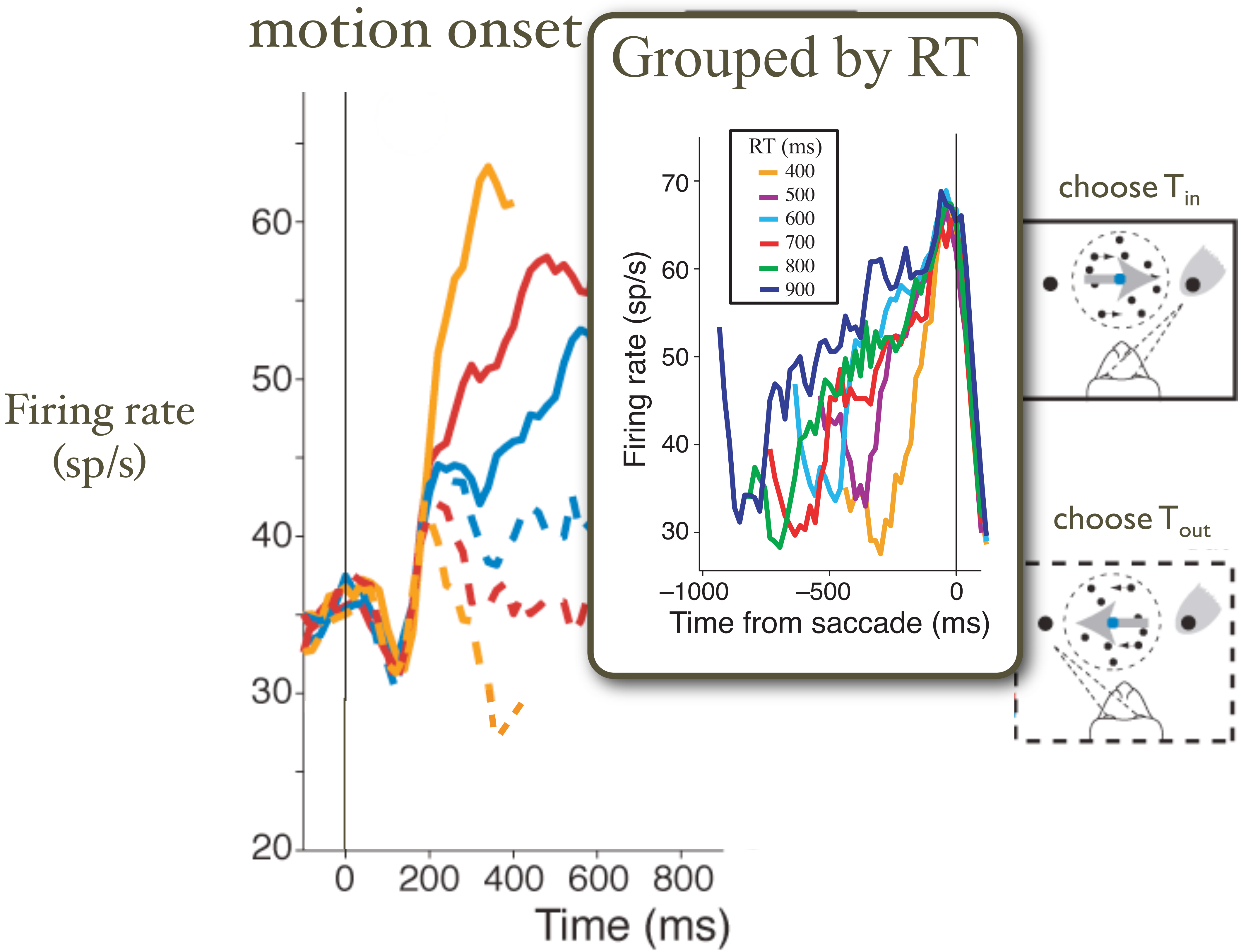
motion onset







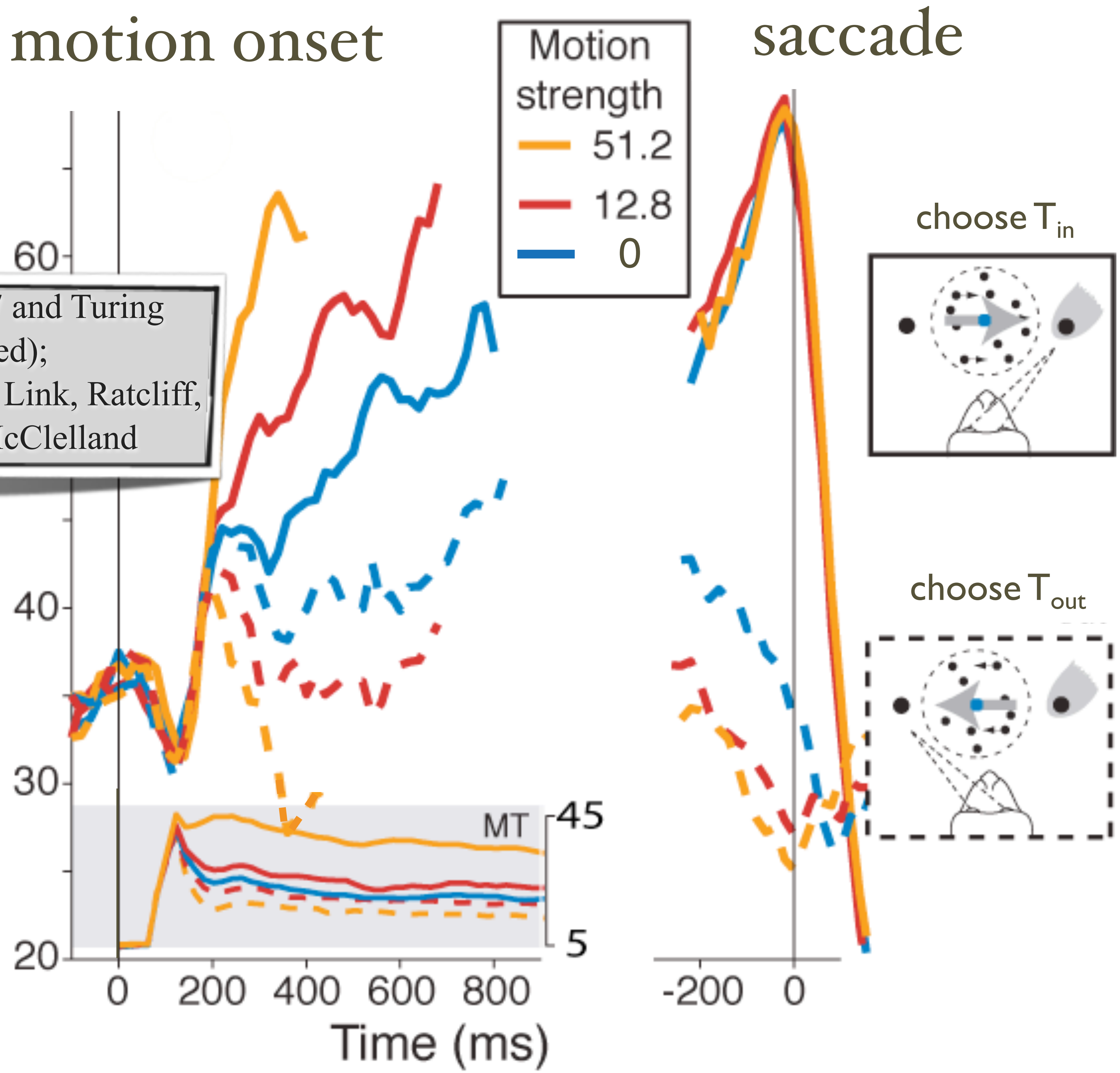
Roitman & Shadlen, 2002 J. Neurosci.



Roitman & Shadlen, 2002 J. Neurosci.

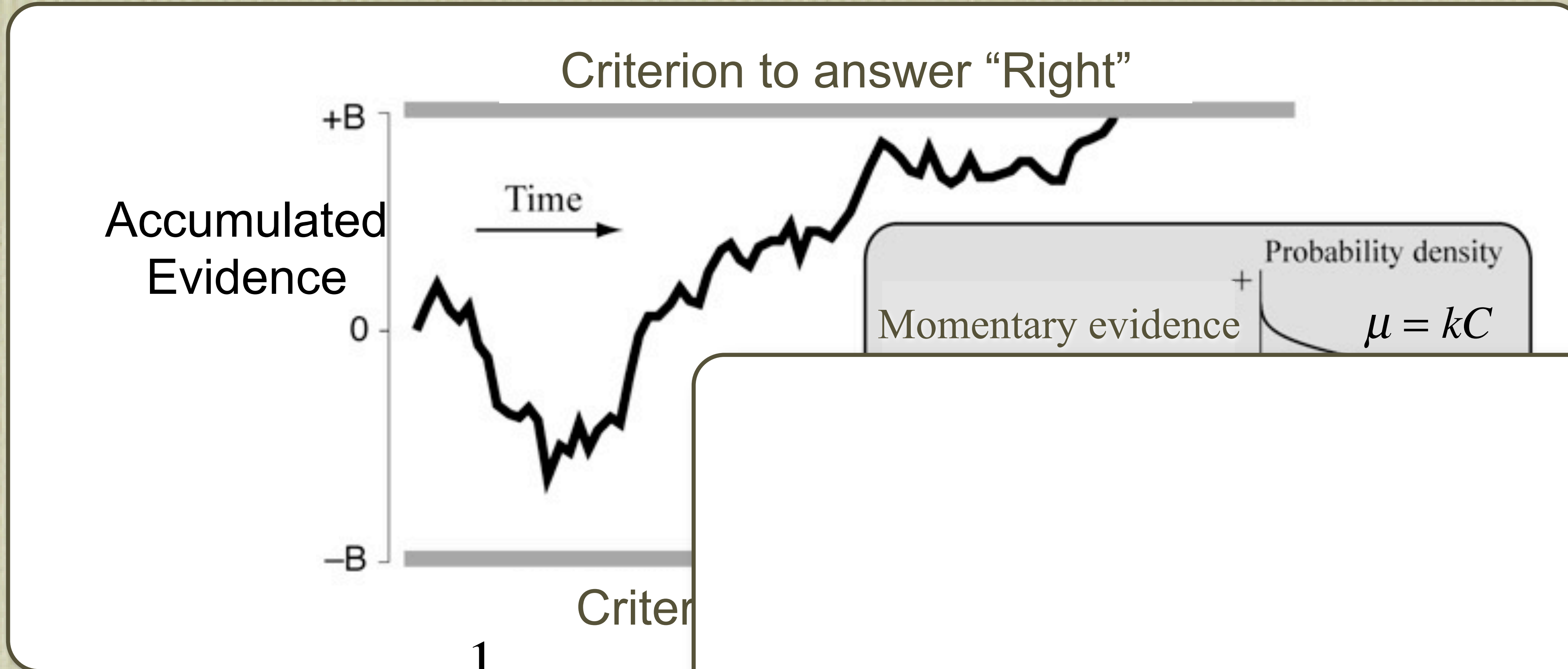
Proposed by Wald, 1947 and Turing (WW II, classified); Stone, 1960; then Laming, Link, Ratcliff, Vickers, P., Usher & McClelland

Bounded Evidence Accumulation



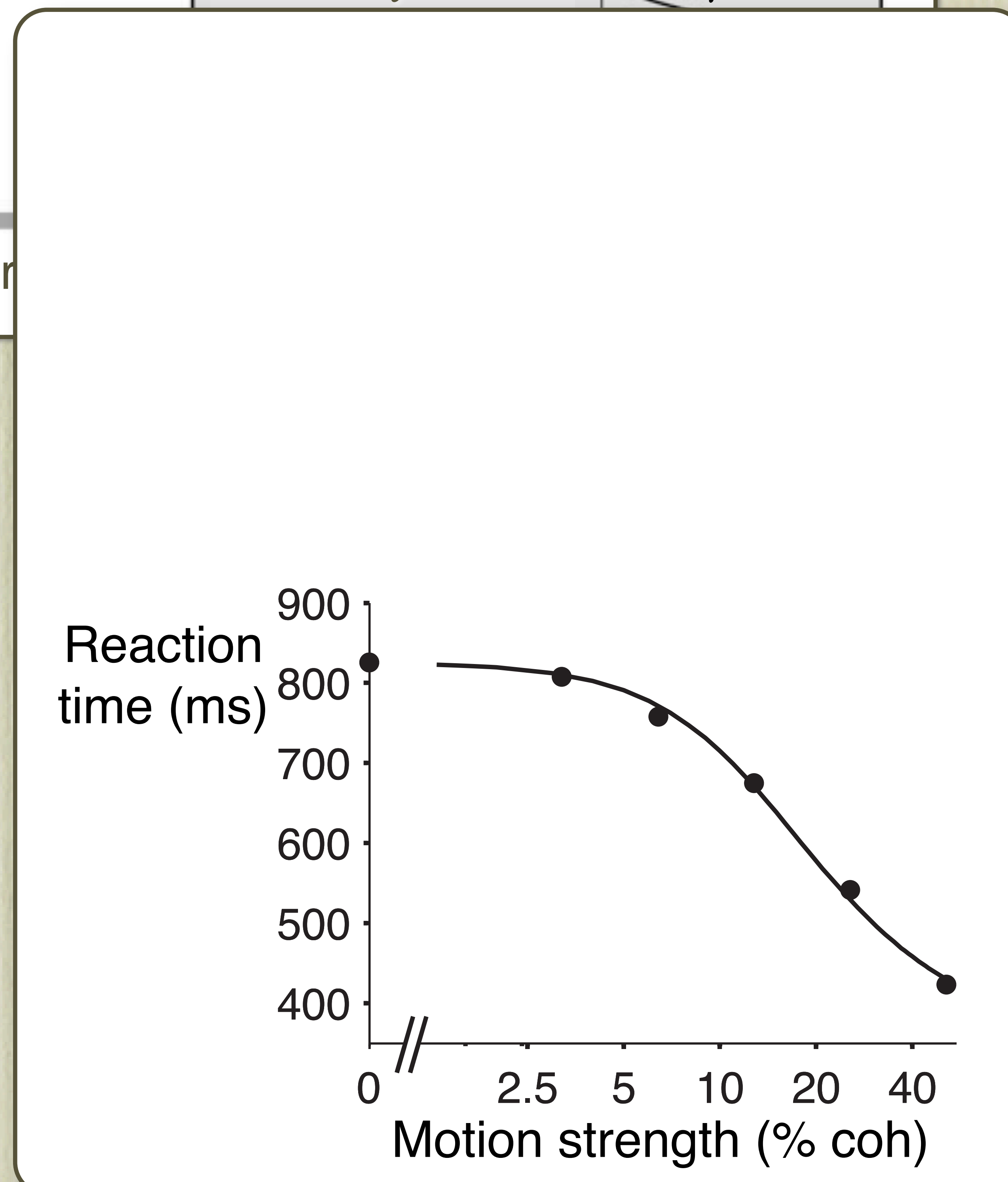
Roitman & Shadlen, 2002 J. Neurosci.

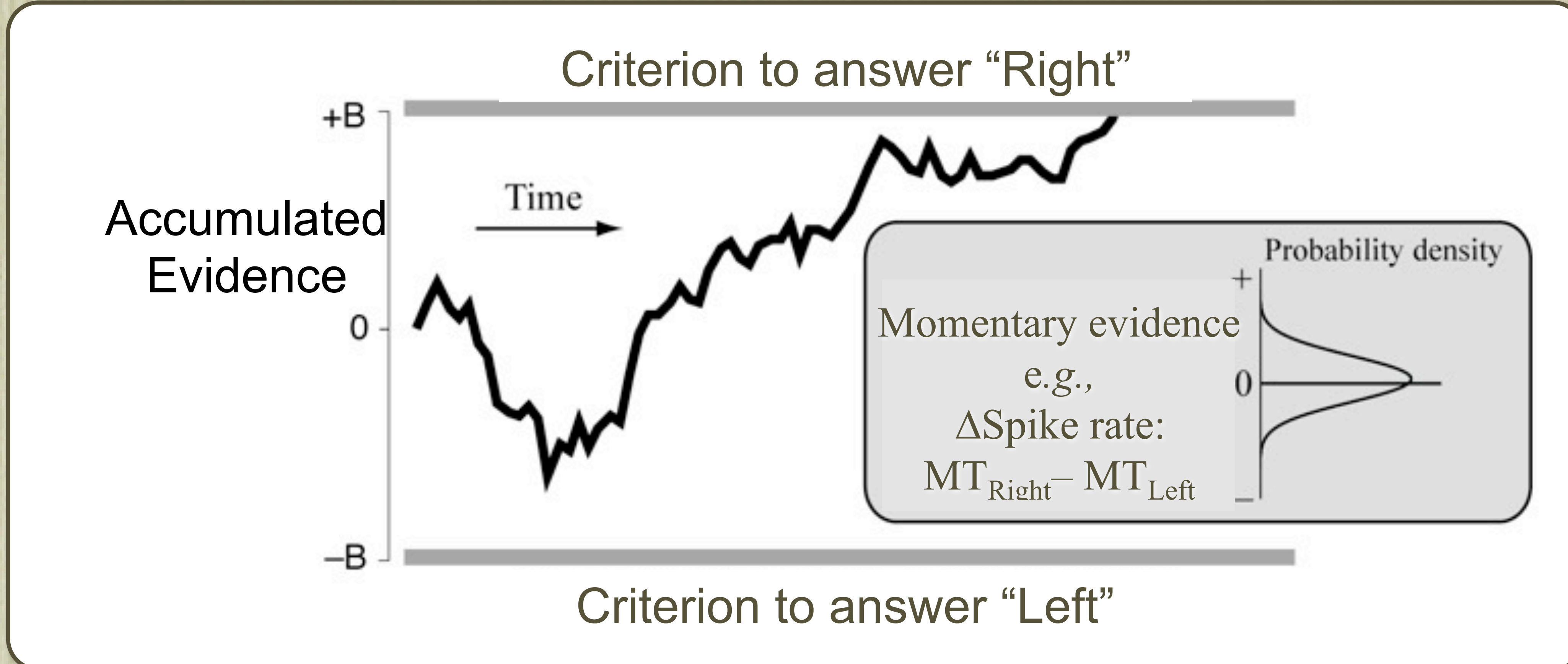
Bounded Evidence Accumulation

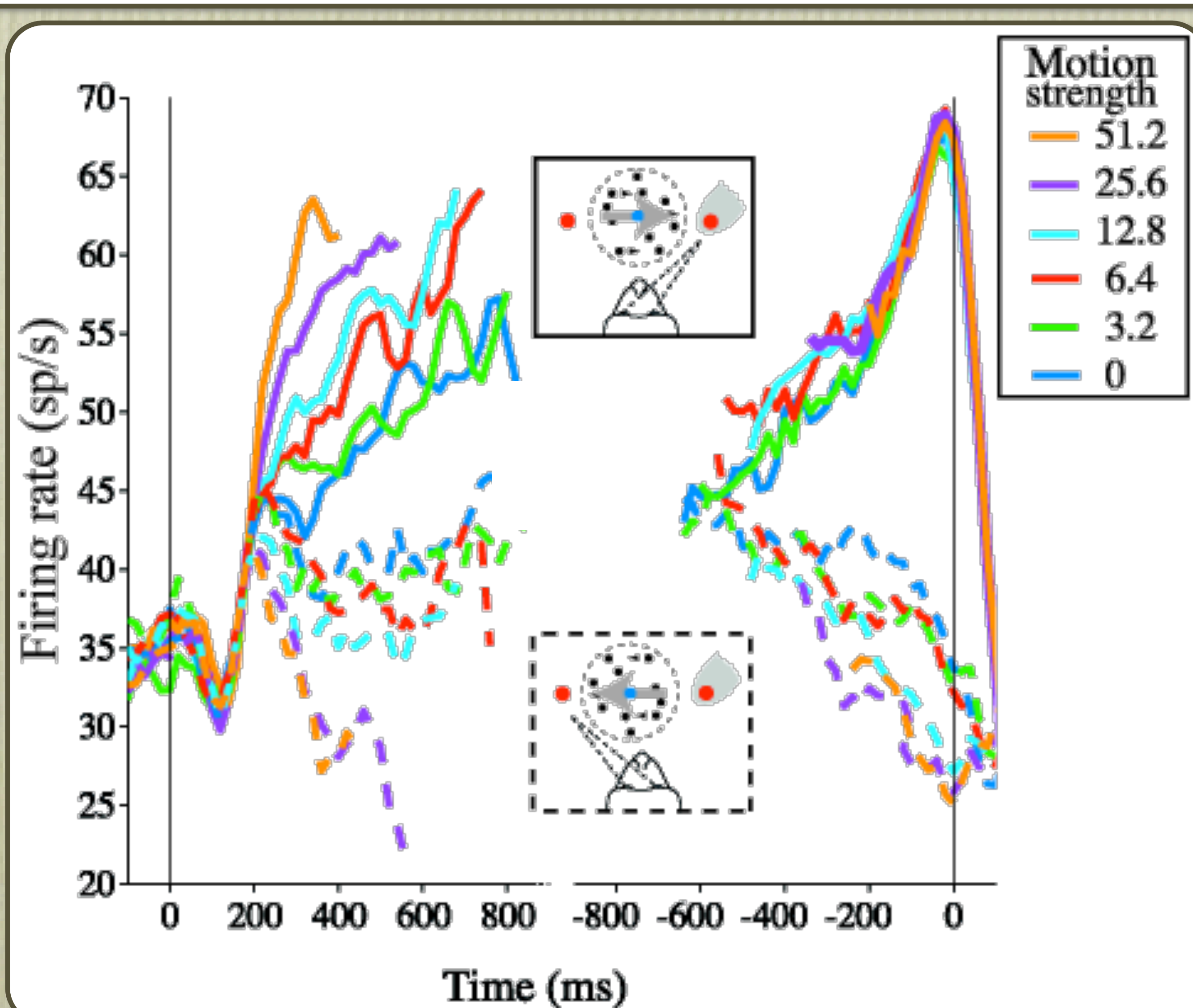


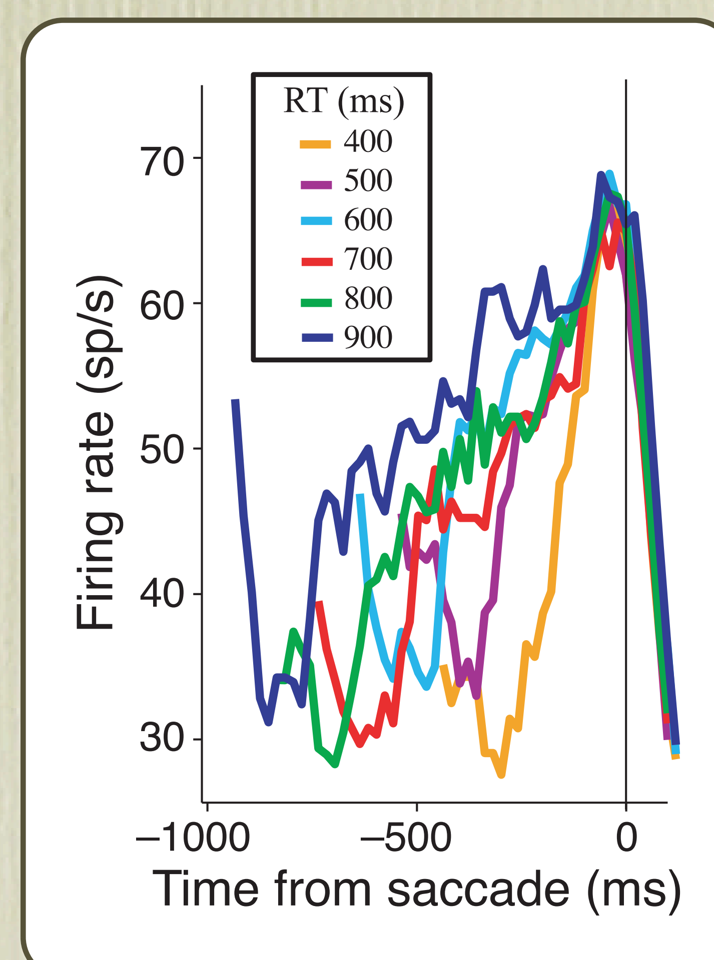
$$P = \frac{1}{1 + e^{-2k|C|B}}$$

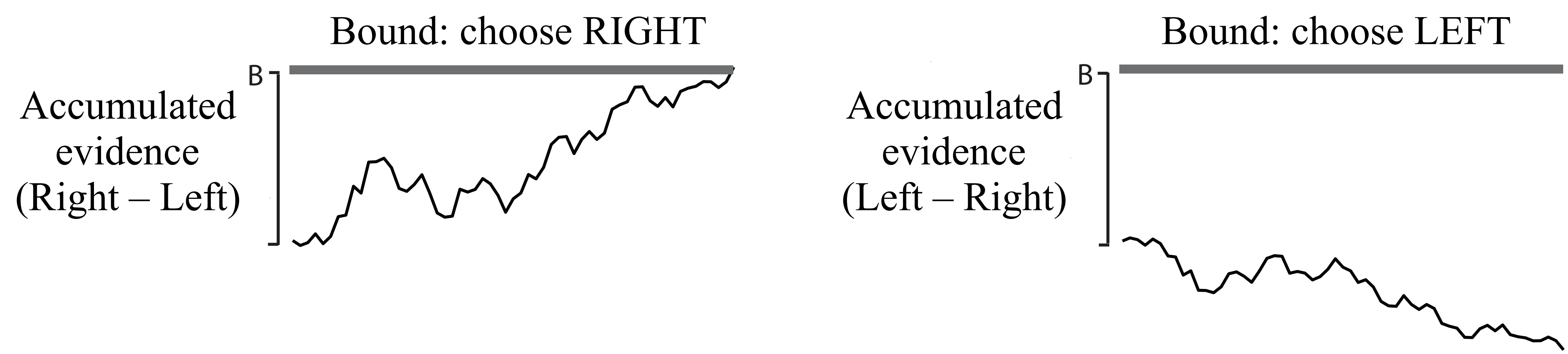
$$t(C) = \frac{B}{kC}\tanh(BkC) + t_{nd}$$

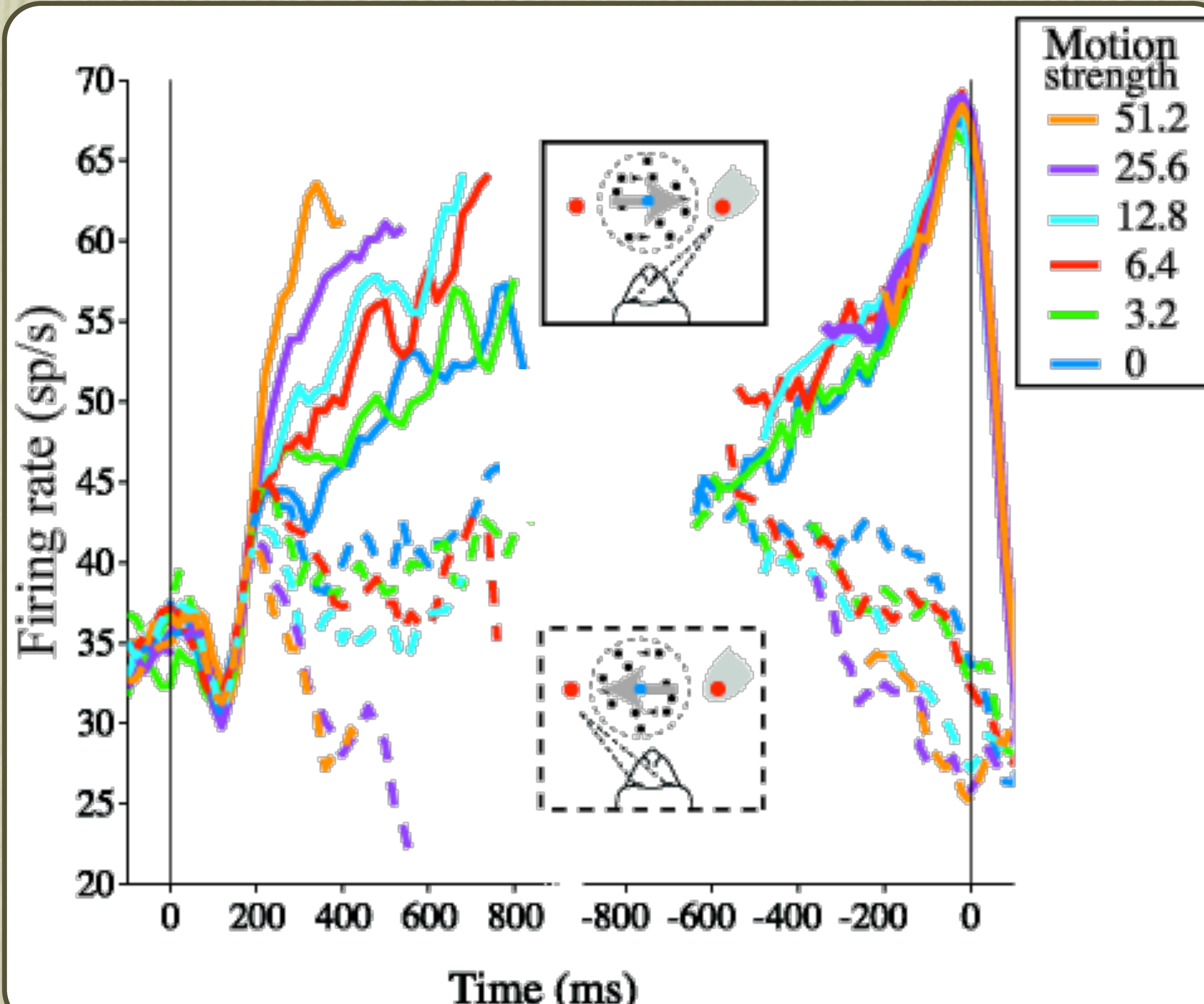



Criterion to answer "Right"
+B
Accumulated Evidence
0
Time
Probability density
+
Momentary evidence
e.g.,
ΔSpike rate:
MTRight– MTLeft
0
–
–B
Criterion to answer "Left"

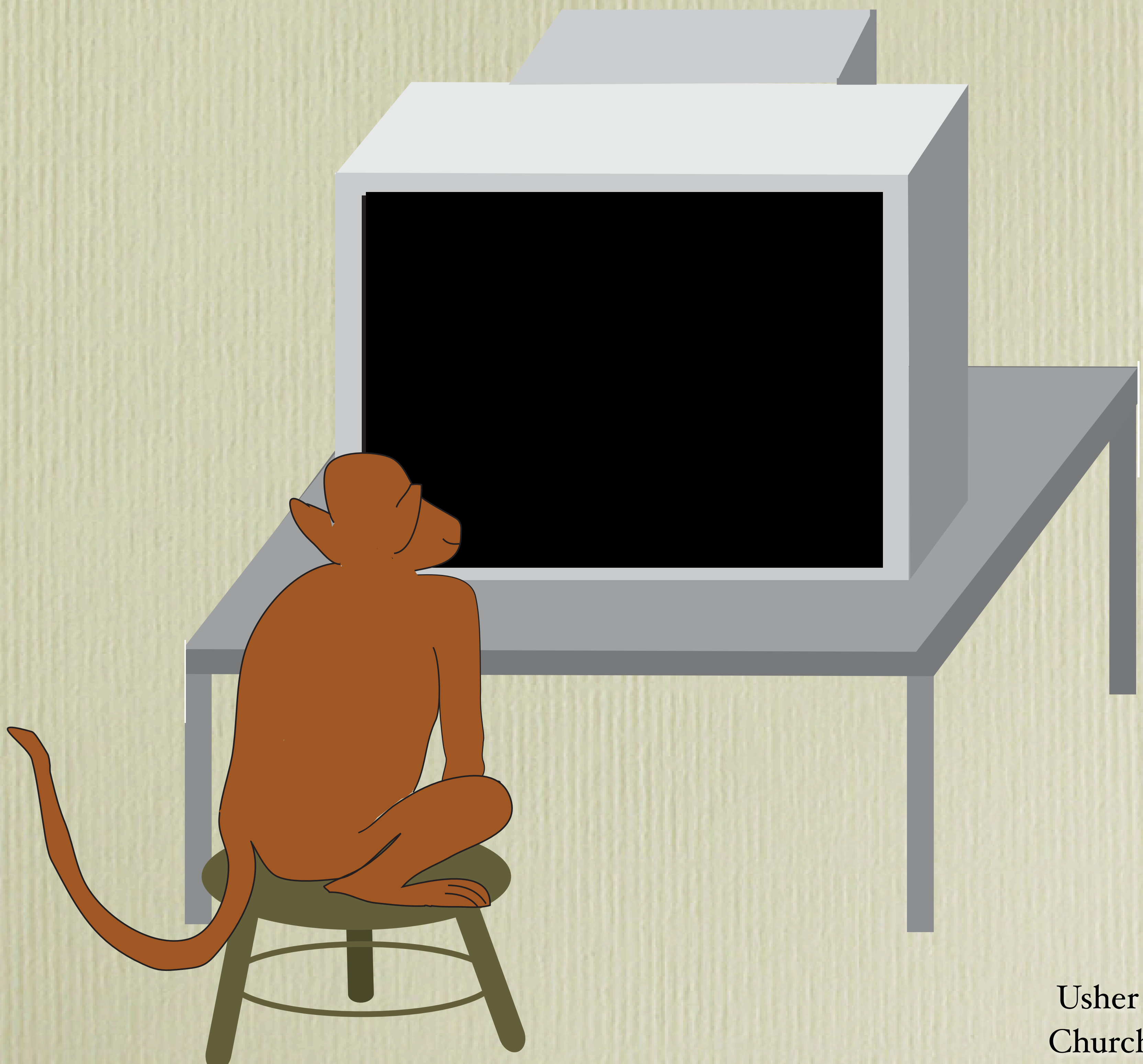

Motion strength
51.2
25.6
12.8
6.4
3.2
0
Firing rate (sp/s)
70
65
60
55
50
45
40
35
30
25
20
0
200
400
600
800
-800
-600
-400
-200
0
Time (ms)


RT (ms)
400
500
600
700
800
900
Firing rate (sp/s)
70
60
50
40
30
−1000
−500
0
Time from saccade (ms)


Bound: choose RIGHT
Accumulated evidence (Right – Left)
B
Bound: choose LEFT
Accumulated evidence (Left – Right)
B


Firing rate (sp/s)
70
65
60
55
50
45
40
35
30
25
20
0
200
400
600
800
-800
-600
-400
-200
0
Time (ms)
Motion strength
51.2
25.6
12.8
6.4
3.2
0

4-choice decisions



Anne Churchland

Usher & McClelland, 2001
Churchland, Kiani & Shadlen, 2008

Outline

1. Probabilistic reasoning
2. Perceptual decisions: speed and accuracy
3. *Optional*: Sequential analysis, Wald's martingale, logistic choice function
4. Variance and covariance as signatures of neural computation
5. Confidence
6. *Optional*: Integration of prior probability & evidence

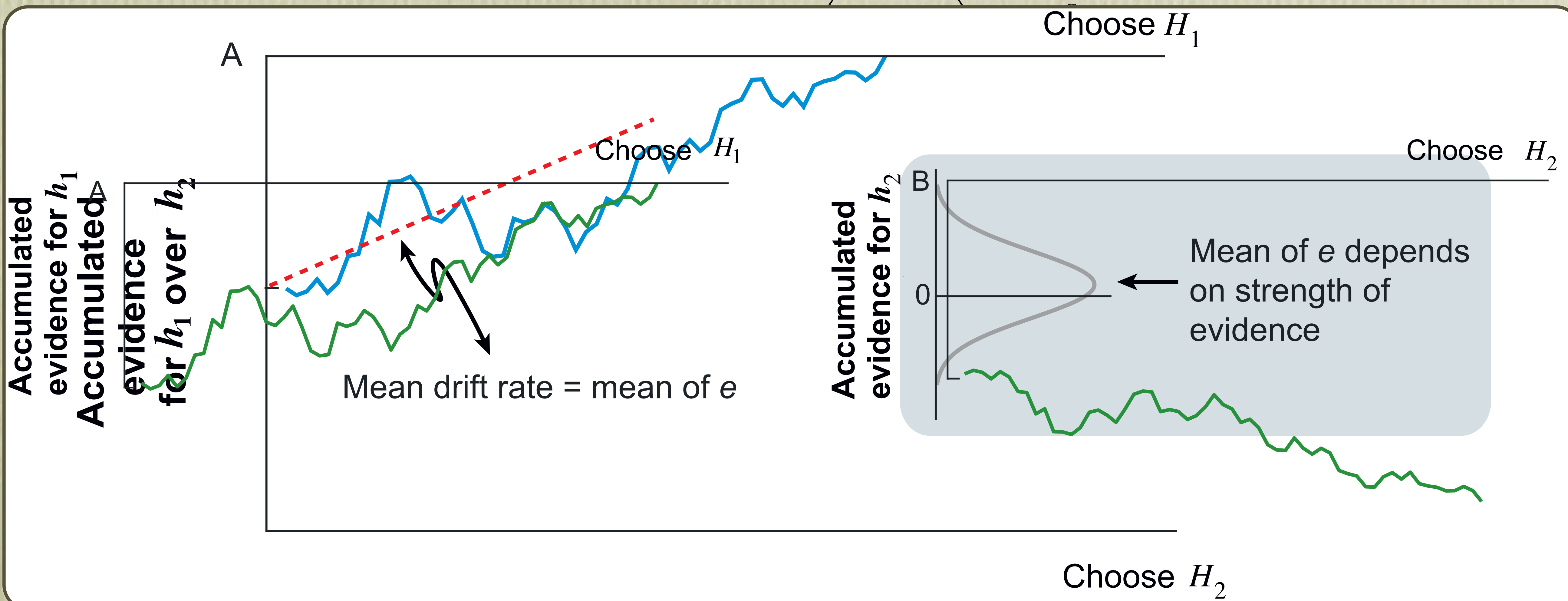
Sequential analysis framework

$$e_0 \rightarrow f_0\ e_0 \Rightarrow \begin{matrix} Stop \\ or \end{matrix}$$

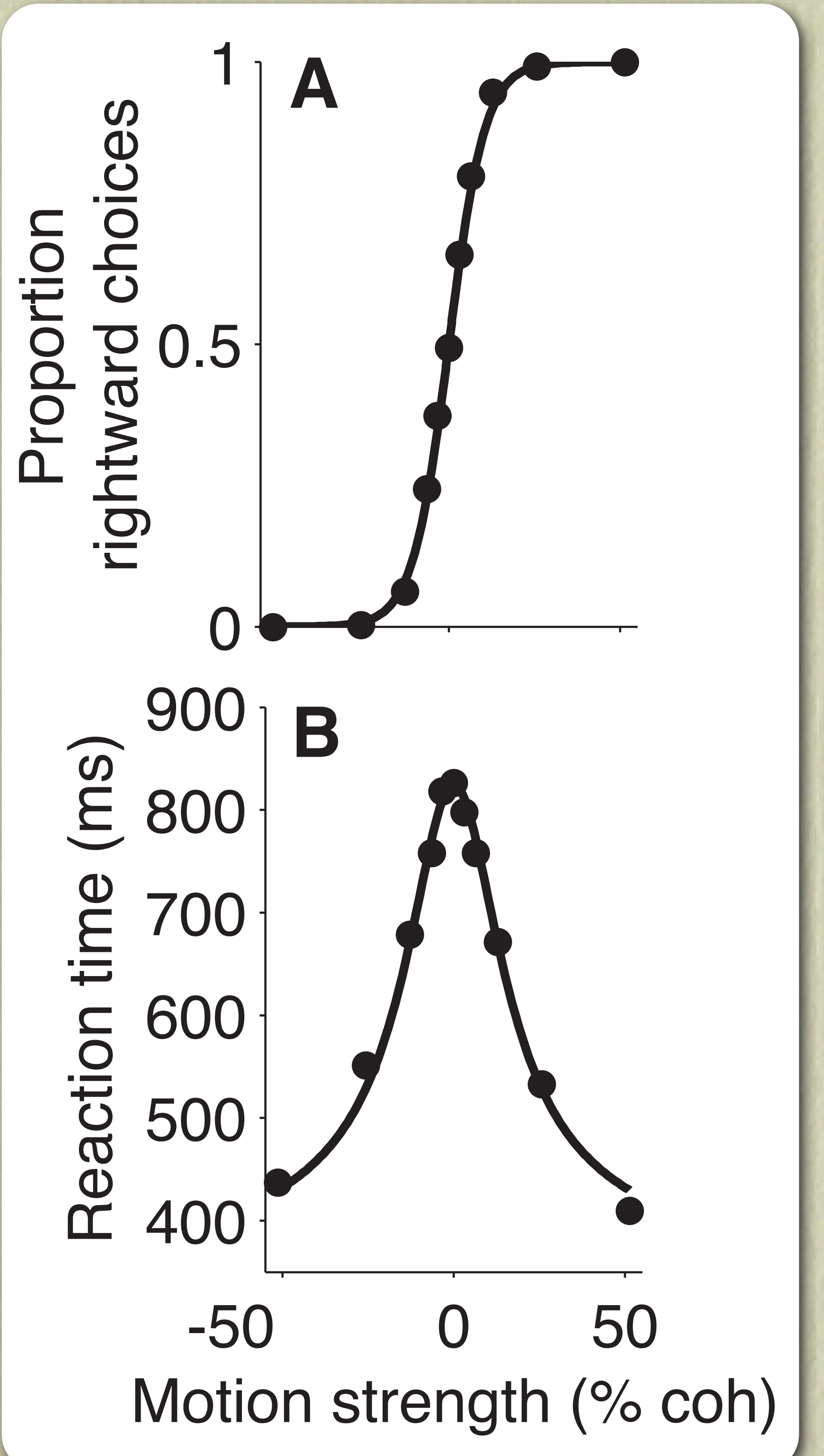
$$\swarrow$$

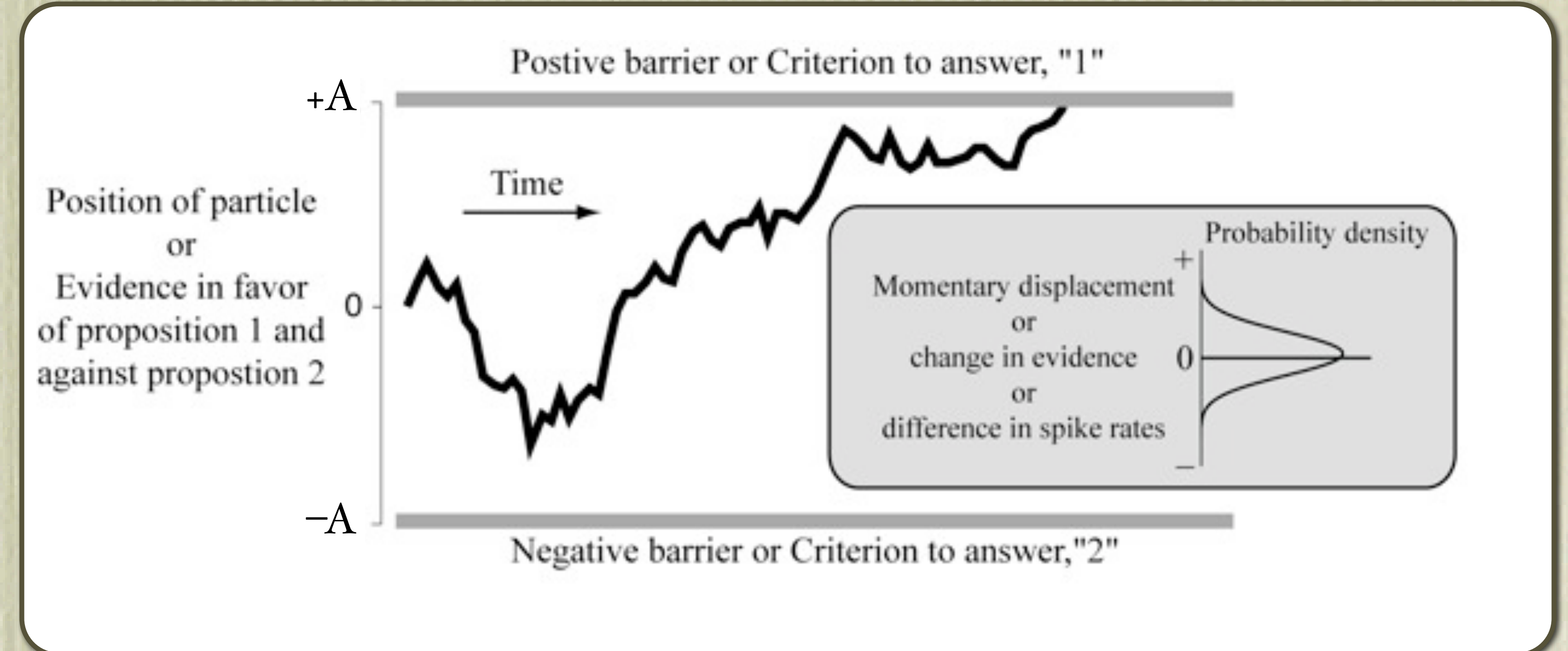
$$e_1 \rightarrow f_1(e_0, e_1) \Rightarrow \begin{matrix} Stop \\ or \end{matrix}$$

$$\swarrow$$



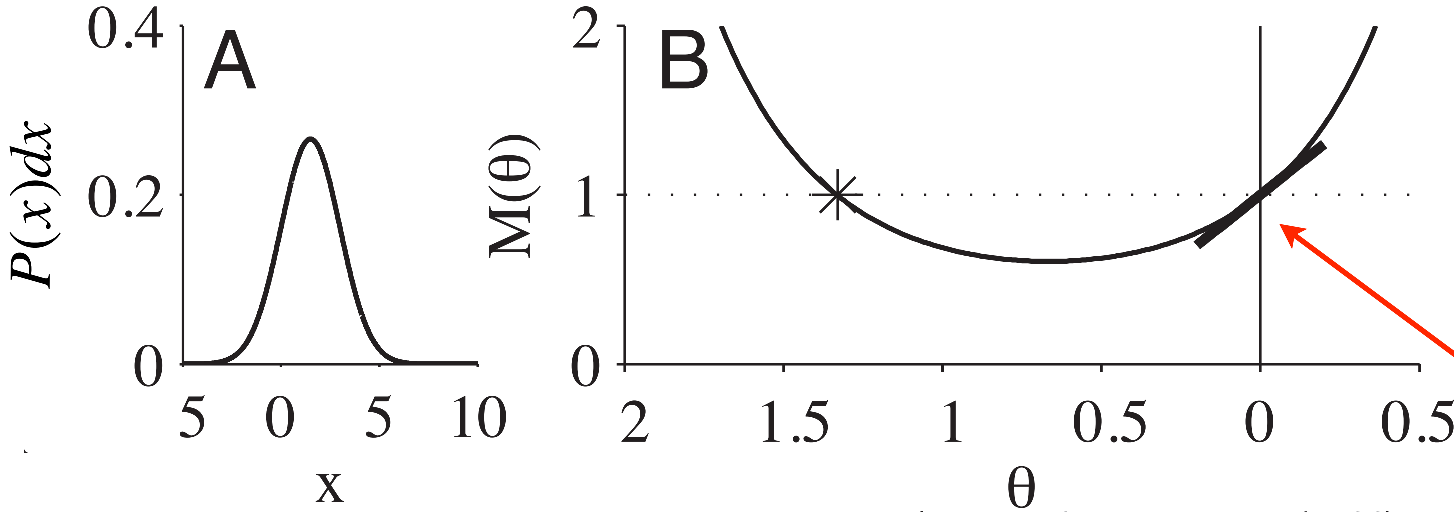
Choice probability & decision time from bounded accumulation





Moment Generating Function

$$M_X(\theta) \equiv E\left[e^{\theta X}\right] = \int_{-\infty}^{\infty} f(x) e^{\theta x} dx$$



$$M'_X(\theta) = \frac{d}{d\theta} E\left[e^{\theta X}\right]$$

$$= \frac{d}{d\theta} \int_{-\infty}^{\infty} f(x) e^{\theta x} dx$$

$$= \int_{-\infty}^{\infty} x f(x) e^{\theta x} dx$$

$$\boxed{M'_X(0) = \int_{-\infty}^{\infty} x f(x) dx = E[x]}$$

Normal distribution

$$M_X(\theta) = e^{\theta\mu + \frac{1}{2}\theta^2\sigma^2}$$

$$\theta_1 = -\frac{2\mu}{\sigma^2}$$

$$\approx 2kC \quad (k > 0)$$

Wald's Martingale

Accumulation	Wald's Martingale
$Y_0 = 0$	$Z_0 = 1$
$+X_1$	$\times \dfrac{e^{\theta X_1}}{M(\theta)}$
Y_1	Z_1
$+X_2$	$\times \dfrac{e^{\theta X_2}}{M(\theta)}$
Y_2	Z_2
$\vdots \quad \vdots$	$\vdots \quad \vdots$
Y_{n-1}	Z_{n-1}
$+X_n$	$\times \dfrac{e^{\theta X_n}}{M(\theta)}$
$Y_n = \sum_{i=1}^{n} X_n$	$Z_n = \dfrac{e^{\theta Y_n}}{M^n(\theta)}$

Wald's Martingale

Accumulation		Wald's Martingale	
$Y_0 = 0$		$Z_0 = 1$	
	$+X_1$		$\times \dfrac{e^{\theta X_1}}{M(\theta)}$
Y_1		Z_1	
	$+X_2$		$\times \dfrac{e^{\theta X_2}}{M(\theta)}$
Y_2		Z_2	
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Y_{n-1}		Z_{n-1}	
	$+X_n$		$\times \dfrac{e^{\theta X_n}}{M(\theta)}$
$Y_n = \sum_{i=1}^{n} X_n$		$Z_n = \dfrac{e^{\theta Y_n}}{M^n(\theta)}$	

$$E\left[Z_{n+1} \middle| Y_0, Y_1, Y_2, \ldots, Y_n\right] = E\left[M_X^{-(n+1)}(\theta) e^{\theta Y_{n+1}} \middle| Y_0, Y_1, Y_2, \ldots, Y_n\right]$$

$$= E\left[M_X^{-(n+1)}(\theta) e^{\theta(Y_n + X_{n+1})}\right]$$ by the rule for generating Y_{n+1}

$$= E[M_X^{-1}(\theta) M_X^{-n}(\theta) e^{\theta Y_n} e^{\theta X_{n+1}}]$$

$$= E[M_X^{-1}(\theta) Z_n e^{\theta X_{n+1}}]$$ using the definition of Z_n

$$= M_X^{-1}(\theta) Z_n E\left[e^{\theta X_{n+1}}\right]$$ because Z_n and $M_X(\theta)$ are known

$$= Z_n$$

$$E[Z_n] = E\left[M_X^{-n}(\theta) e^{\theta Y_n}\right]$$

$$= M_X^{-n}(\theta) E\left[e^{\theta Y_n}\right]$$

$$= M_X^{-n}(\theta) M_{Y_n}(\theta)$$

$$= 1$$

MGF of the bounded accumulation

Calculate two ways:

(i) by brute force from 2 possible values

$$M_{\tilde{Y}}(\theta) = E[e^{\theta \tilde{Y}}]$$

$$= P_+ e^{\theta A} + (1 - P_+) e^{-\theta A}$$

(ii) using Wald's Identity $E\left[M_X^{-n}(\theta) e^{\theta Y_n}\right.$

Define the stopped accum

$\tilde{Z} = M_X^{-\tilde{n}}(\theta) e^{\theta \tilde{Y}}$ Wald's martingale, whe accumulation stops

$E\left[\tilde{Z}\right] = E\left[Z_n\right]$ optional stopping theo

$$E\left[M_X^{-\tilde{n}}(\theta) e^{\theta \tilde{Y}}\right] = 1$$

$E\left[e^{\theta_1 \tilde{Y}}\right] = 1$ simplify at the special r

So, at special root of mgf

$$M_{\tilde{Y}}(\theta_1) = E[e^{\theta_1 \tilde{Y}}]$$

$$= P_+ e^{\theta_1 A} + (1 - P_+) e^{-\theta_1 A}$$

$$= 1$$

$$P_+ = \frac{1 - e^{-\theta_1 A}}{e^{\theta_1 A} - e^{-\theta_1 A}}$$

$$= \frac{1 - e^{-\theta_1 A}}{e^{-\theta_1 A}\left(e^{\theta_1 A} + 1\right)\left(e^{\theta_1 A} - 1\right)}$$

$$= \frac{1}{1 + e^{\theta_1 A}}$$

Decision time

$$E\left[M_X^{-\tilde{n}}(\theta)e^{\theta\tilde{Y}}\right]=1$$ Wald's identity

$$\frac{d}{d\theta}E\left[M_X^{-\tilde{n}}(\theta)e^{\theta\tilde{Y}}\right]=0$$

$$=E\left[e^{\theta\tilde{Y}}\tilde{Y}M_X^{-\tilde{n}}(\theta)-e^{\theta\tilde{Y}}\tilde{n}M_X^{-1-\tilde{n}}(\theta)M_X'(\theta)\right]$$

$$=E\left[\tilde{Y}-\tilde{n}\mu\right]$$ holds for $\theta=0$

$$E[\tilde{n}]=\frac{E[\tilde{Y}]}{\mu}\quad (\text{for } \mu\neq 0)$$

$$=\frac{(2P_+-1)A}{\mu}$$ recall that $P_+=\frac{1}{1+e^{-2kCA}}$

$$=\frac{A}{\mu}\left(\frac{2}{1+e^{\theta_1 A}}-1\right)$$

$$=\frac{A}{\mu}\left(\frac{1-e^{\theta_1 A}}{1+e^{\theta_1 A}}\right)$$

$$=\frac{A}{\mu}\left(\frac{e^{-\frac{\theta_1 A}{2}}-e^{\frac{\theta_1 A}{2}}}{e^{-\frac{\theta_1 A}{2}}+e^{\frac{\theta_1 A}{2}}}\right)$$

$$=\frac{A}{\mu}\tanh\left(-\frac{\theta_1 A}{2}\right)$$

for the dots task $E[t]=\frac{A}{kC}\tanh(kCA)$

$$\lim_{C\to 0}\frac{A}{kC}\tanh(kCA)=A^2$$

Outline

1. Probabilistic reasoning
2. Perceptual decisions: speed and accuracy
3. *Optional*: Sequential analysis, Wald's martingale, logistic choice function
4. Variance and covariance as signatures of neural computation
5. Confidence
6. *Optional*: Integration of prior probability & evidence

Doubly stochastic point processes

Law of total variance

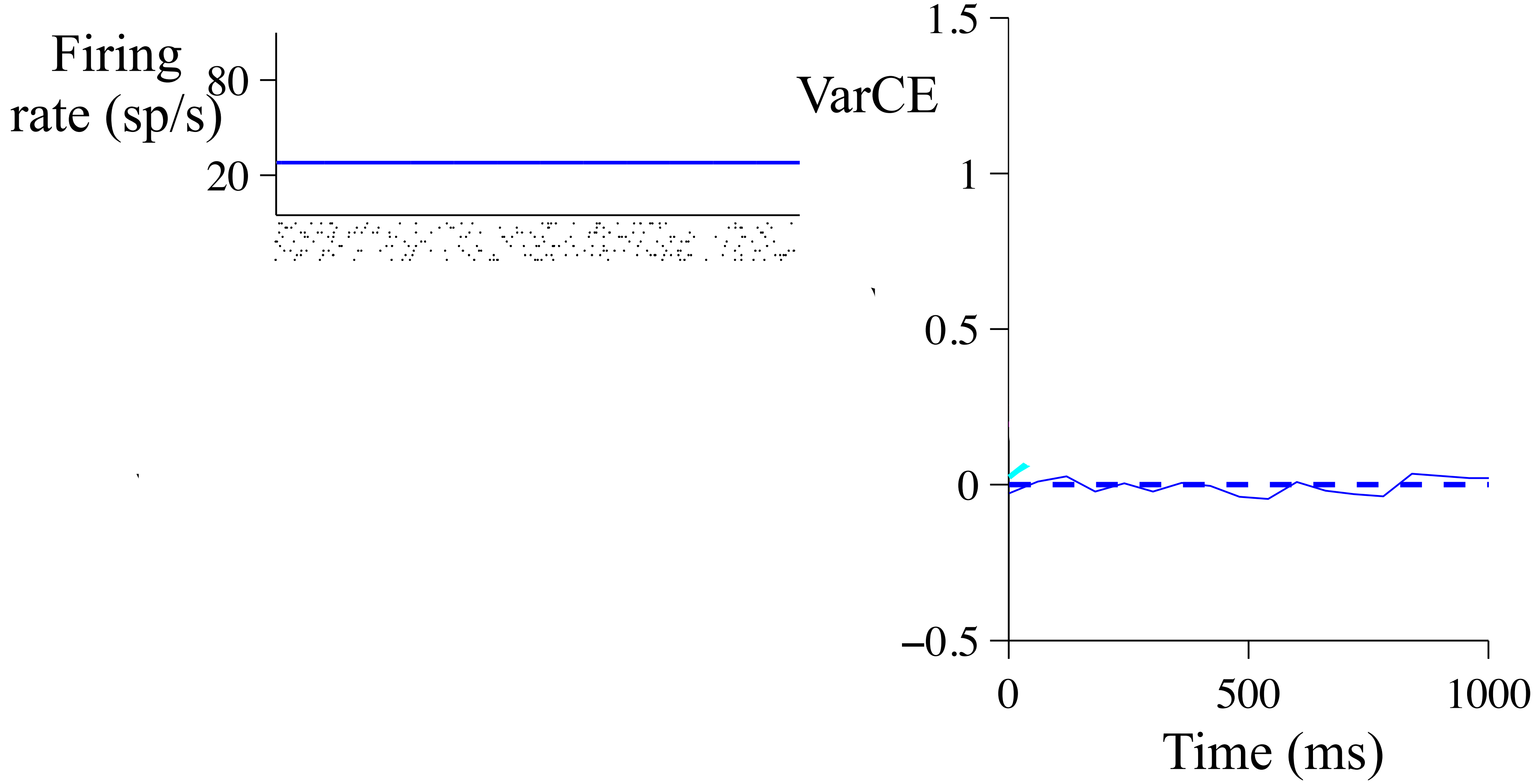
$$Var[X] = \underbrace{Var\left[\left\langle X|Y \right\rangle\right]}_{\text{variance of conditional expectation (VCE)}} + \underbrace{\left\langle Var\left[X|Y\right]\right\rangle}_{\text{expectation of conditional variance}}$$

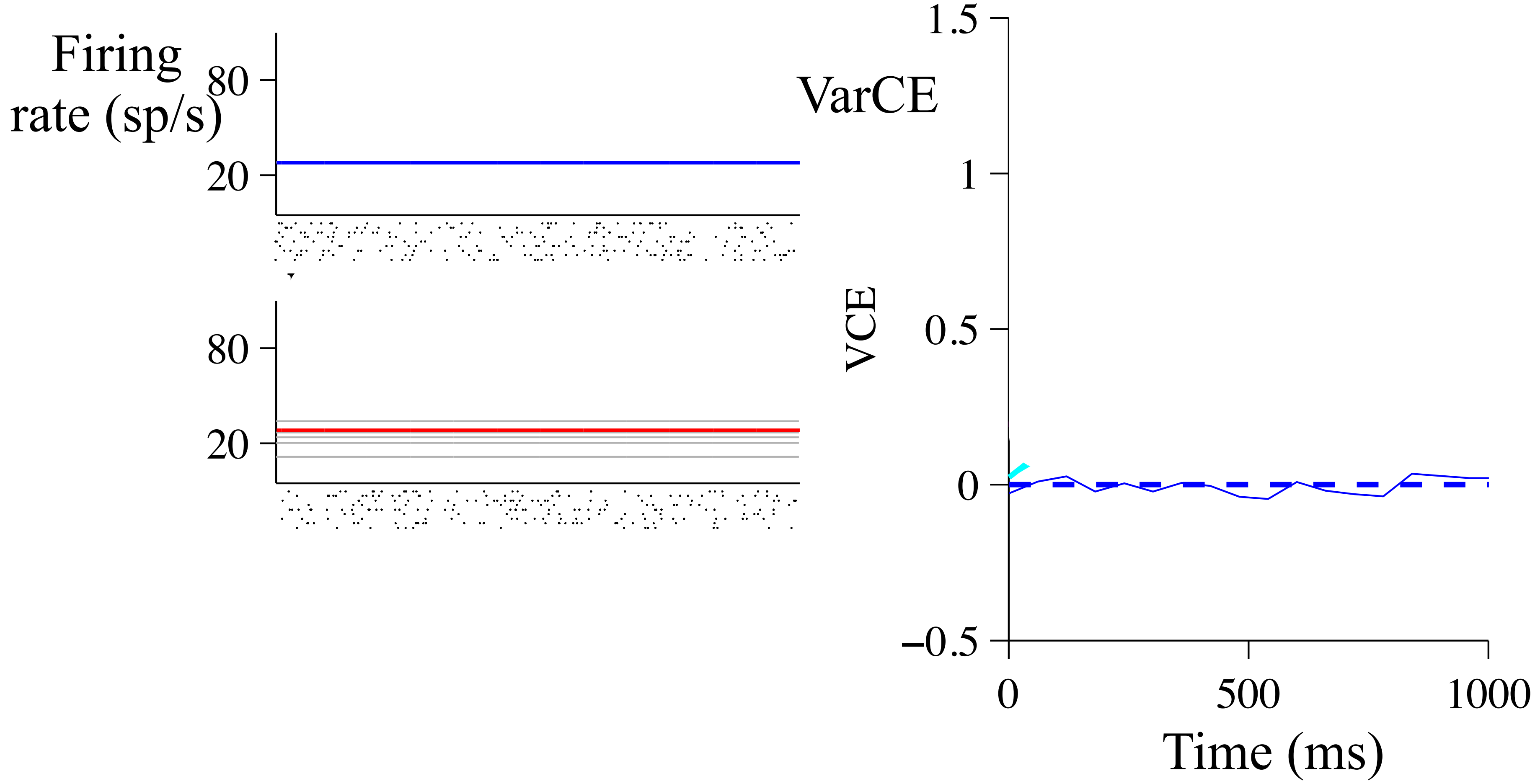
Applied to DSPPs

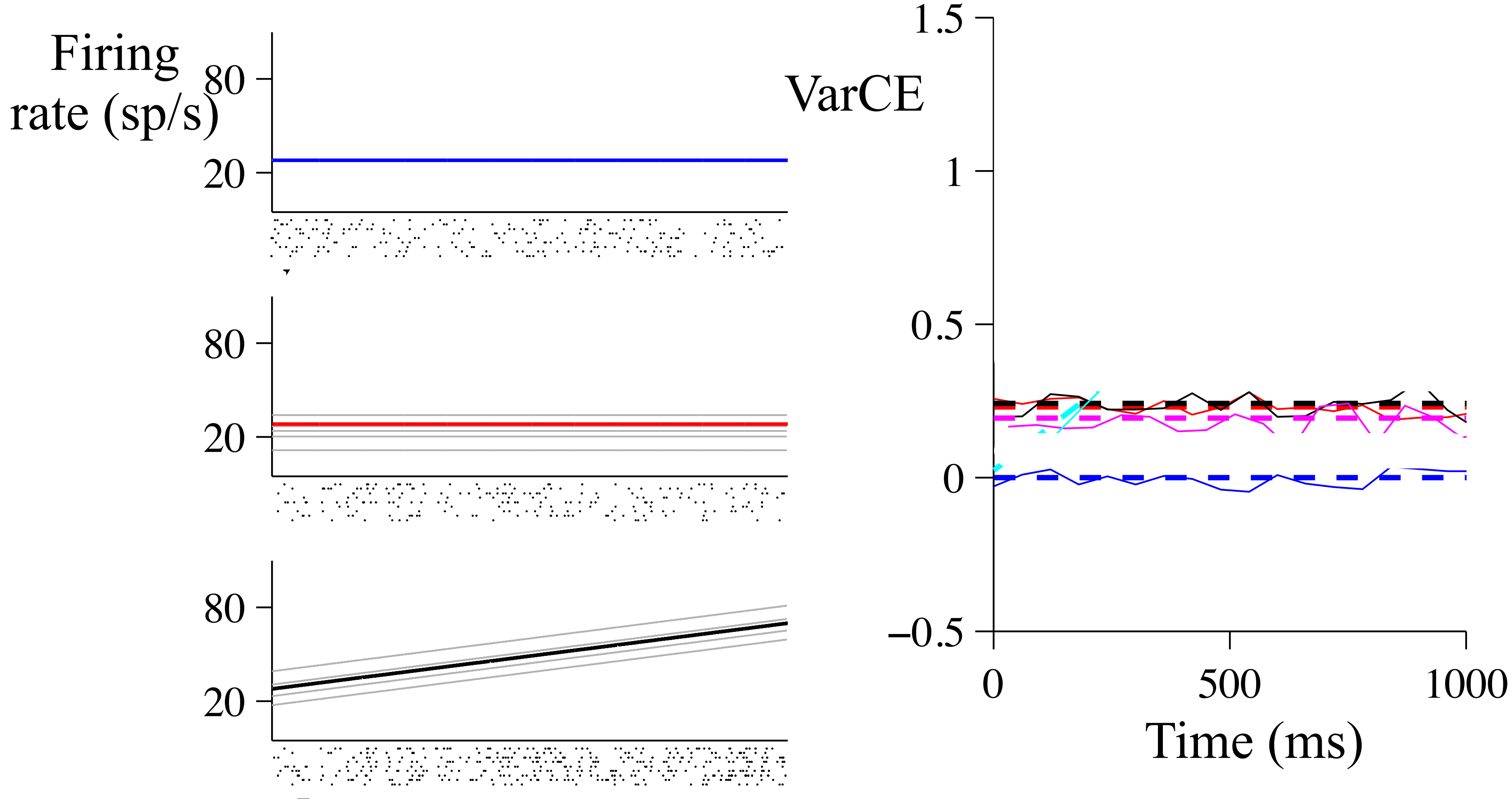
$$\underbrace{\sigma^2_{N_i}}_{\text{Total measured variance}} = \underbrace{\sigma^2_{\langle N_i \rangle}}_{VCE} + \underbrace{\left\langle \sigma^2_{N|\lambda_i} \right\rangle}_{\text{Point process variance (PPV)}}$$

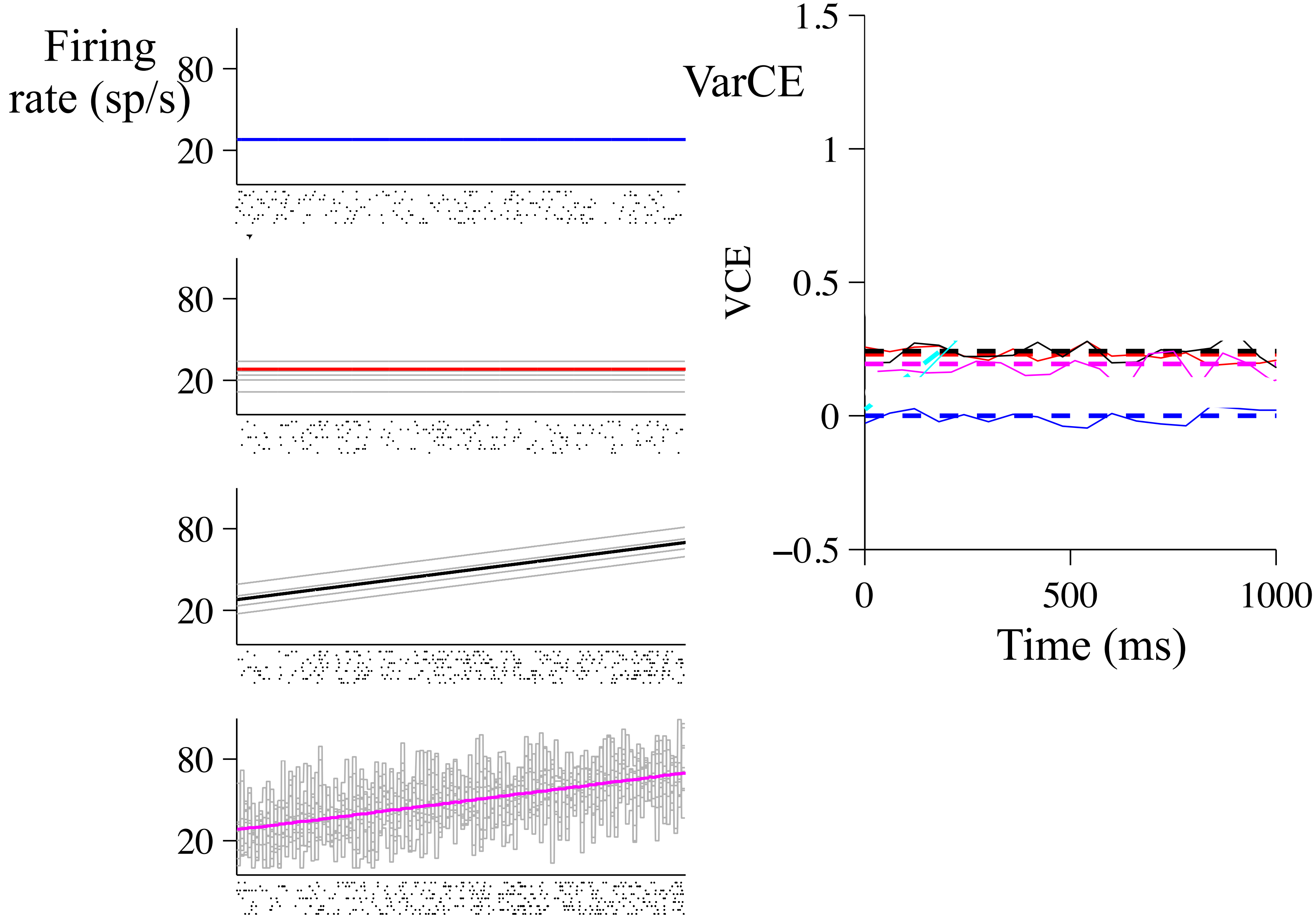
Estimator of VCE

$$s^2_{\langle N_i \rangle} = s^2_{N_i} - \phi \overline{N_i}$$

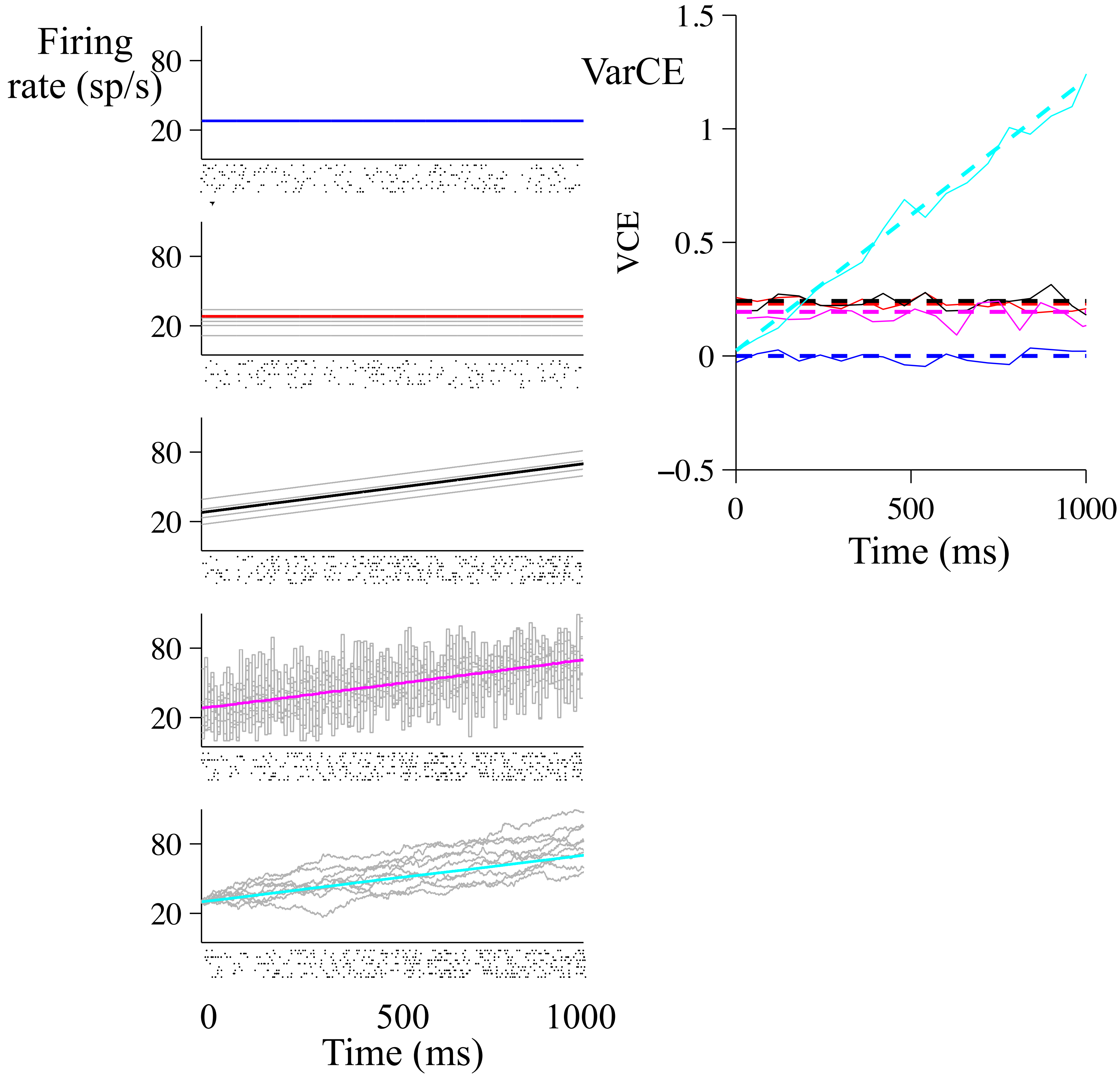

Firing rate (sp/s)
80
20
VarCE
1.5
1
0.5
0
−0.5
0
500
1000
Time (ms)


Firing rate (sp/s)
80
20
80
20
VarCE
1.5
1
0.5
0
−0.5
VCE
0
500
1000
Time (ms)

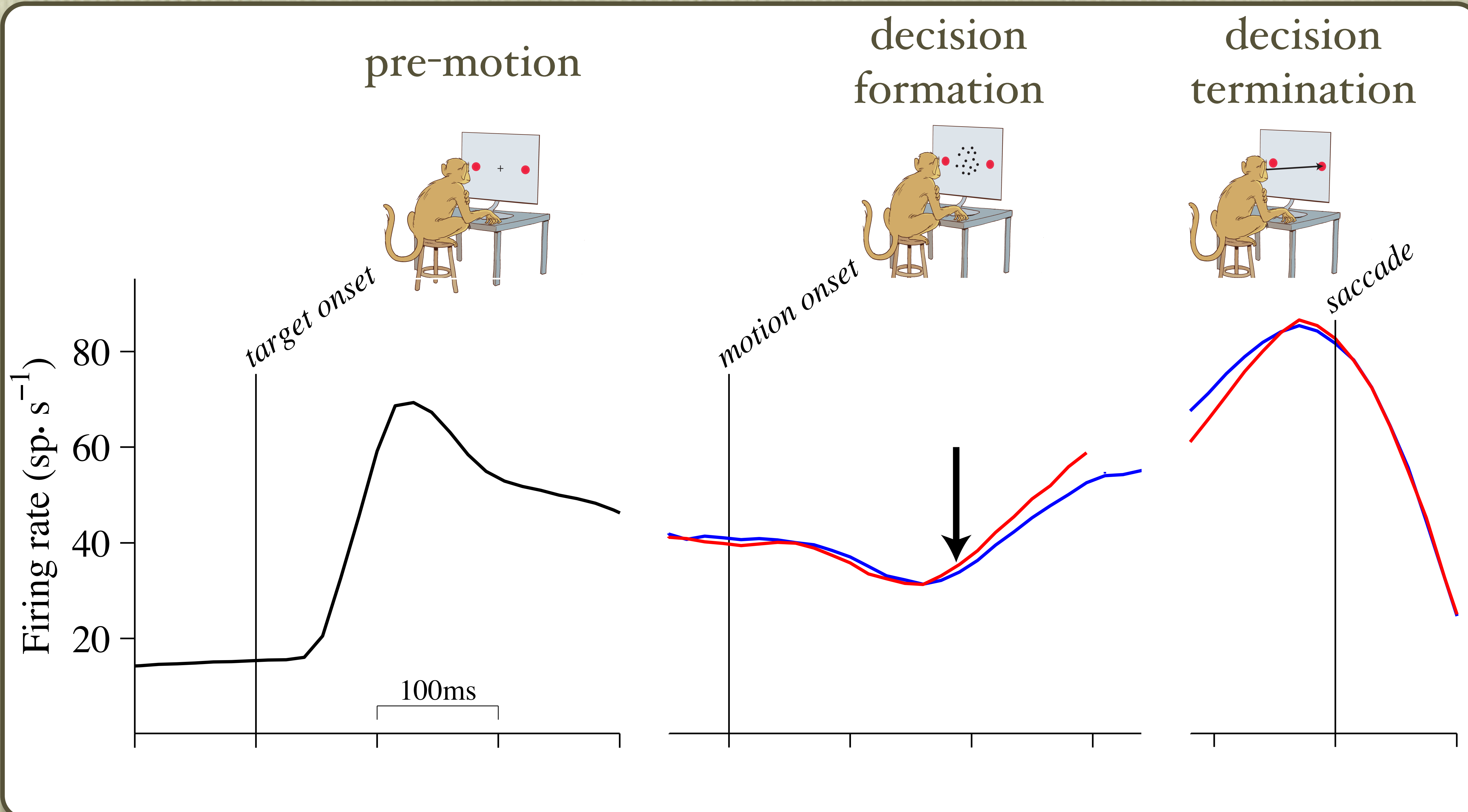

Firing
rate (sp/s)
80
20
80
20
80
20
VarCE
1.5
1
0.5
0
−0.5
0
500
1000
Time (ms)



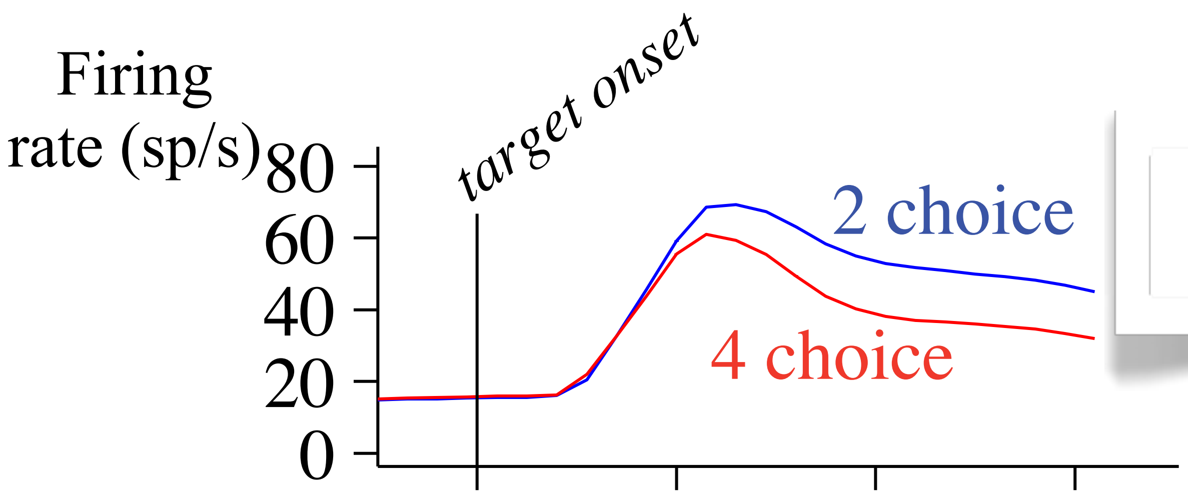
Firing rate (sp/s)
80
20
80
20
80
20
80
20
VarCE
1.5
1
0.5
0
−0.5
VCE
0
500
1000
Time (ms)


Firing rate (sp/s)
80
20
80
20
80
20
80
20
80
20
0
500
1000
Time (ms)
VarCE
1.5
1
0.5
0
−0.5
VCE
0
500
1000
Time (ms)

Three analysis epochs

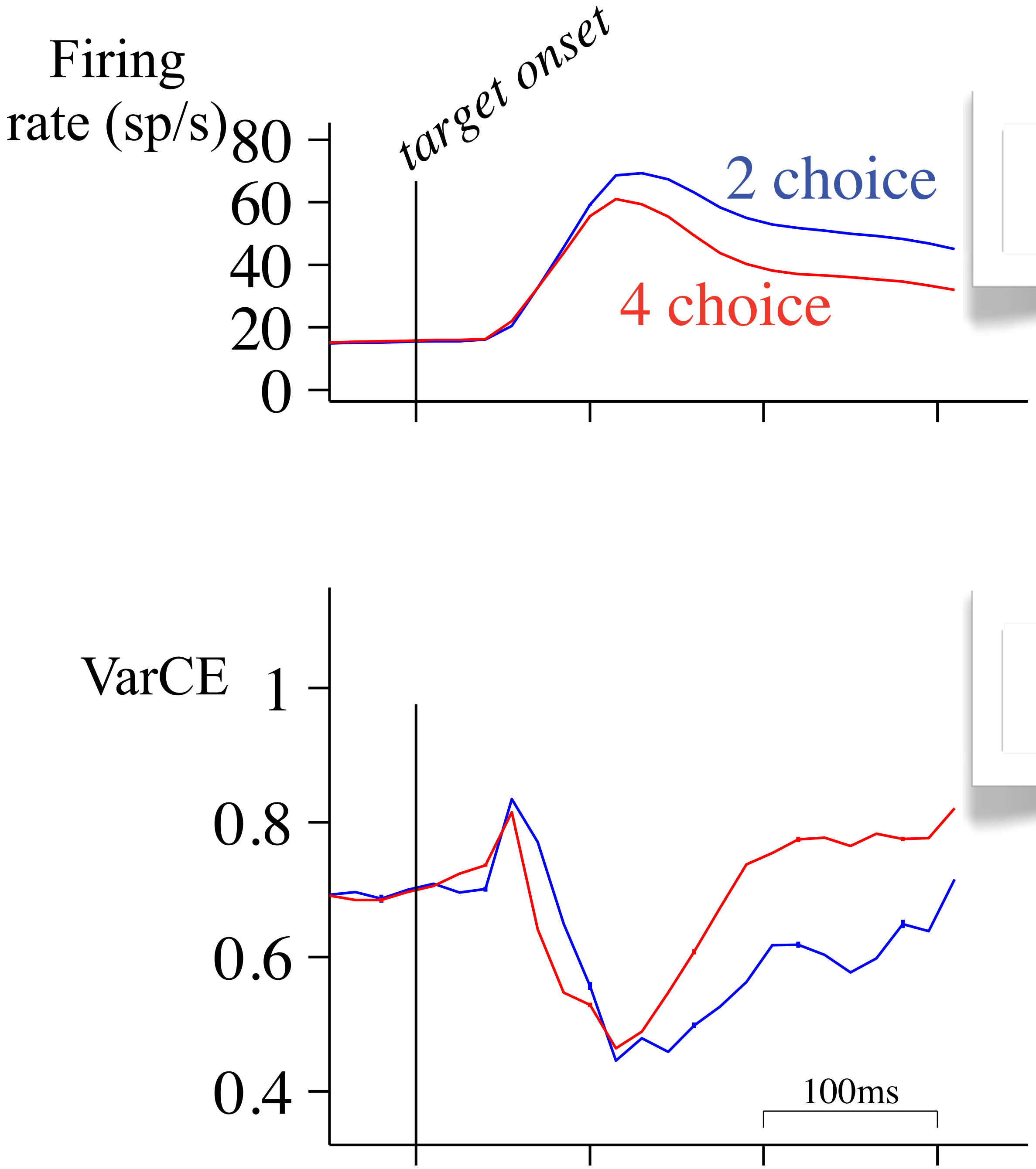


Pre-motion epoch



consistent with
Basso & Wurtz (1997, 1998)

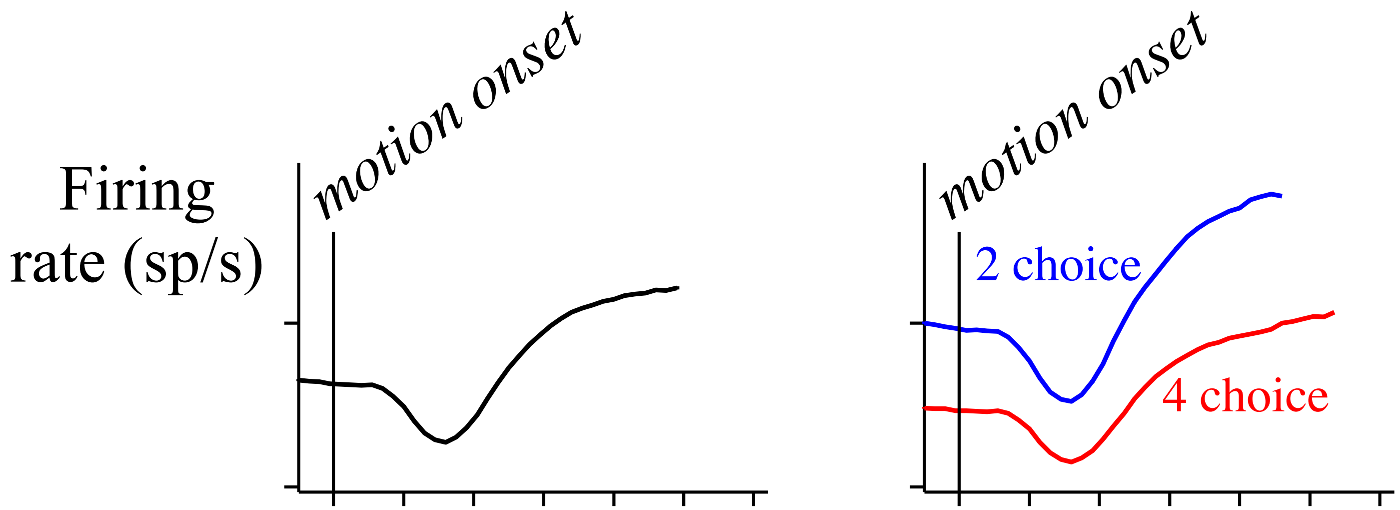
Pre-motion epoch



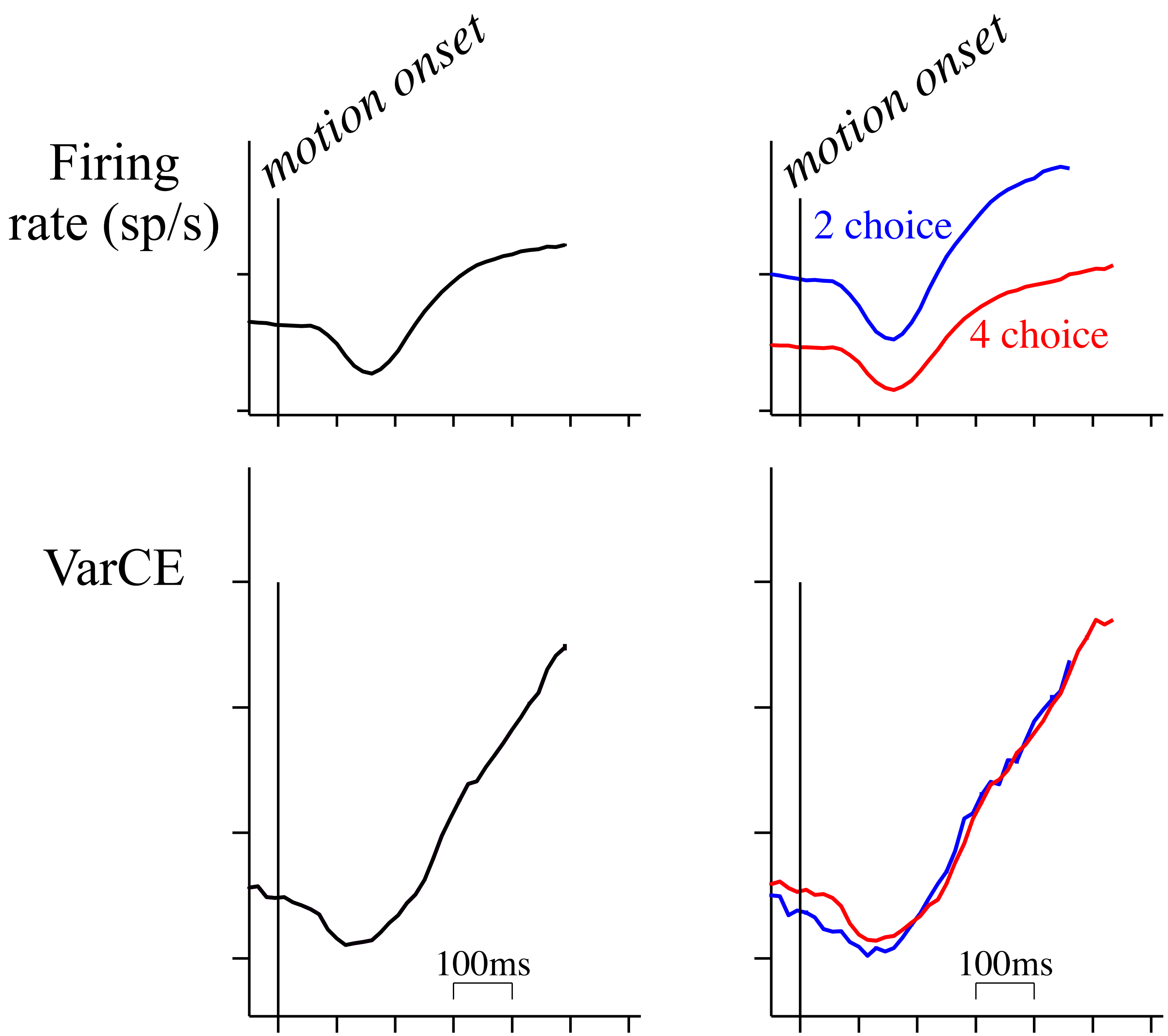
consistent with
Basso & Wurtz (1997, 1998)

Lower FR explained by
mixture of states

Early motion viewing



Early motion viewing



Doubly stochastic point processes

Law of total covariance

$$Cov\left[N_i, N_j\right] = \underbrace{Cov\left[\left\langle N_i, N_j \middle| \lambda_i, \lambda_j \right\rangle\right]}_{\substack{\text{covariance of} \\ \text{conditional expectation}}} + \underbrace{\left\langle Cov\left[N_i, N_j \middle| \lambda_i, \lambda_j \right]\right\rangle}_{\substack{\text{expectation of} \\ \text{conditional covariance}}}$$

Doubly stochastic point processes

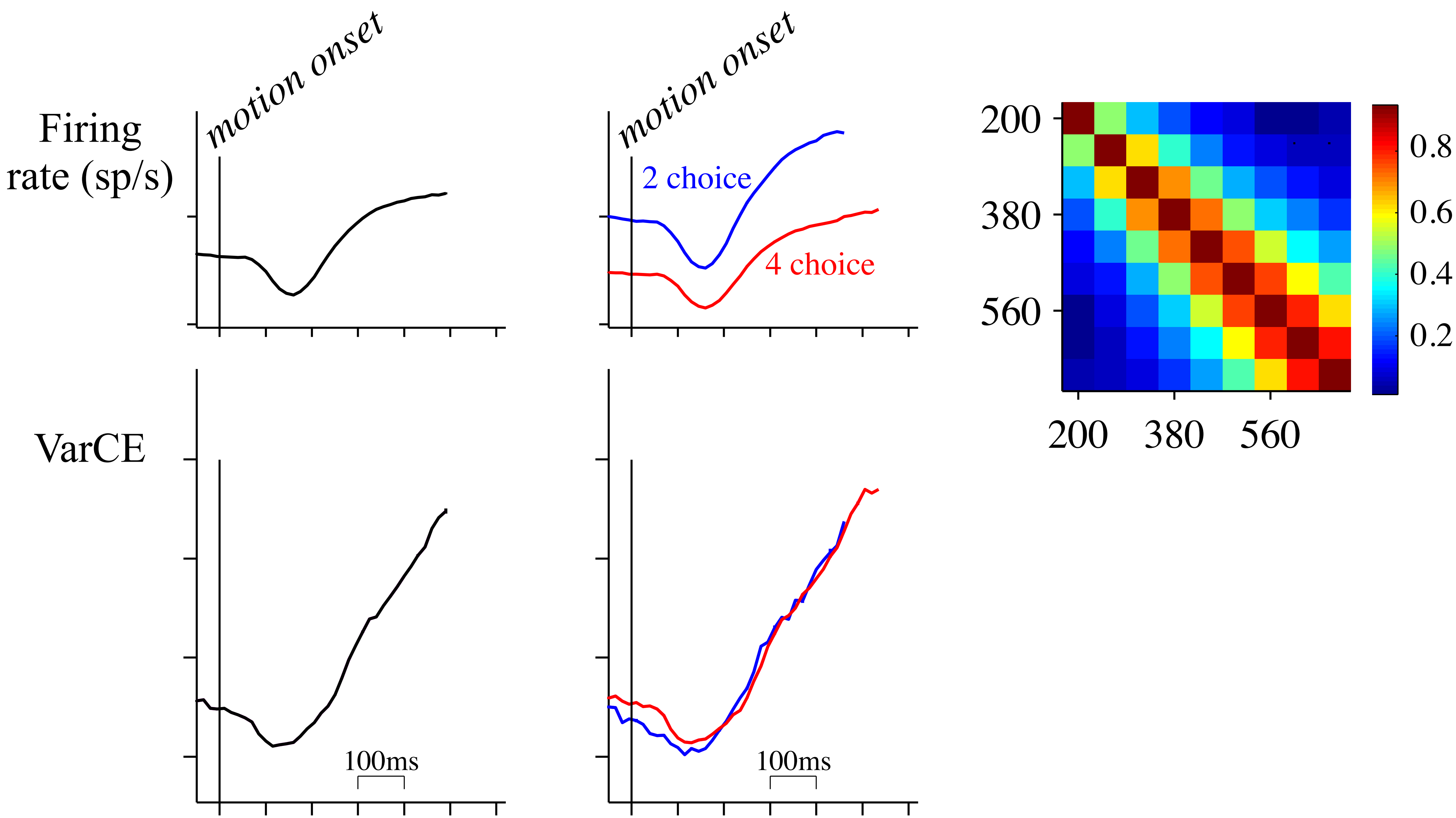
Law of total covariance

$$Cov\left[N_i, N_j\right] = \underbrace{Cov\left[\left\langle N_i, N_j \middle| \lambda_i, \lambda_j \right\rangle\right]}_{\text{covariance of conditional expectation}} + \underbrace{\left\langle Cov\left[N_i, N_j \middle| \lambda_i, \lambda_j \right]\right\rangle}_{\text{expectation of conditional covariance}}$$

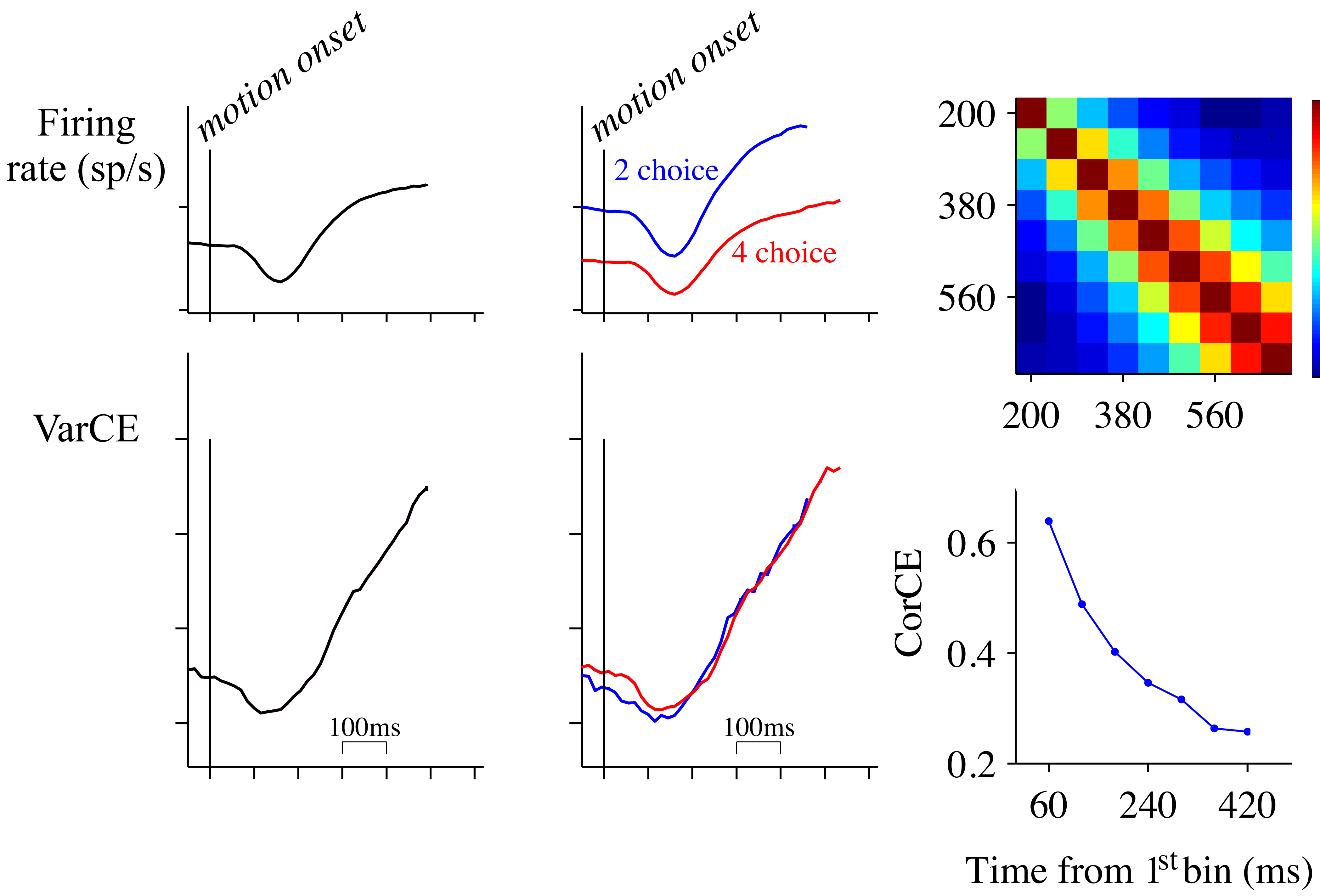
Applied to DSPPs

$$= \left.\begin{array}{rr} VCE + & PPV \\ CovCE + & 0 \end{array}\right\} \begin{array}{l} i = j \\ i \neq j \end{array}$$

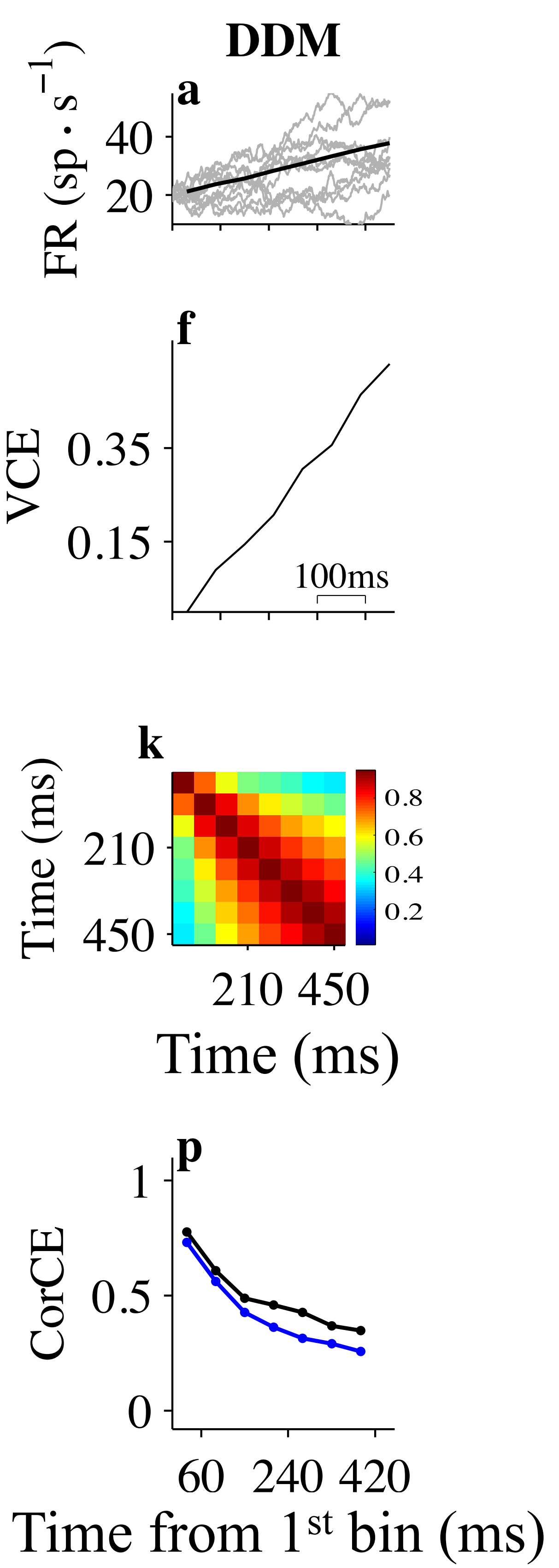
Early motion viewing



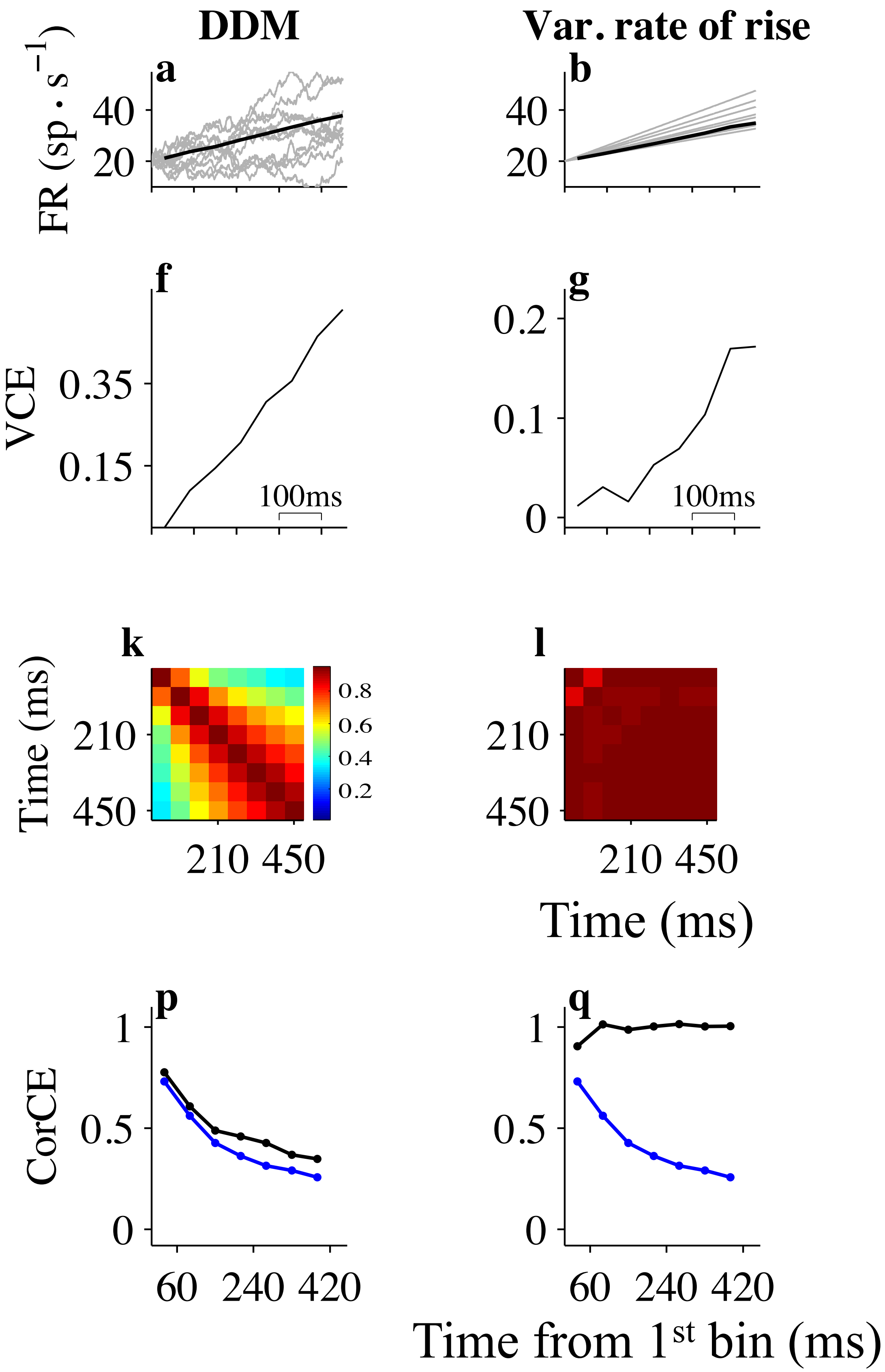
Early motion viewing



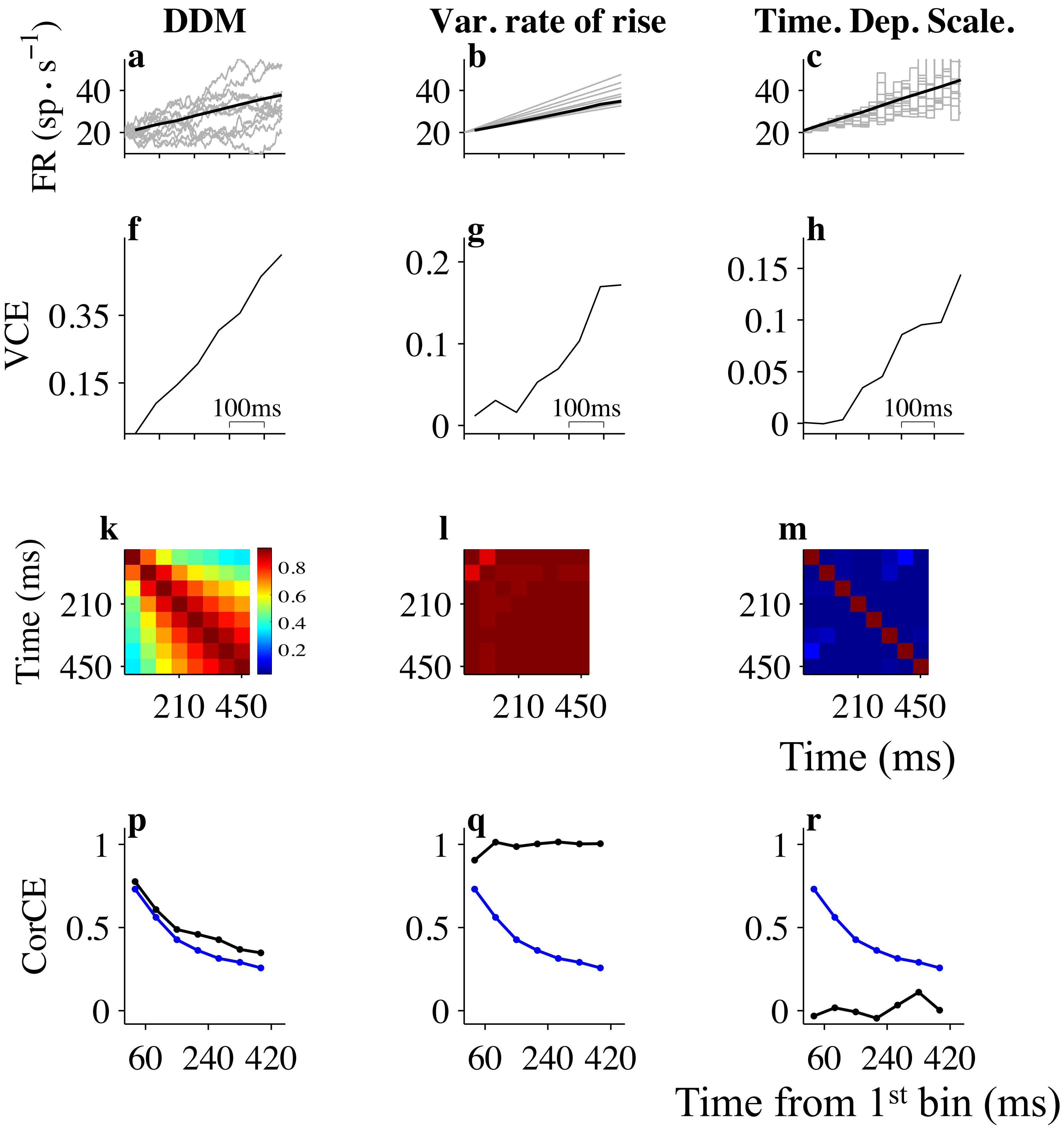
Saturday, June 4, 2011

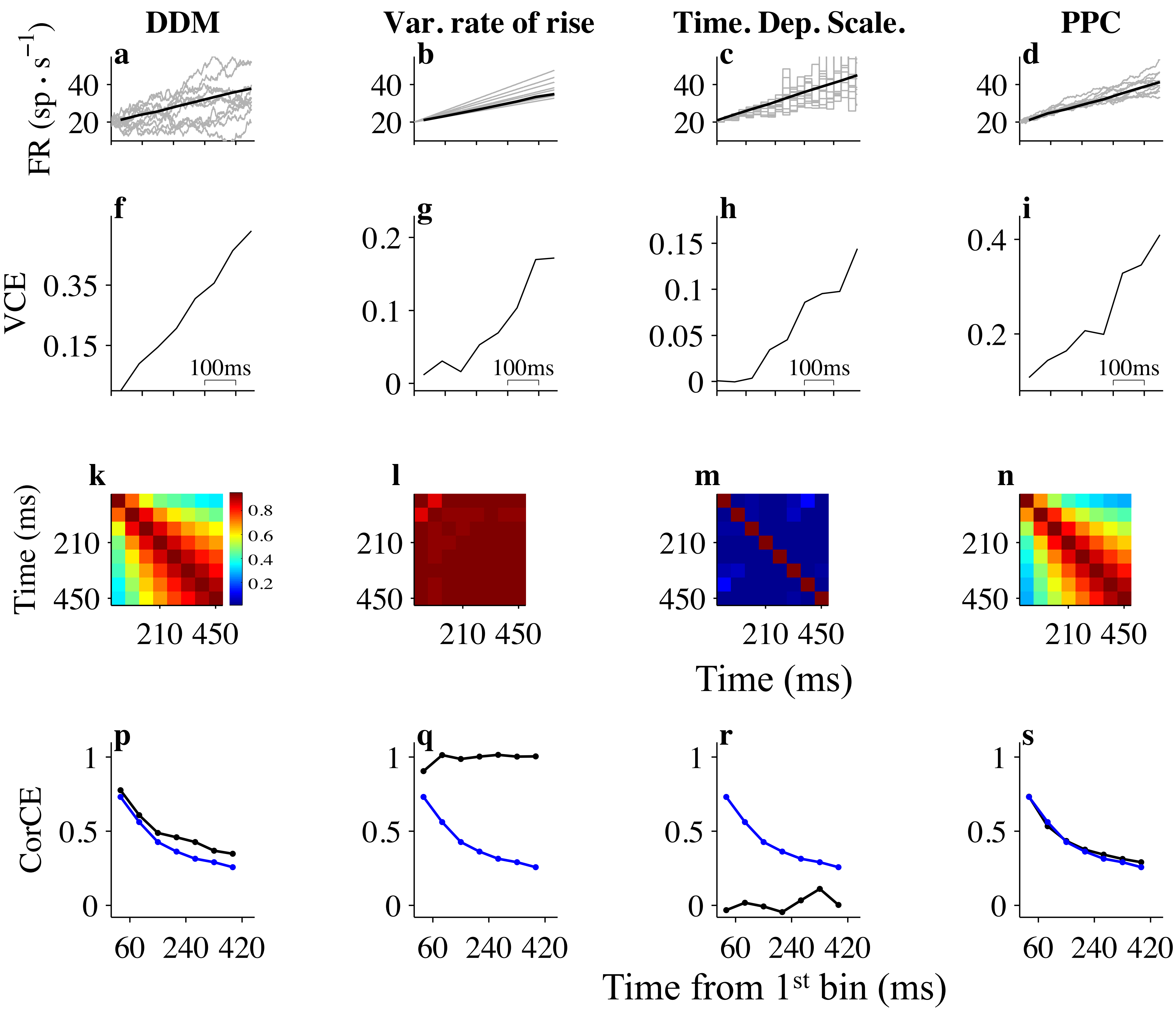


DDM
a
FR (sp · s⁻¹)
40
20
f
VCE
0.35
0.15
100ms
k
Time (ms)
210
450
0.8
0.6
0.4
0.2
210 450
Time (ms)
p
CorCE
1
0.5
0
60 240 420
Time from 1st bin (ms)

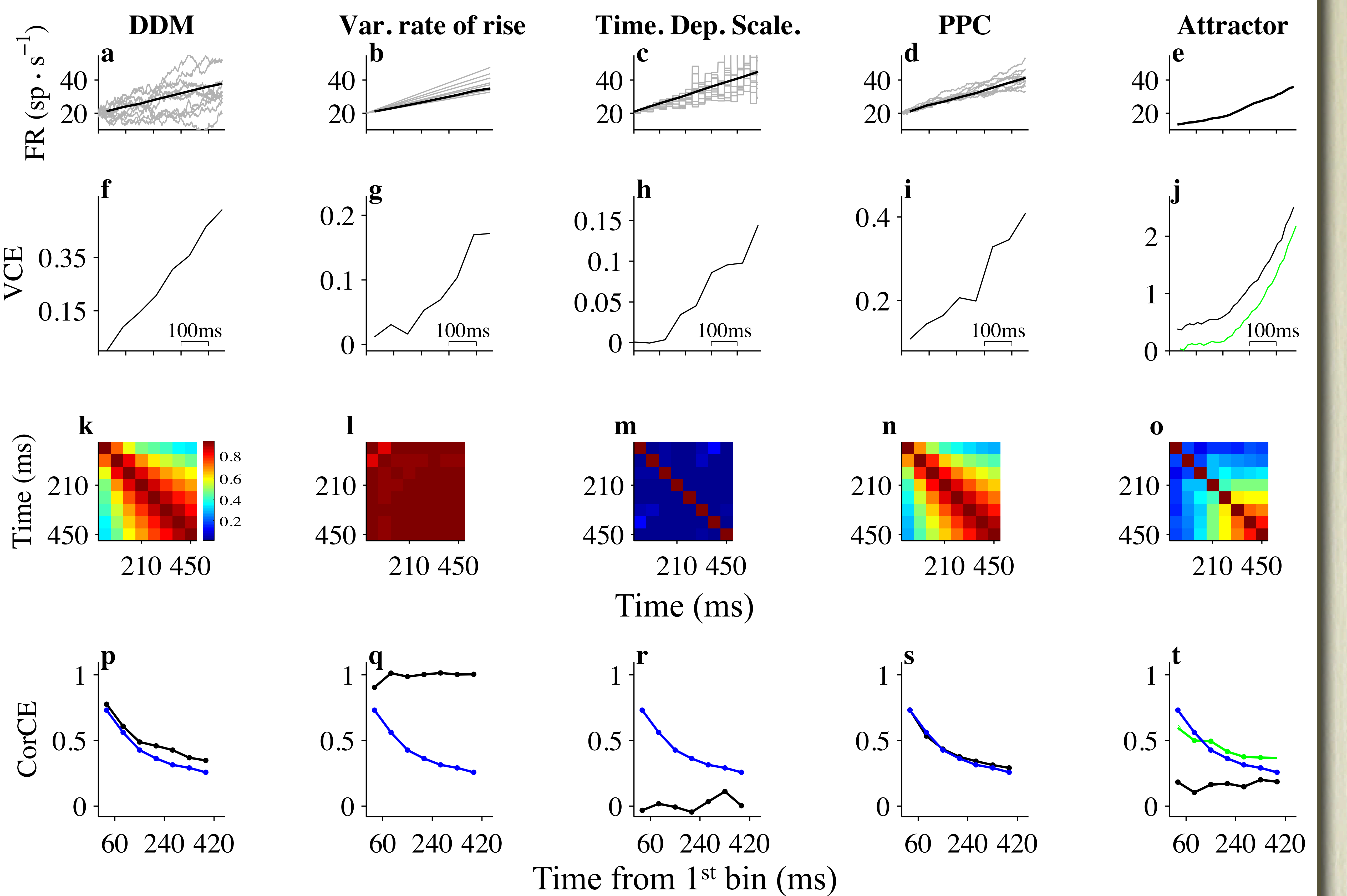


DDM
Var. rate of rise
a
b
FR (sp · s⁻¹)
40
20
f
g
VCE
0.35
0.15
0.2
0.1
0
100ms
k
l
Time (ms)
210
450
0.8
0.6
0.4
0.2
210 450
Time (ms)
p
q
CorCE
1
0.5
0
60 240 420
Time from 1st bin (ms)

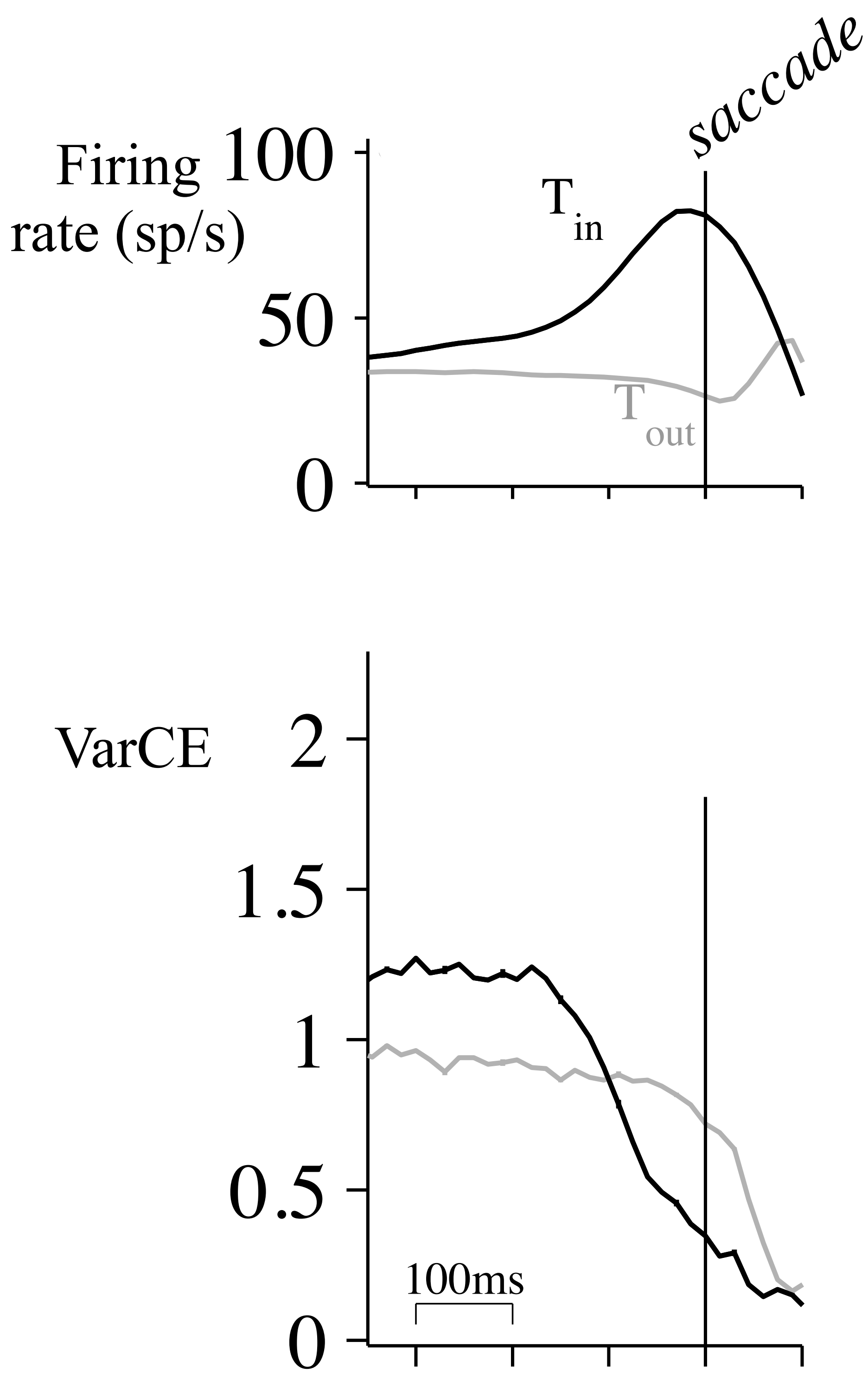

DDM
Var. rate of rise
Time. Dep. Scale.
a
b
c
FR (sp · s −1)
40
20
f
g
h
VCE
0.35
0.15
0.2
0.1
0
0.15
0.1
0.05
100ms
k
l
m
Time (ms)
210
450
0.8
0.6
0.4
0.2
210 450
Time (ms)
p
q
r
CorCE
1
0.5
0
60 240 420
Time from 1st bin (ms)



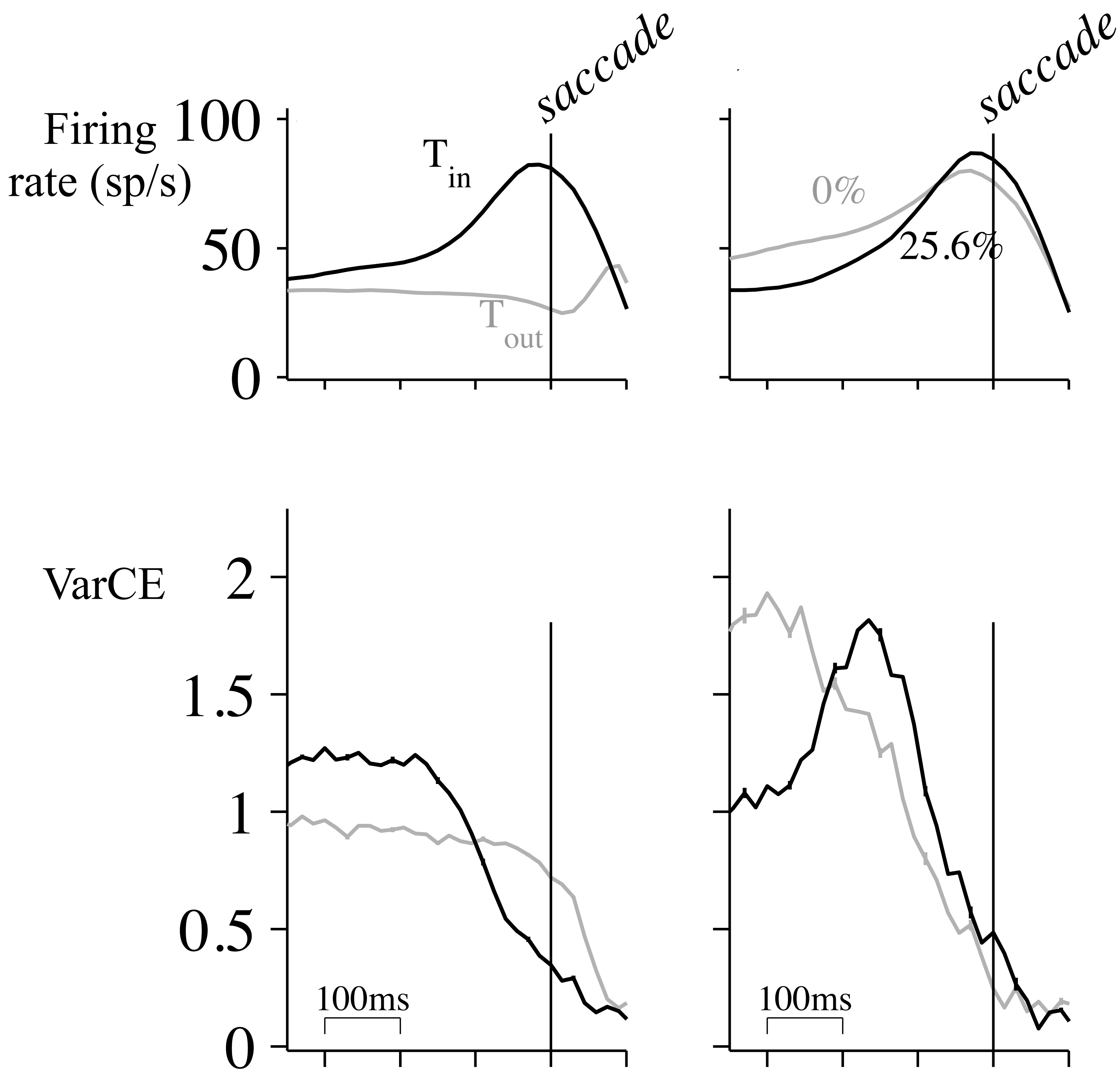
DDM
Var. rate of rise
Time. Dep. Scale.
PPC
FR (sp · s^{-1})
VCE
Time (ms)
CorCE
100ms
Time (ms)
Time from 1st bin (ms)


DDM
Var. rate of rise
Time. Dep. Scale.
PPC
Attractor
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
FR (sp · s⁻¹)
VCE
Time (ms)
CorCE
100ms
Time (ms)
Time from 1st bin (ms)

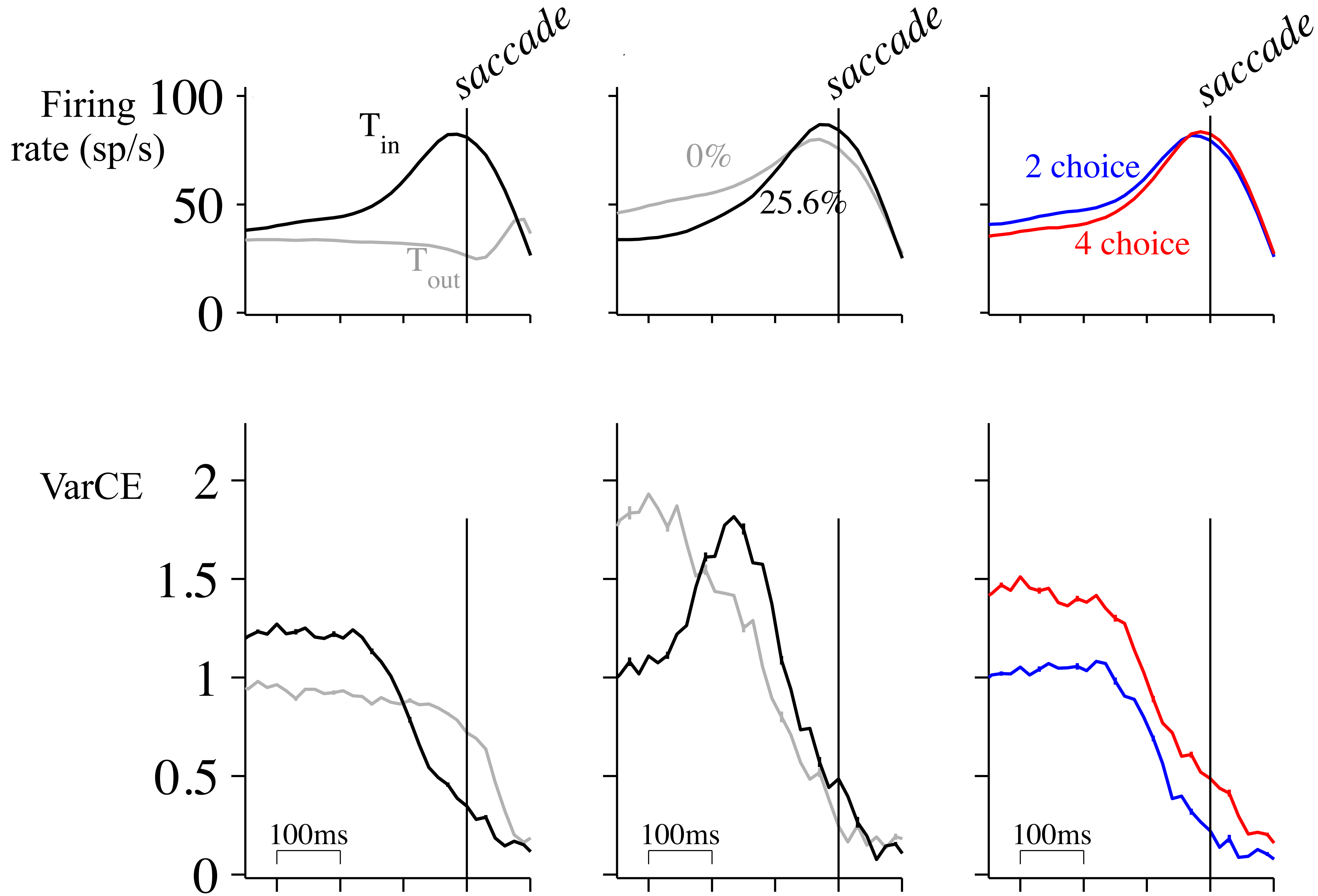
Decision termination



Decision termination



Decision termination



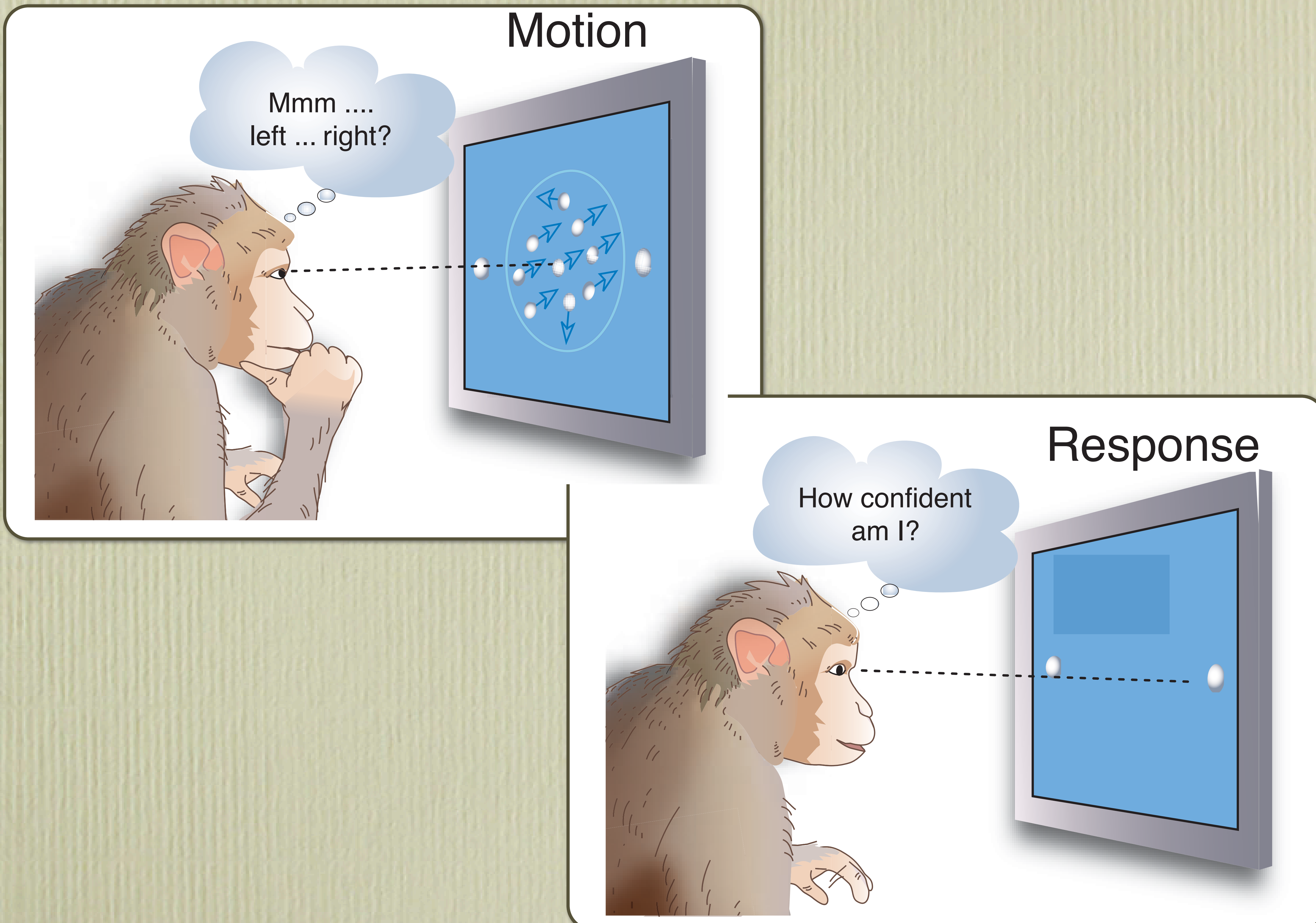
Summary of section

- VarCE and CorCE are useful tools
  - Capture "variation in what is computed"
  - Expose features of neural computations in decision making

    e.g., integration, mixtures, termination bound, refutes change point and several plausible alternative models
- The main limitation is in estimating ϕ

Outline

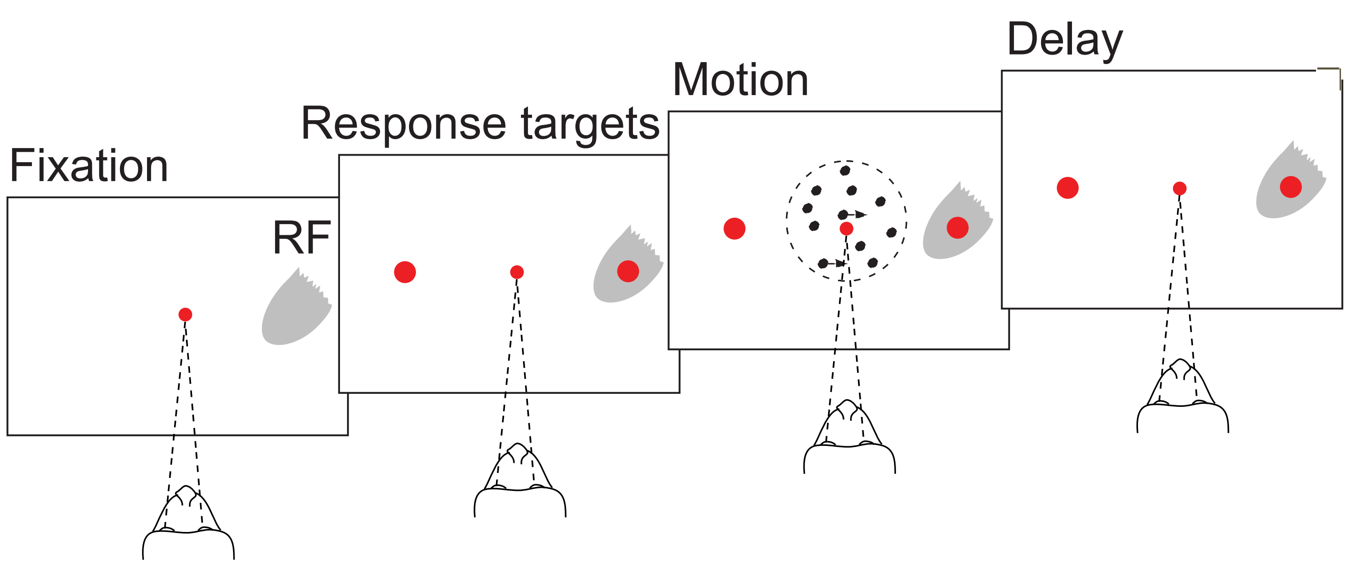
1. Probabilistic reasoning
2. Perceptual decisions: speed and accuracy
3. *Optional*: Sequential analysis, Wald's martingale, logistic choice function
4. Variance and covariance as signatures of neural computation
5. Confidence
6. *Optional*: Integration of prior probability & evidence


Motion
Mmm ....
left ... right?
Response
How confident
am I?

Post-decision wagering

based on Hampton (2001) PNAS

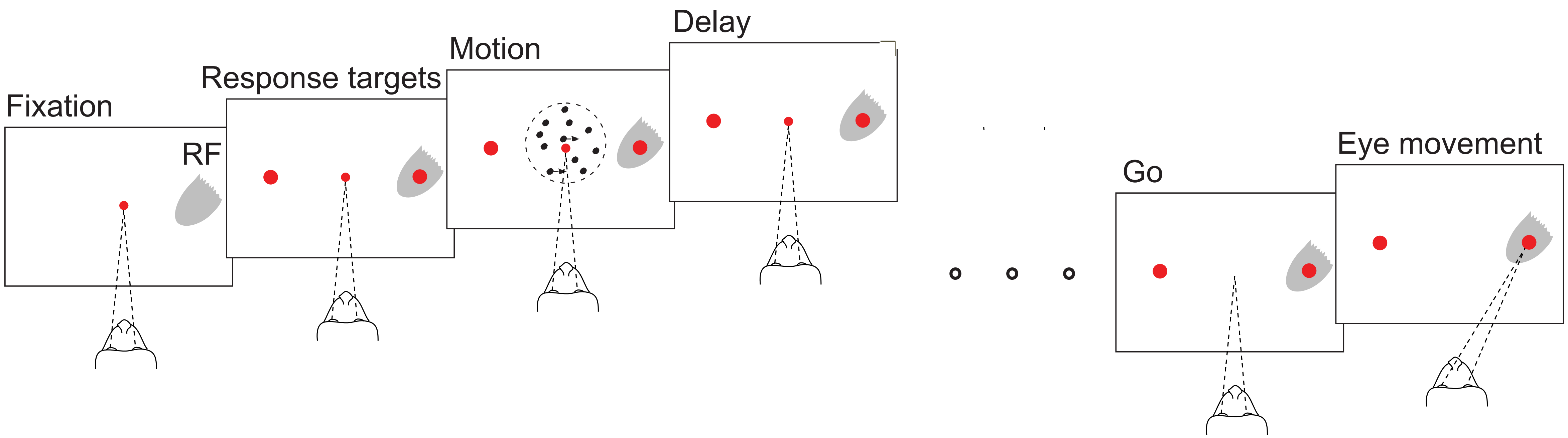
Roozbeh Kiani



Kiani & Shadlen (2009) Science 324:759-764.

Post-decision wagering

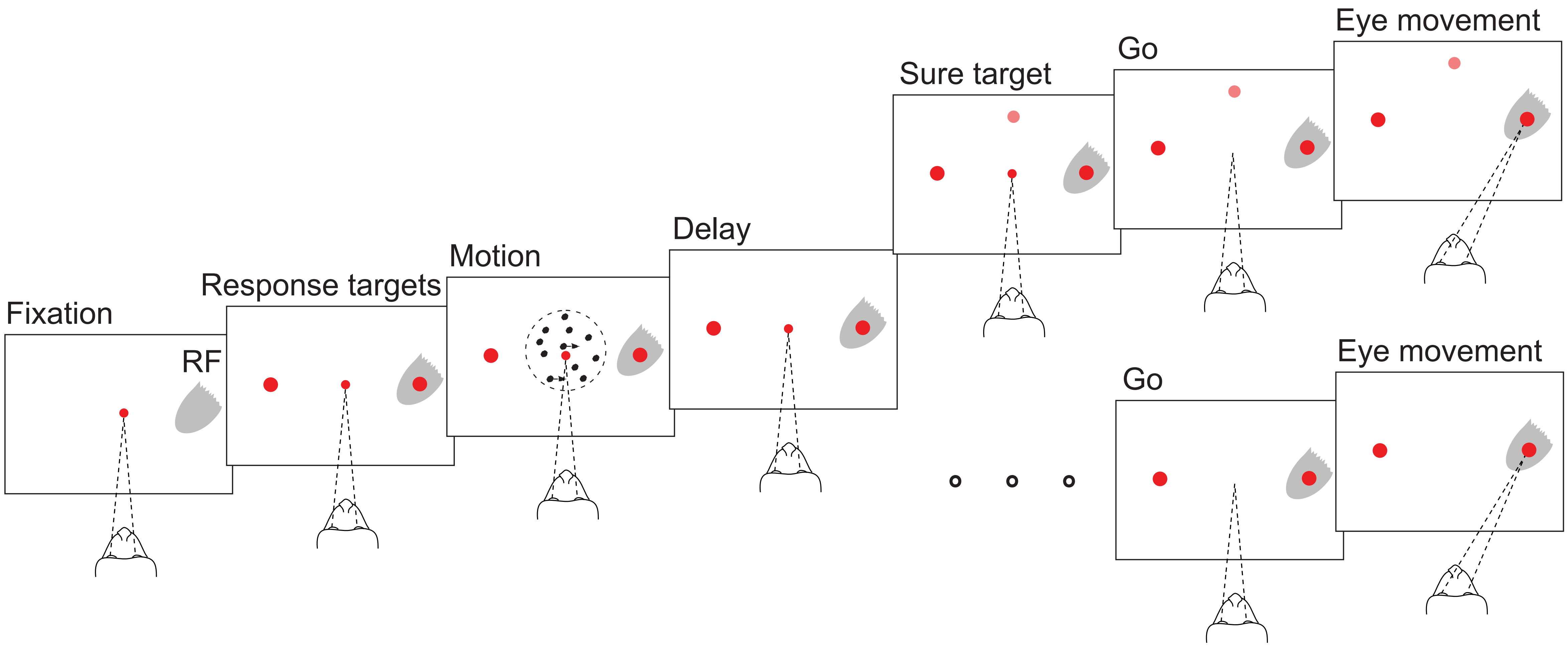
Roozbeh Kiani



Kiani & Shadlen (2009) Science 324:759-764.

Post-decision wagering

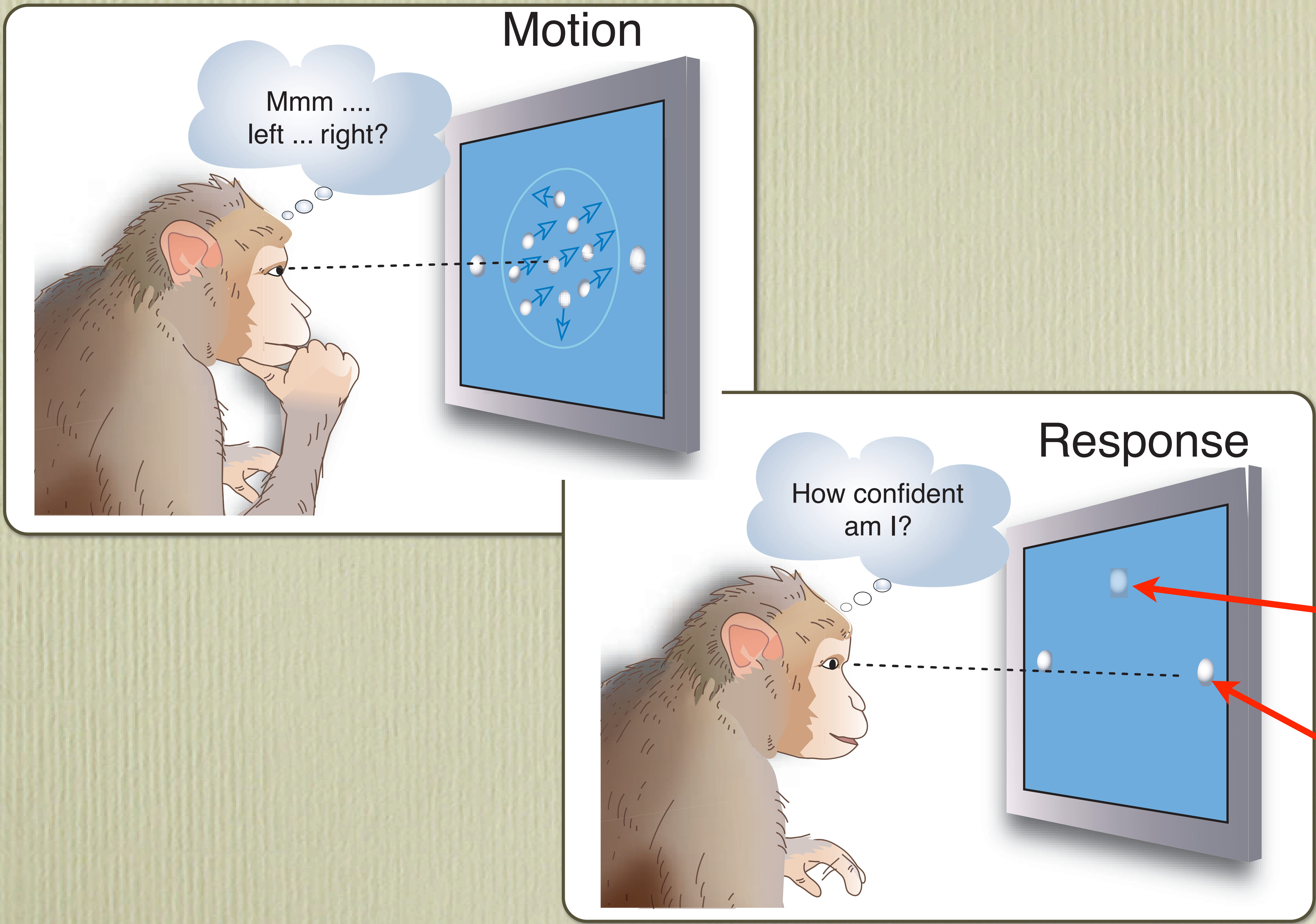
Roozbeh Kiani

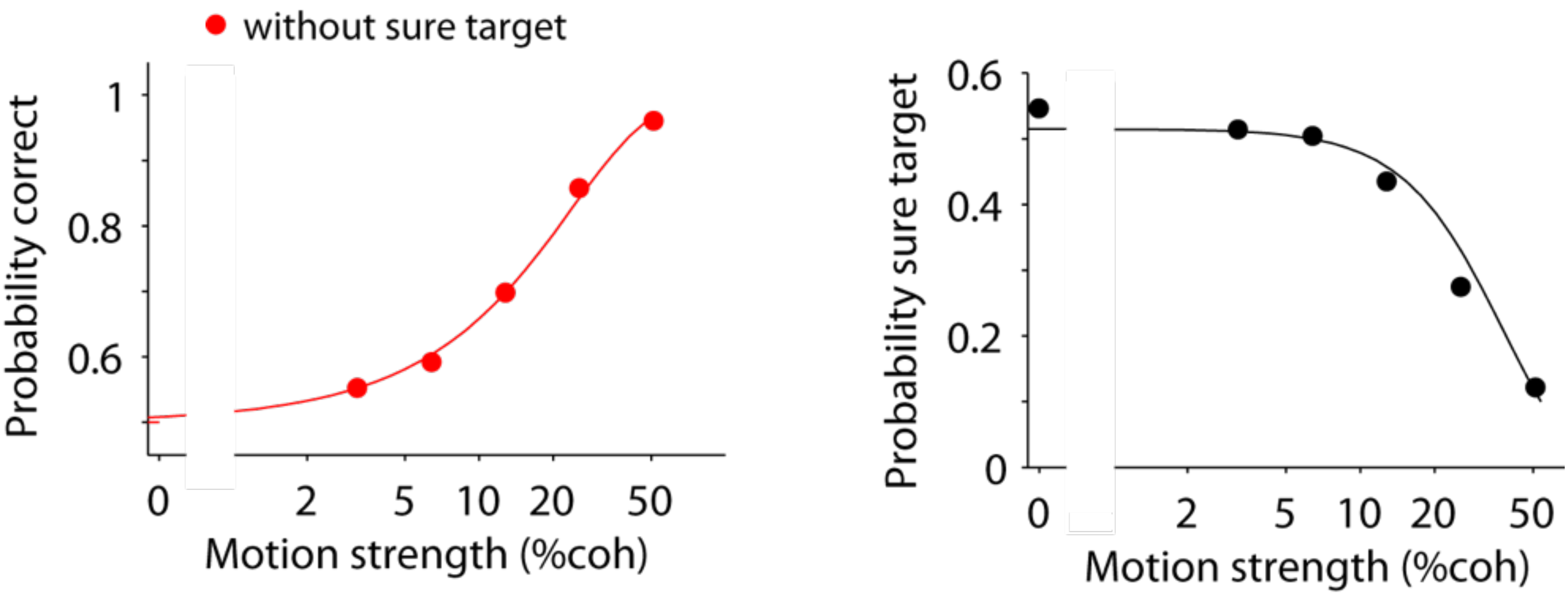


Kiani & Shadlen (2009) Science 324:759-764.

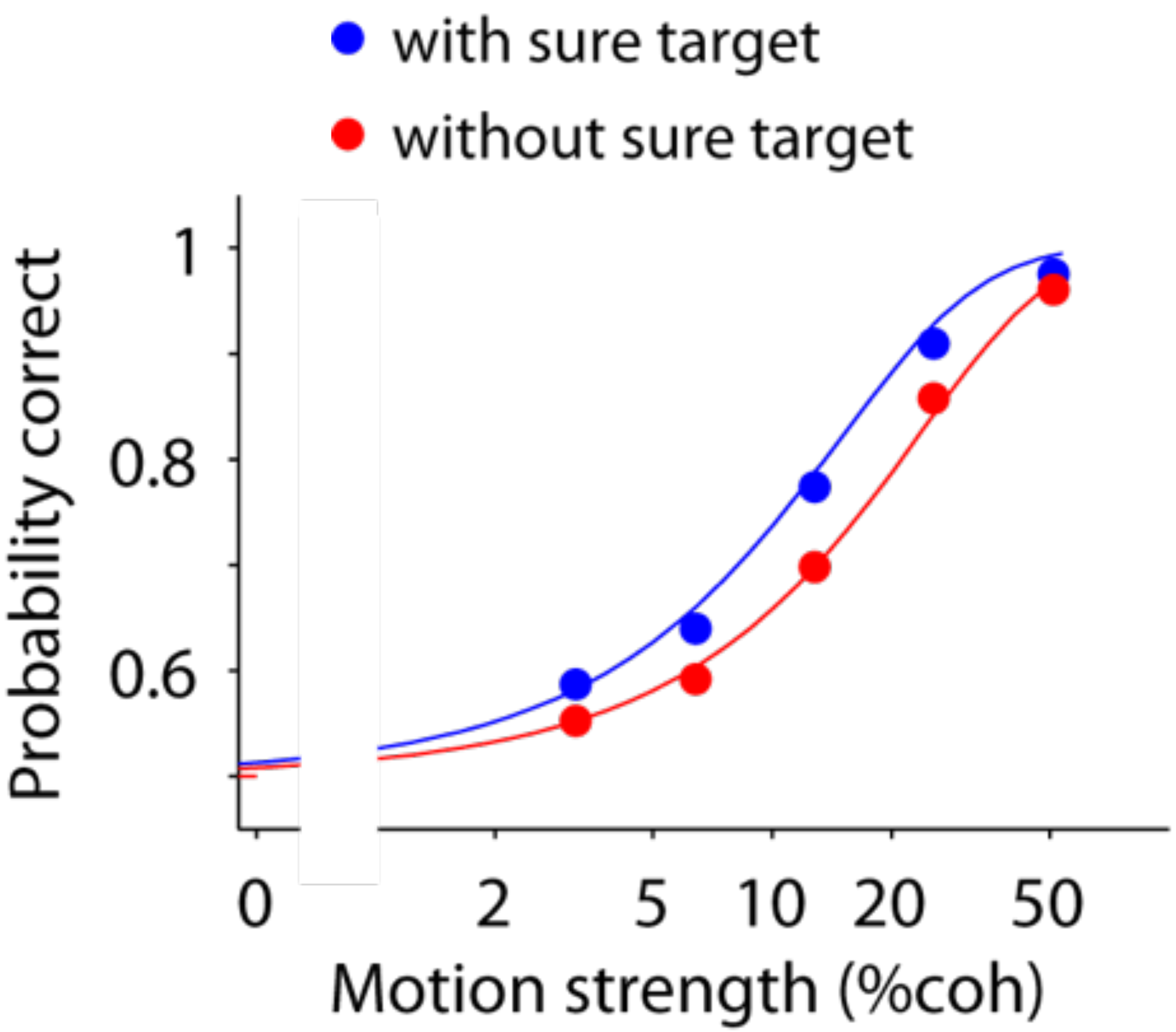
Post-decision wagering

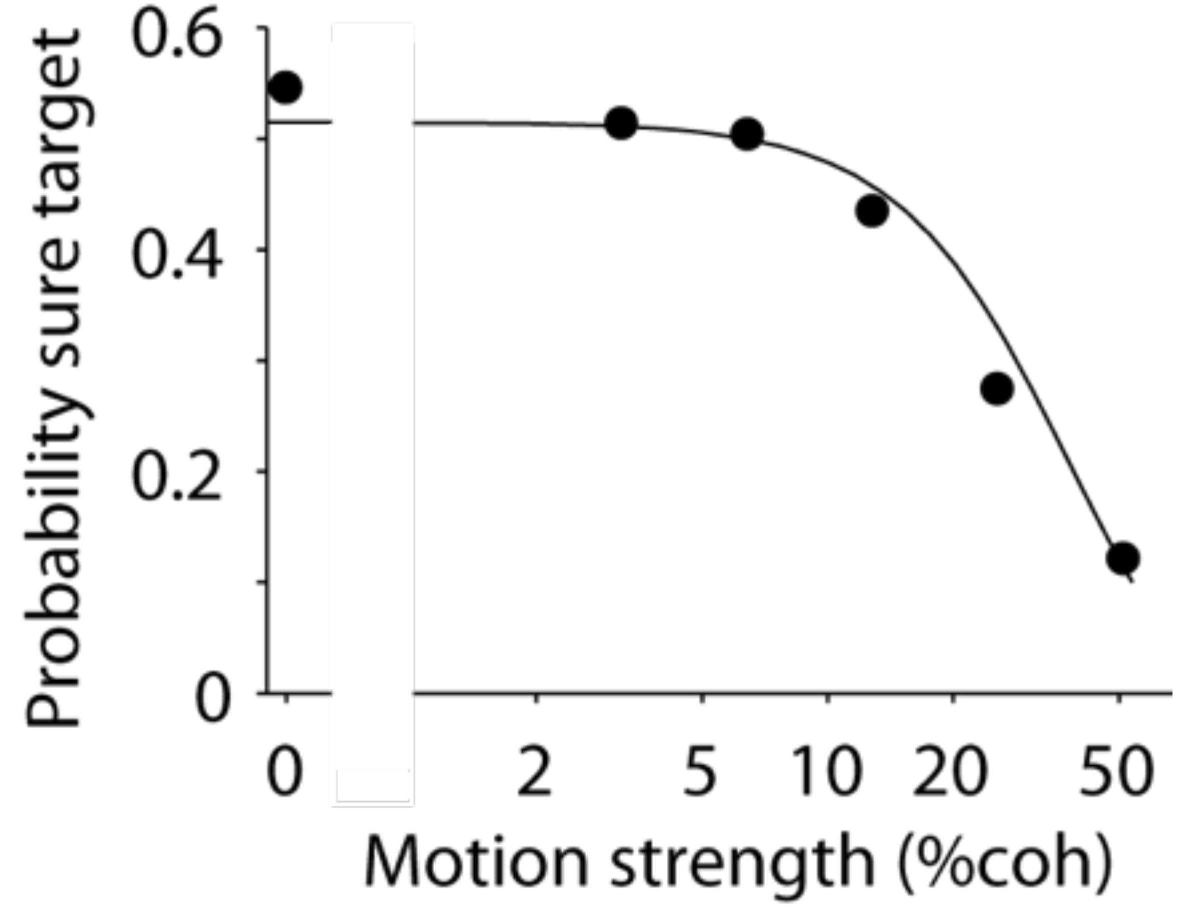

Roozbeh Kiani



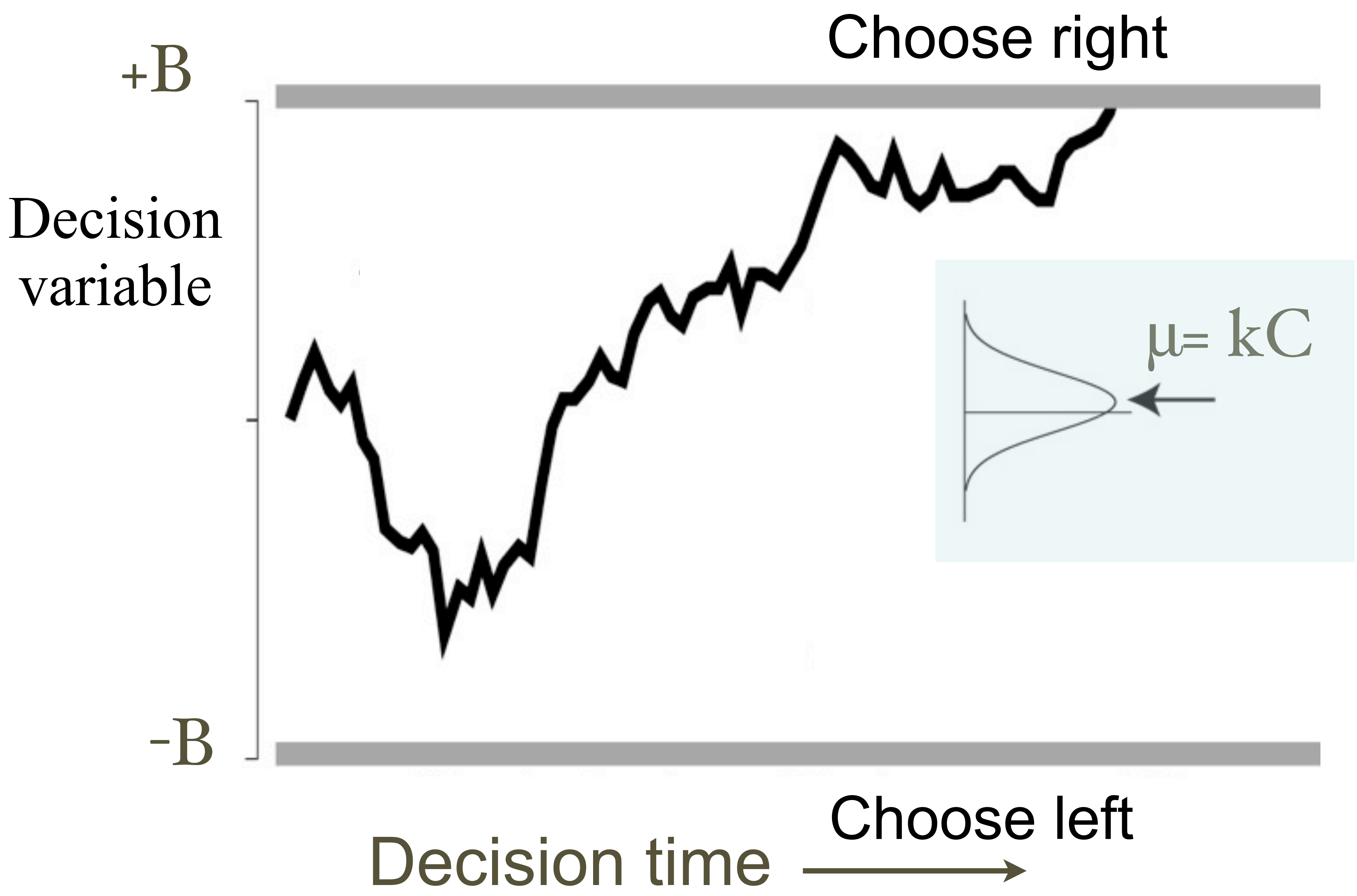


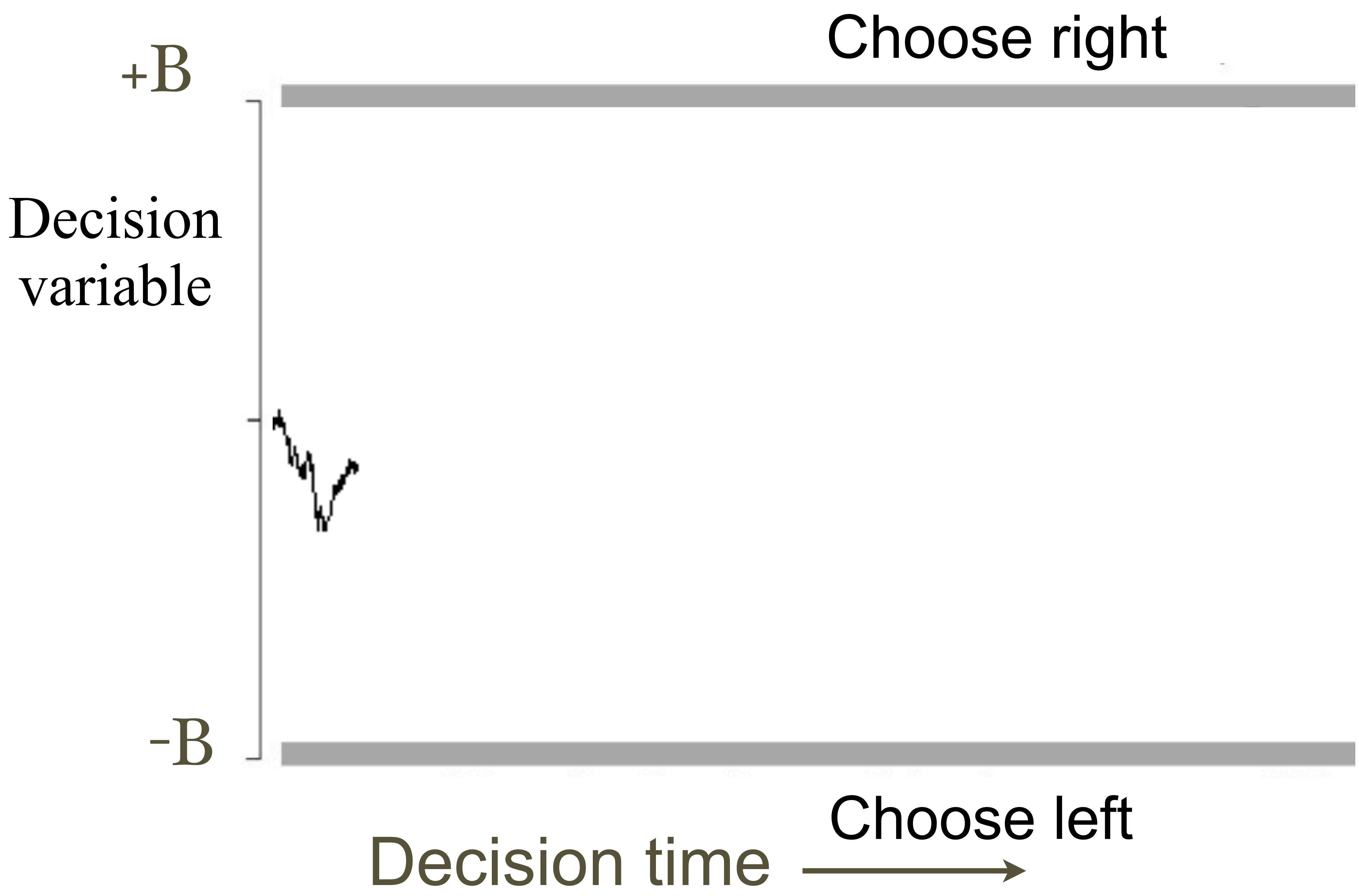
Kiani & Shadlen (2009) Science 324:759-764.



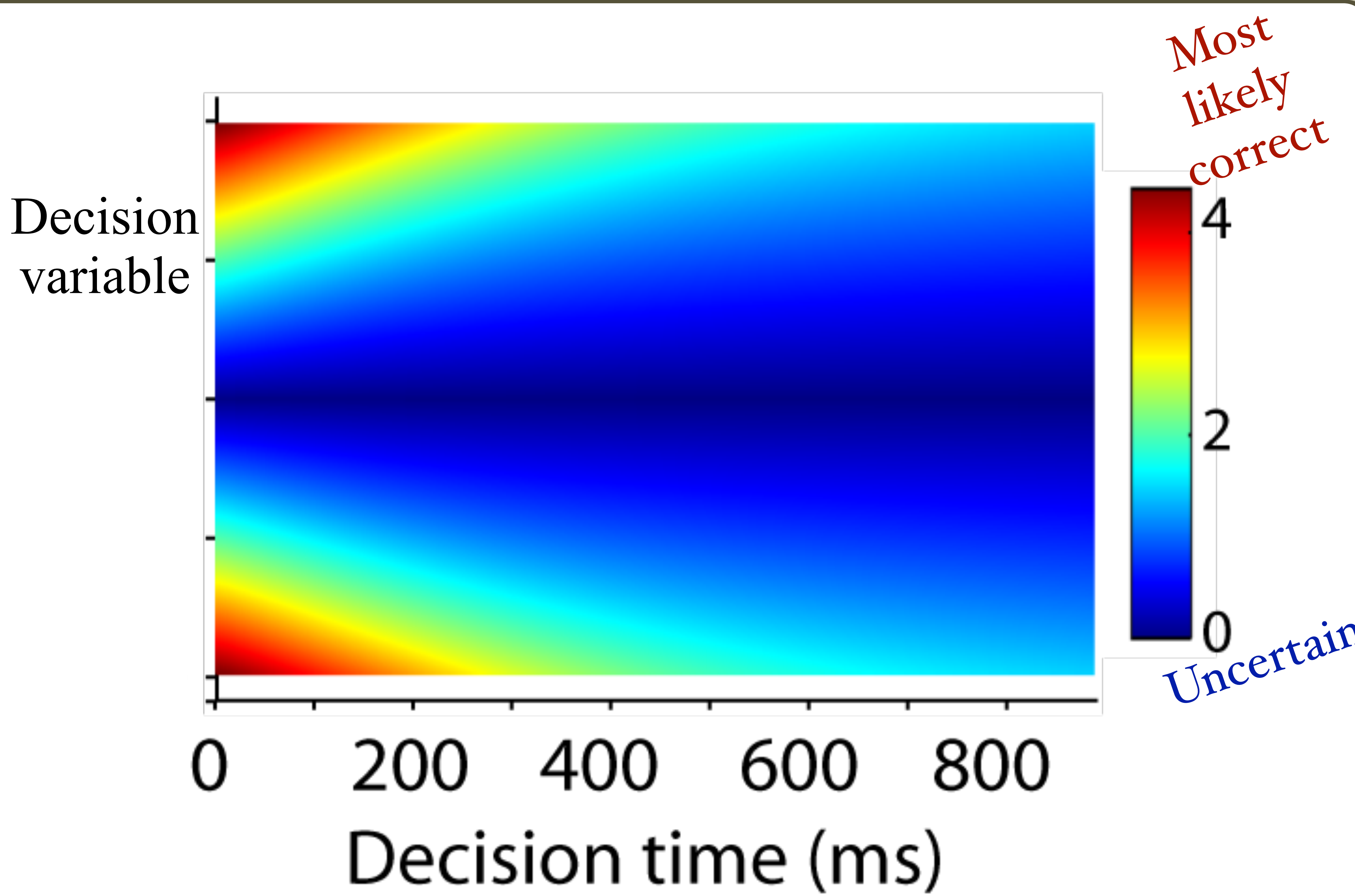


Kiani & Shadlen (2009) Science 324:759-764.


Choose right
+B
Decision variable
μ= kC
-B
Choose left
Decision time

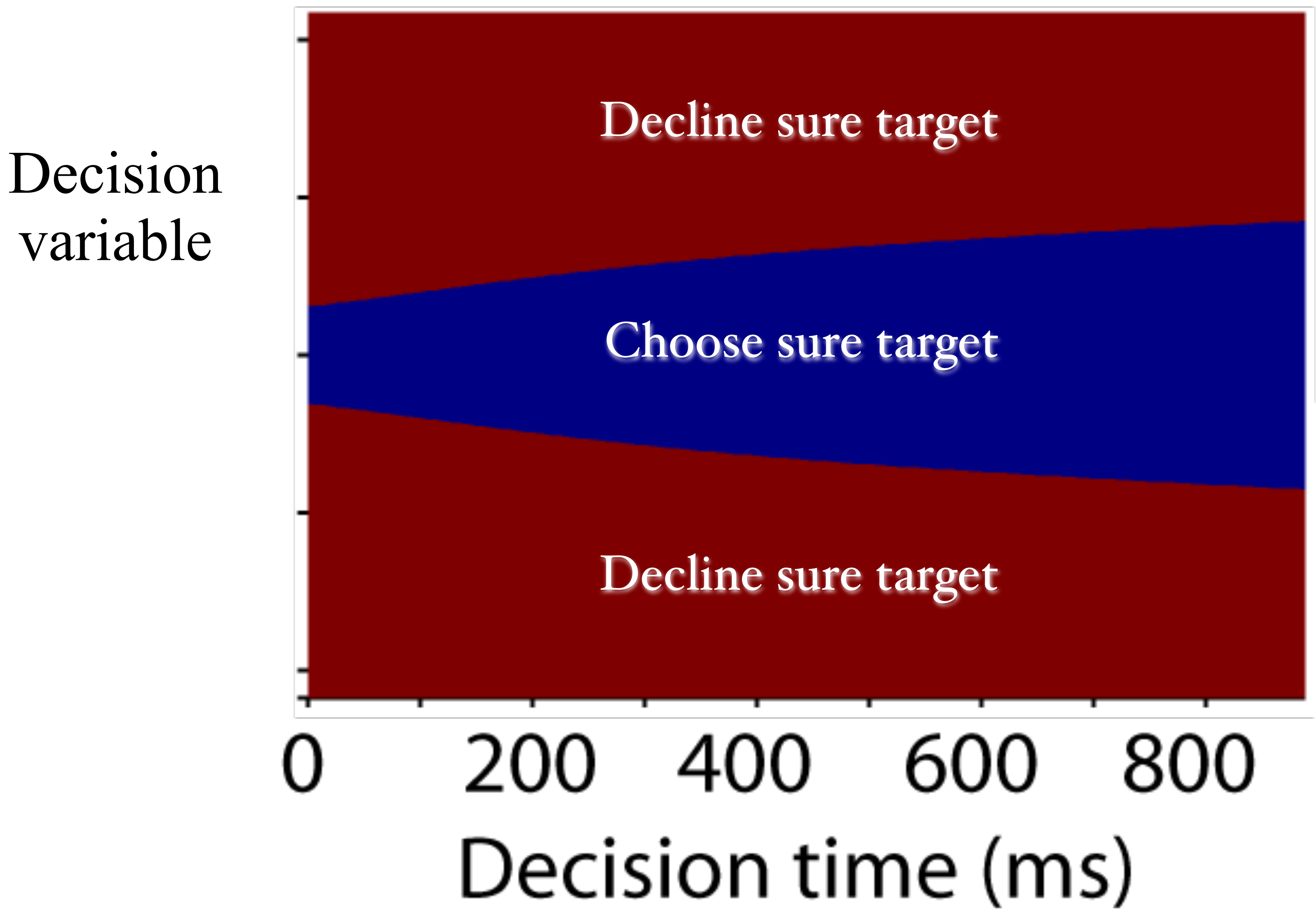

Choose right
+B
Decision
variable
-B
Choose left
Decision time

Log odds of making the correct choice

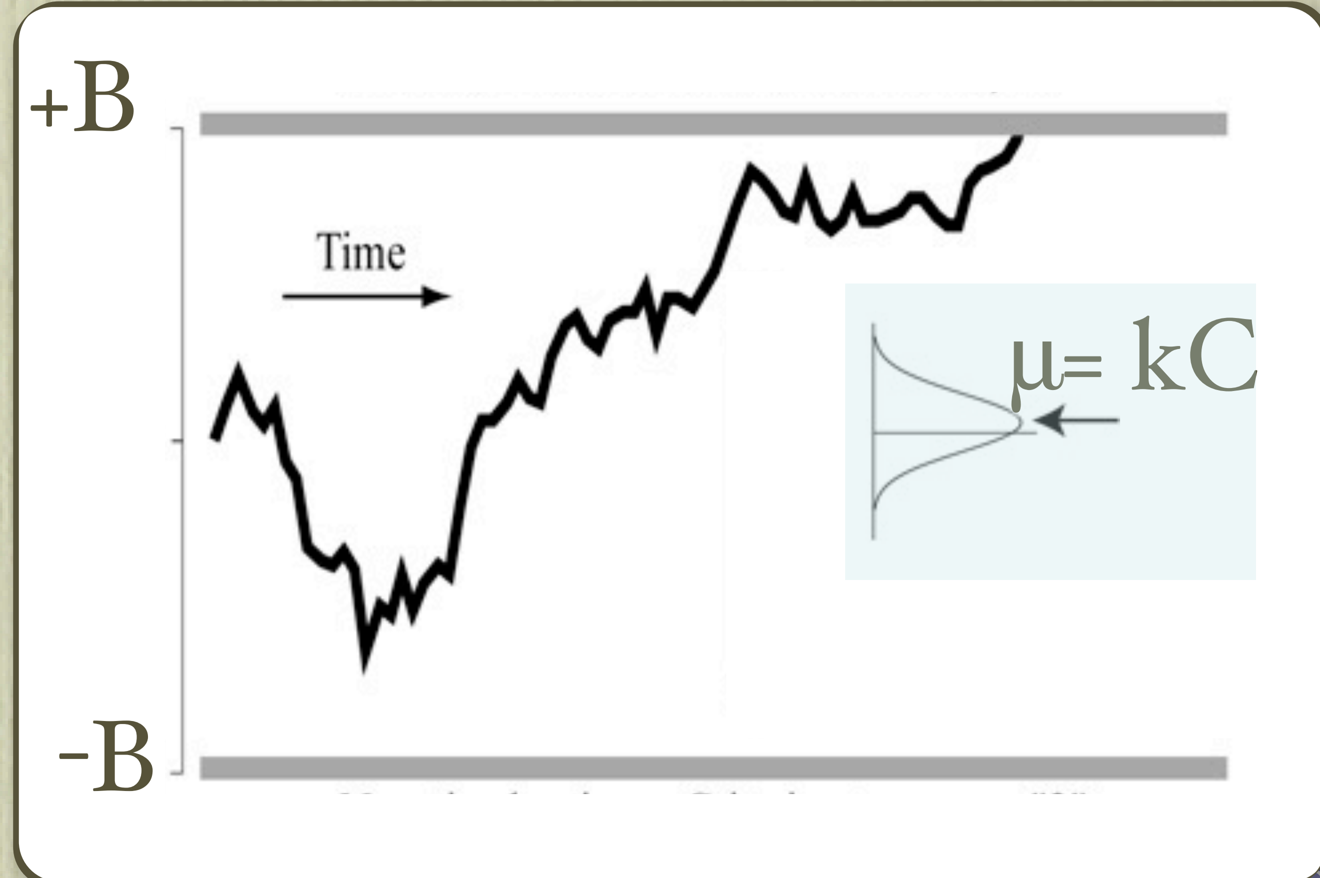


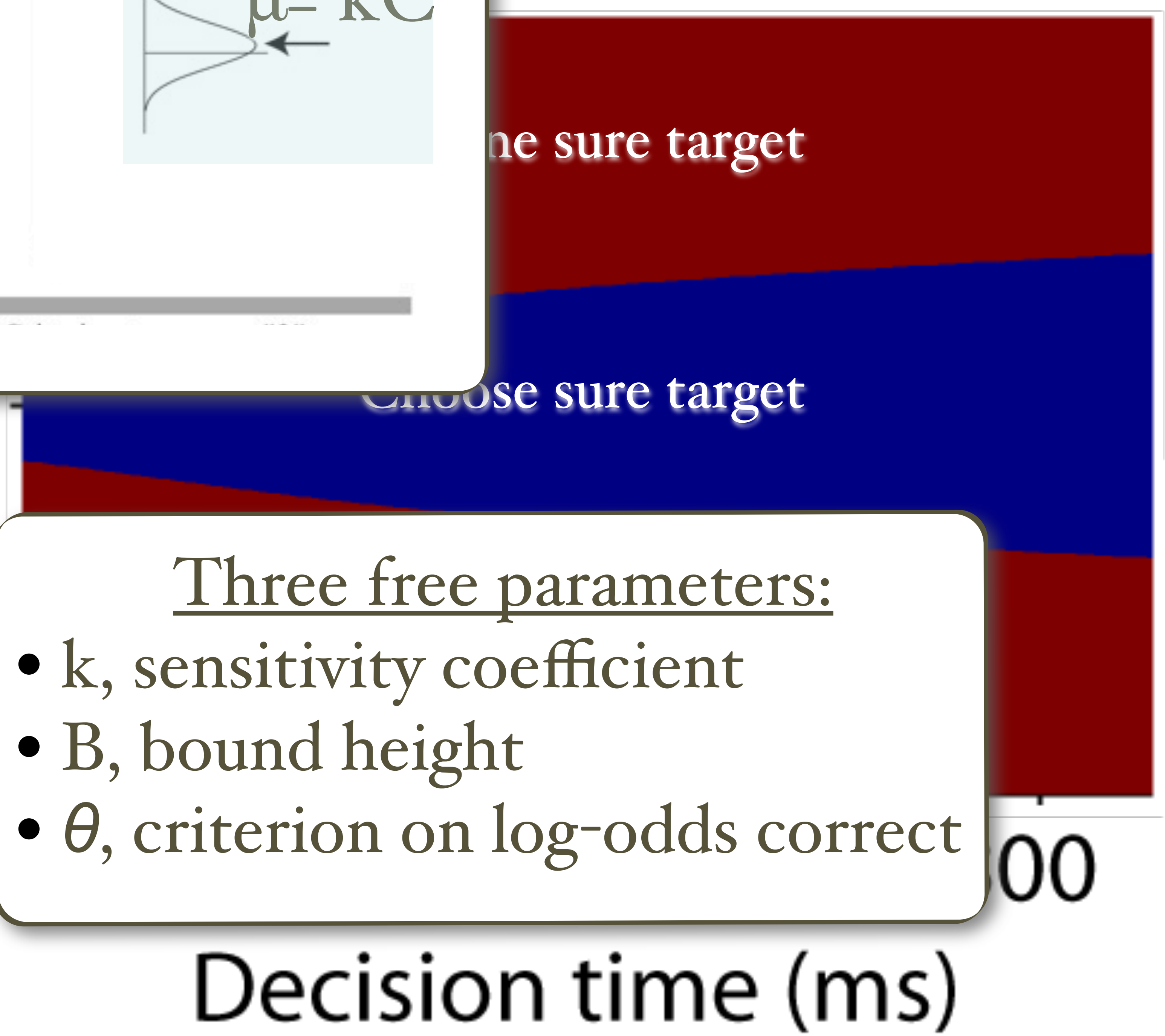
Kiani & Shadlen, 2009

Log odds of making the correct choice



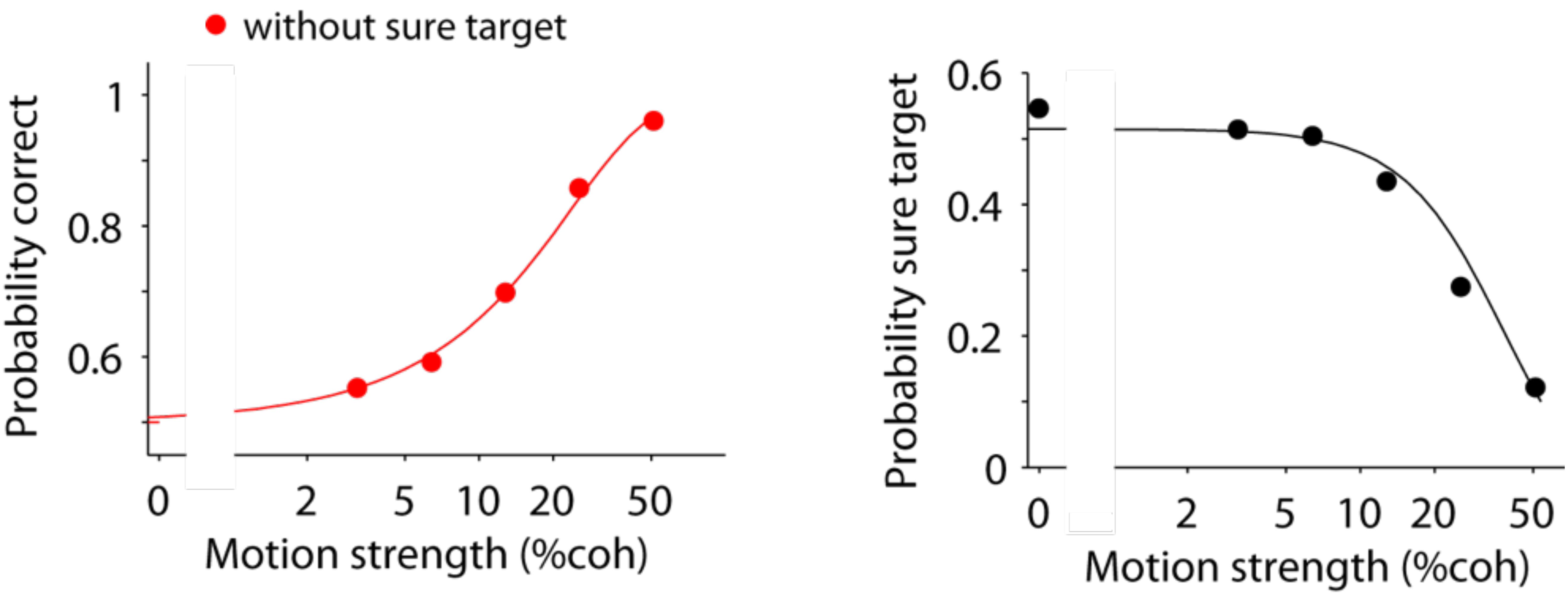
Kiani & Shadlen, 2009



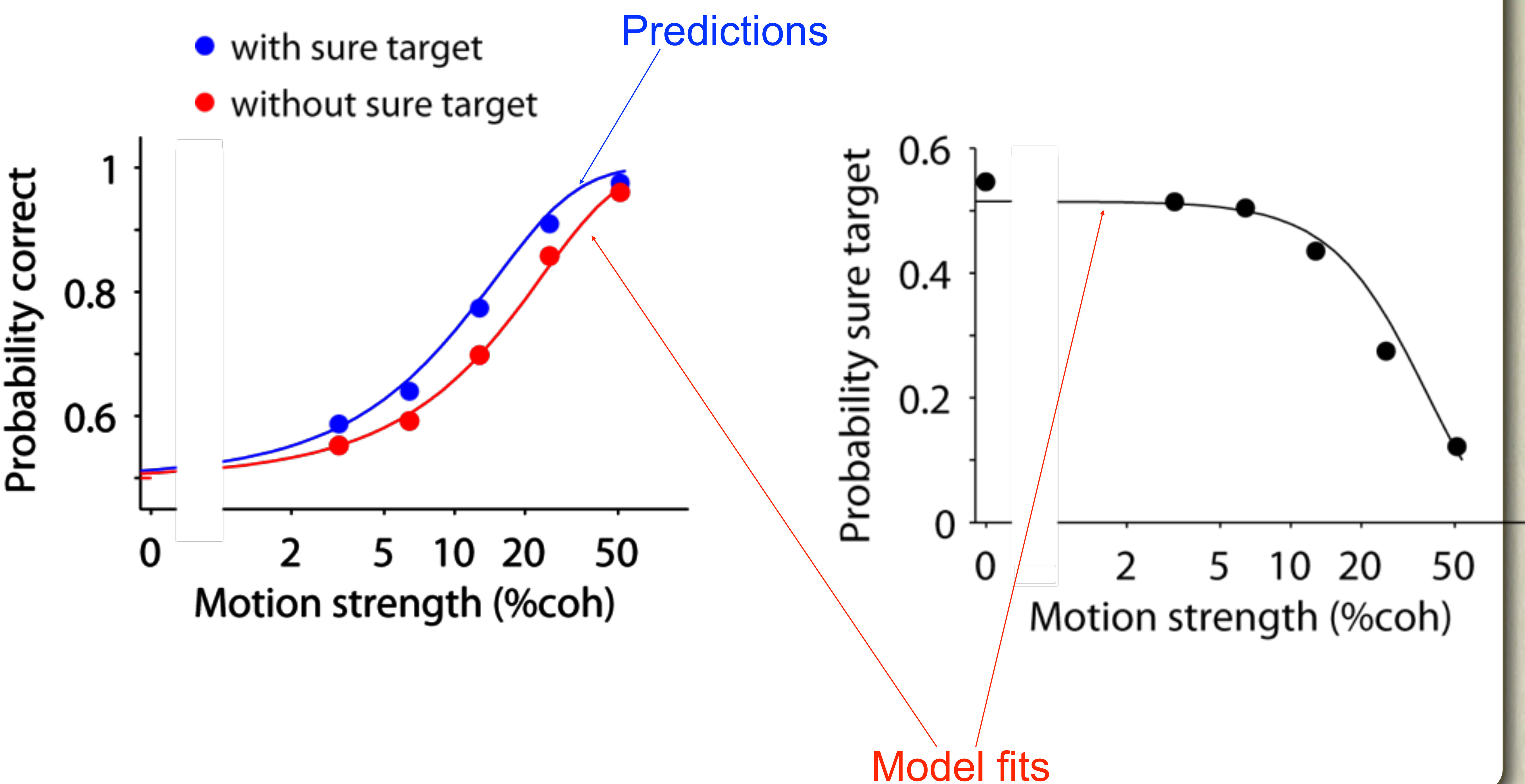


Three free parameters:

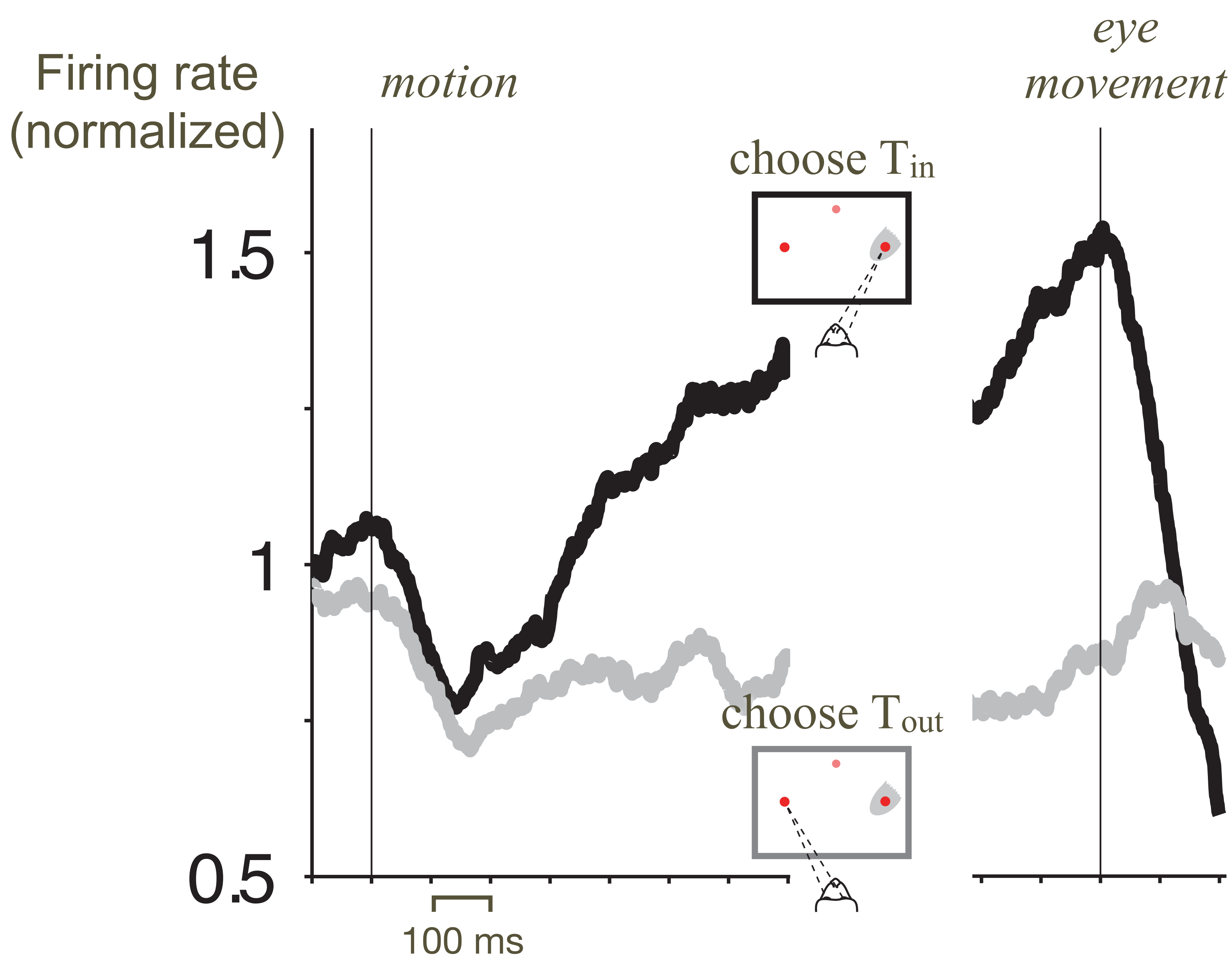
- k, sensitivity coefficient
- B, bound height
- θ, criterion on log-odds correct



Kiani & Shadlen, 2009

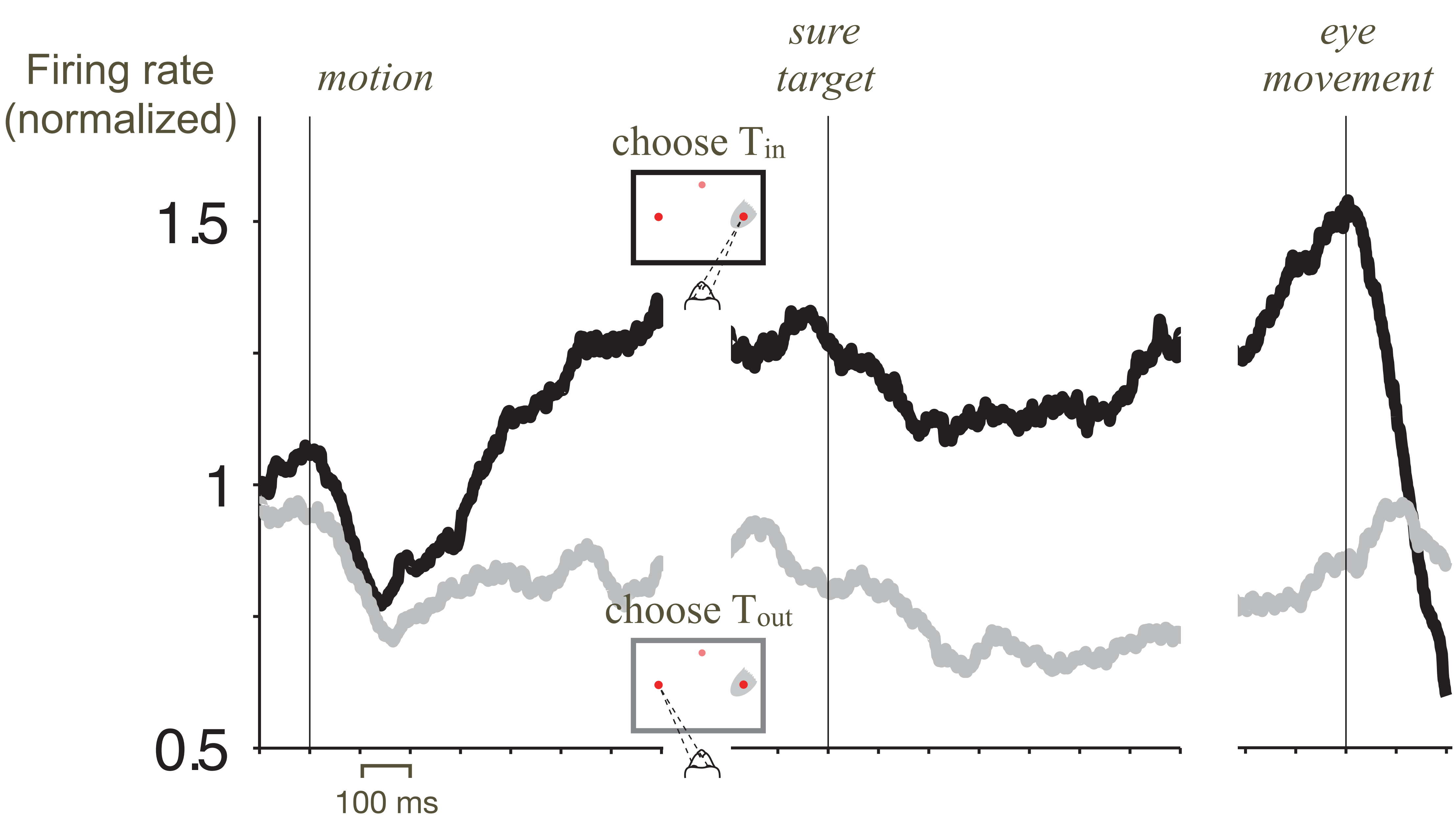

Predictions
with sure target
without sure target
Probability correct
1
0.8
0.6
0
2
5
10
20
50
Motion strength (%coh)
Probability sure target
0.6
0.4
0.2
0
0
2
5
10
20
50
Motion strength (%coh)
Model fits

weak motion strength
N = 70 neurons



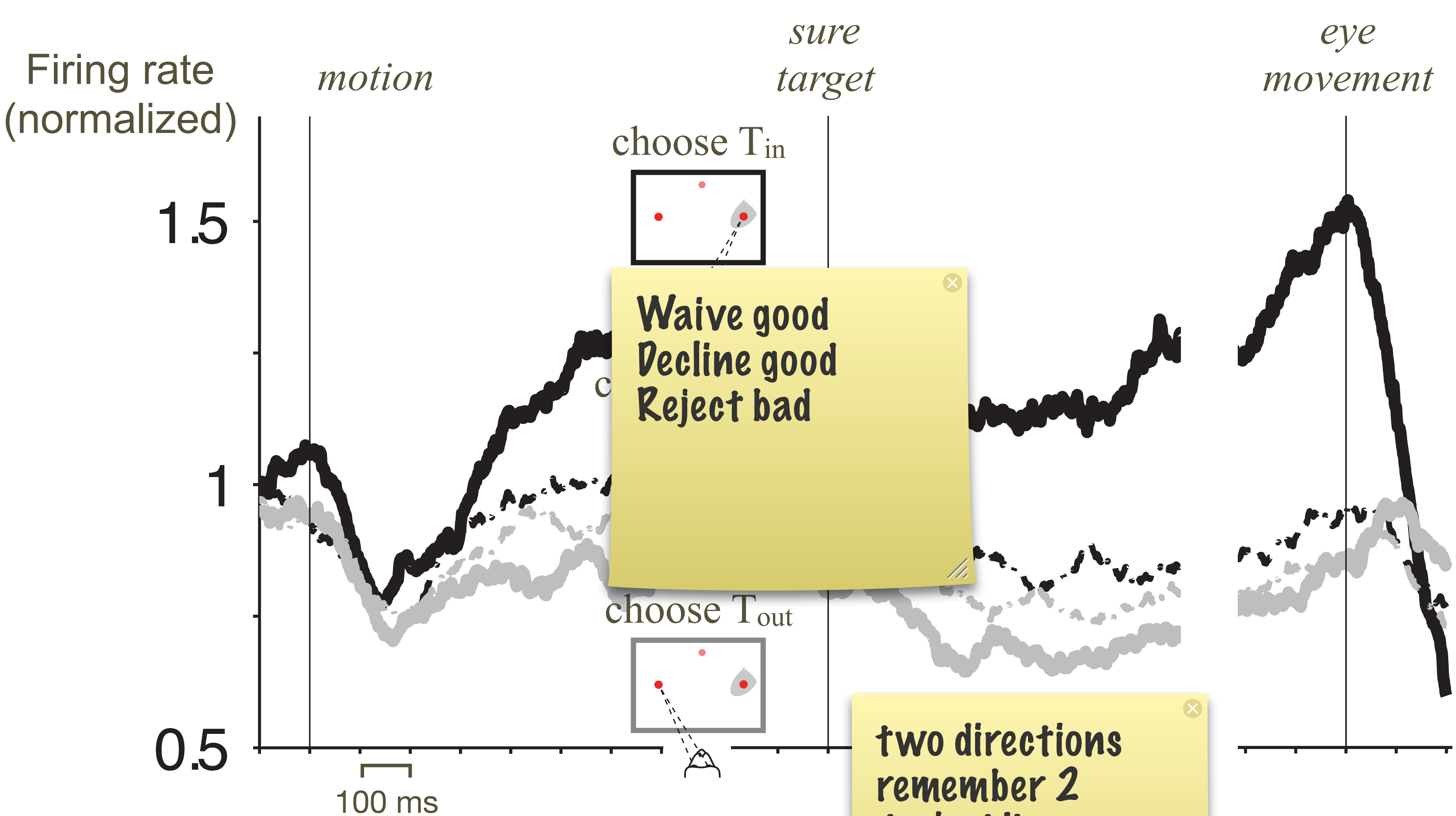
Kiani & Shadlen (2009) Science 324:759-764.

weak motion strength
N = 70 neurons



Kiani & Shadlen (2009) Science 324:759-764.

weak motion strength
N = 70 neurons



Waive good
Decline good
Reject bad

two directions
remember 2
dashed lines.

Kiani & Shadlen (2009) Science 324:759-764.

Conclusions from confidence experiment

- It is possible to study "degree of belief" in neurophysiology
- Bounded evidence accumulation unites 3 fundamental measures of choice behavior:

    accuracy, response time, confidence

- Suggests probability is represented by firing rate & elapsed time

FORMAT

Work on

Possible t
'deg of be
wagering