The drift-diffusion model of decision making

Ariel Zylberberg
Columbia University

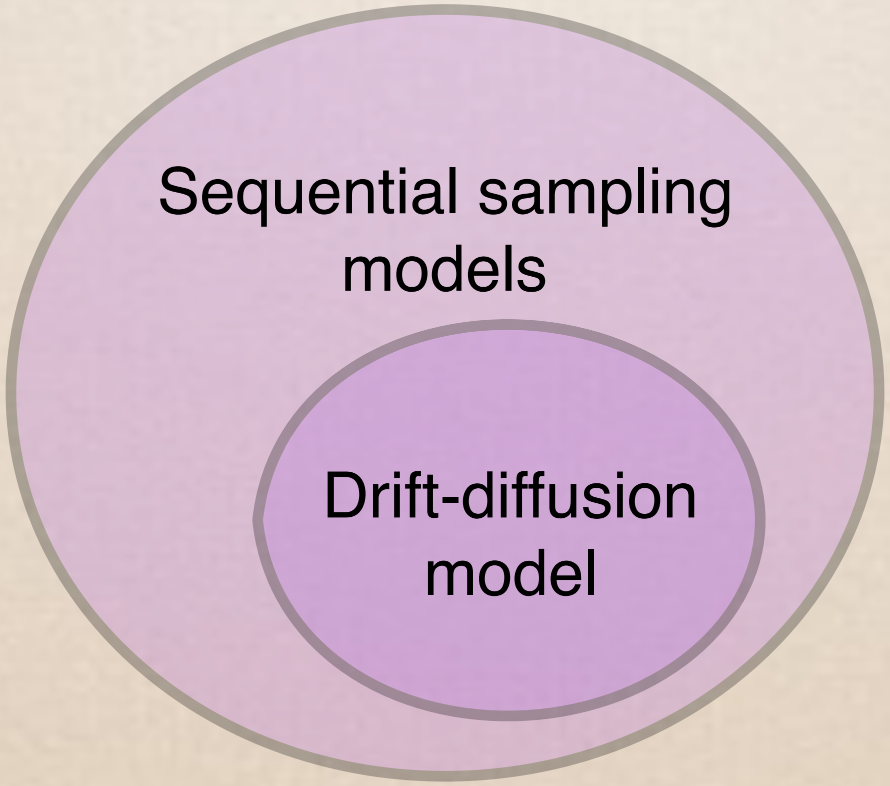
The drift-diffusion model of decision making

Ariel Zylberberg
Columbia University

Drift-diffusion model

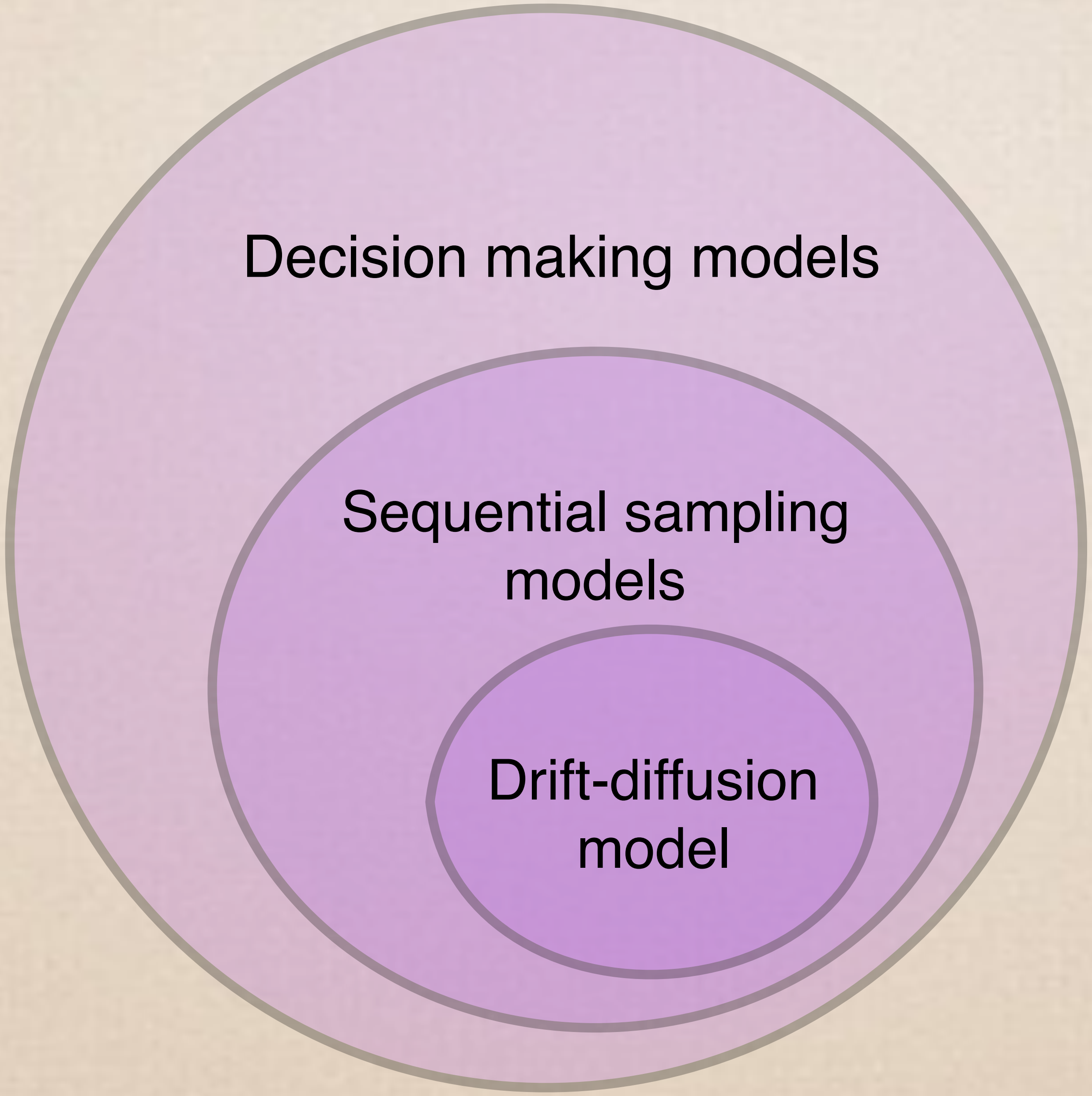
The drift-diffusion model of decision making

Ariel Zylberberg
Columbia University



The drift-diffusion model of decision making

Ariel Zylberberg
Columbia University



Speed, accuracy and difficulty

Speed, accuracy and difficulty

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

» Easy

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

- Easy
- Fast

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

- Easy
- Fast
- Consistent / accurate

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

- » More difficult

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

- More difficult
- Slower

Speed, accuracy and difficulty

The garbage plate, Rochester, NY

- » More difficult
- » Slower
- » Less consistent / accurate

Typical decision-making tasks

Typical decision-making tasks

Which snack do you prefer?

Typical decision-making tasks

Which snack do you prefer?

Word or non-word?

piggle

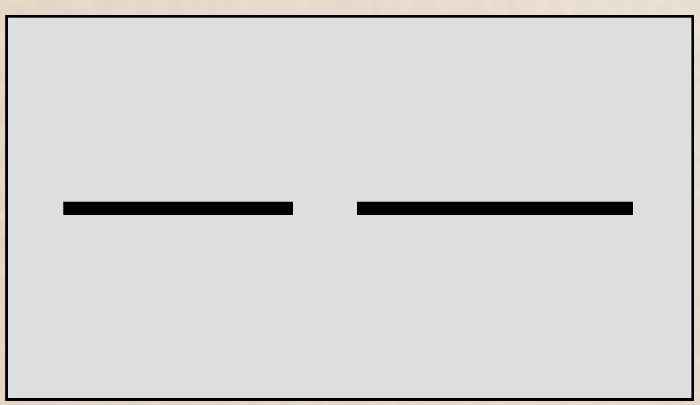
Typical decision-making tasks

Which snack do you prefer?

Word or non-word?

piggle

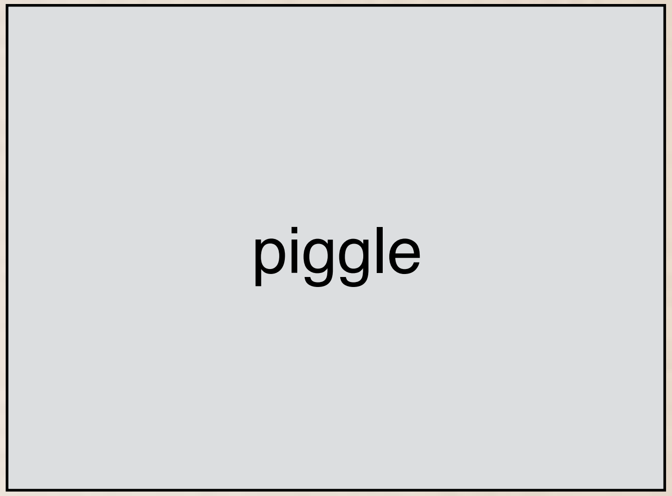
Which line is longer?



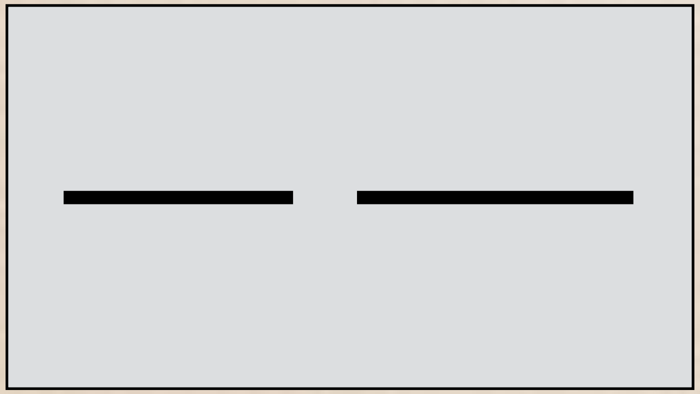
Typical decision-making tasks

Which snack do you prefer?

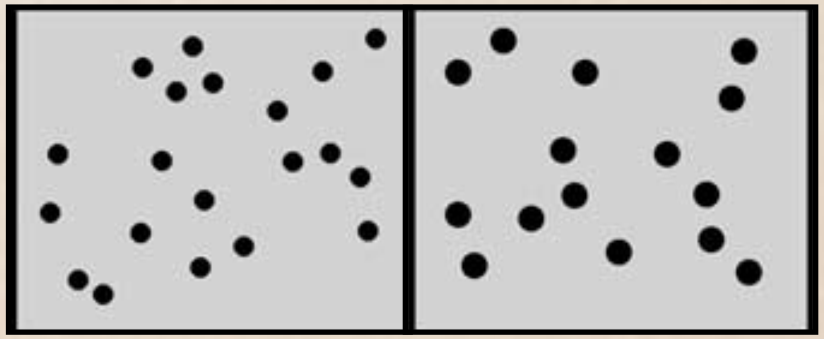
Word or non-word?



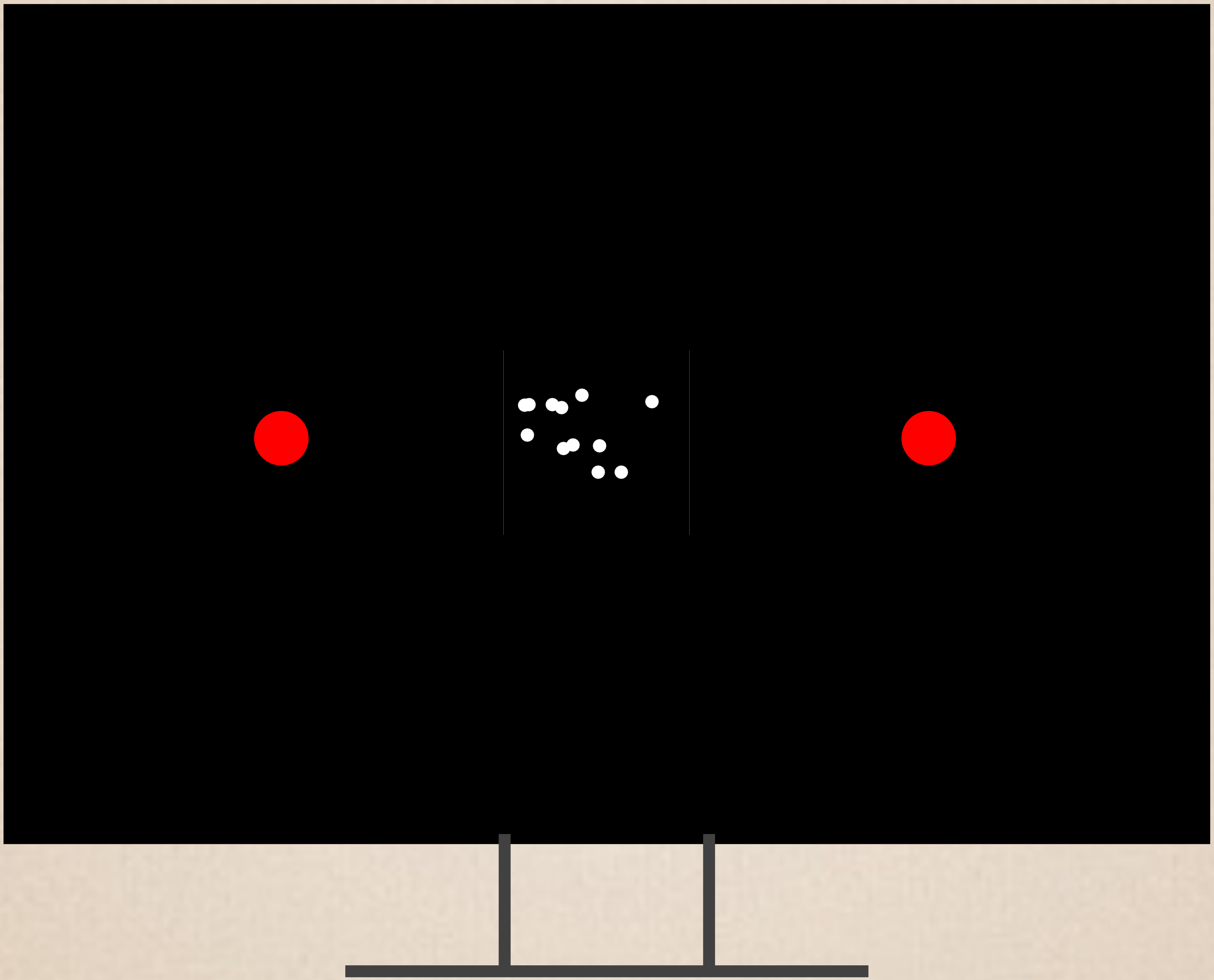
Which line is longer?



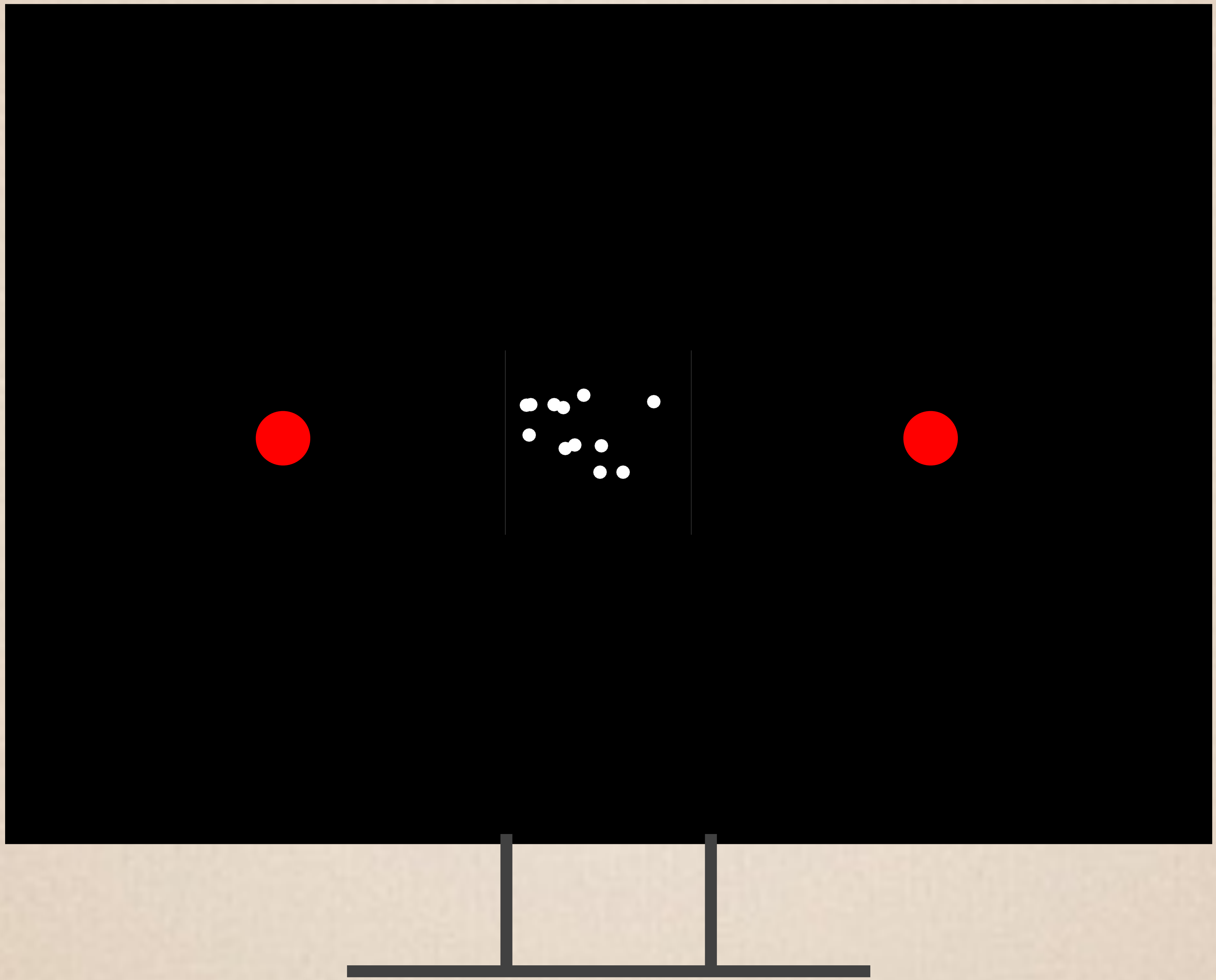
Which display has more dots?



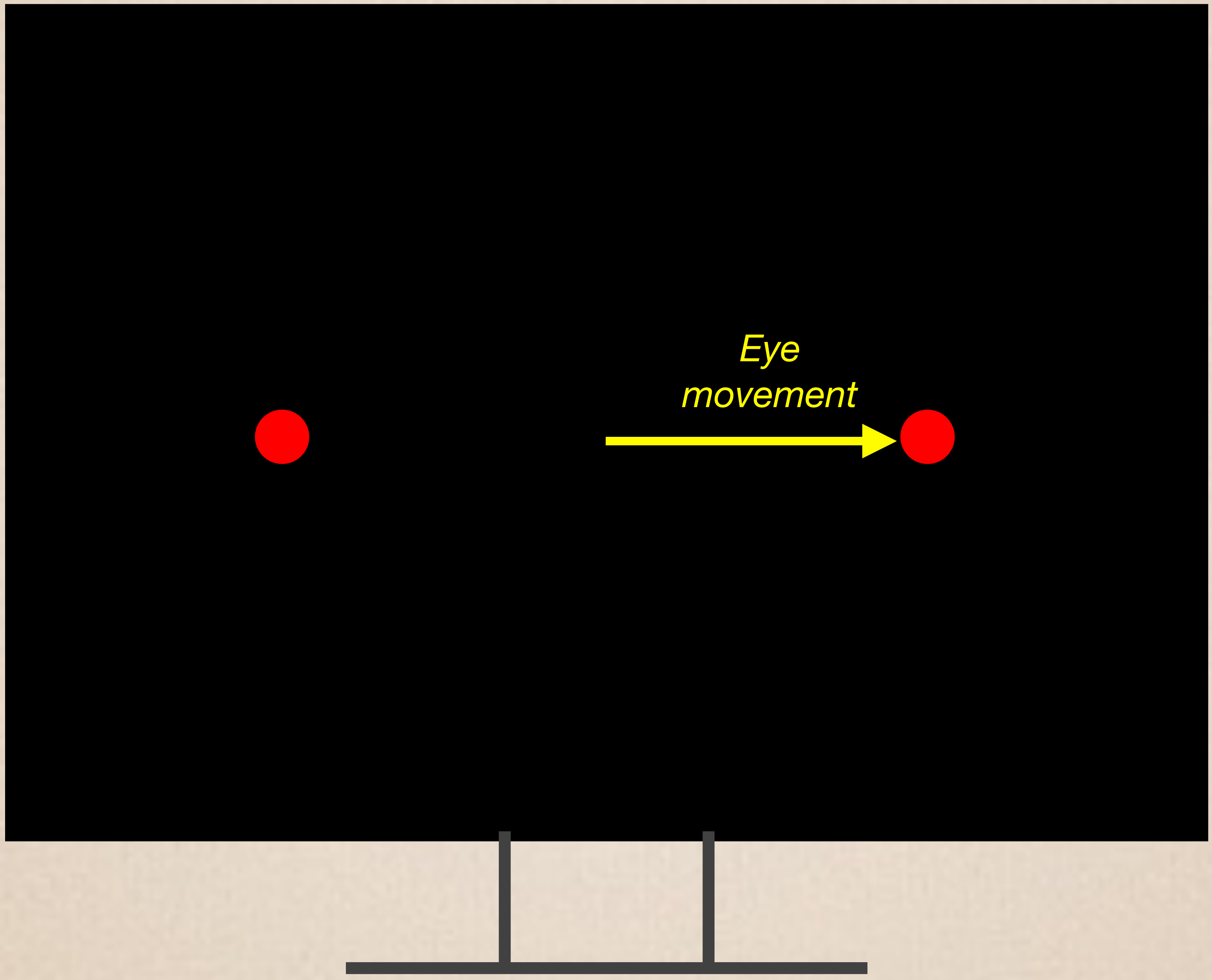
Random dot motion discrimination task



Random dot motion discrimination task

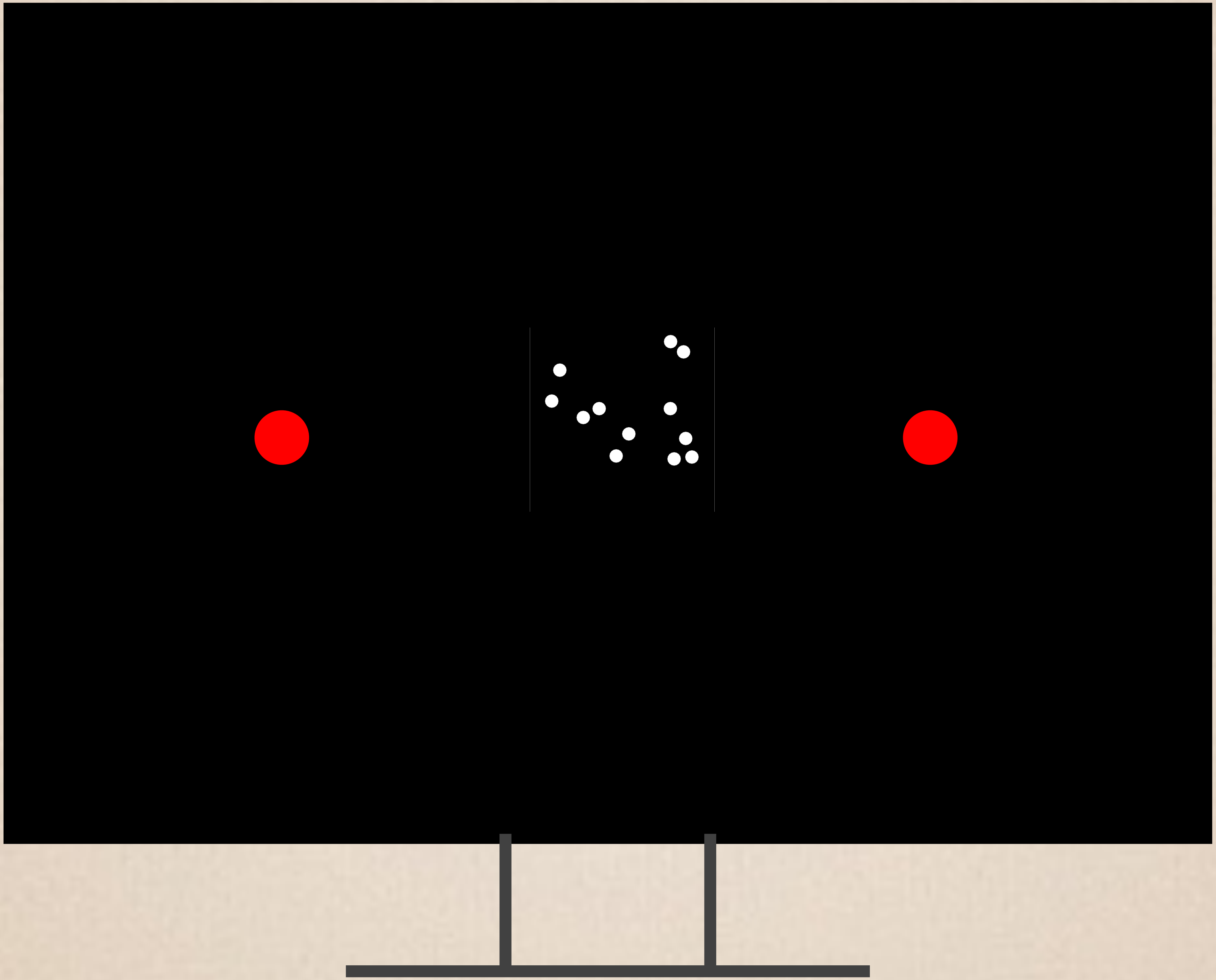


Random dot motion discrimination task



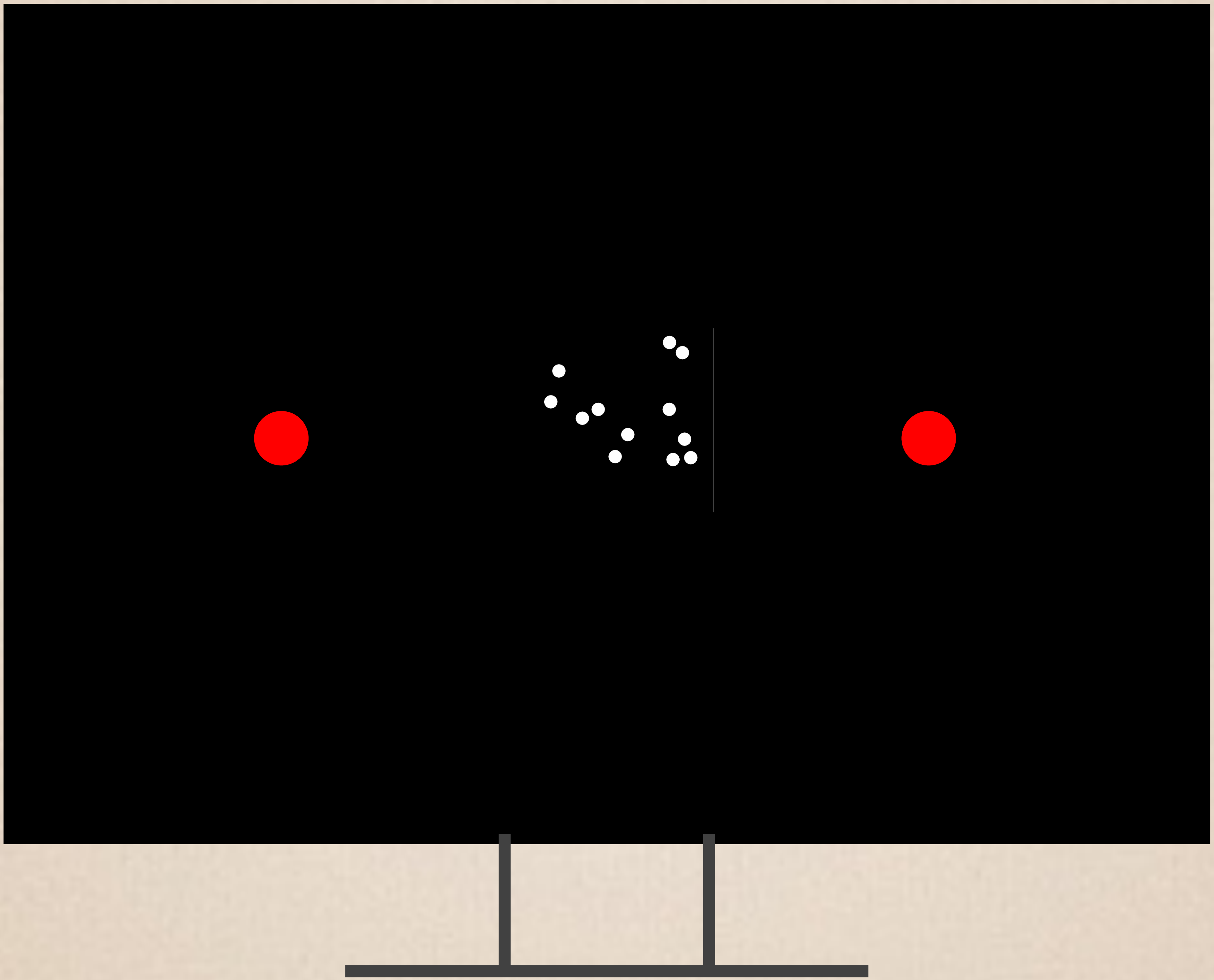
Random dot motion discrimination task

(More difficult)



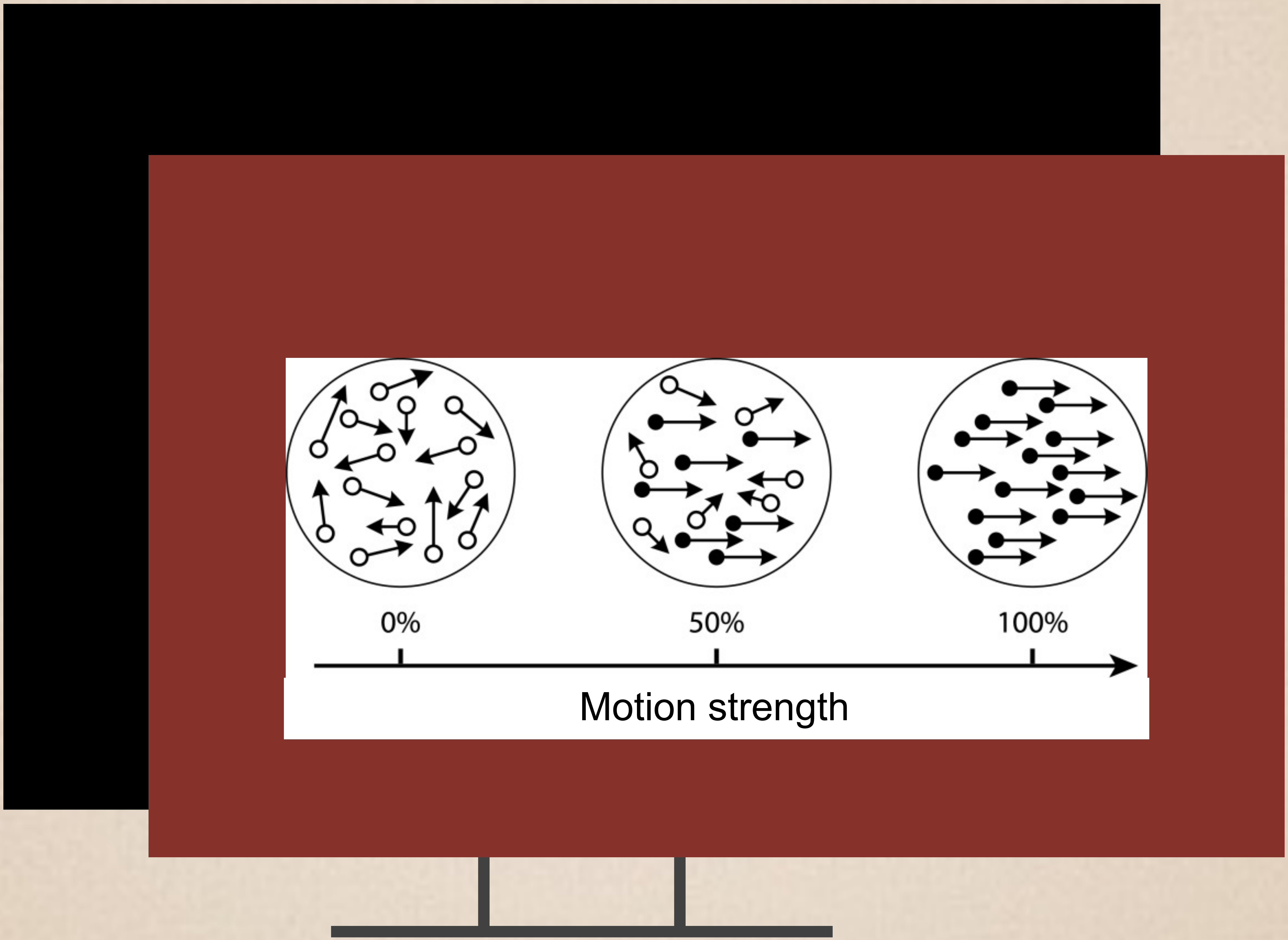
Random dot motion discrimination task

(More difficult)



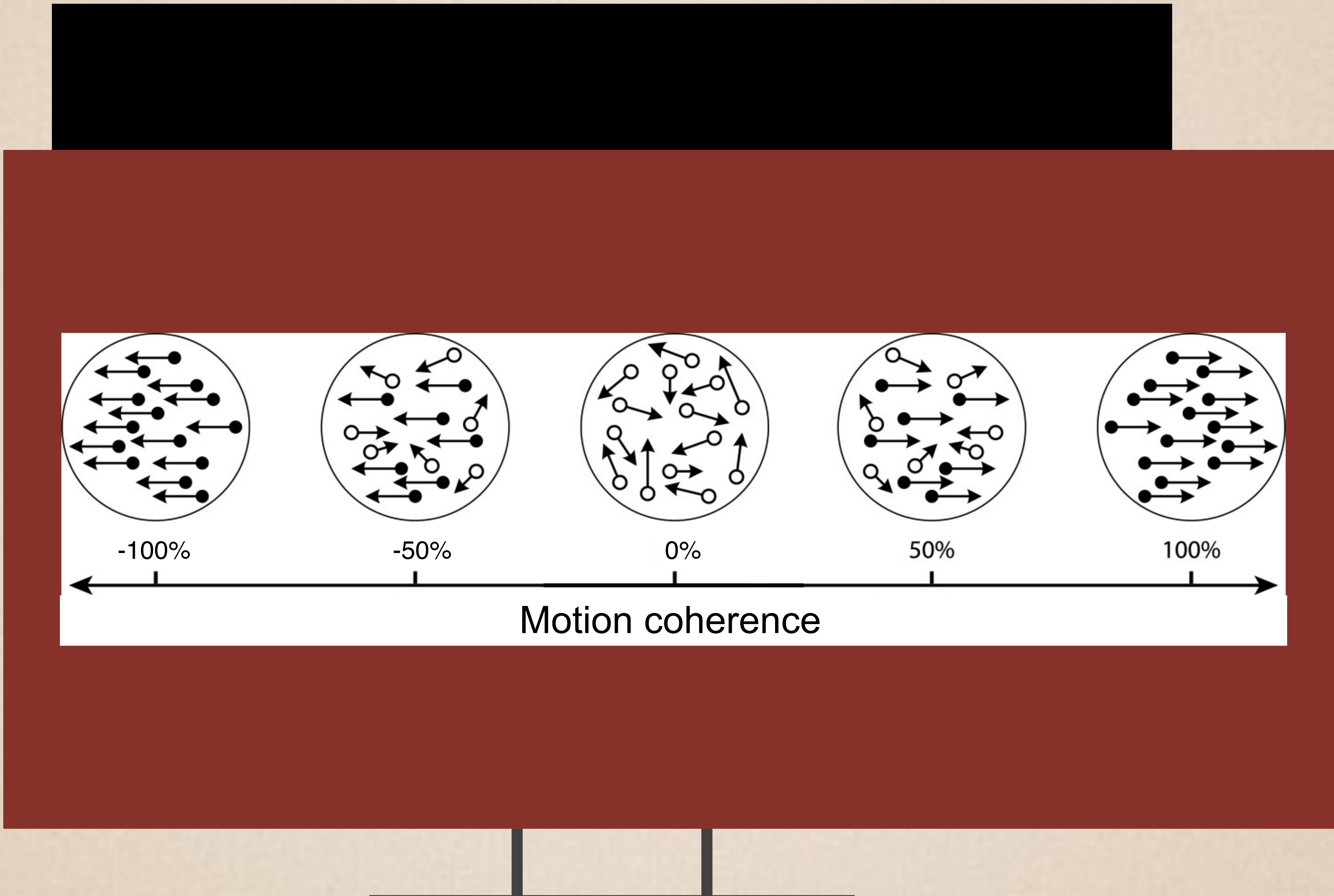
Random dot motion discrimination task

(More difficult)

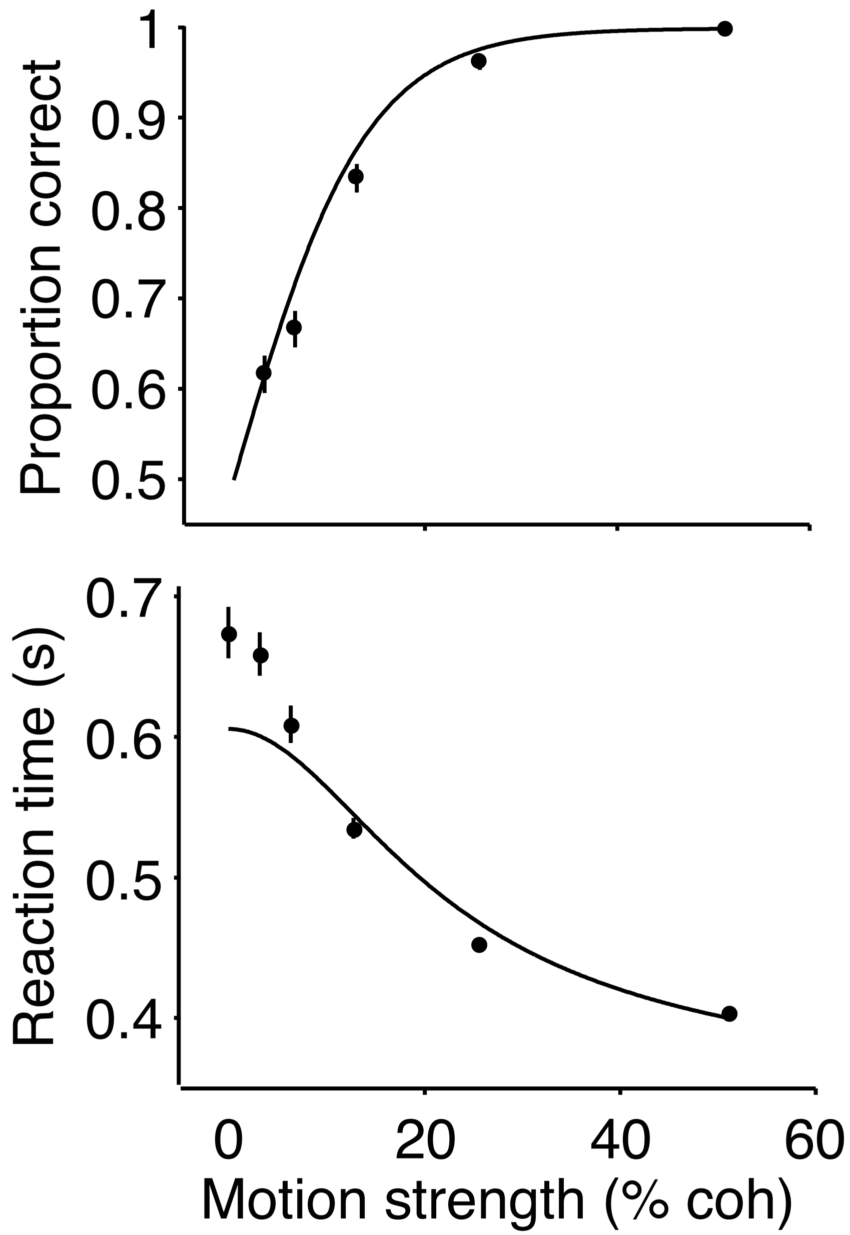


Random dot motion discrimination task

(More difficult)

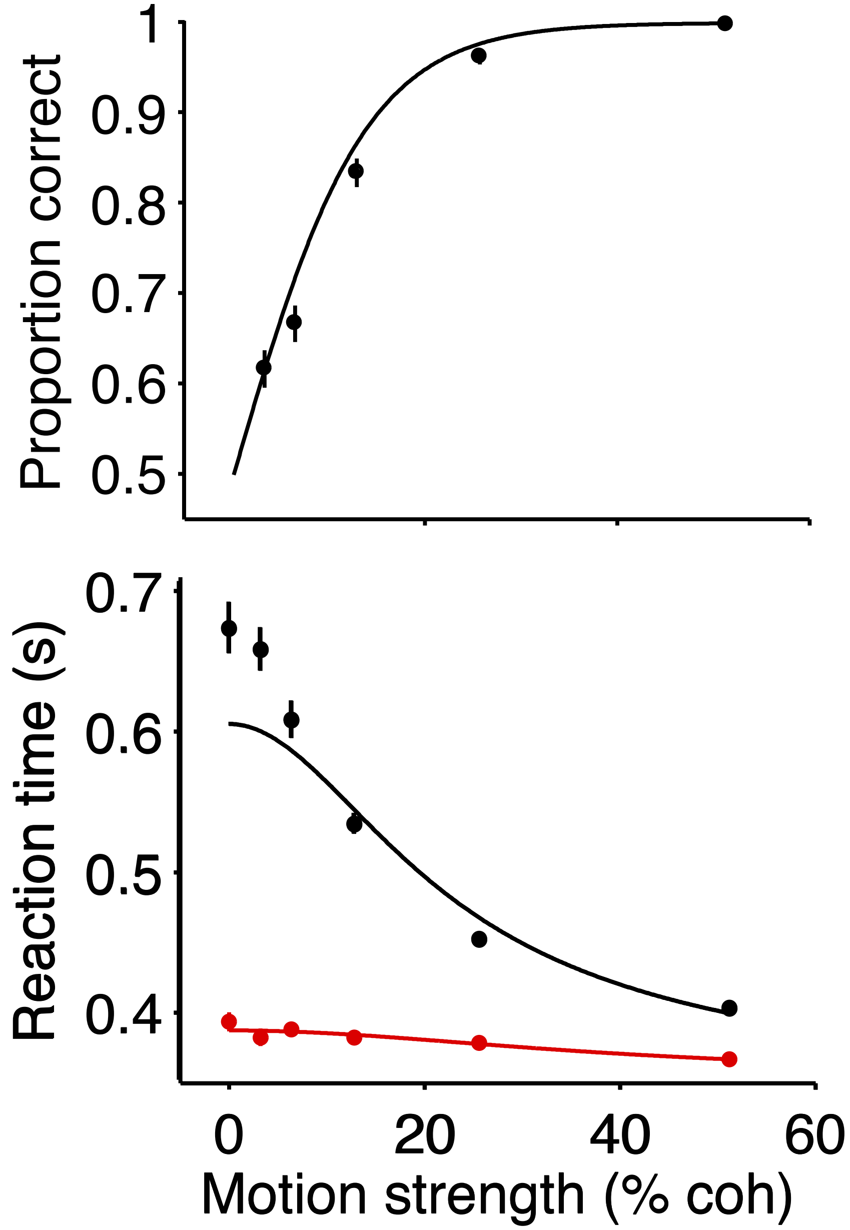


Speed-accuracy tradeoff



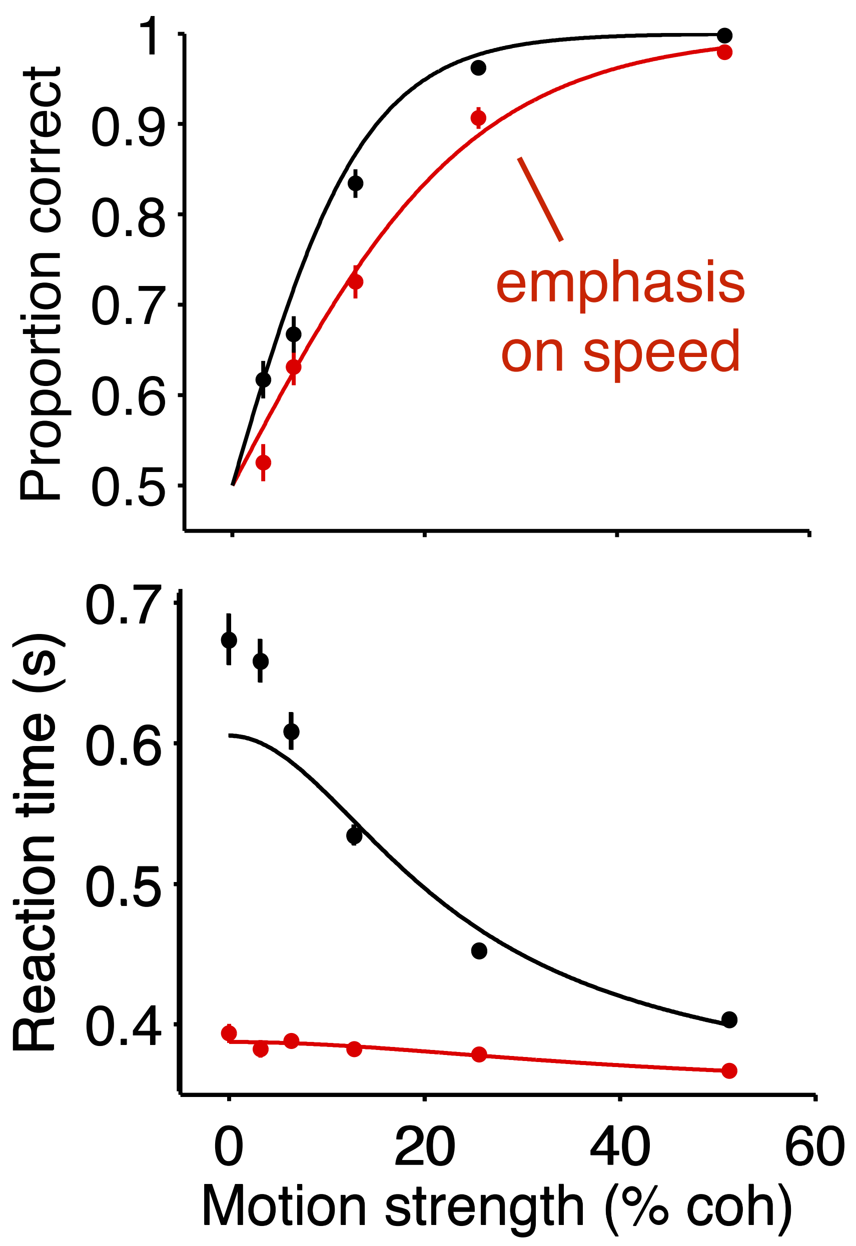
Hanks et al. eLife, 2014

Speed-accuracy tradeoff



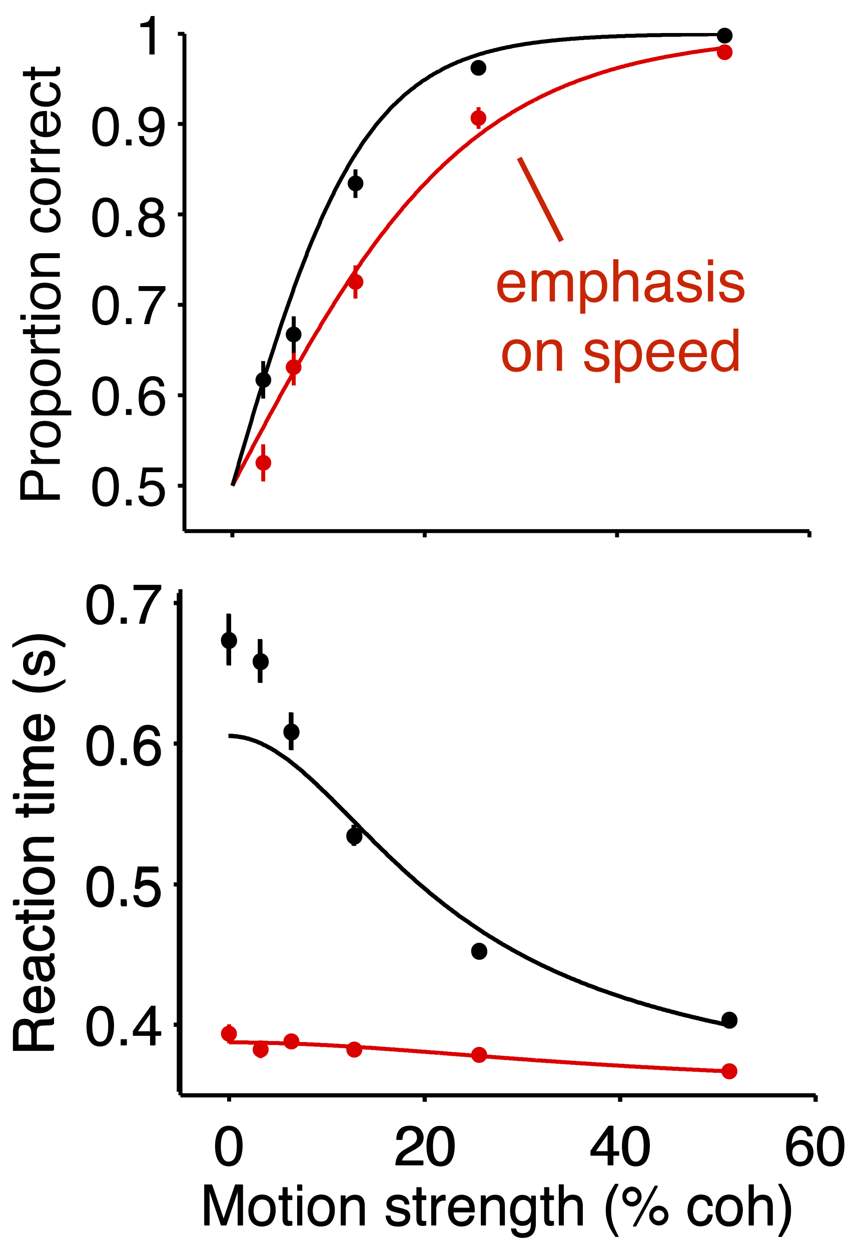
Hanks et al. eLife, 2014

Speed-accuracy tradeoff

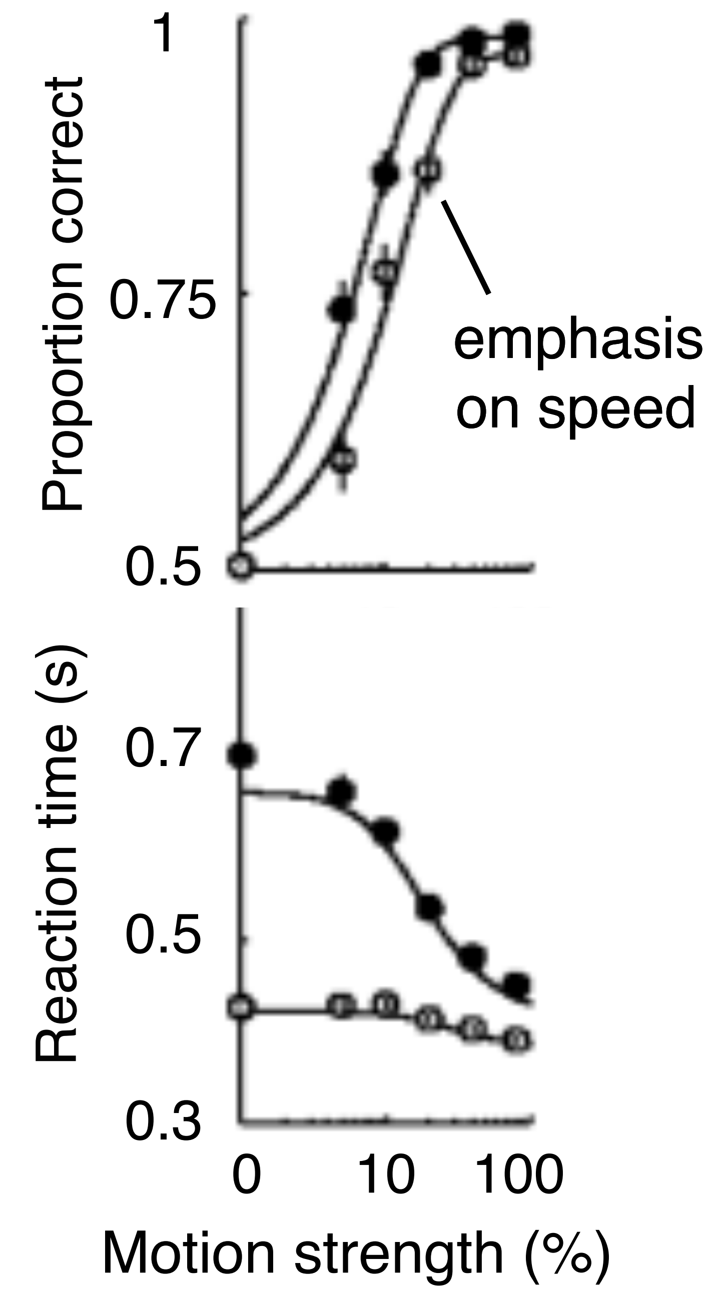


Hanks et al. eLife, 2014

Speed-accuracy tradeoff



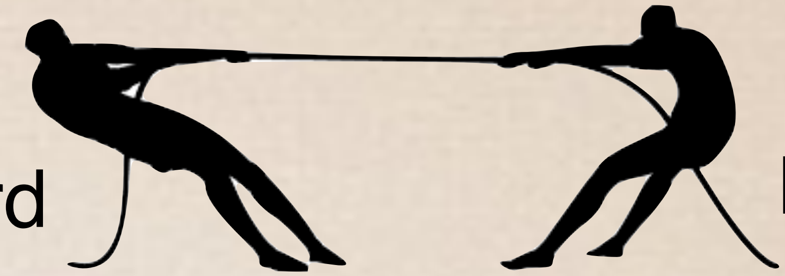
Hanks et al. eLife, 2014



Mulder et al.
Att Percept Psych
2013

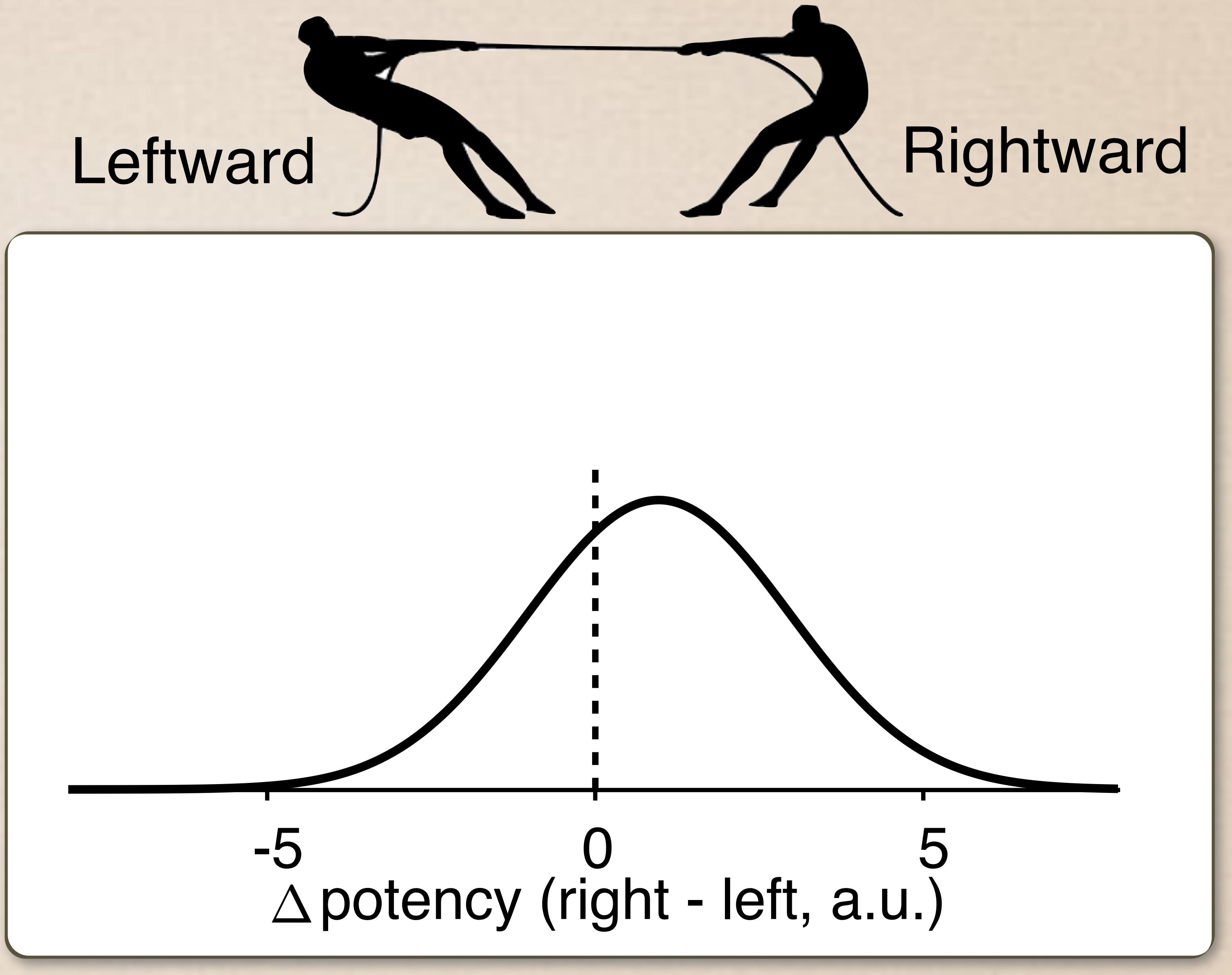
Opponent process theory of discrimination

Cartwright and Festinger (1943), “A quantitative theory of decision”.



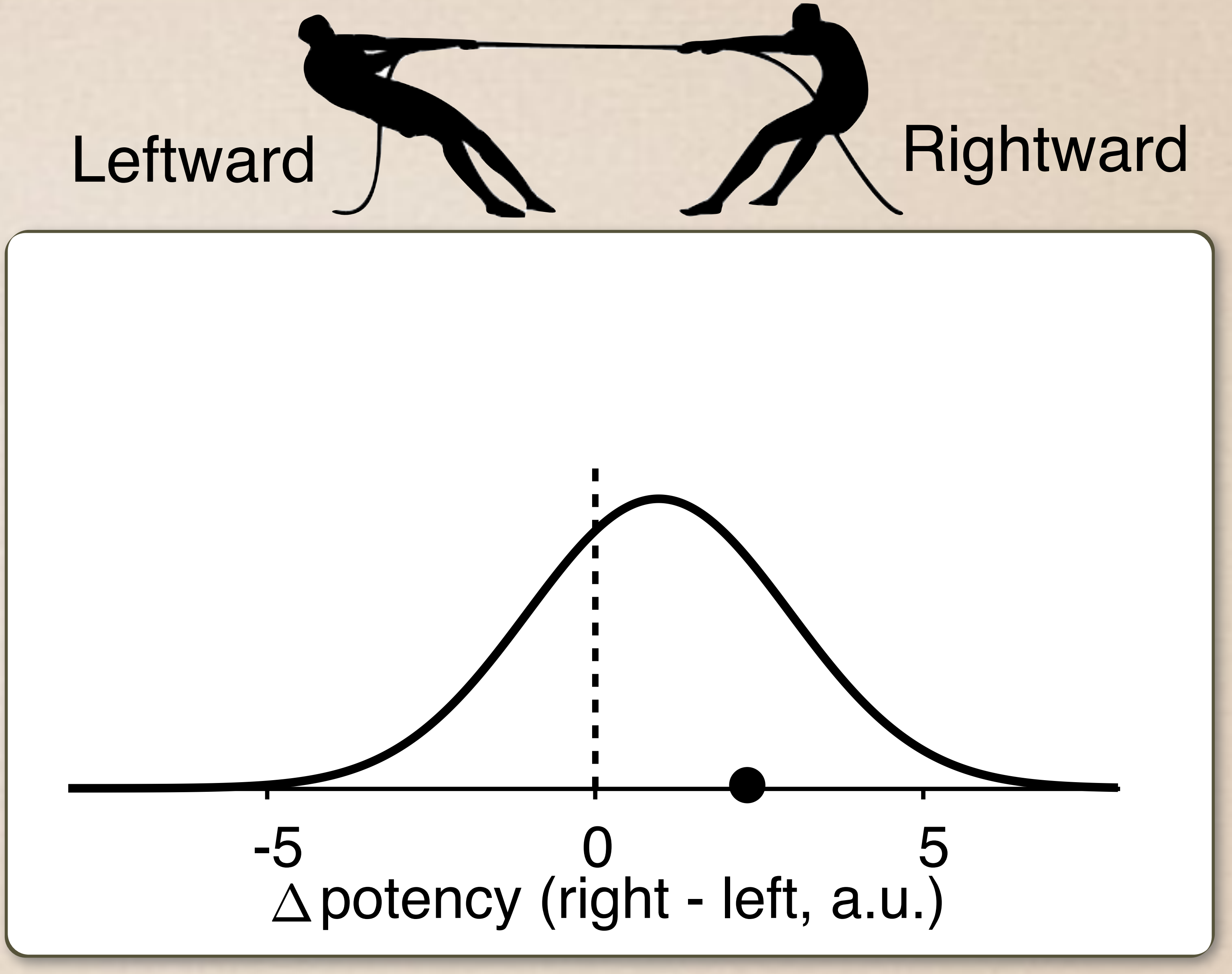
Opponent process theory of discrimination

Cartwright and Festinger (1943), “A quantitative theory of decision”.



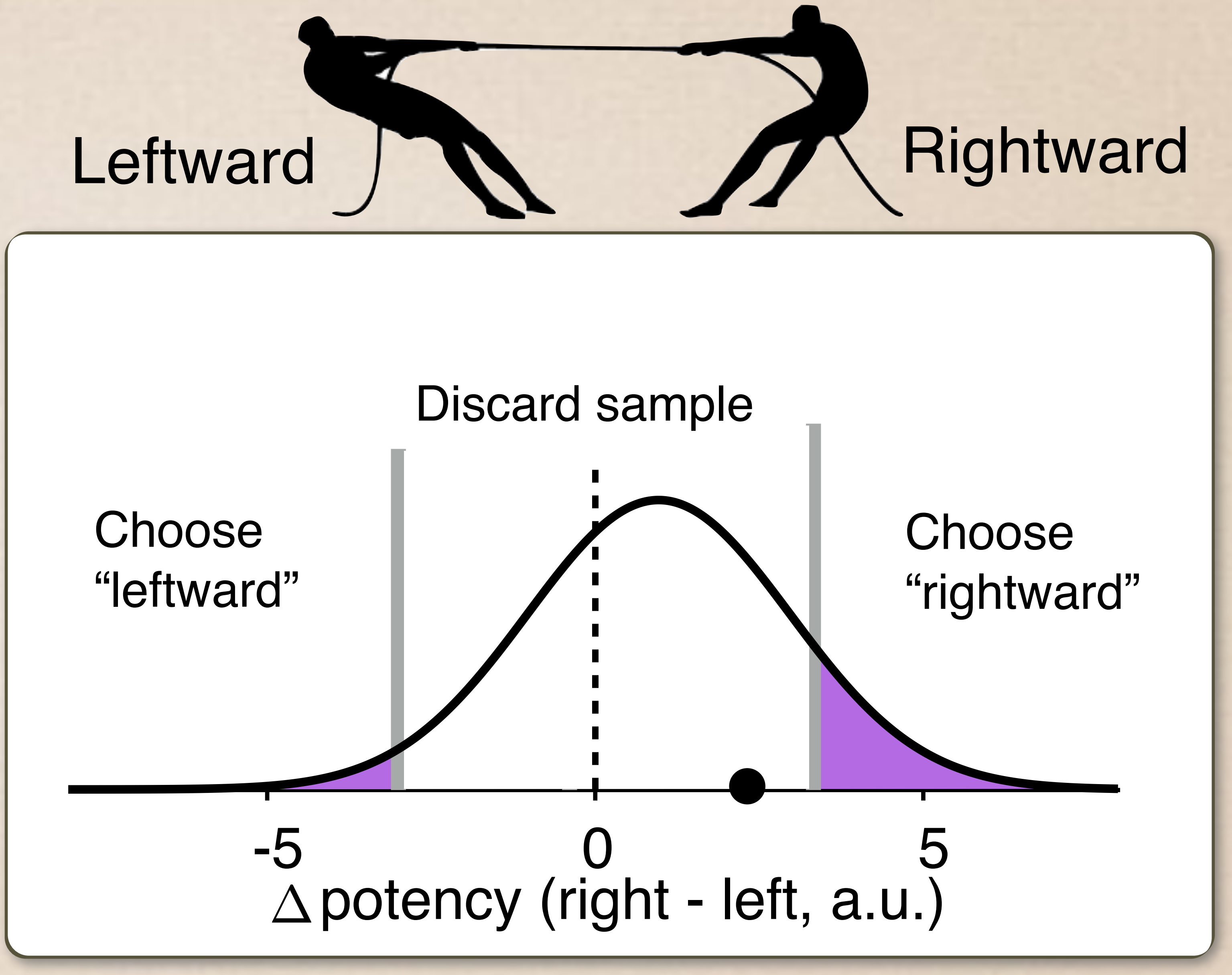
Opponent process theory of discrimination

Cartwright and Festinger (1943), “A quantitative theory of decision”.



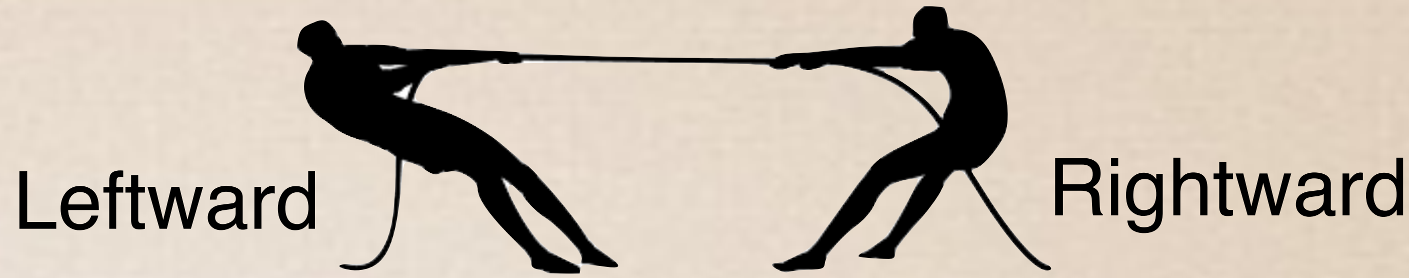
Opponent process theory of discrimination

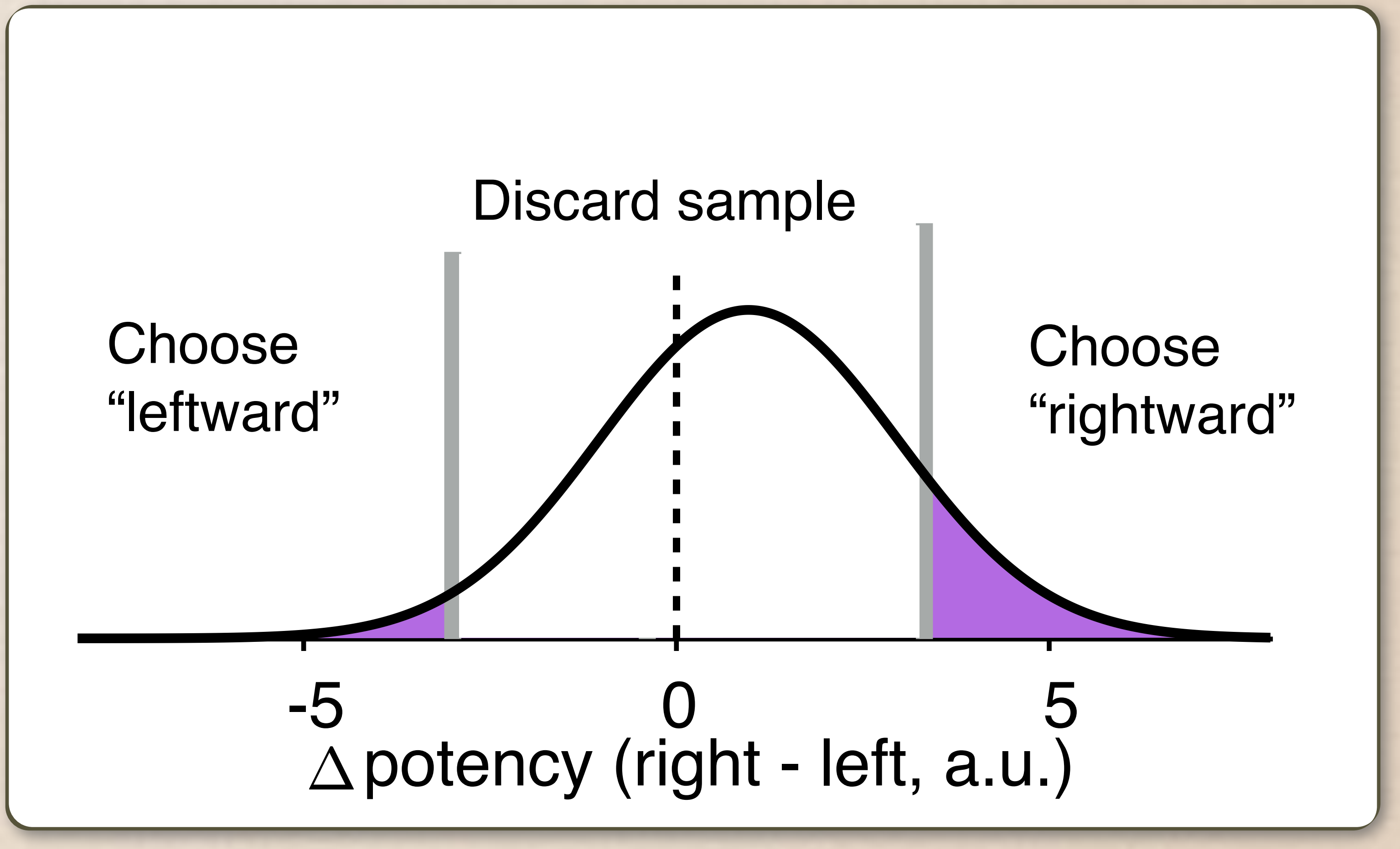
Cartwright and Festinger (1943), “A quantitative theory of decision”.



Opponent process theory of discrimination

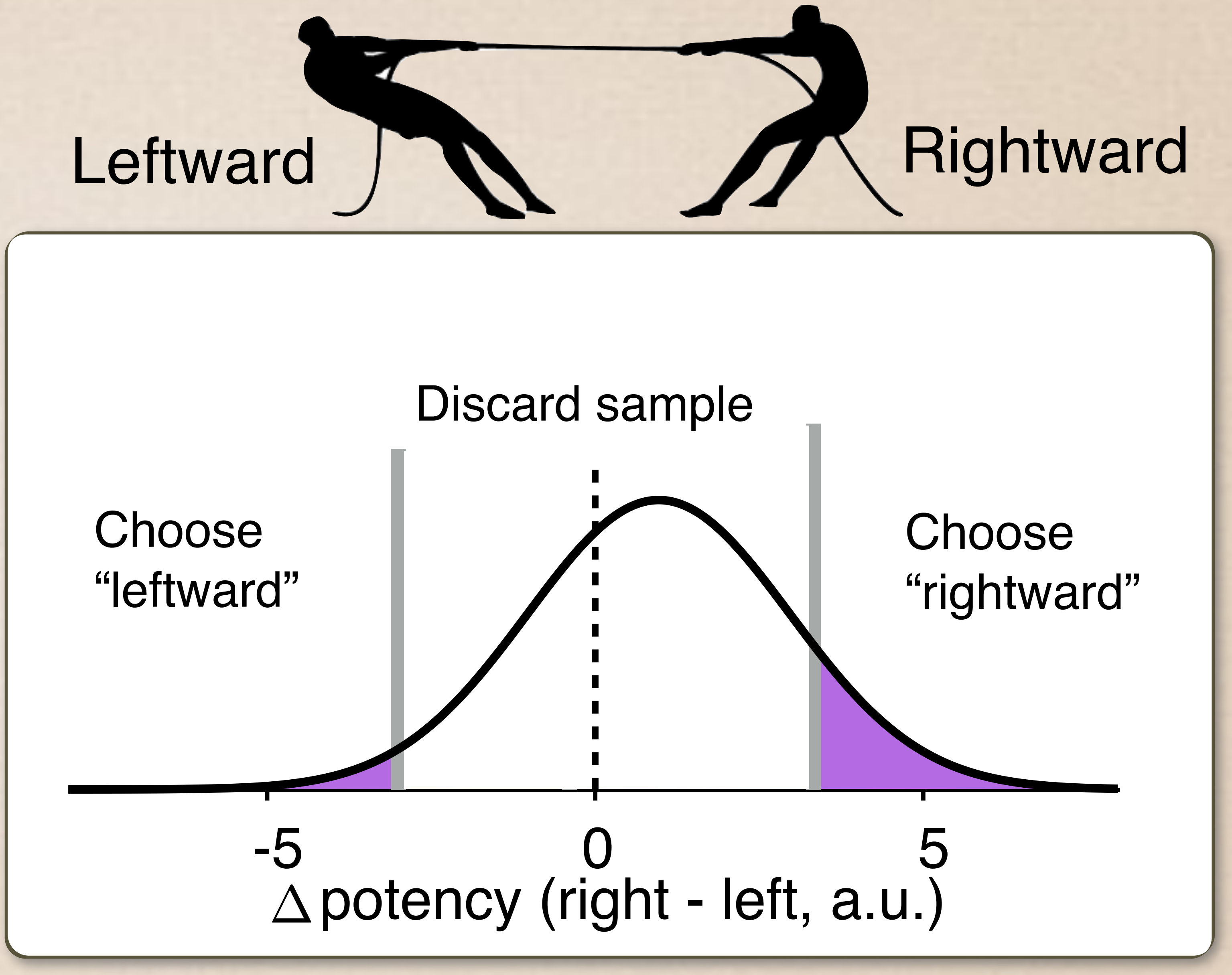
Cartwright and Festinger (1943), "A quantitative theory of decision".





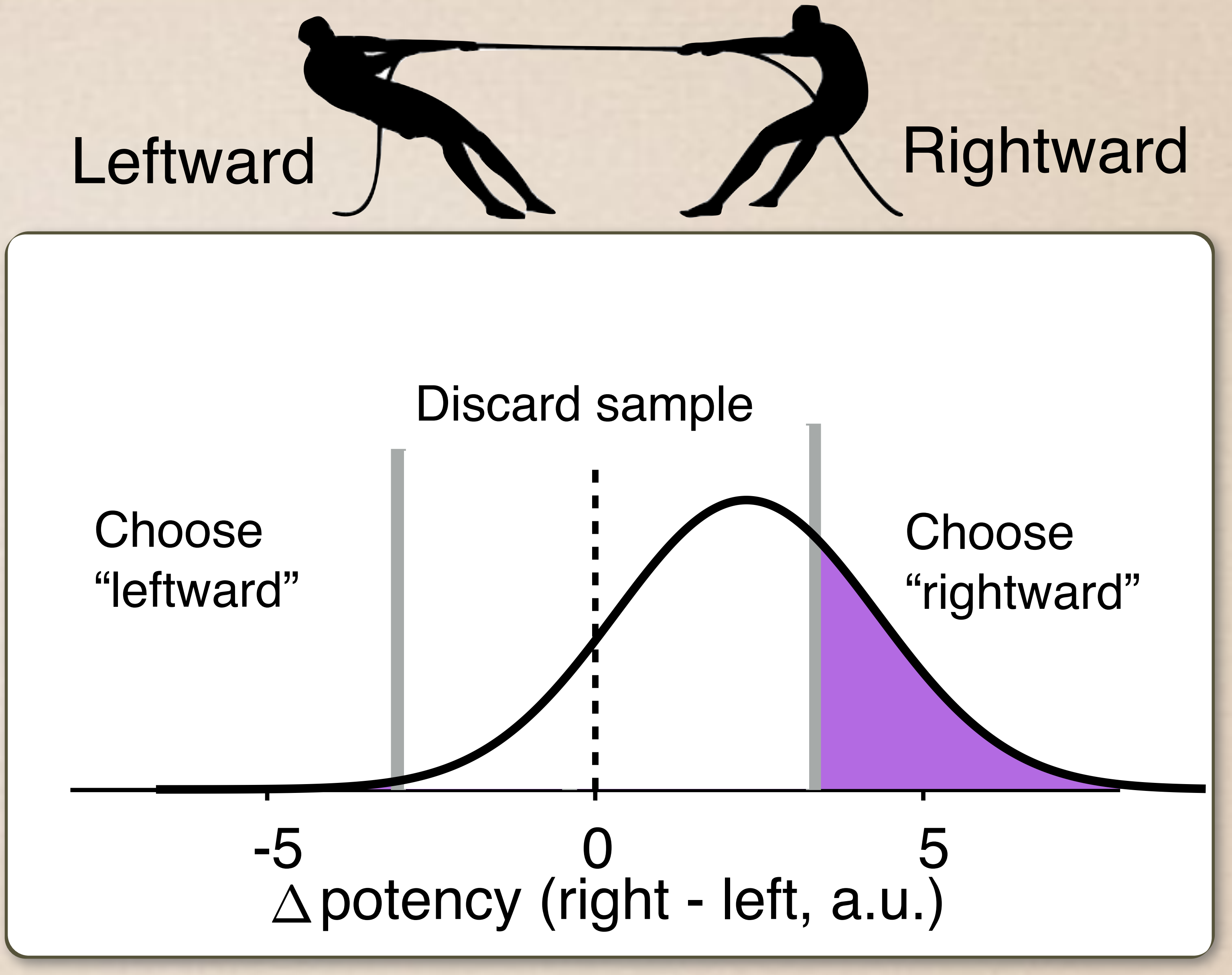
Opponent process theory of discrimination

Cartwright and Festinger (1943), “A quantitative theory of decision”.



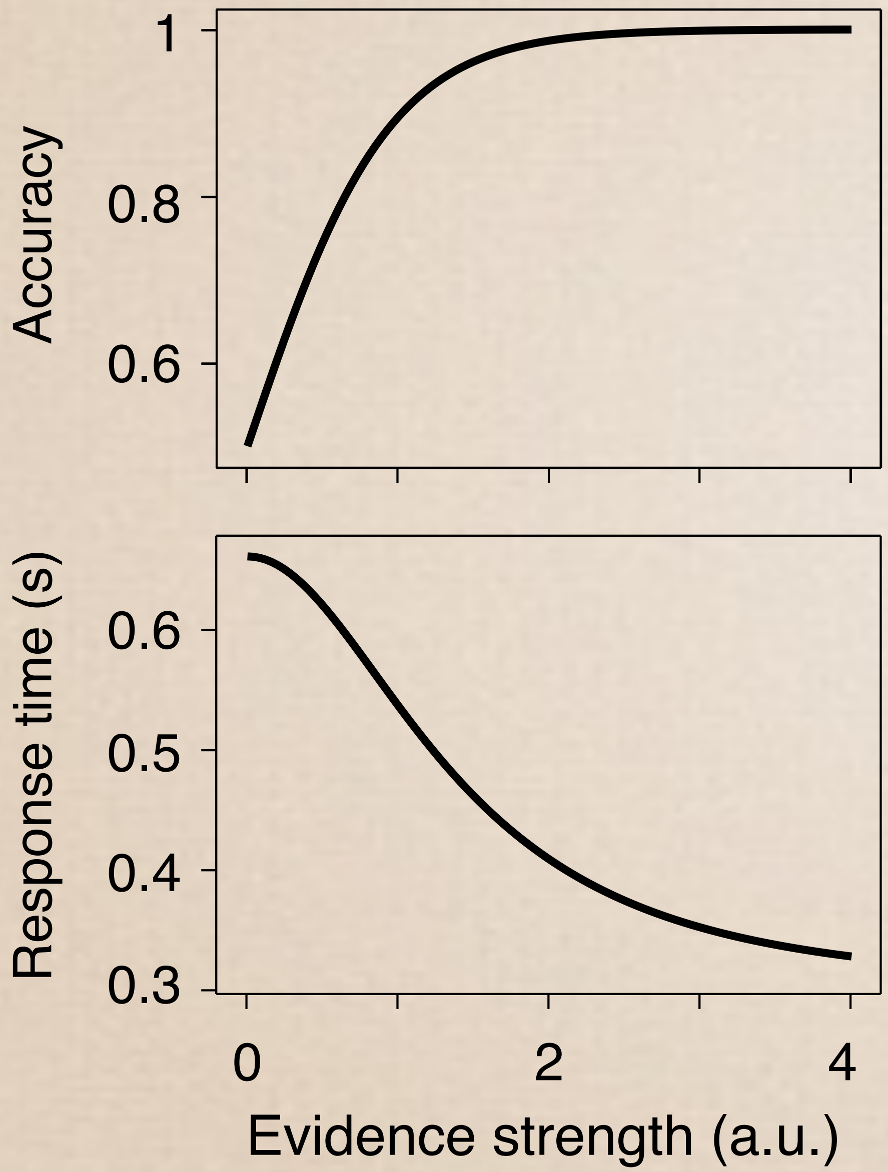
Opponent process theory of discrimination

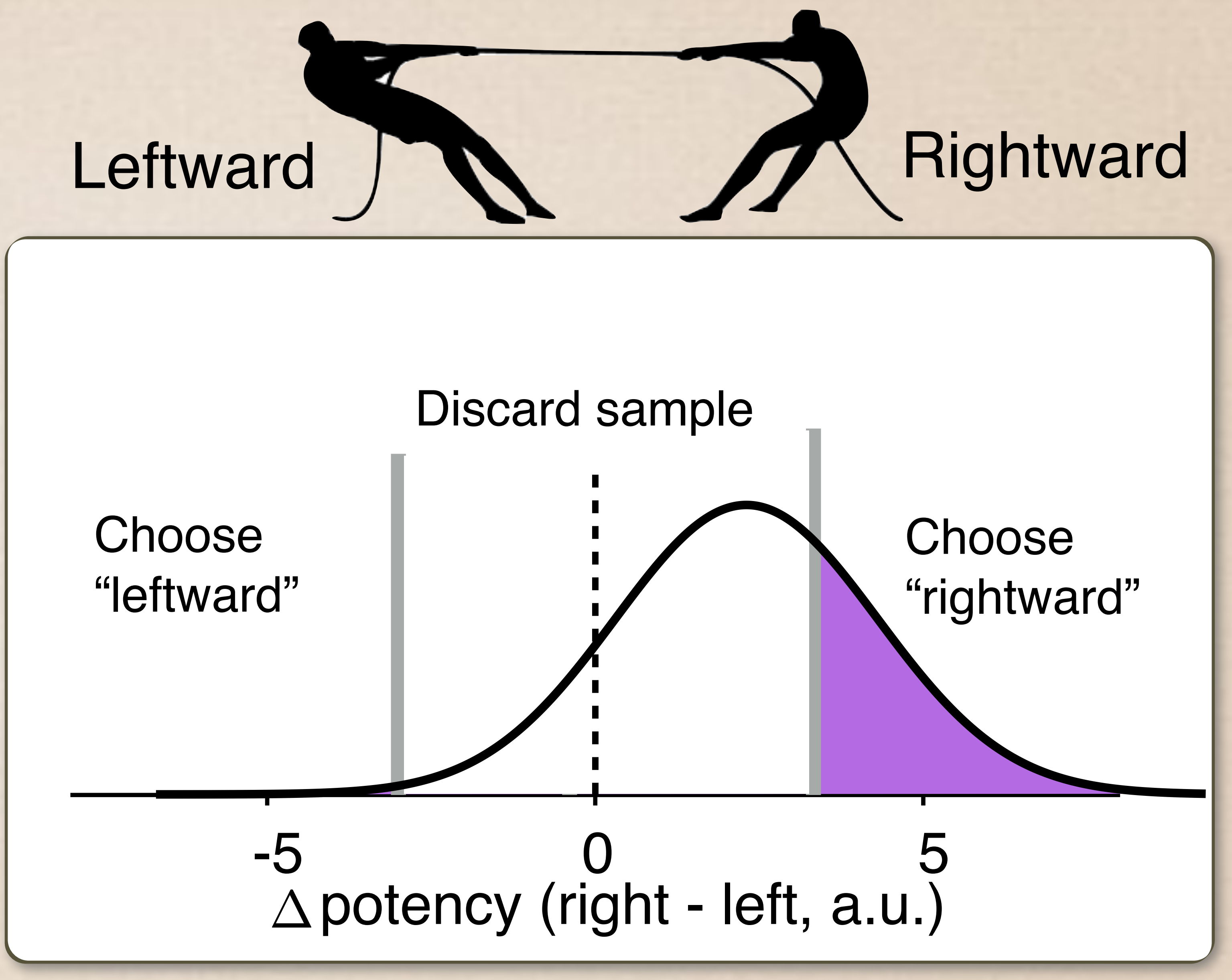
Cartwright and Festinger (1943), “A quantitative theory of decision”.



Opponent process theory of discrimination

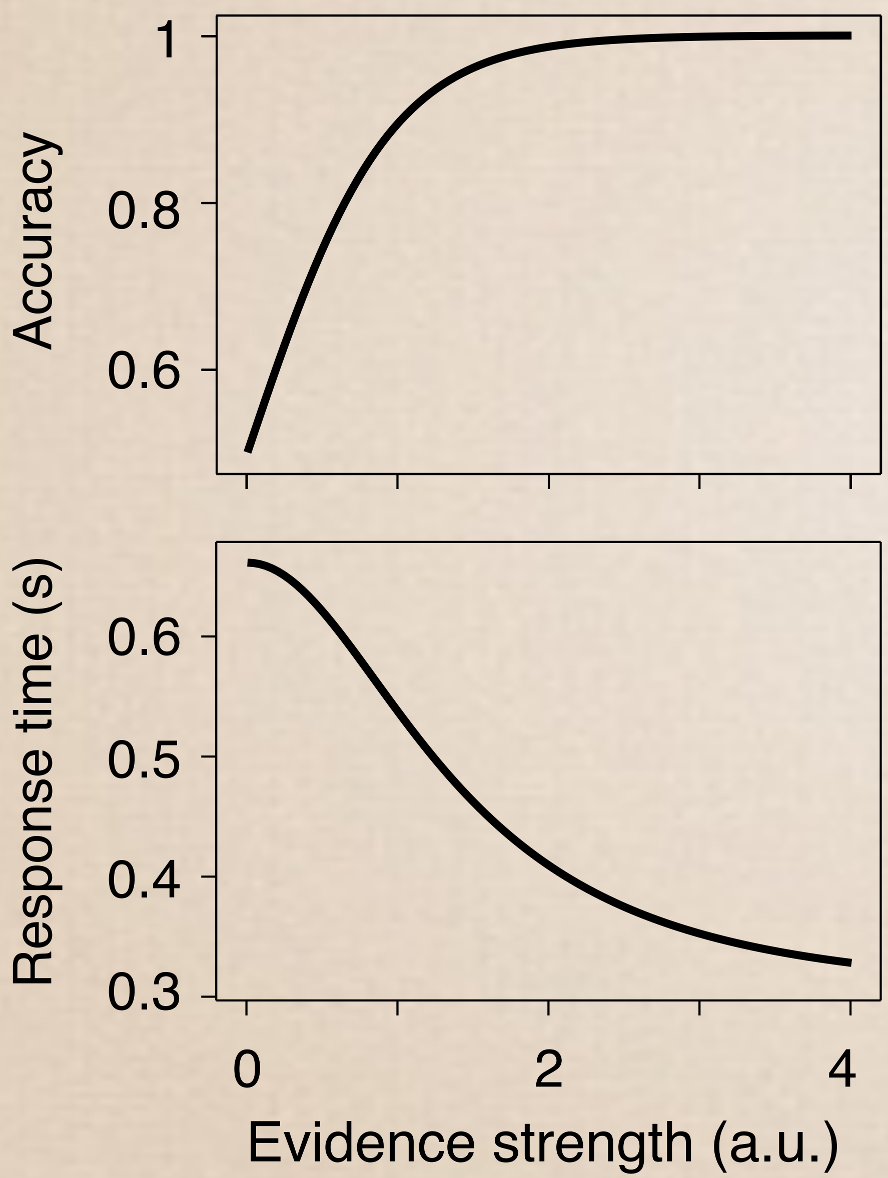
Cartwright and Festinger (1943), “A quantitative theory of decision”.

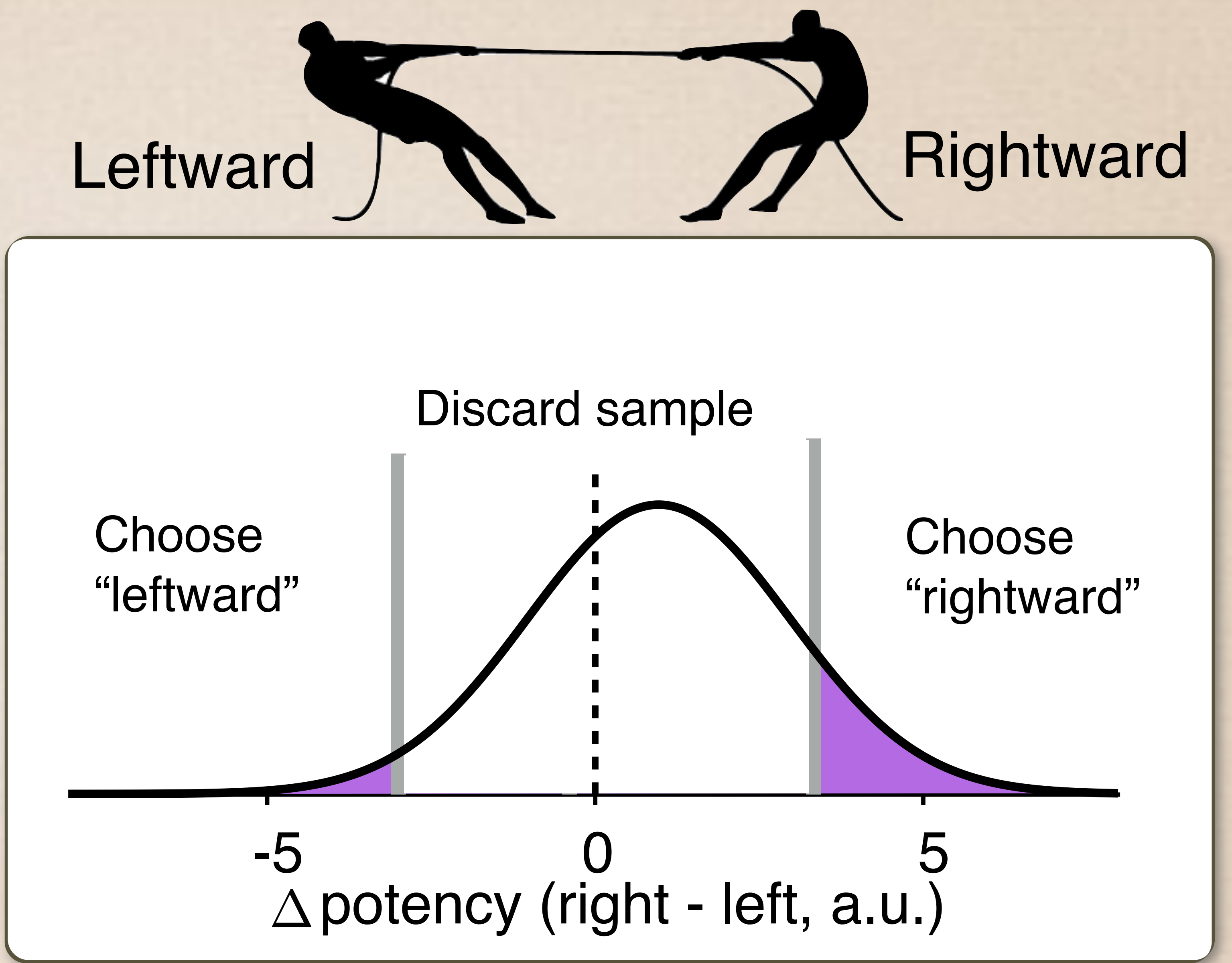




Opponent process theory of discrimination

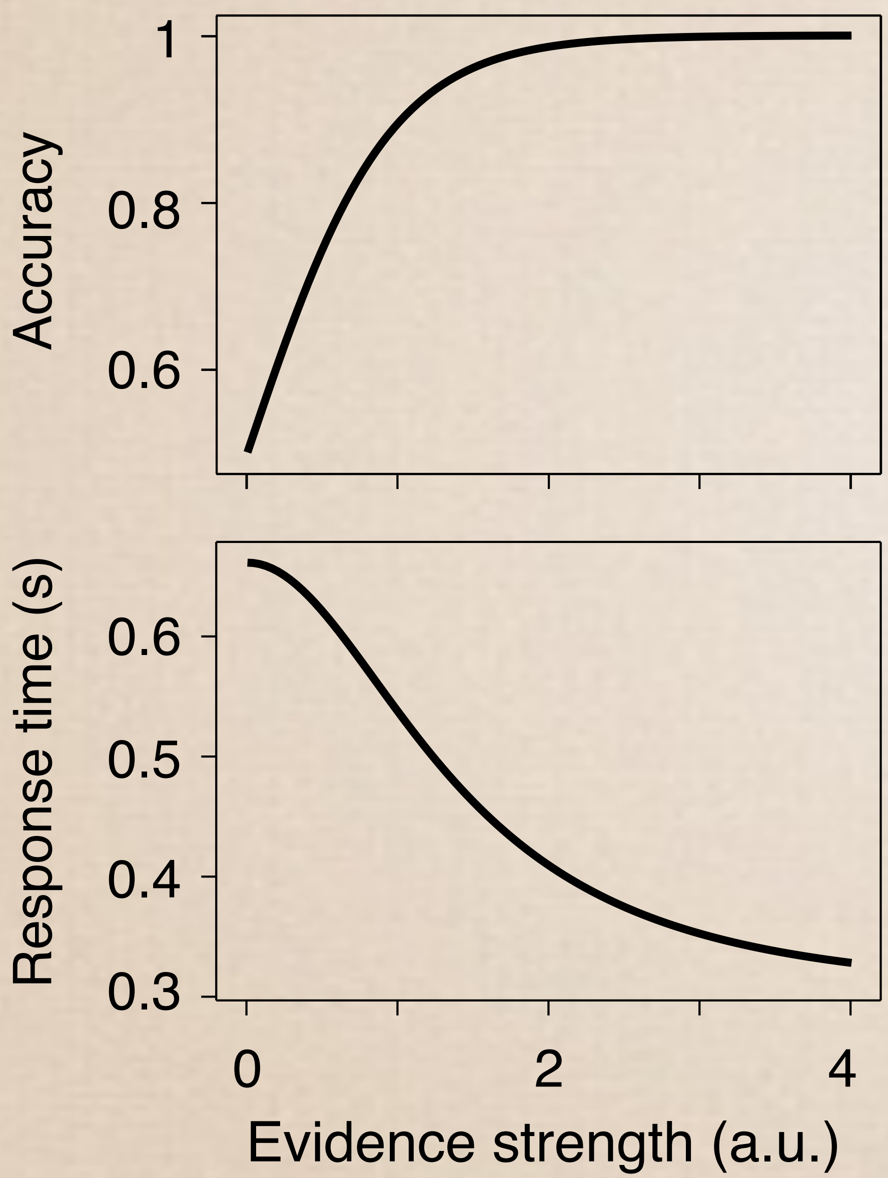
Cartwright and Festinger (1943), “A quantitative theory of decision”.

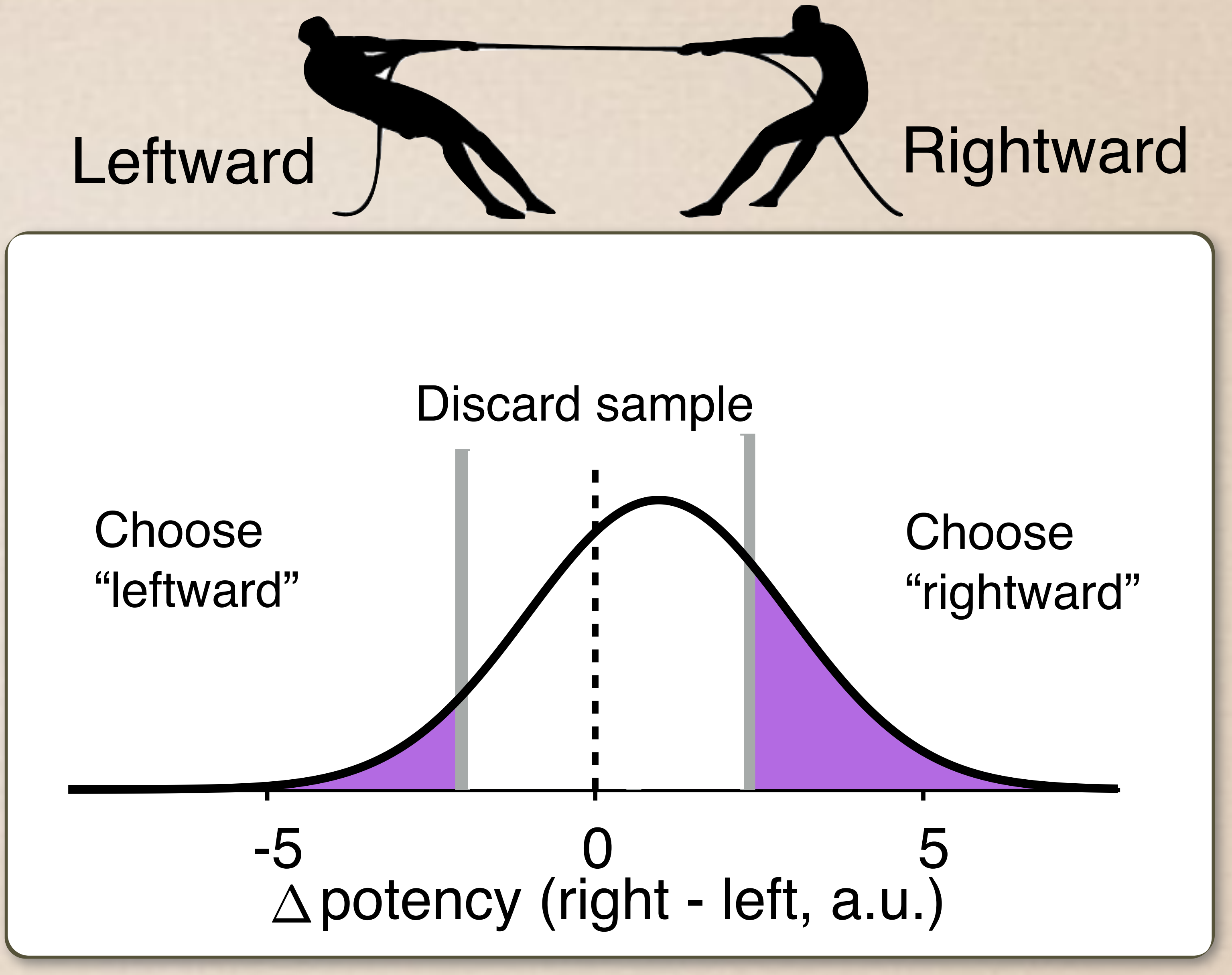




Opponent process theory of discrimination

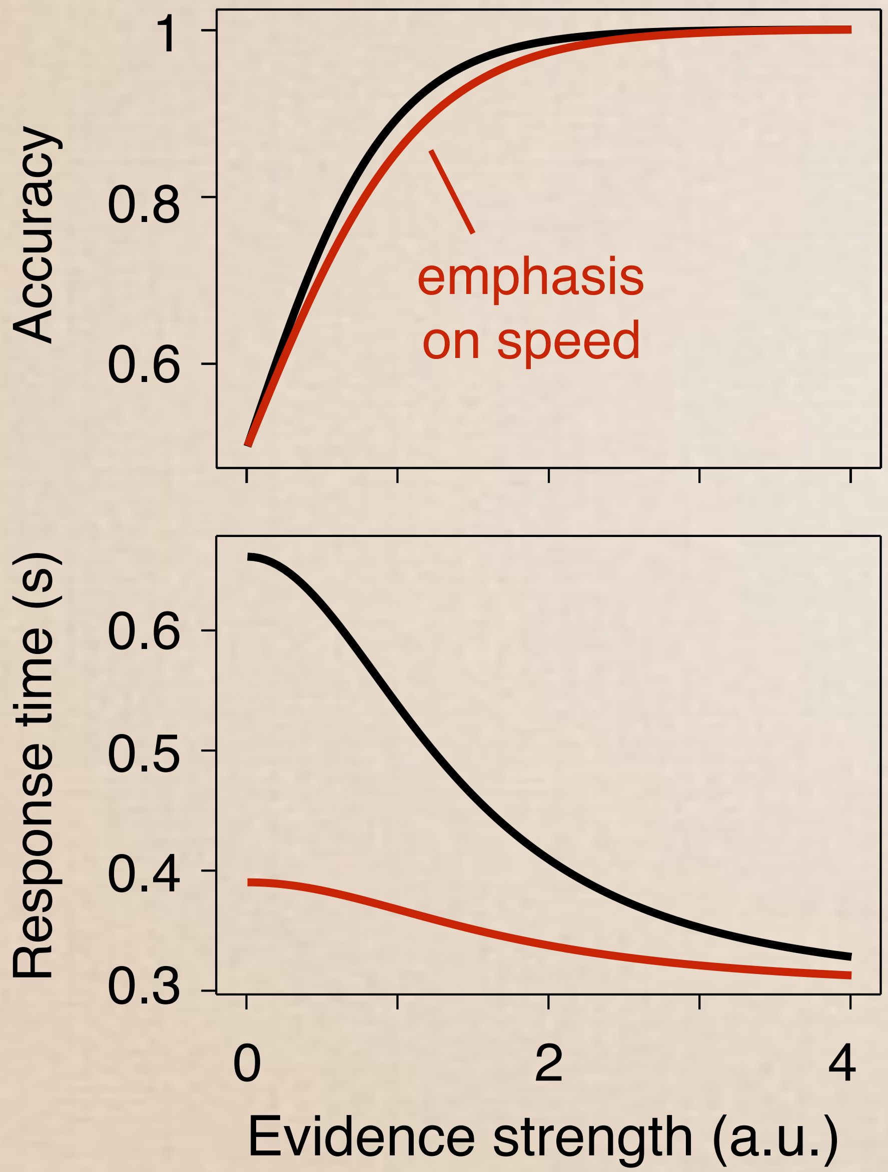
Cartwright and Festinger (1943), "A quantitative theory of decision".

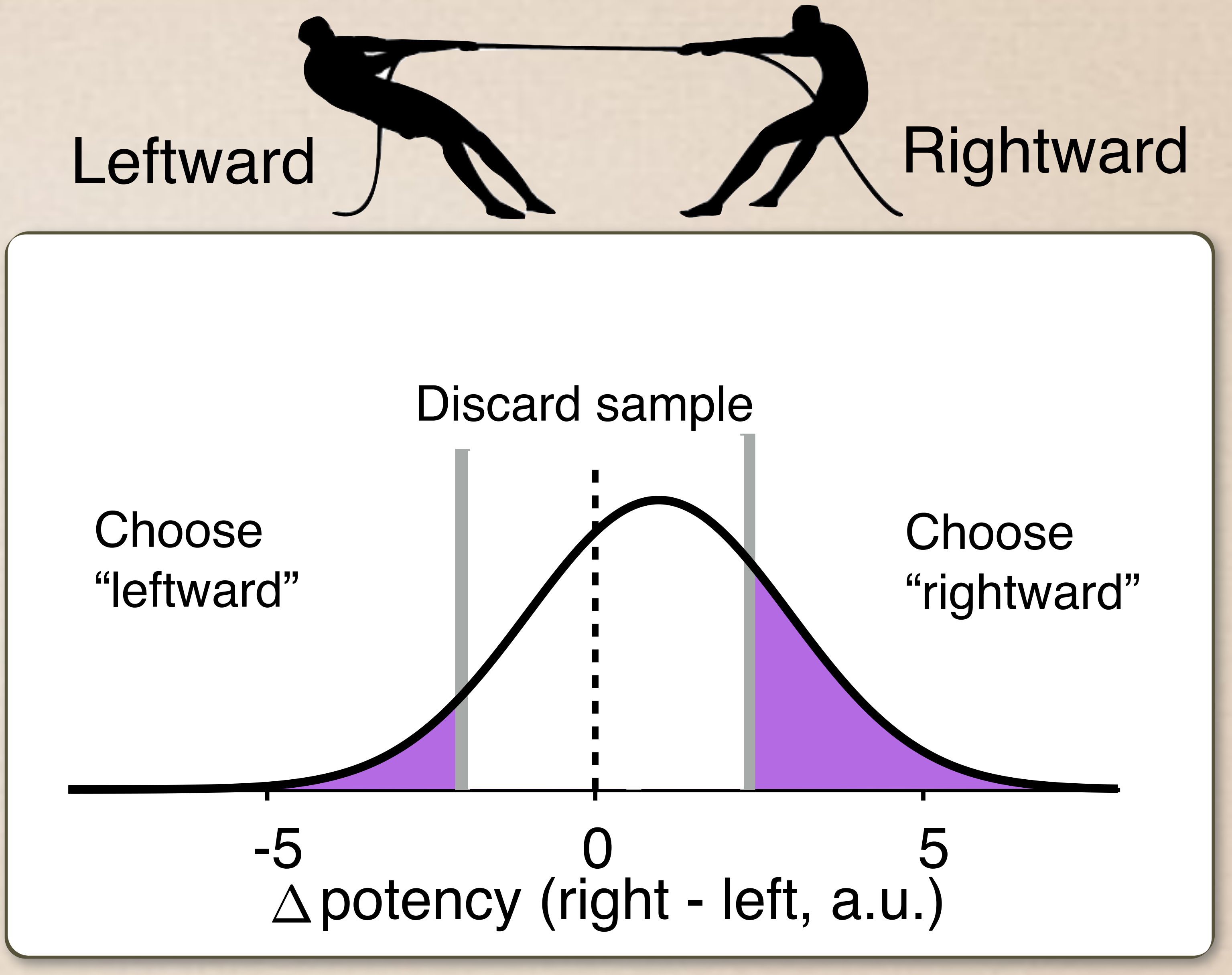




Opponent process theory of discrimination

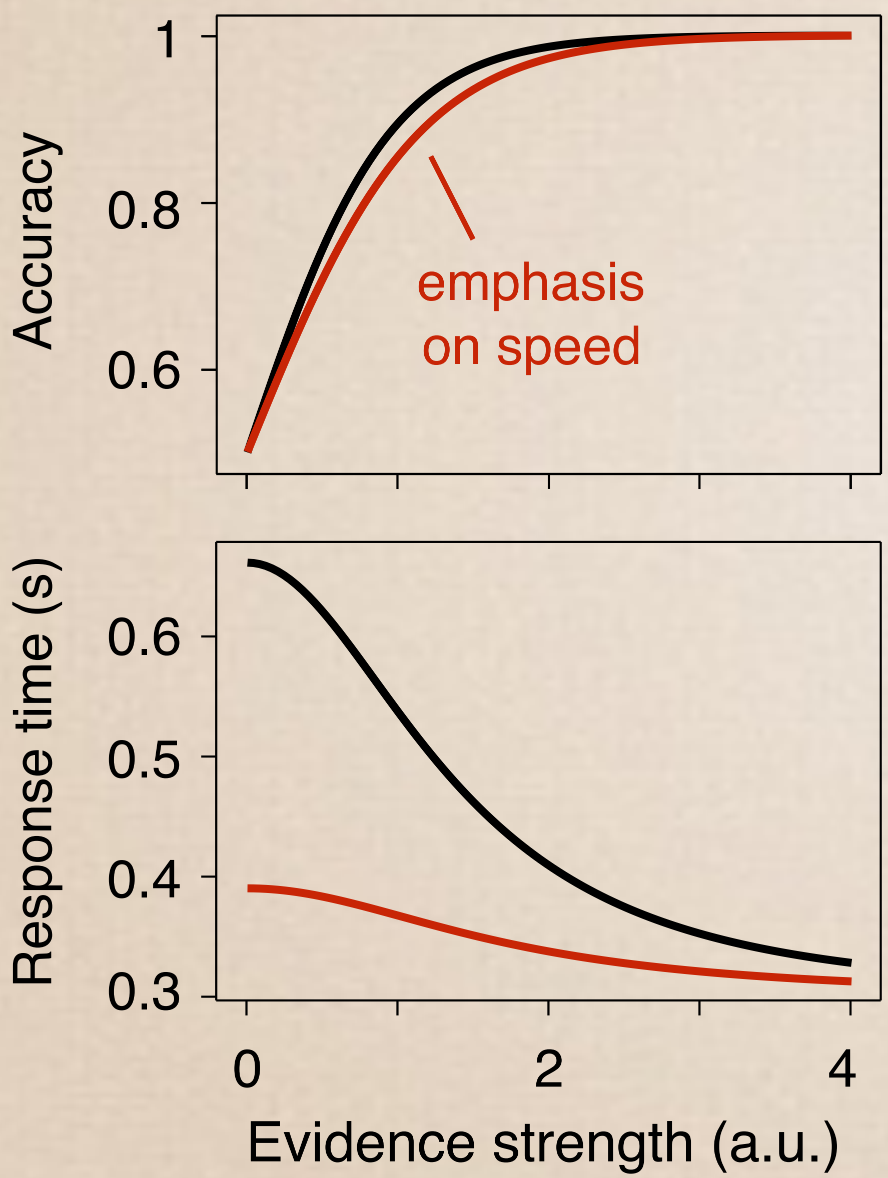
Cartwright and Festinger (1943), “A quantitative theory of decision”.

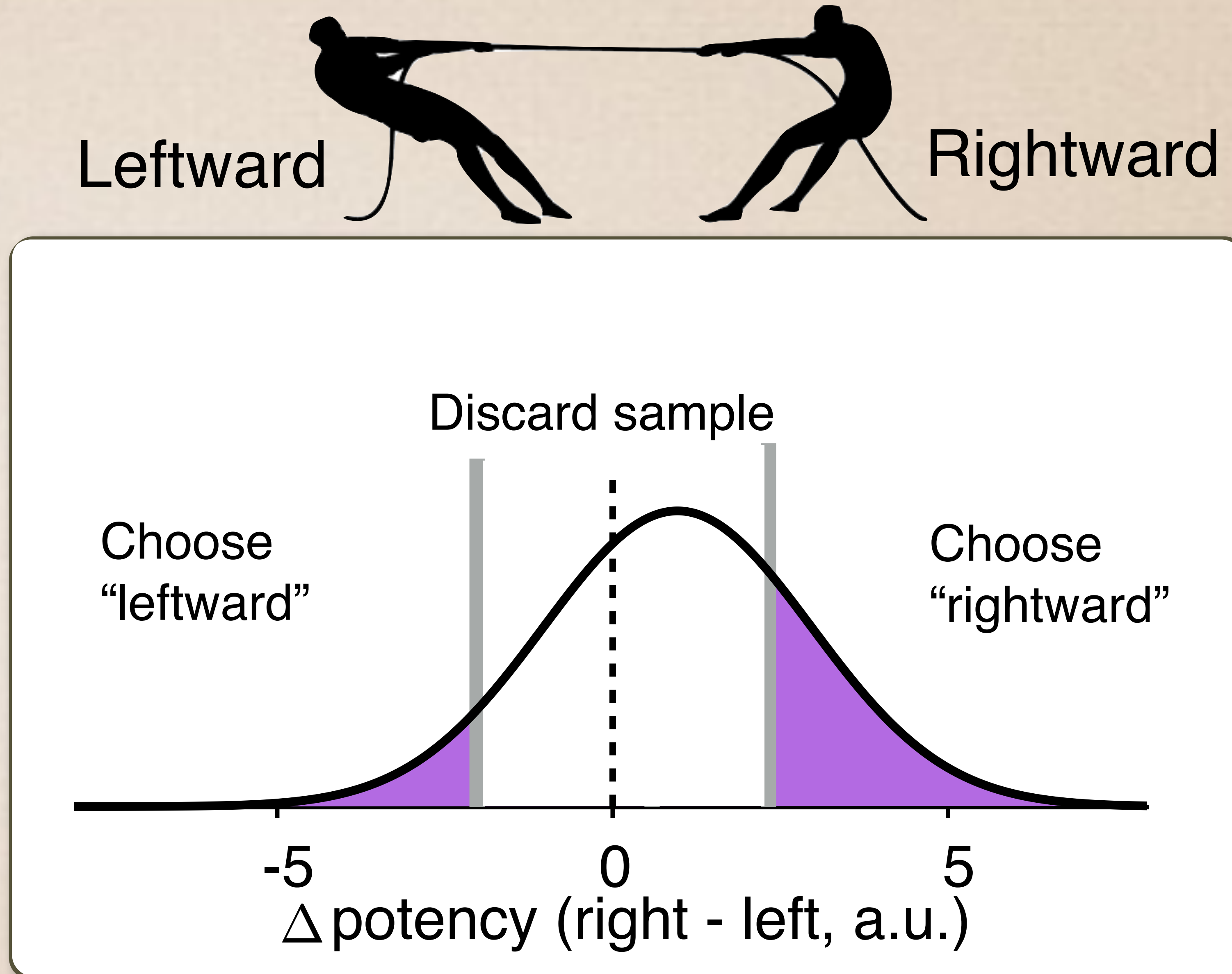




Opponent process theory of discrimination

Cartwright and Festinger (1943), "A quantitative theory of decision".

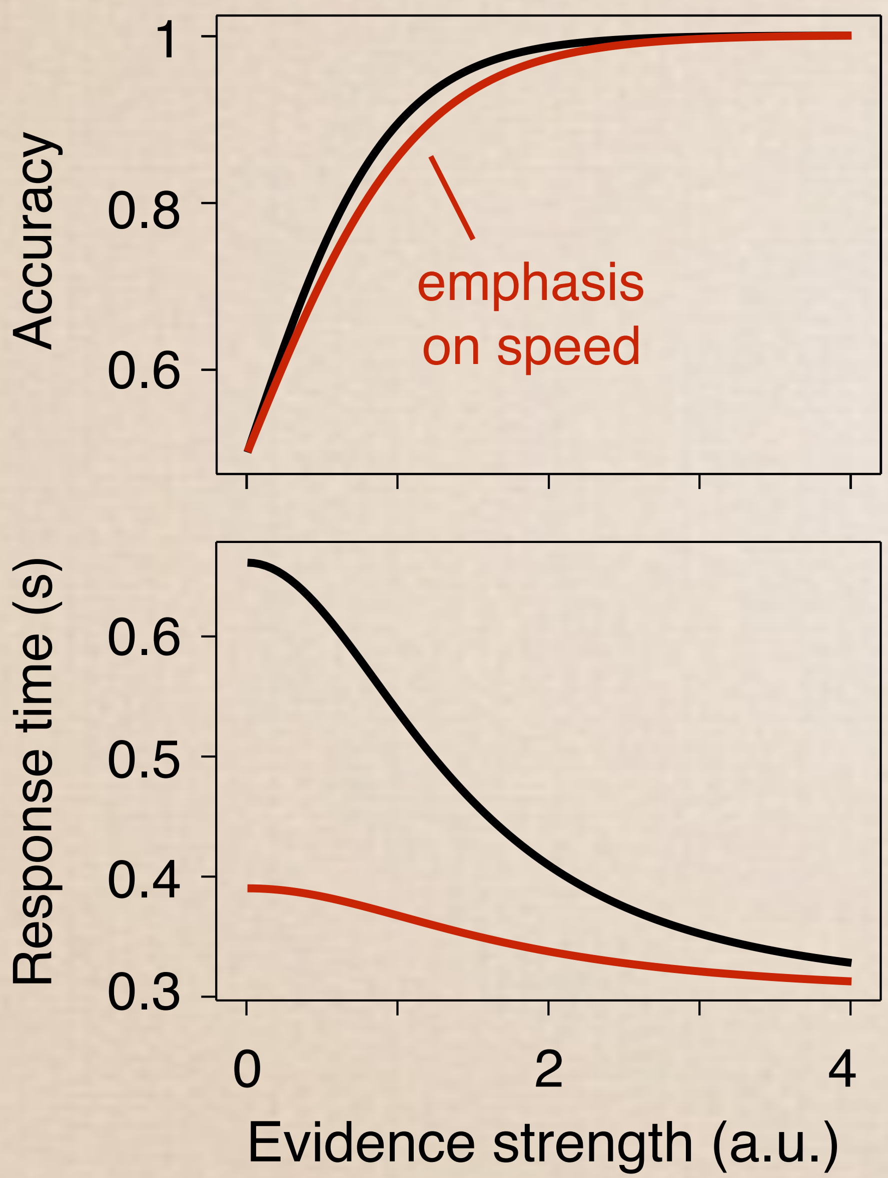


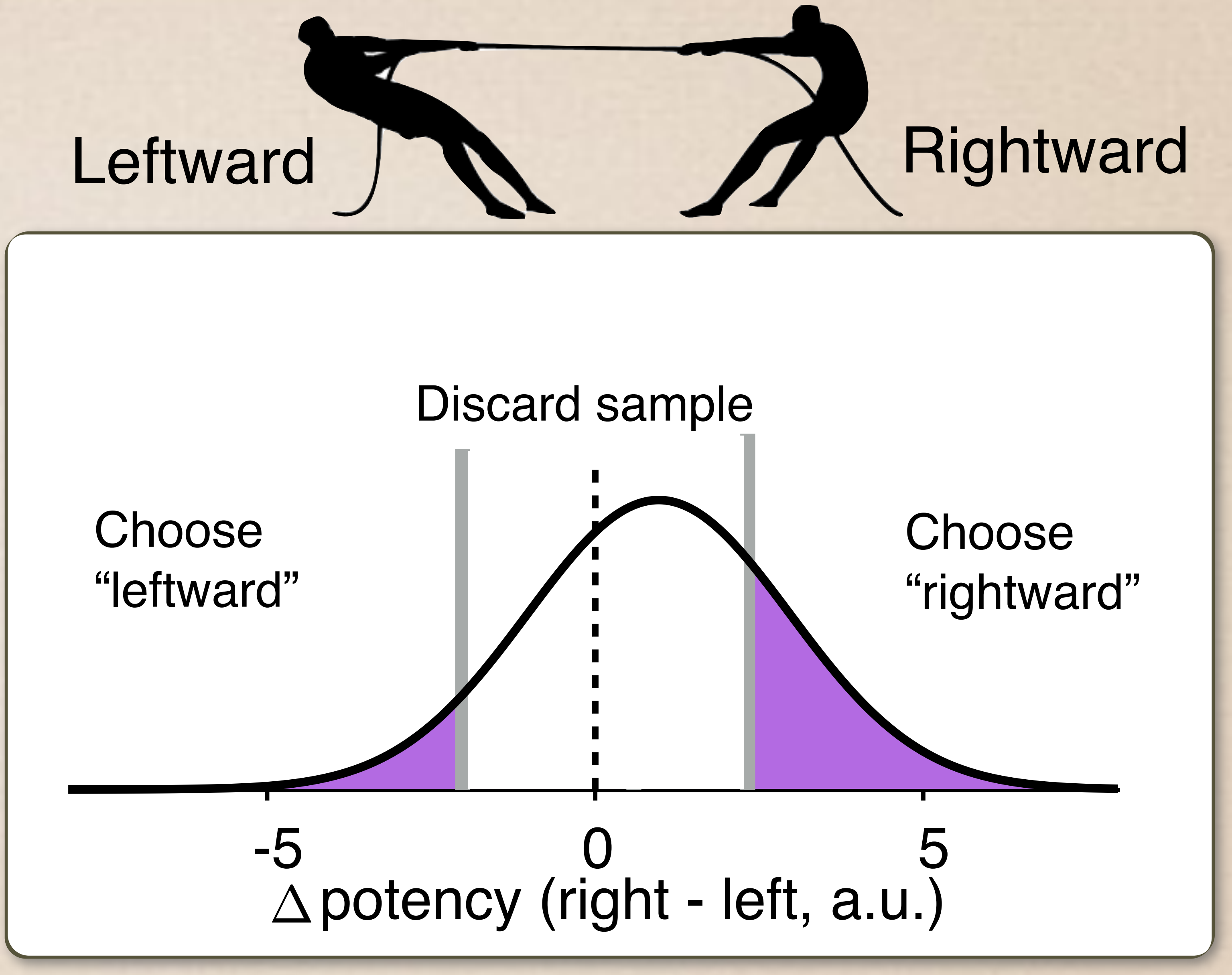


» Decisions may take a long time if high levels of accuracy are required.

Opponent process theory of discrimination

Cartwright and Festinger (1943), "A quantitative theory of decision".

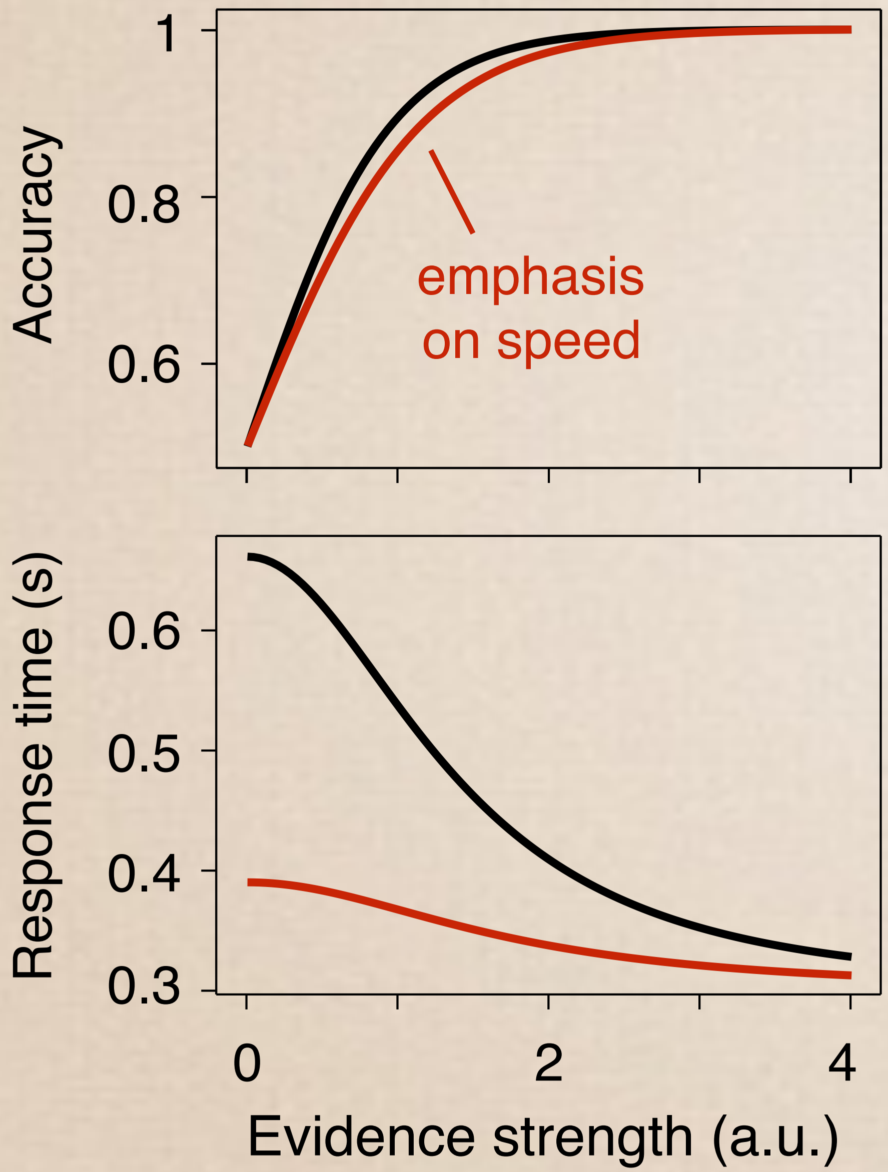


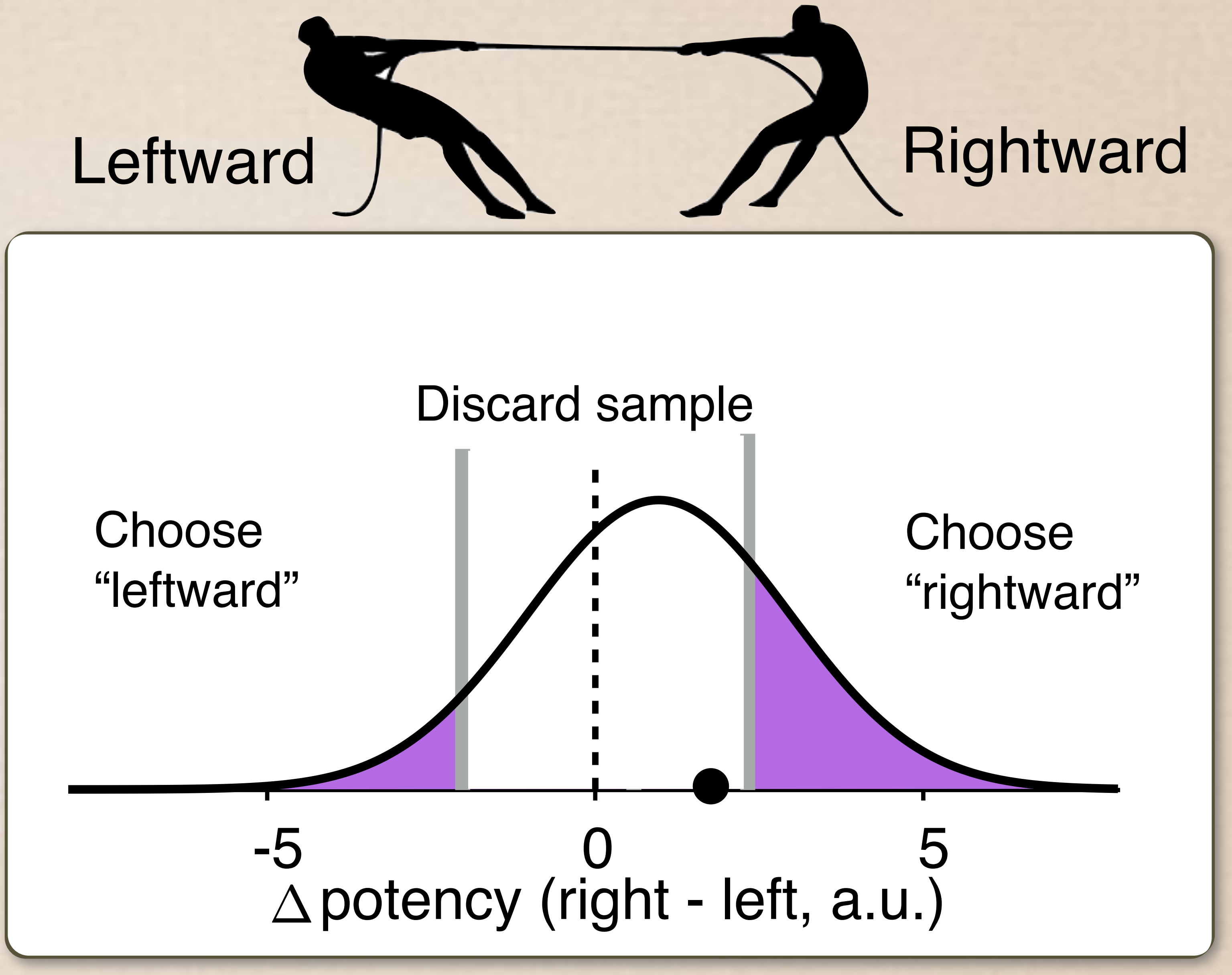


- Decisions may take a long time if high levels of accuracy are required.
- Decision rule is sub-optimal.

Opponent process theory of discrimination

Cartwright and Festinger (1943), “A quantitative theory of decision”.





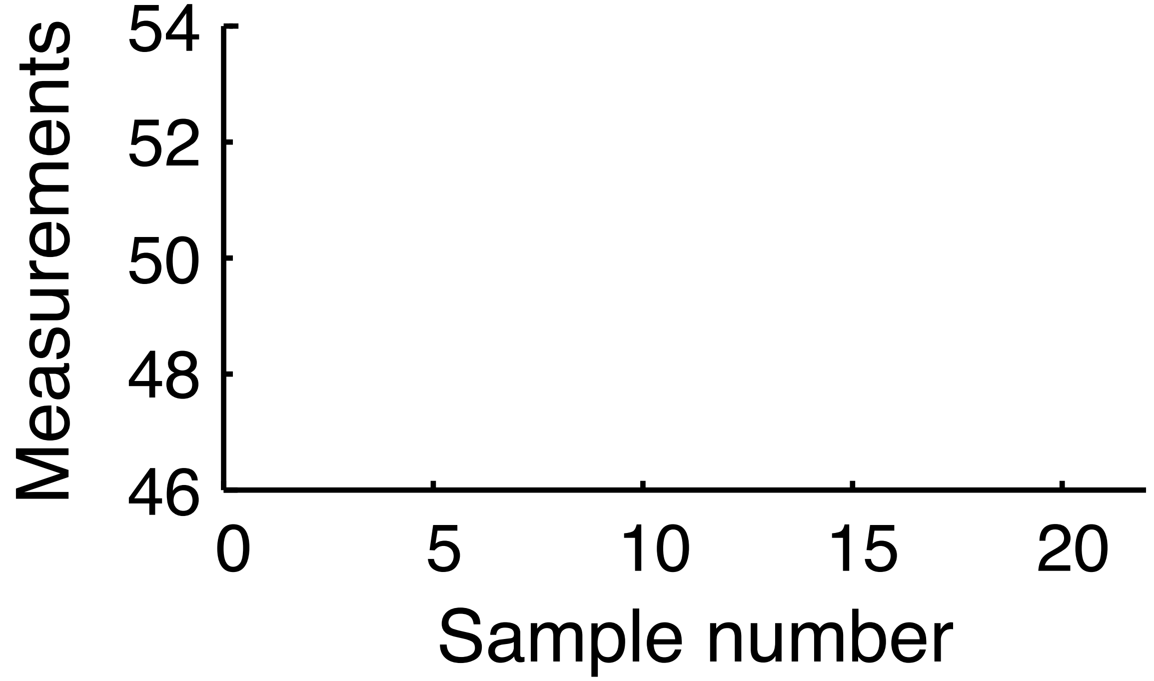
- Decisions may take a long time if high levels of accuracy are required.
- Decision rule is sub-optimal.

Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)

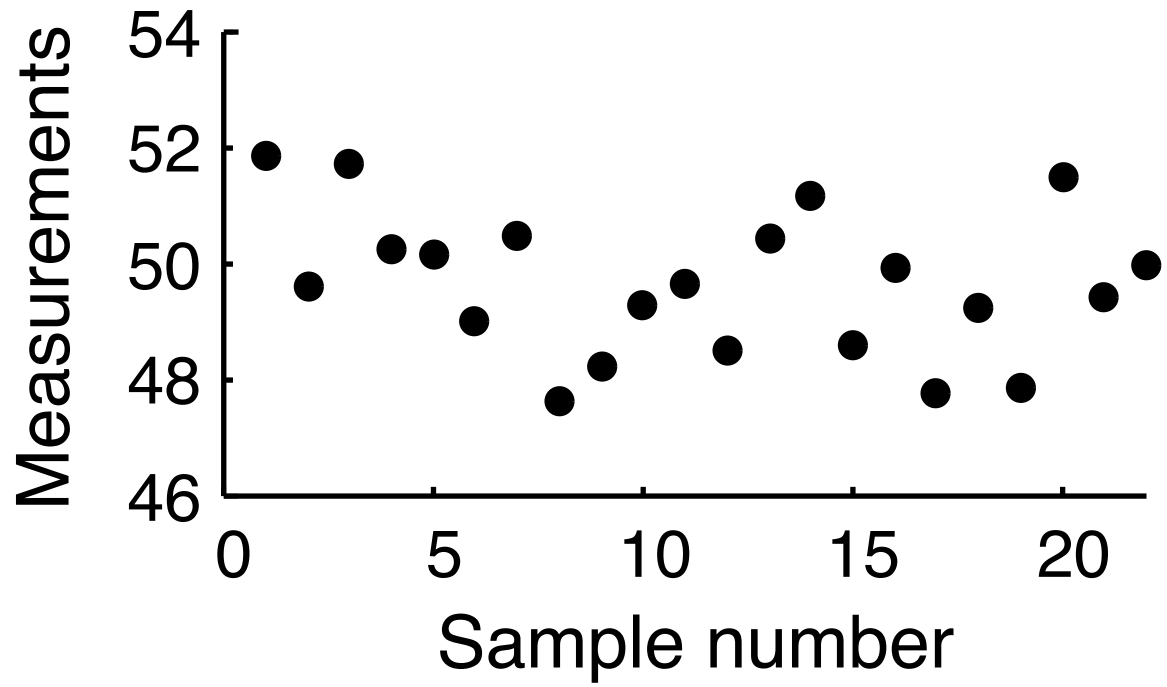
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



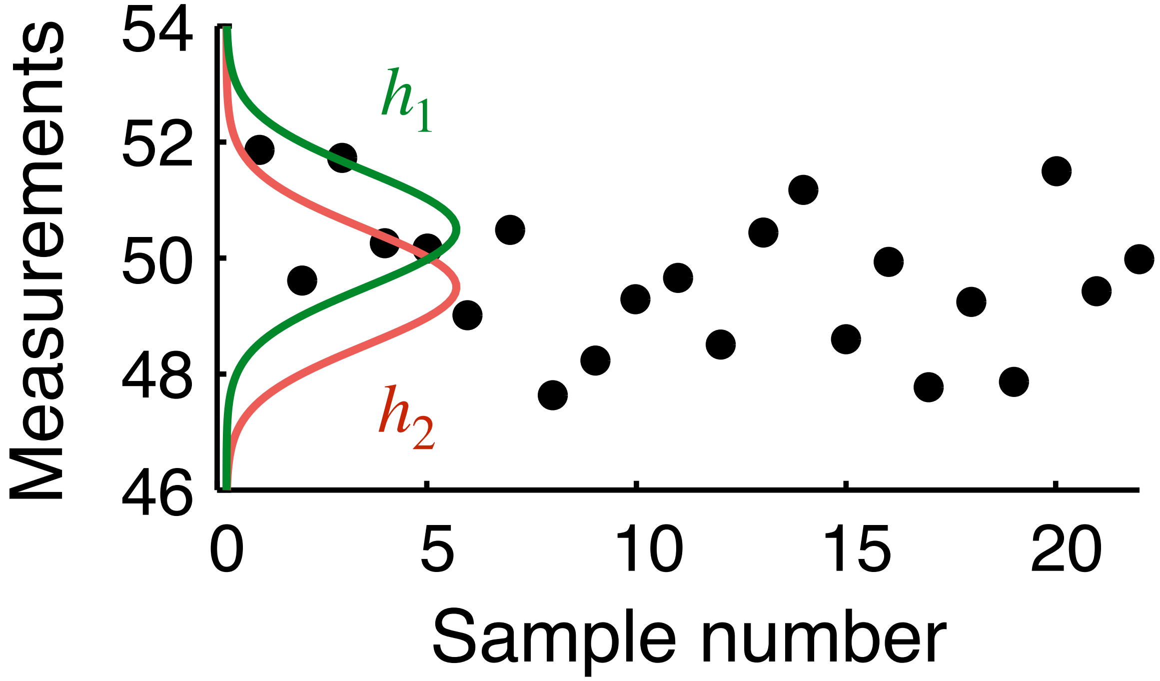
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



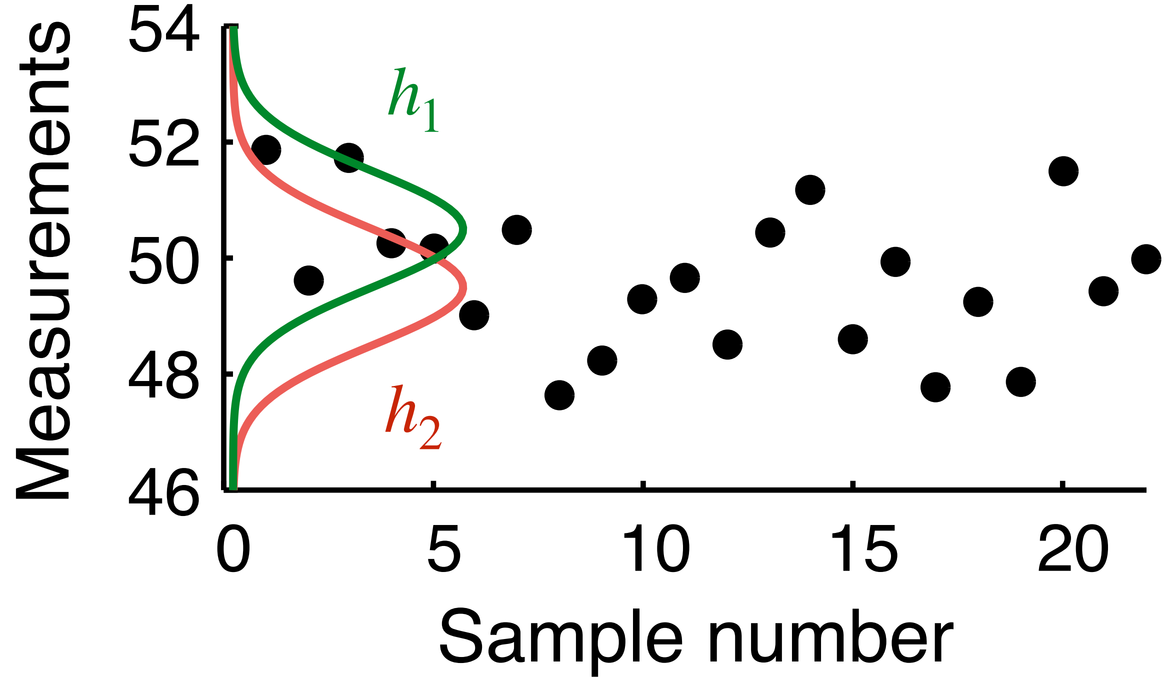
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)

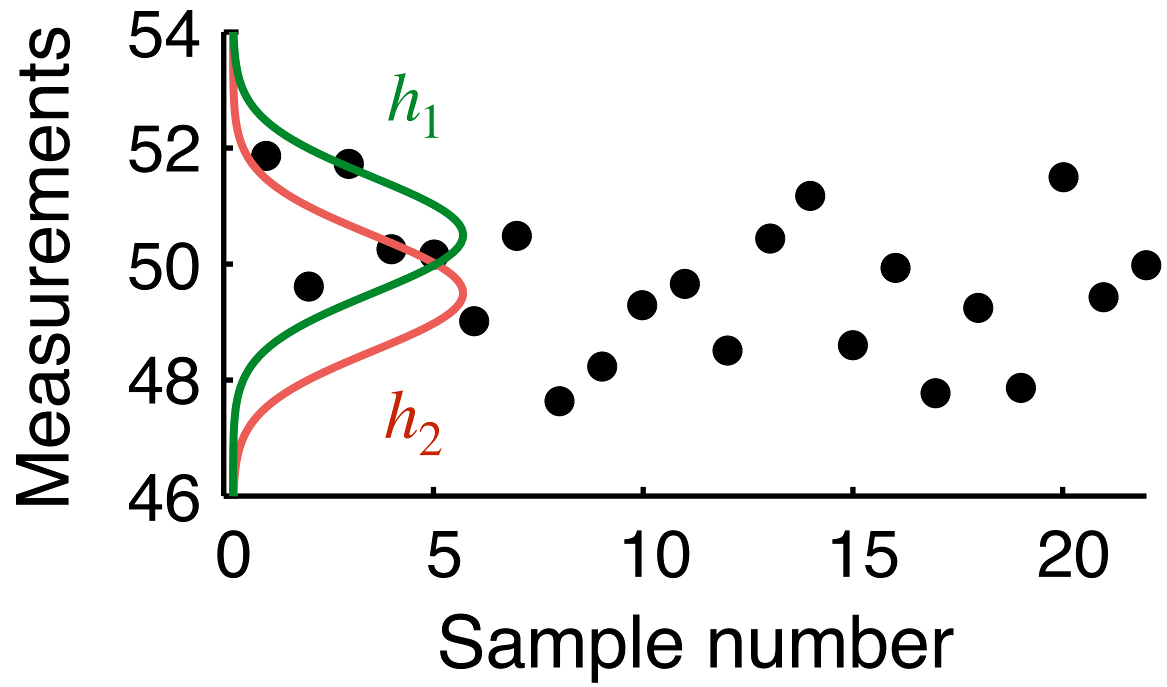


Weight of evidence e in favor of h_1 over h_2 =

$$Log\left[\frac{\text{Probability of } e \text{ if } h_1 \text{ is true}}{\text{Probability of } e \text{ if } h_2 \text{ is true}}\right]$$

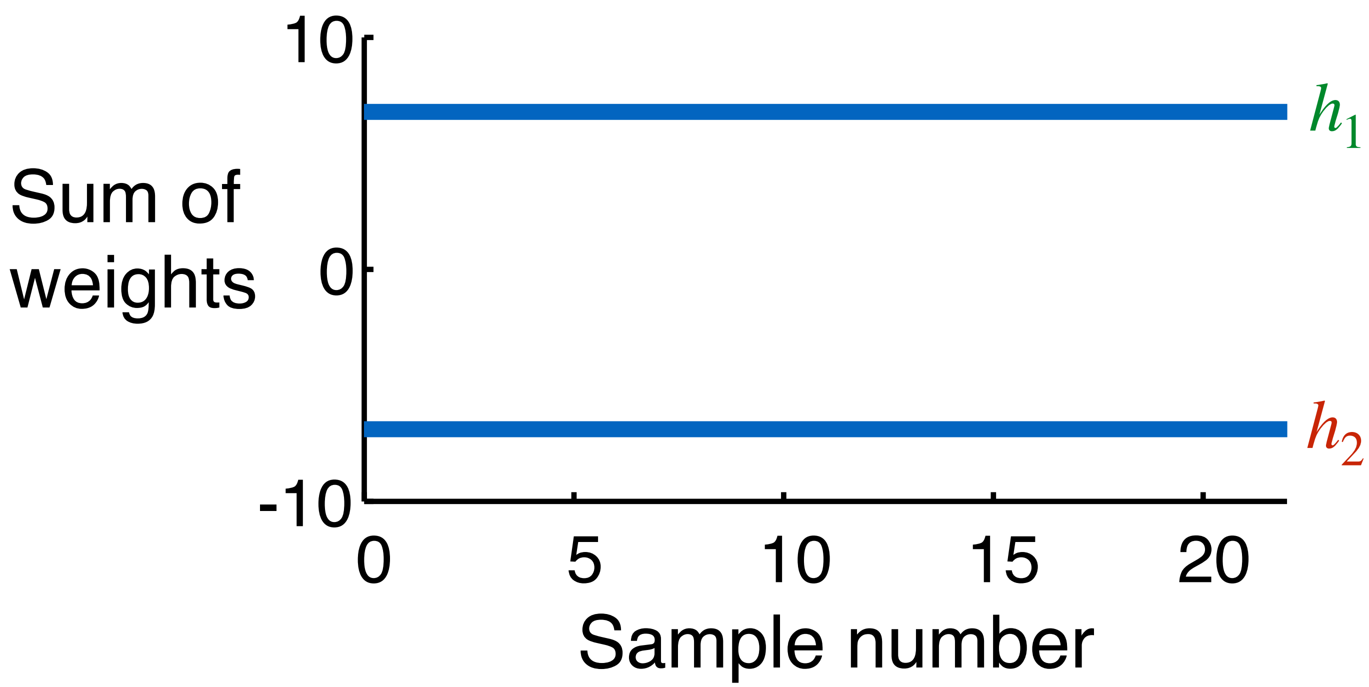
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



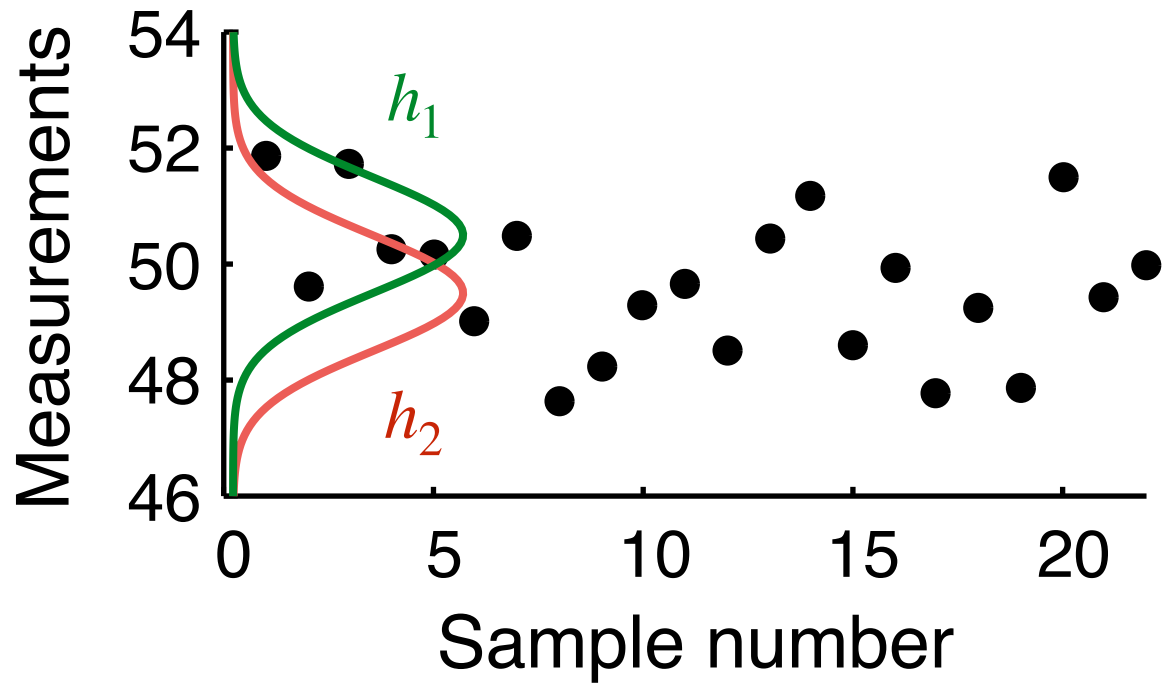
Weight of evidence e in favor of h_1 over h_2 =

$$\mathit{Log}\left[\frac{\text{Probability of } e \text{ if } h_1 \text{ is true}}{\text{Probability of } e \text{ if } h_2 \text{ is true}}\right]$$



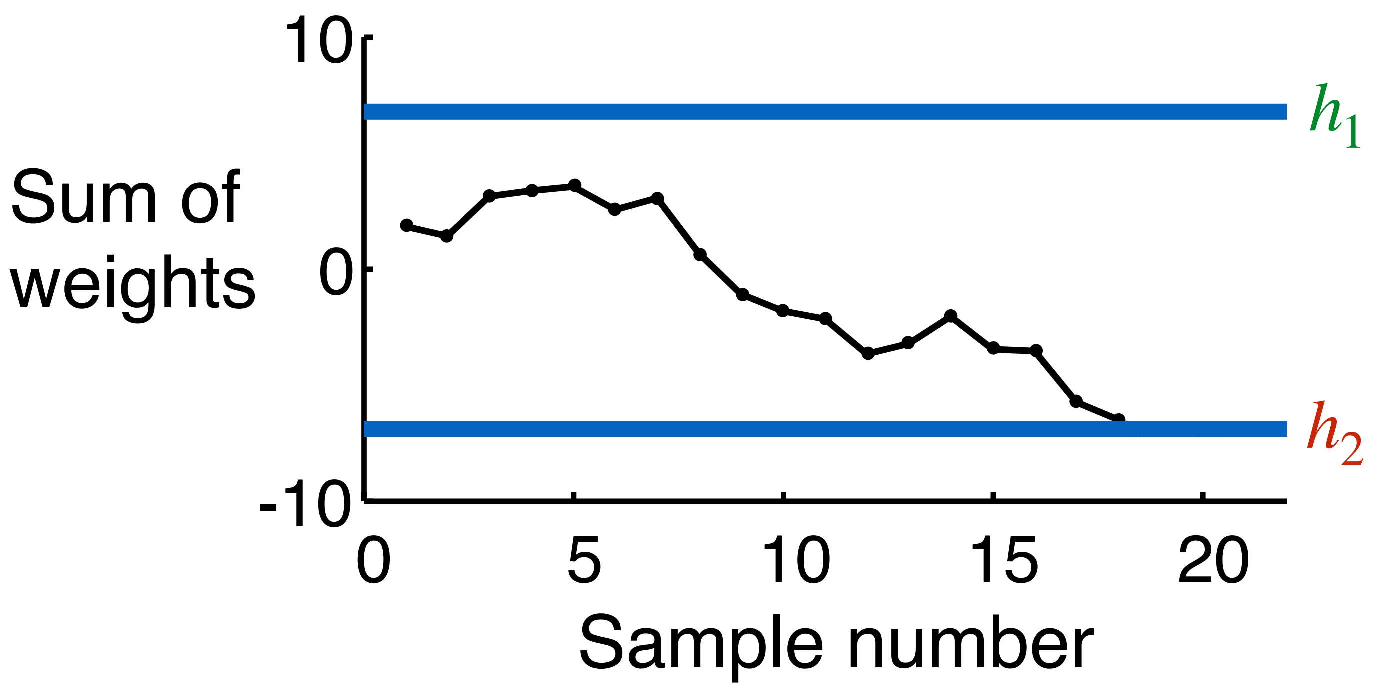
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



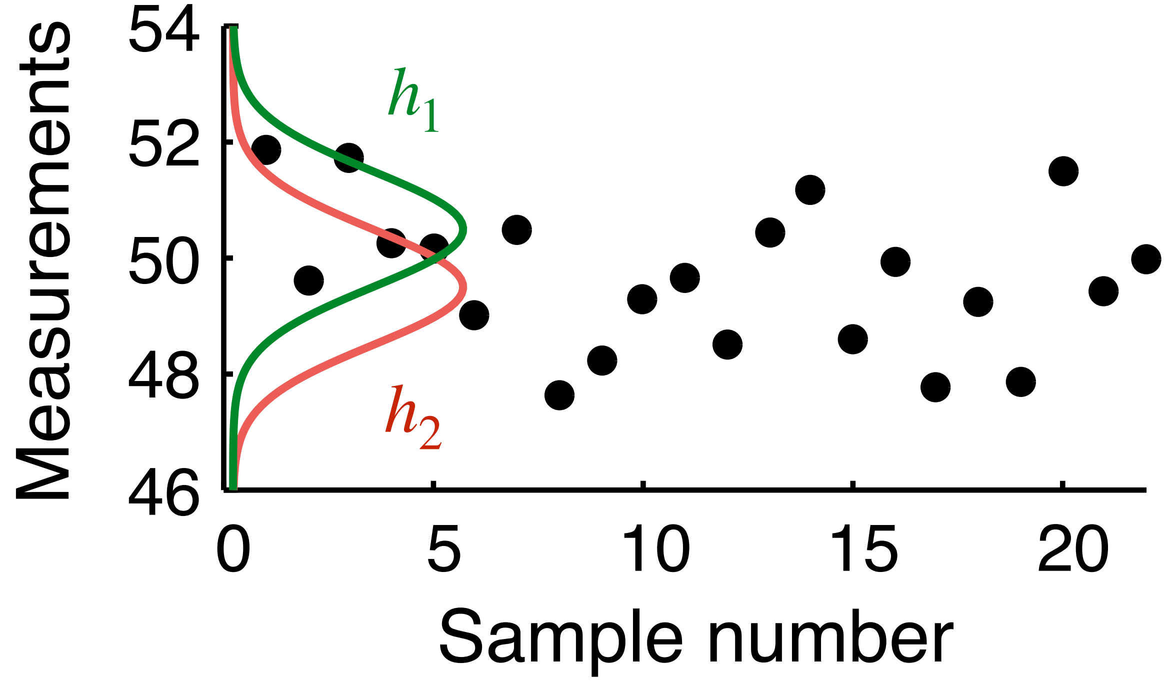
Weight of evidence e in favor of h_1 over h_2 =

$$\mathit{Log}\left[\frac{\textit{Probability of } e \textit{ if } h_1 \textit{ is true}}{\textit{Probability of } e \textit{ if } h_2 \textit{ is true}}\right]$$



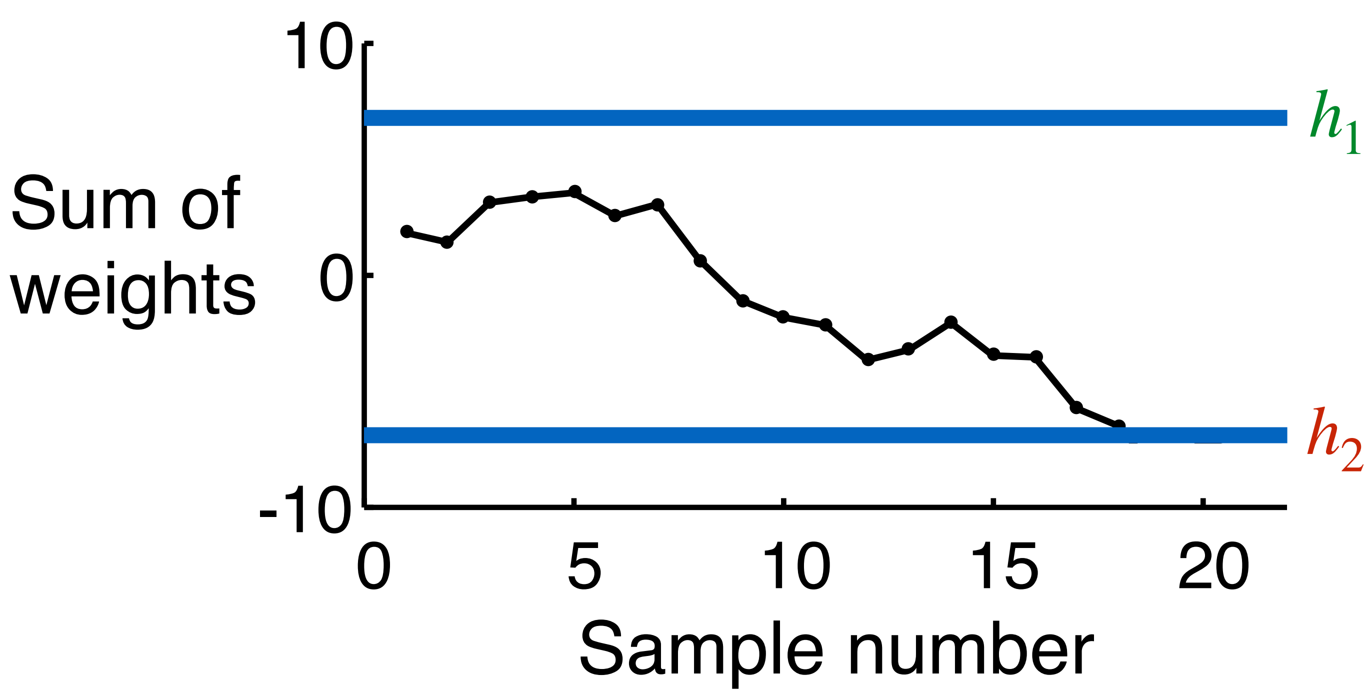
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



Weight of evidence e in favor of h_1 over h_2 =

$$\textit{Log}\left[\frac{\textit{Probability of } e \textit{ if } h_1 \textit{ is true}}{\textit{Probability of } e \textit{ if } h_2 \textit{ is true}}\right]$$



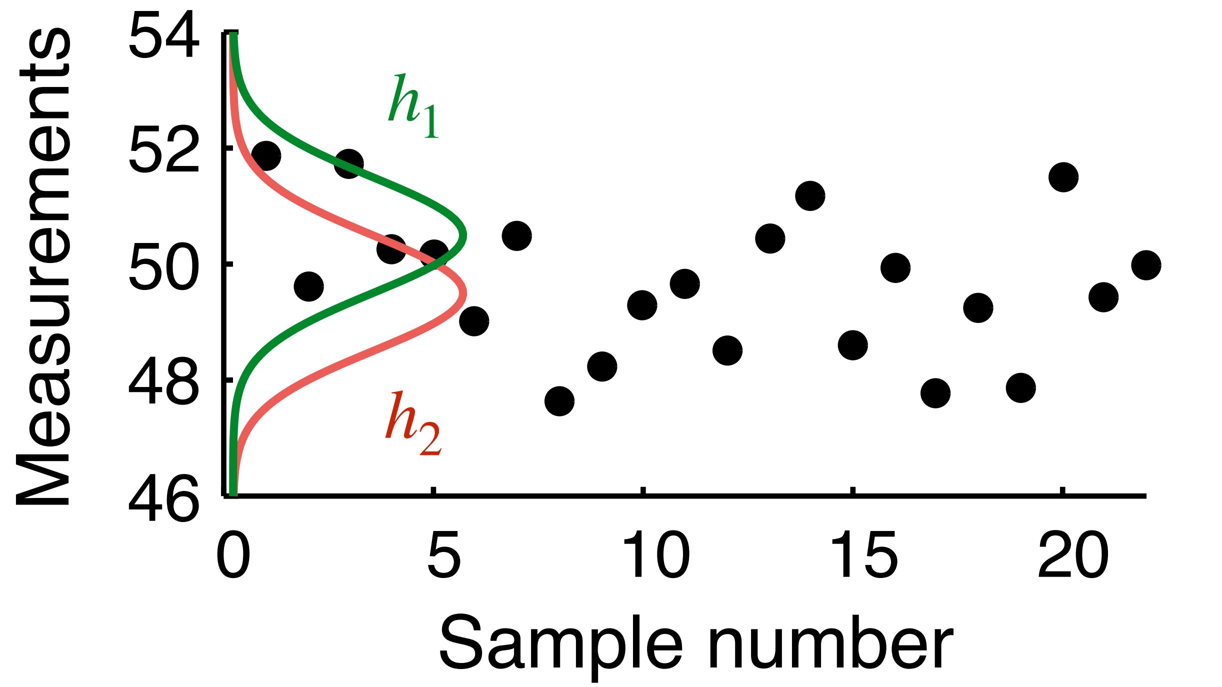
$$B \approx \log\left(\frac{\mathrm{acc}}{1-\mathrm{acc}}\right)$$

bound height

required accuracy

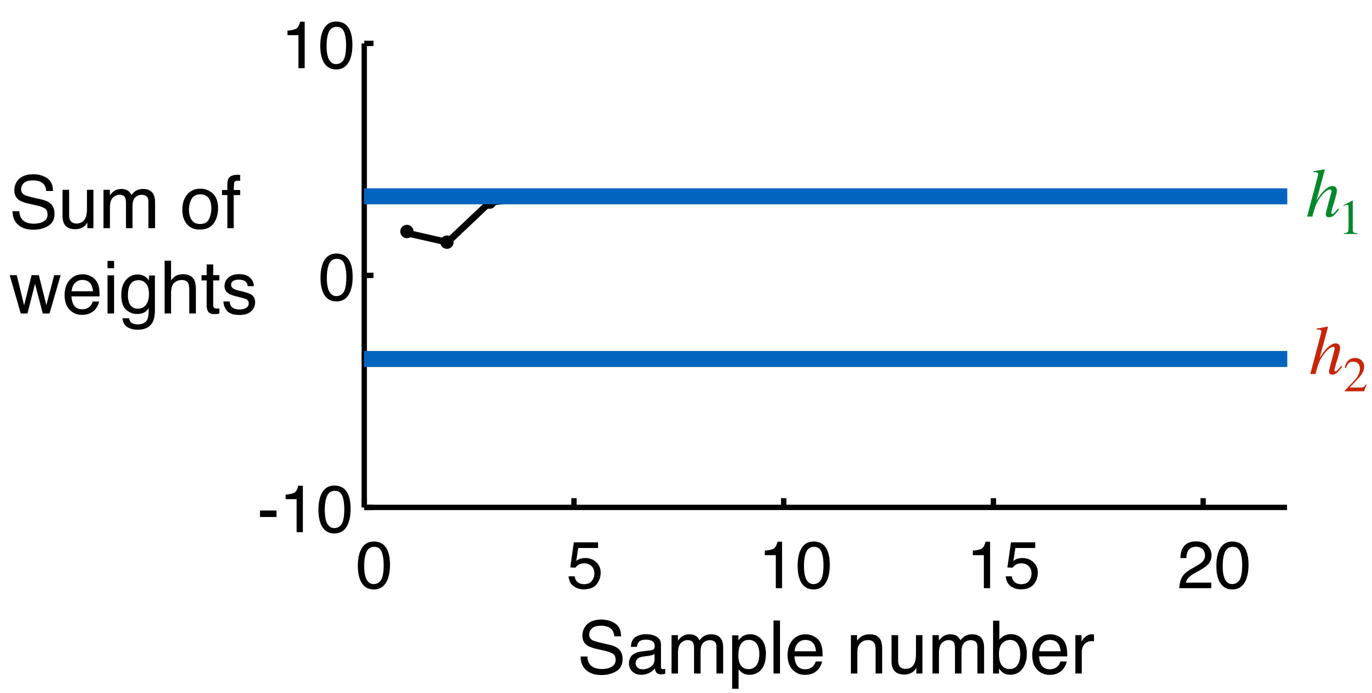
Sequential probability ratio test (SPRT)

Wald & Wolfowitz (1948), A. Turing (194?)



Weight of evidence e in favor of h_1 over h_2 =

$$\textit{Log}\left[\frac{\textit{Probability of } e \textit{ if } h_1 \textit{ is true}}{\textit{Probability of } e \textit{ if } h_2 \textit{ is true}}\right]$$

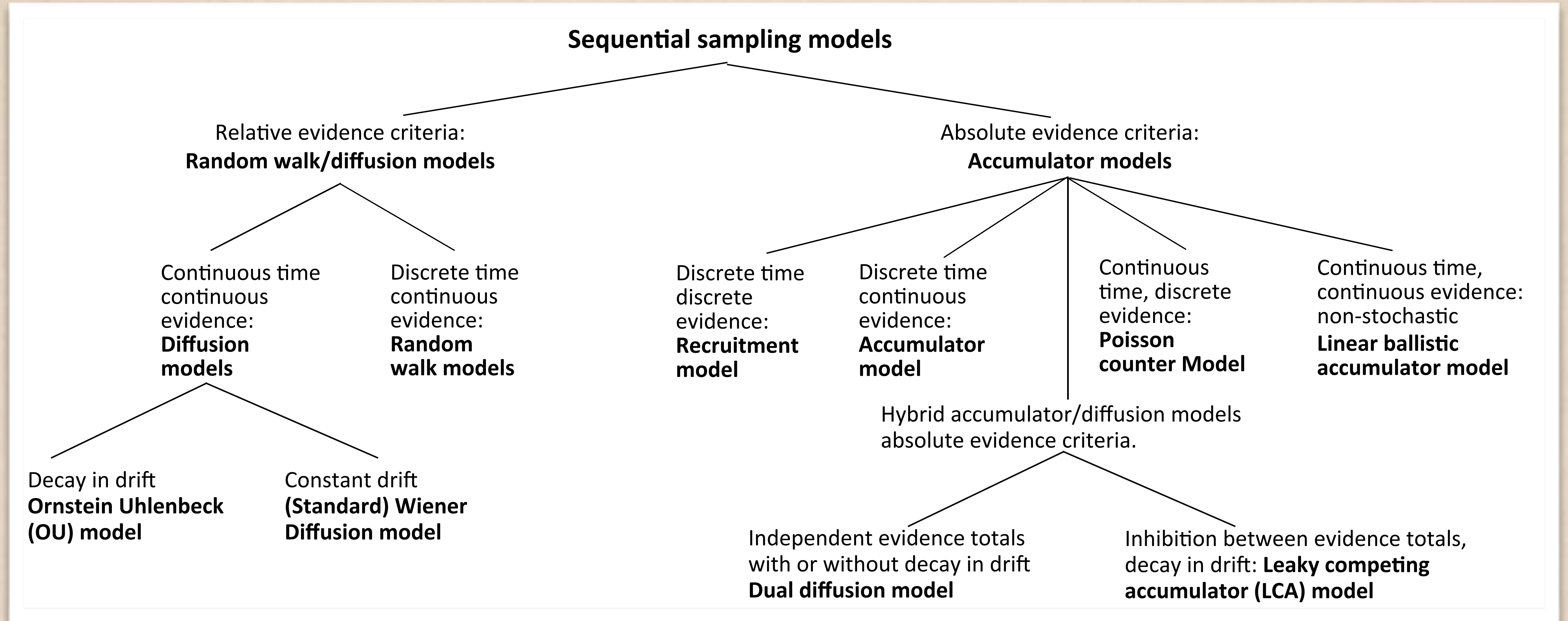


$$B \approx \log\left(\frac{\text{acc}}{1-\text{acc}}\right)$$

bound height

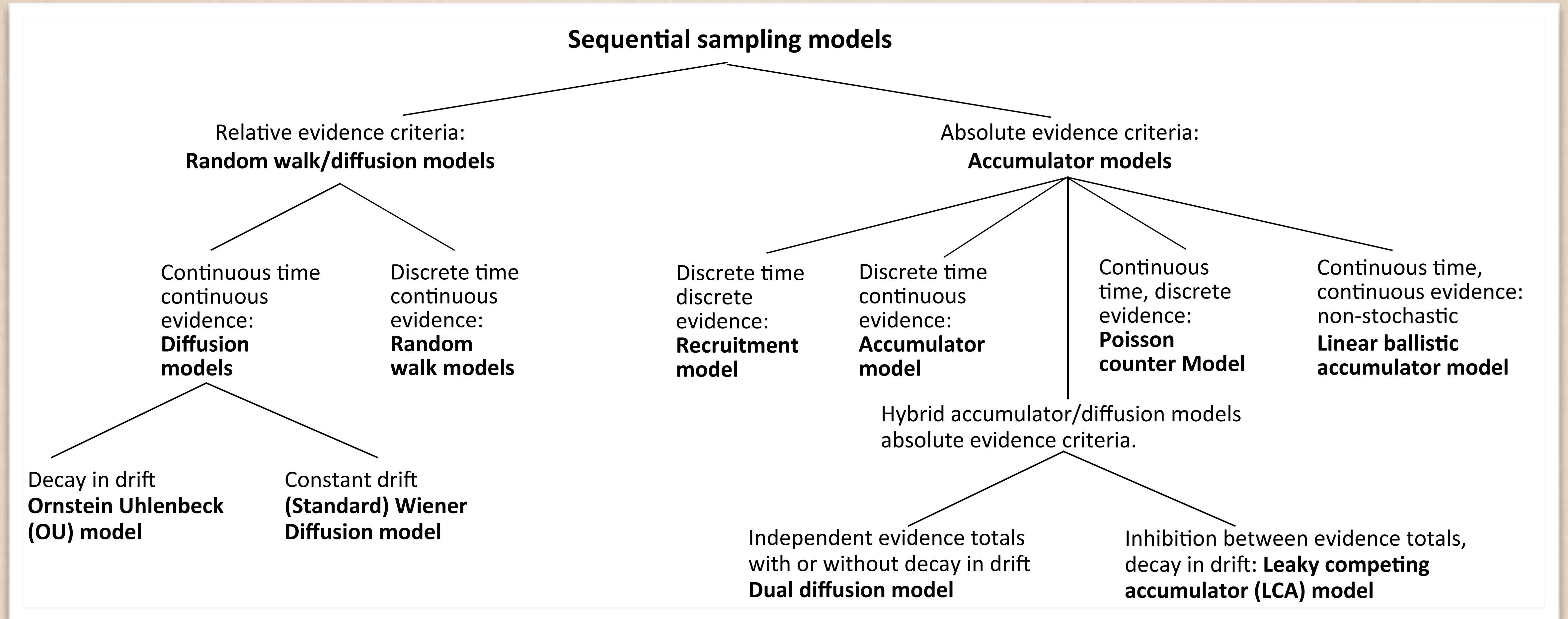
required accuracy

Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

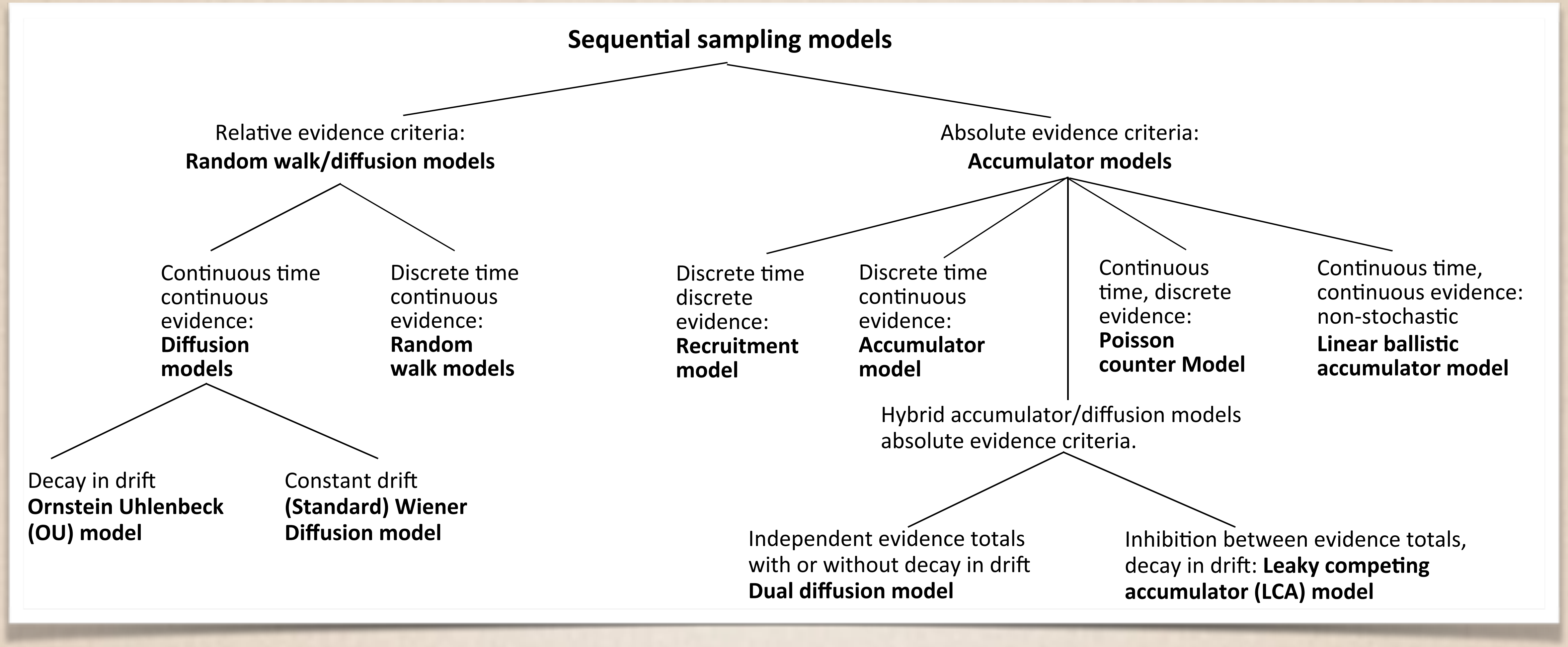
Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

» Time (continuous or discrete)

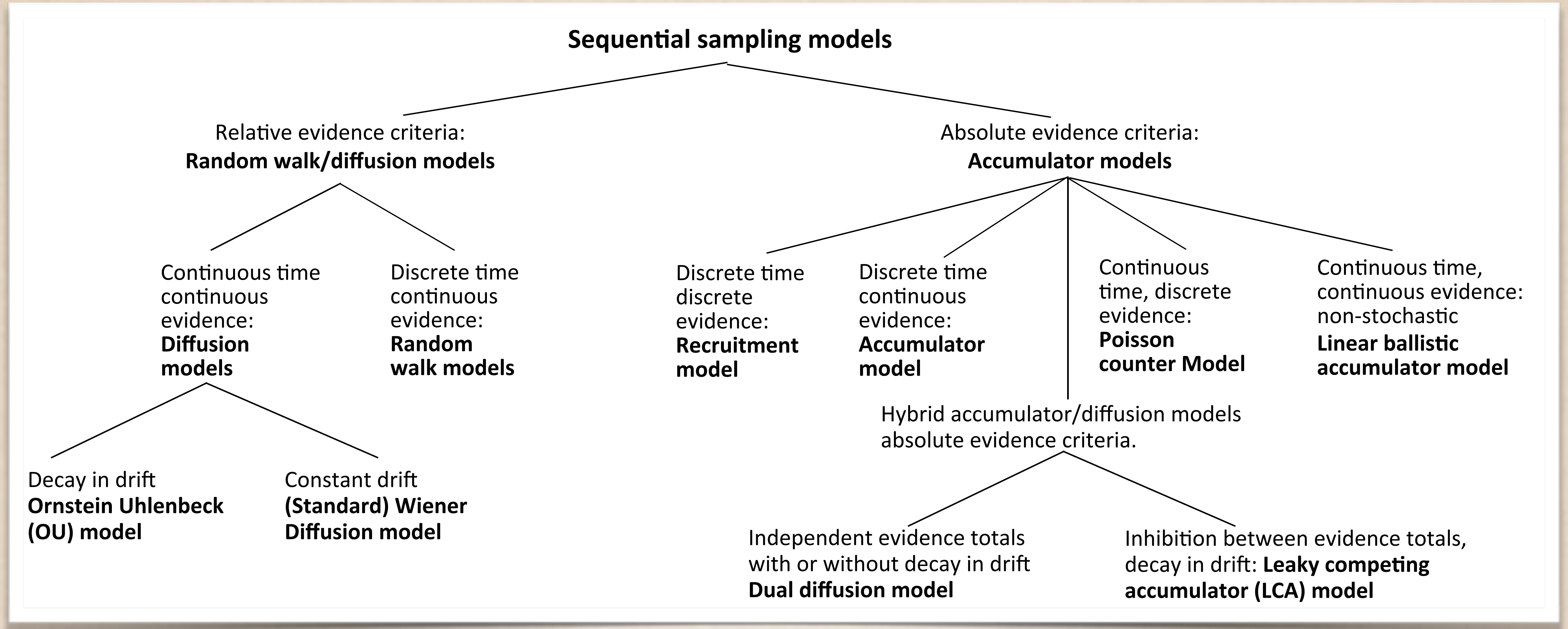
Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

- Time (continuous or discrete)
- Momentary evidence (continuous or discrete)

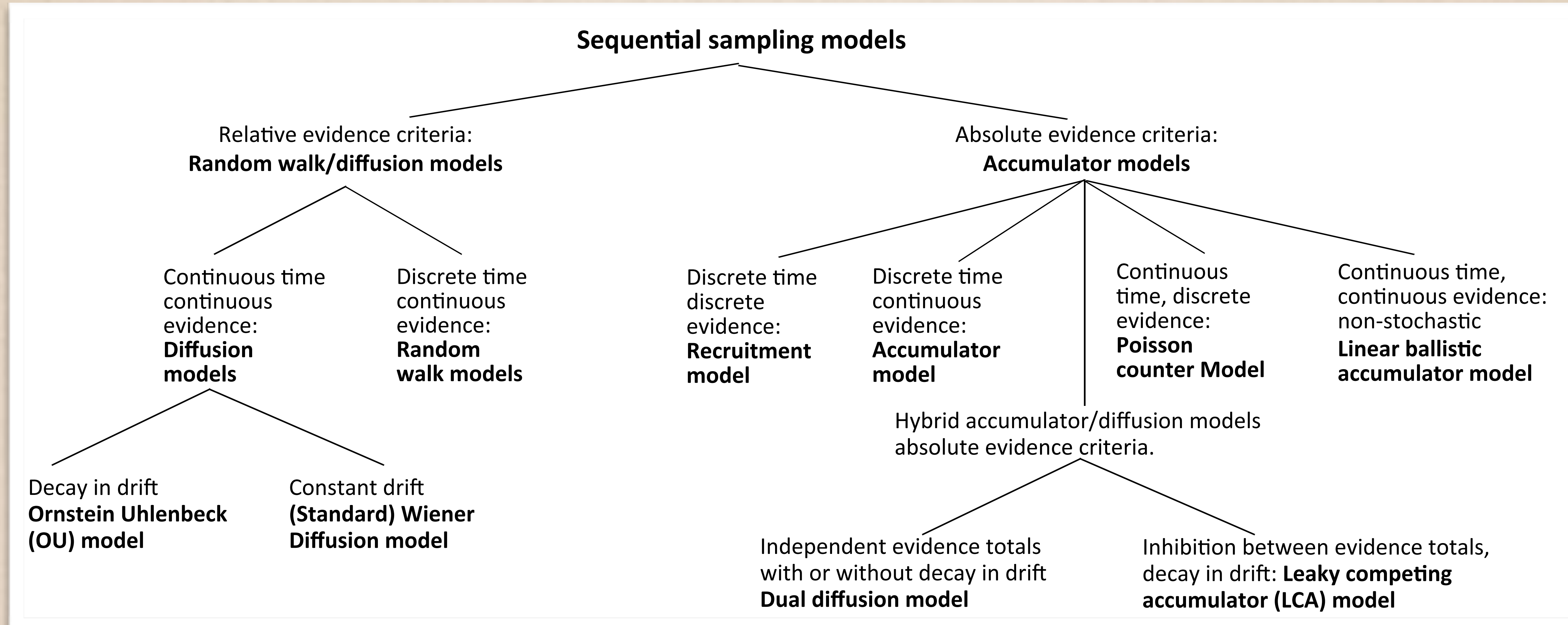
Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

- Time (continuous or discrete)
- Momentary evidence (continuous or discrete)
- Perfect or “leaky” accumulation

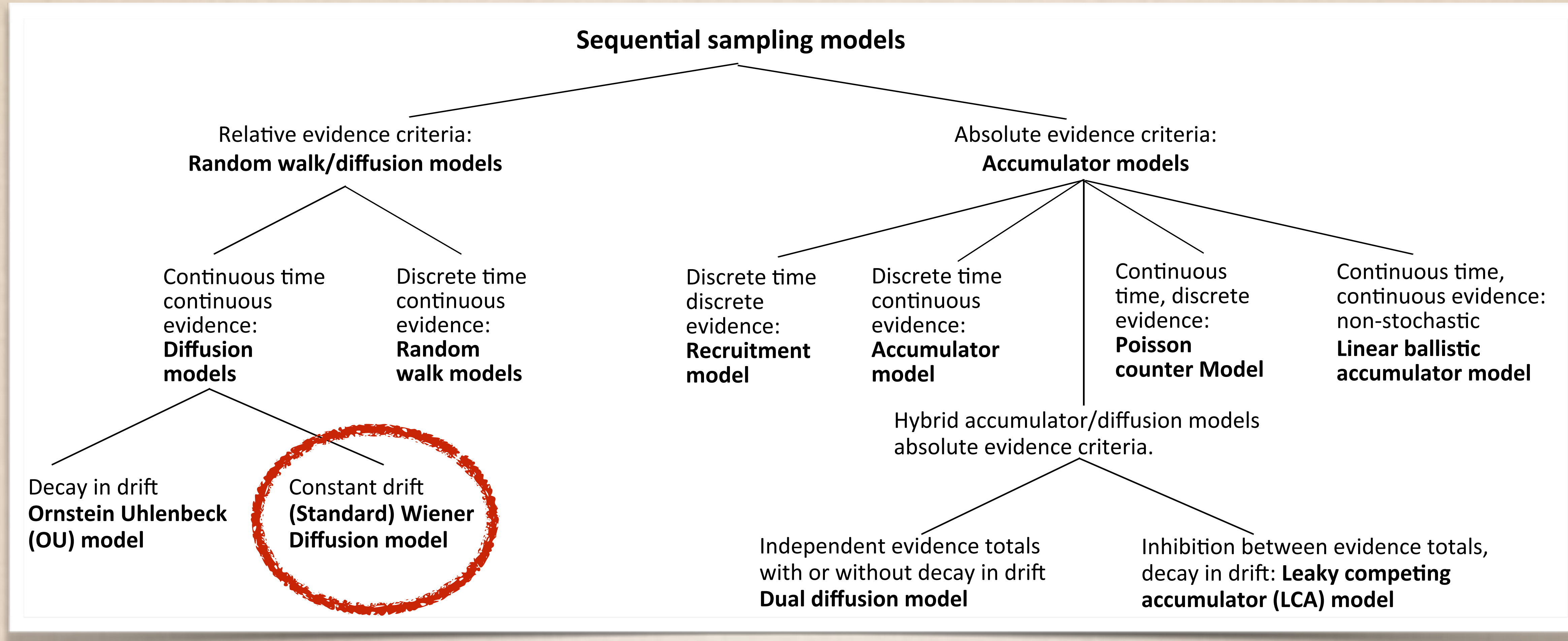
Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

- Time (continuous or discrete)
- Momentary evidence (continuous or discrete)
- Perfect or “leaky” accumulation
- Other factors (number of accumulation processes, with or without inhibition, …)

Many variants of evidence accumulation models for binary decisions



Ratcliff et al. TiCS 2016

- Time (continuous or discrete)
- Momentary evidence (continuous or discrete)
- Perfect or “leaky” accumulation
- Other factors (number of accumulation processes, with or without inhibition, …)

Drift diffusion model

Roger Ratcliff, 1978

Drift diffusion model

- Decision based on the accumulation of samples of evidence arriving sequentially over time

Roger Ratcliff, 1978

Drift diffusion model

» Decision based on the accumulation of samples of evidence arriving sequentially over time

decision variable

$$x_{i+1} = x_i + e_i$$

evidence

i: iteration

Roger Ratcliff, 1978

Drift diffusion model

» Decision based on the accumulation of samples of evidence arriving sequentially over time

decision variable

$$x_{i+1} = x_i + e_i$$

i: iteration

evidence

» Evidence samples e_i are normally distributed

time step

$$e_i \sim \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

drift coefficient

diffusion coefficient

Roger Ratcliff, 1978

Drift diffusion model

» Decision based on the accumulation of samples of evidence arriving sequentially over time

decision variable

$$x_{i+1} = x_i + e_i$$

evidence

i: iteration

» Evidence samples e_i are normally distributed

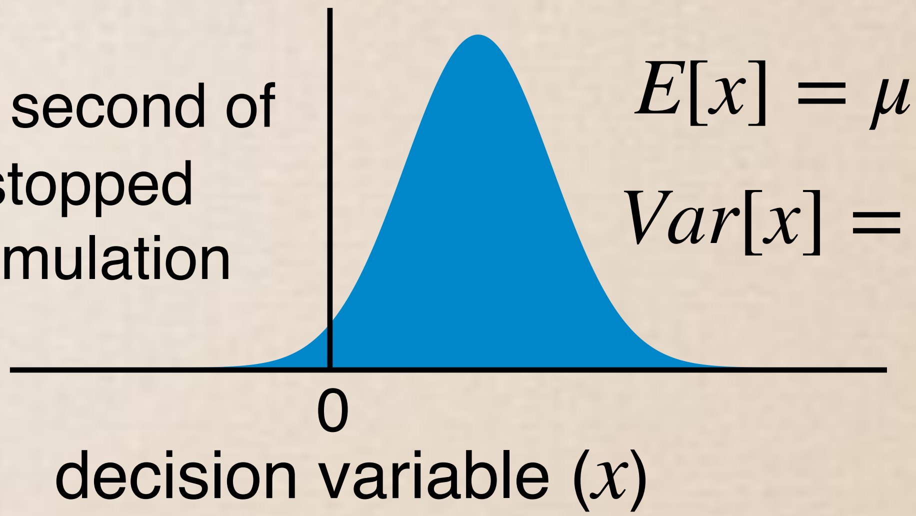
time step

$$e_i \sim \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t)$$

drift coefficient

diffusion coefficient

After t second of unstopped accumulation



$$E[x] = \mu t$$

$$Var[x] = \sigma^2 t$$

Roger Ratcliff, 1978

Drift diffusion model

» Decision based on the accumulation of samples of evidence arriving sequentially over time

decision variable

$$x_{i+1} = x_i + e_i$$

evidence

i: iteration

» Evidence samples e_i are normally distributed

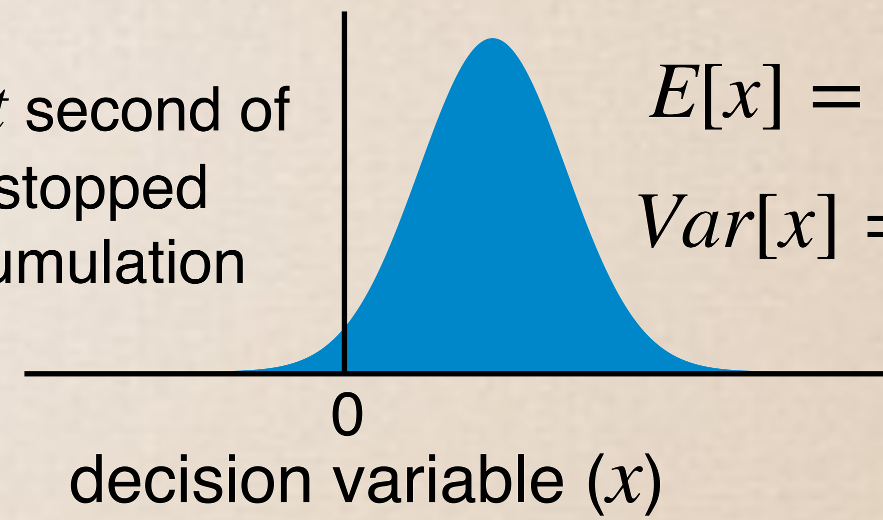
time step

$$e_i \sim \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t)$$

drift coefficient

diffusion coefficient

After t second of unstopped accumulation



$E[x] = \mu t$

$Var[x] = \sigma^2 t$

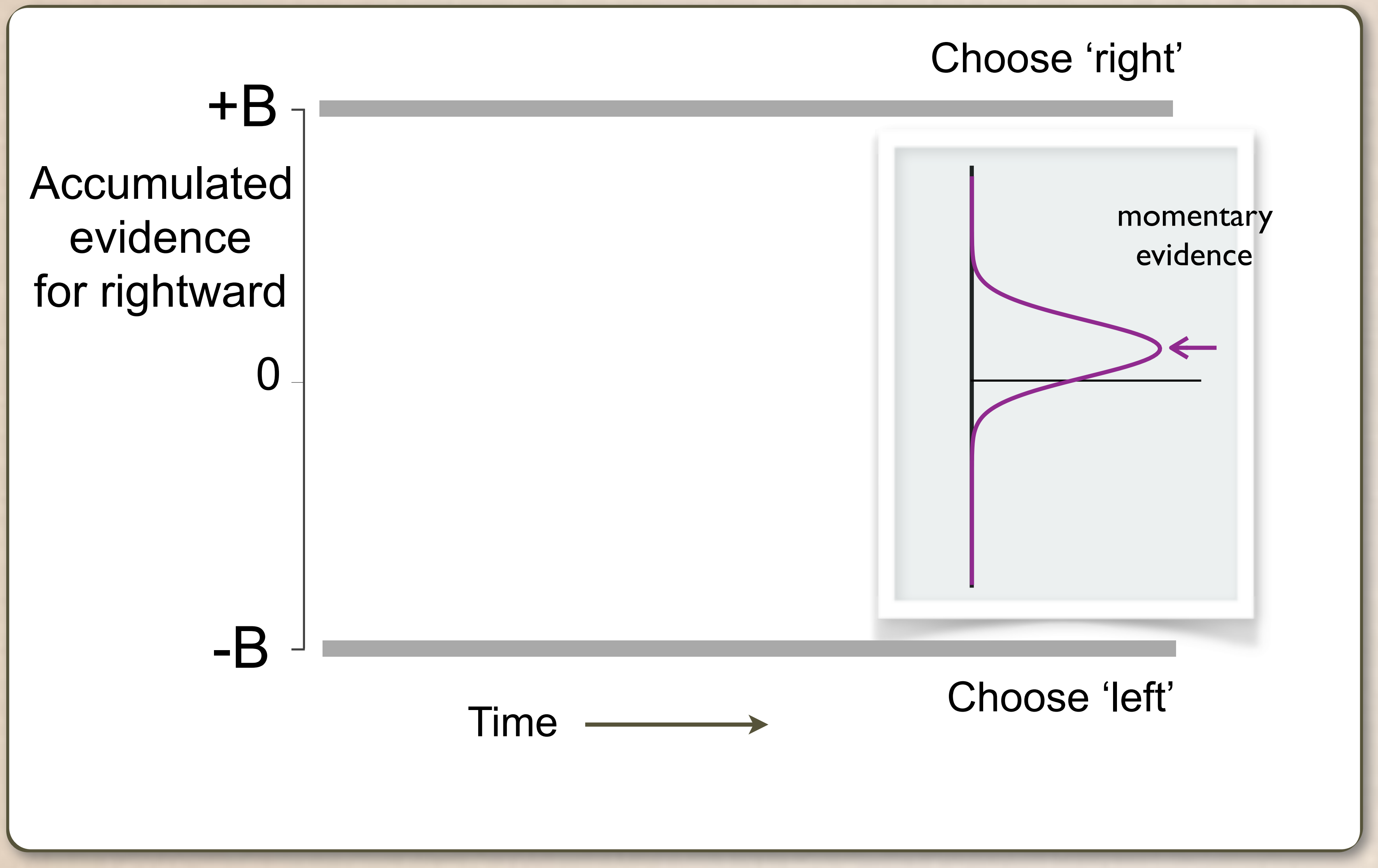
» In the limit $\Delta t \to 0$, the decision process is described by a Wiener process with drift:

decision variable

$$\frac{dx(t)}{dt} = \underbrace{\mu}_{\text{signal}} + \underbrace{\sigma W(t)}_{\text{noise}}$$

Roger Ratcliff, 1978

Drift diffusion model



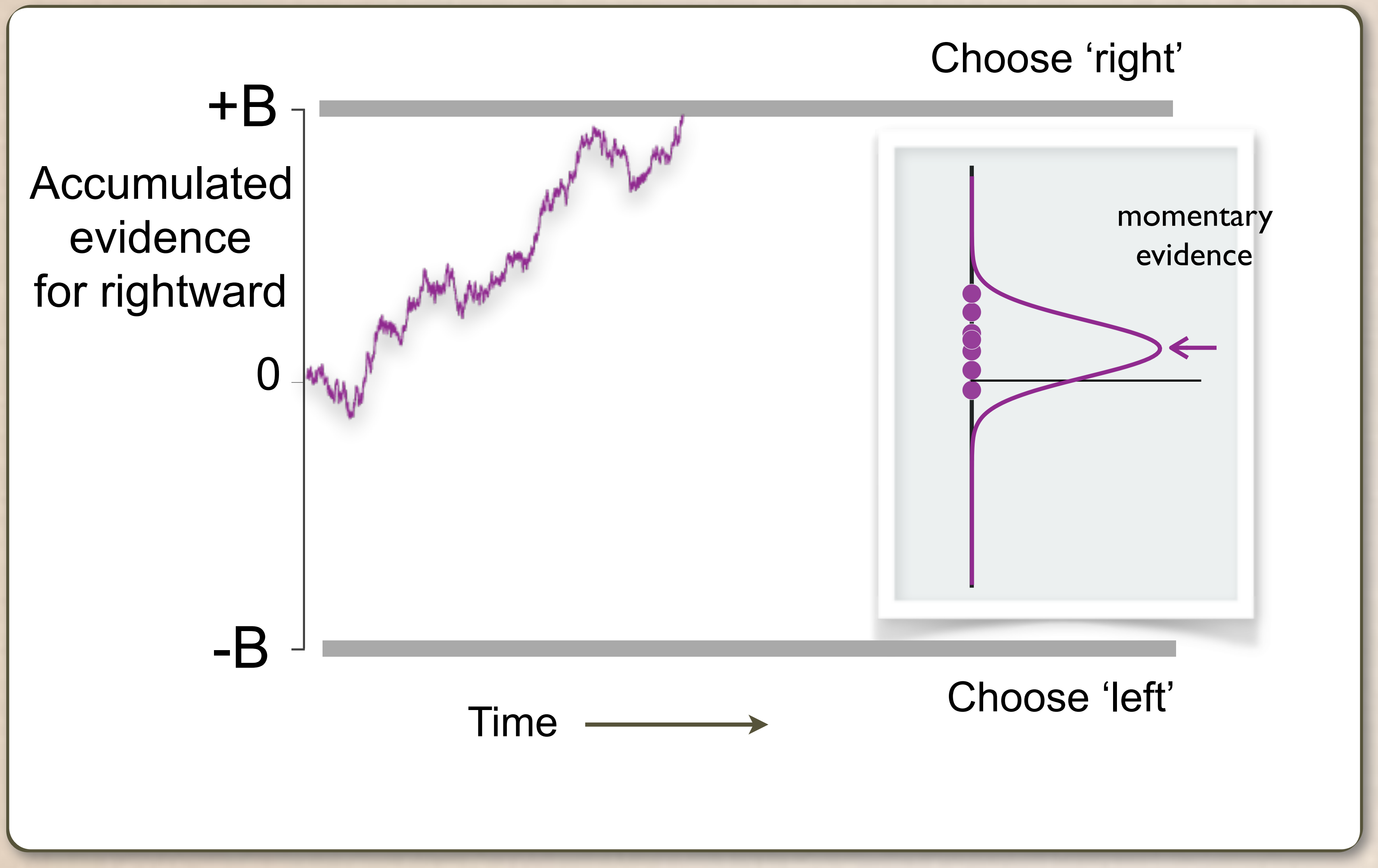
Sample paths:

$$e_i = \mathcal{N}(\mu \Delta t, \sigma^2 \Delta t)$$

$$x_{i+1} = x_i + e_i$$

$$x_0 = 0$$

Drift diffusion model



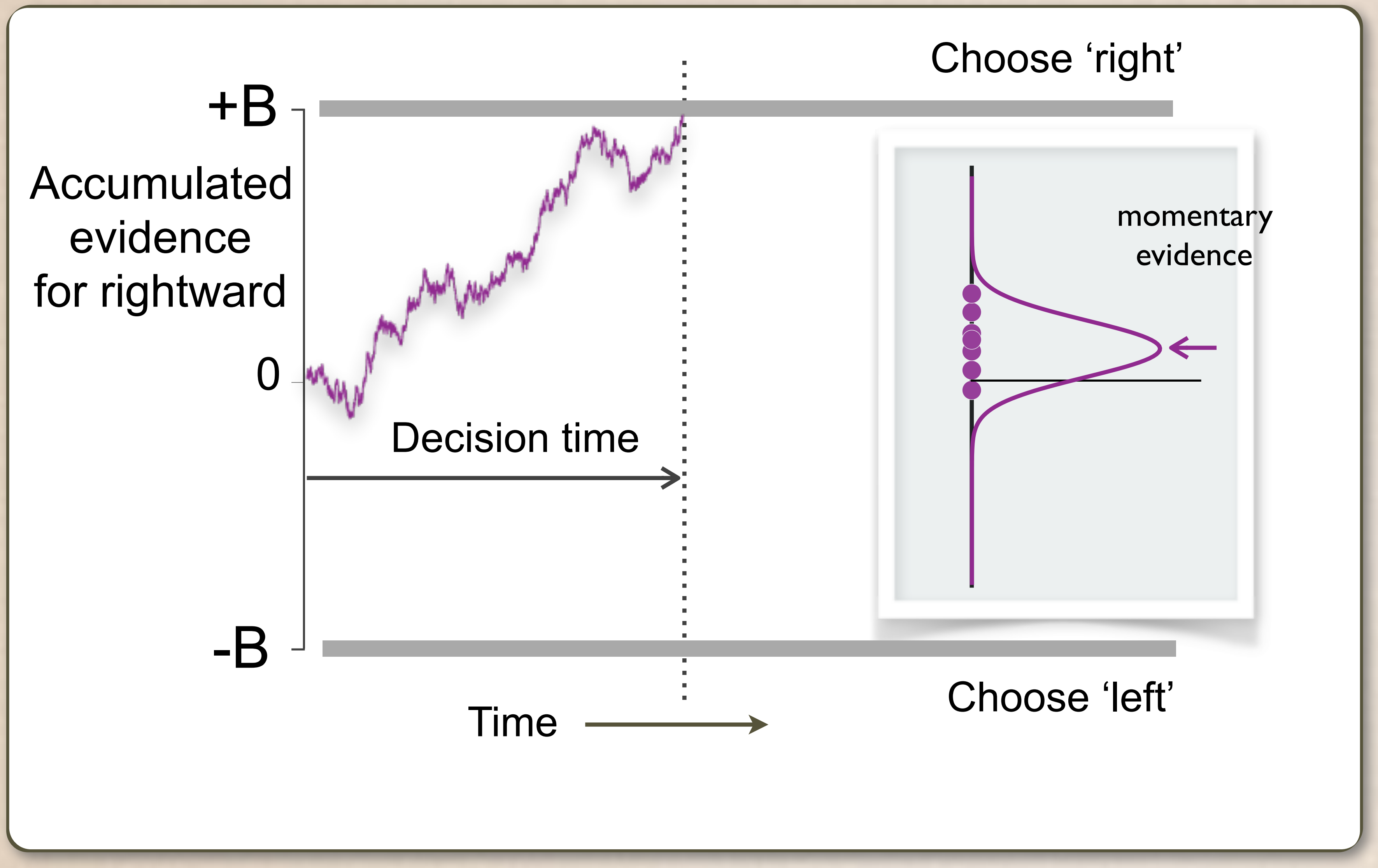
Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

$$x_0 = 0$$

Drift diffusion model



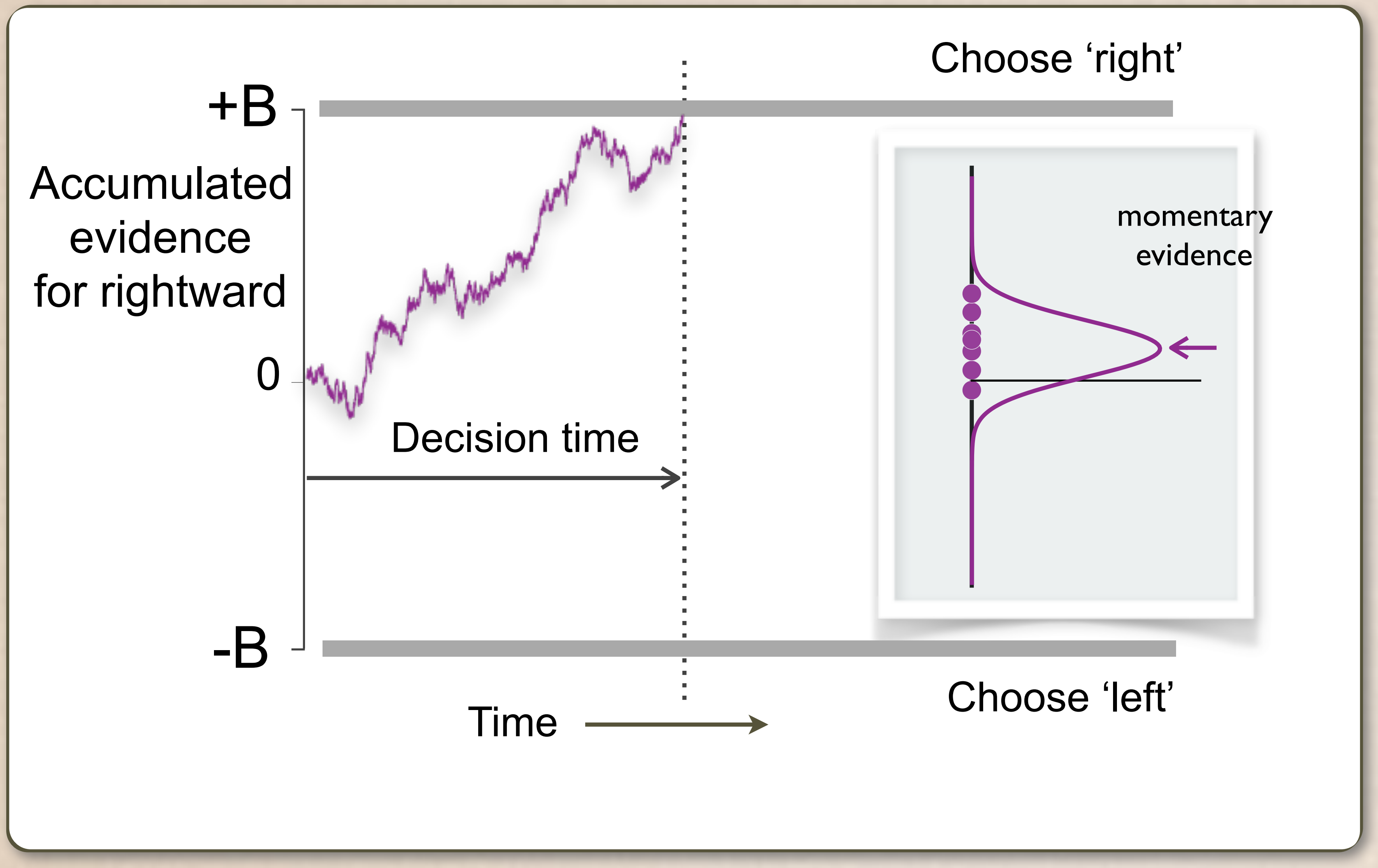
Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

$$x_0 = 0$$

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

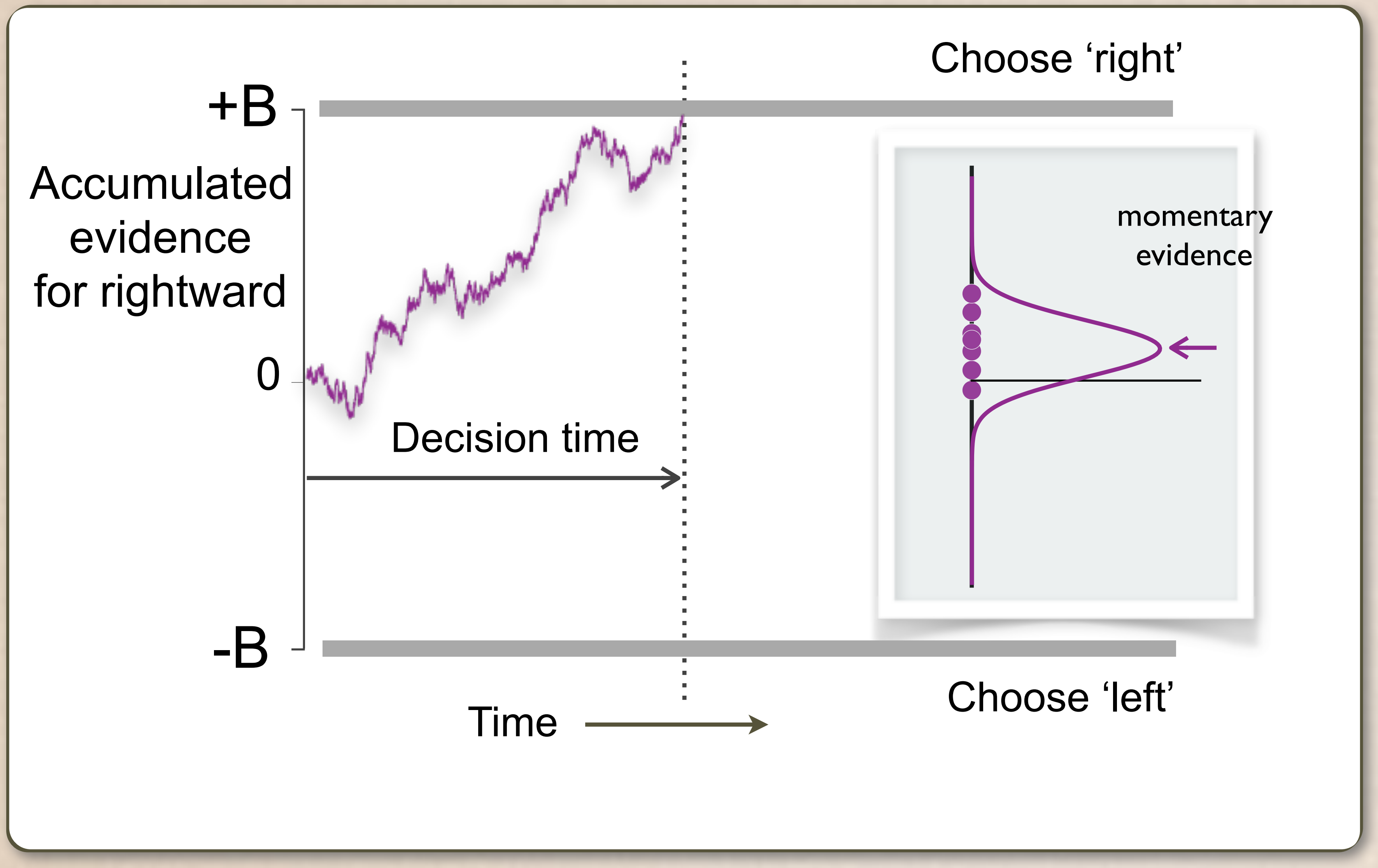
$$x_{i+1} = x_i + e_i$$

$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

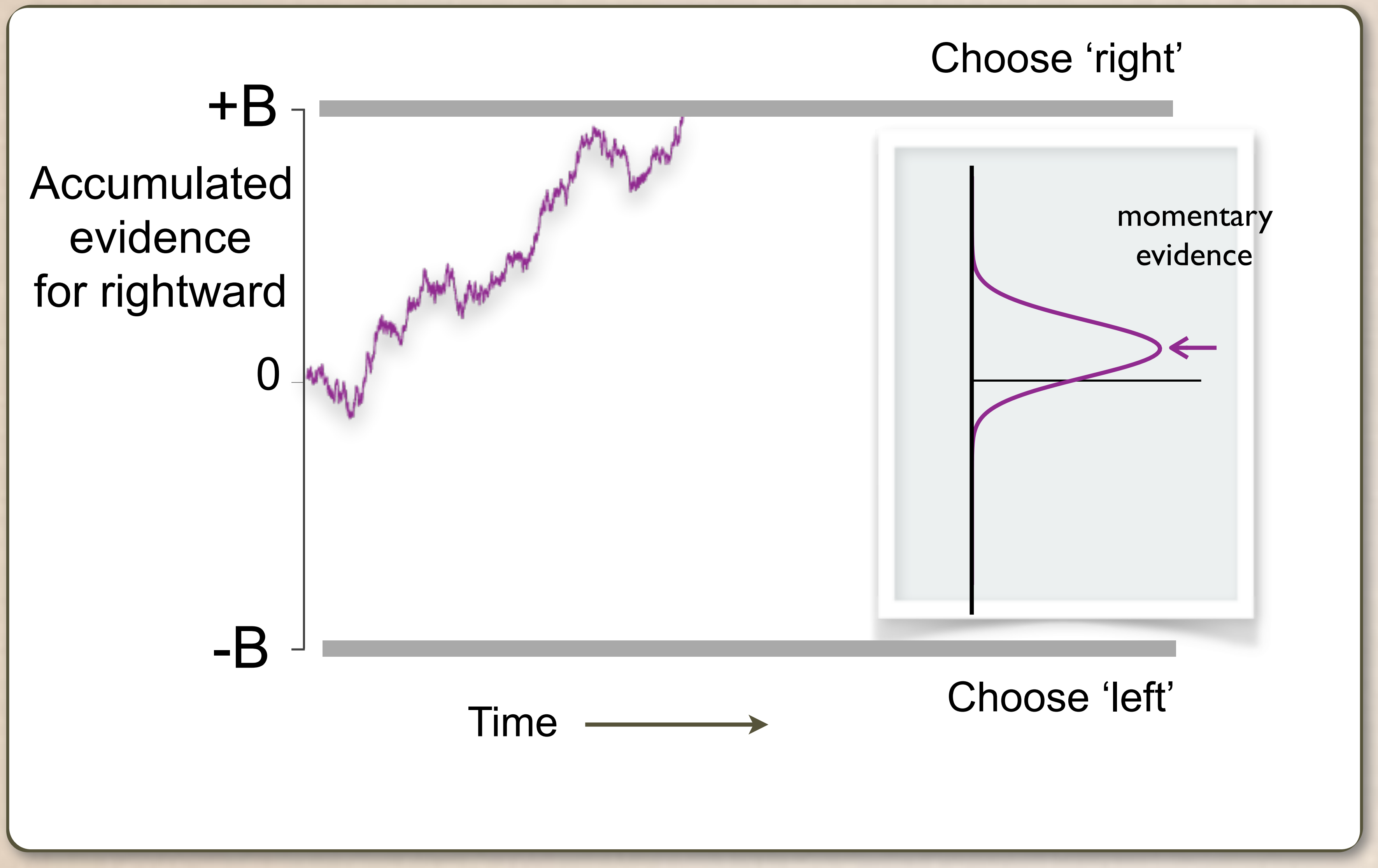
$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

$|\mu|$ increases with motion strength

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

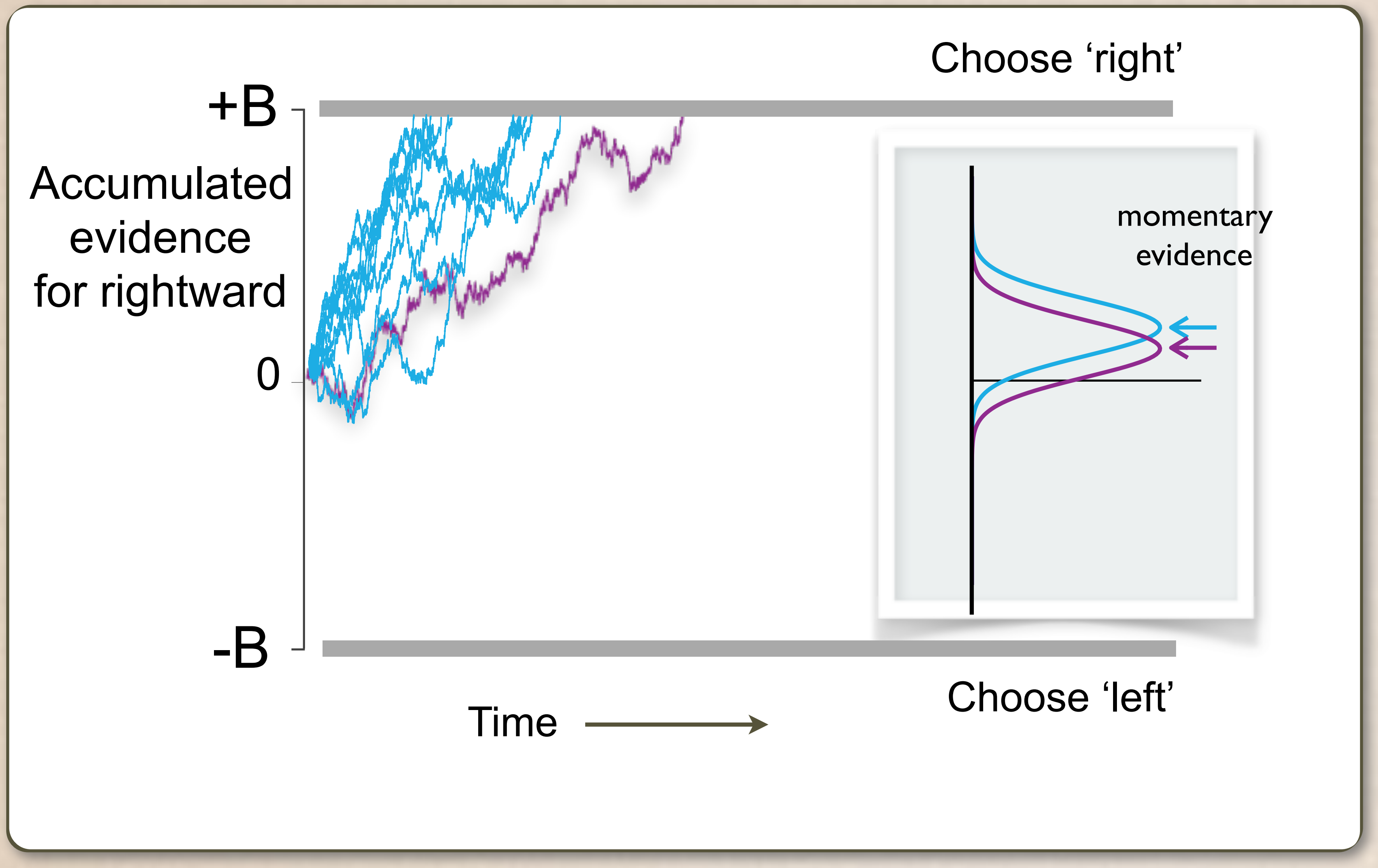
$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

$|\mu|$ Increases with motion strength

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

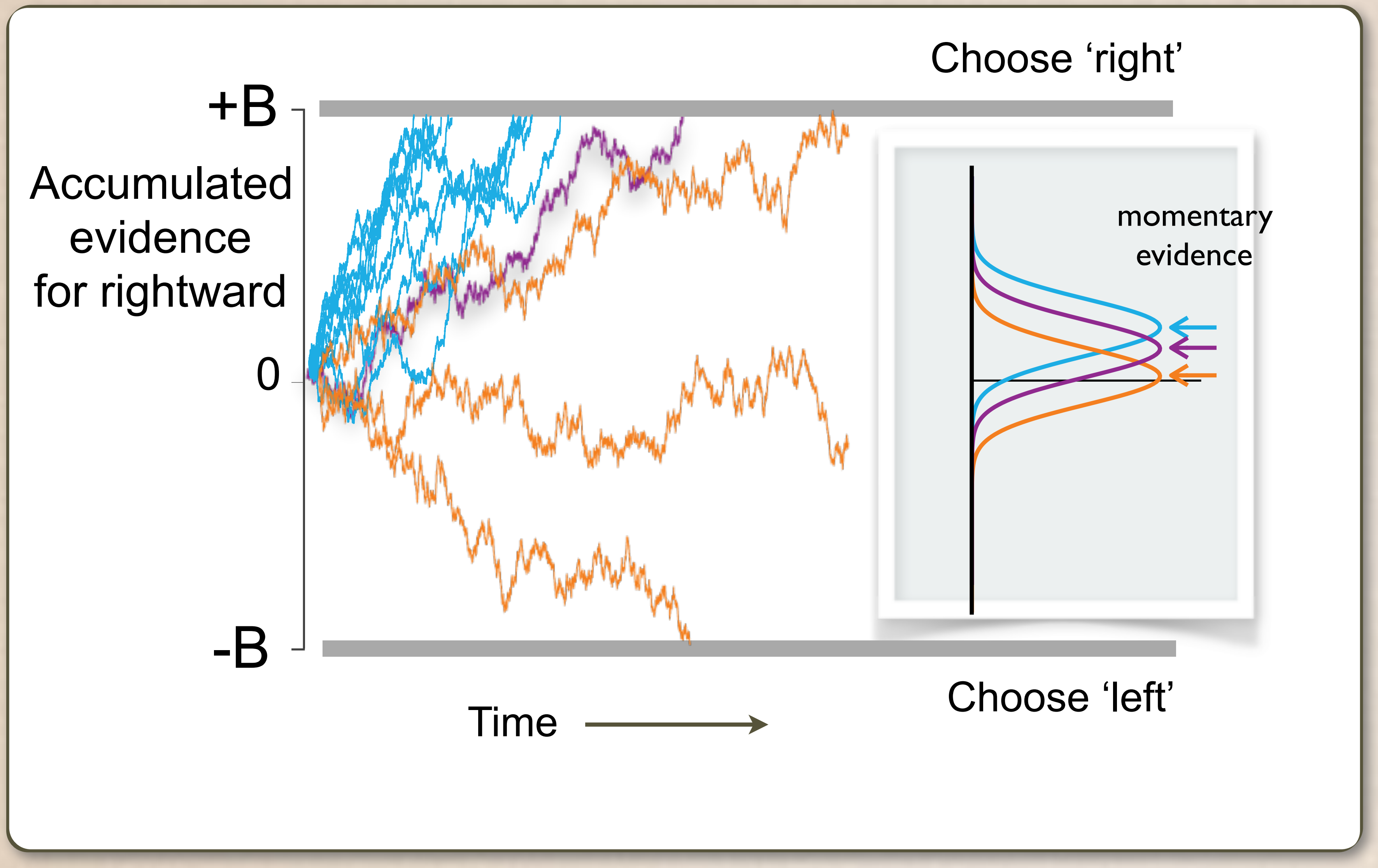
$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

$|\mu|$ Increases with motion strength

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

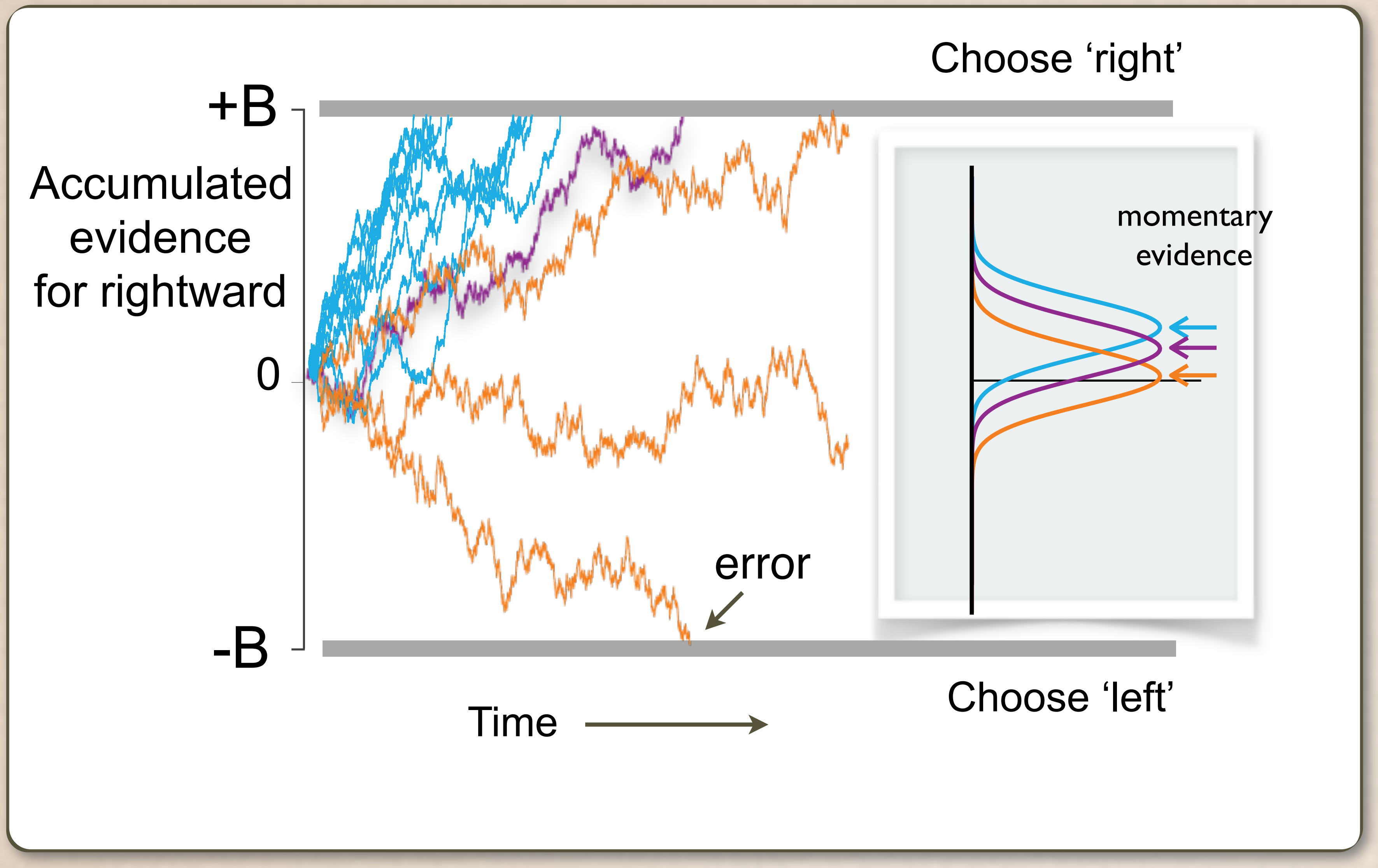
$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

$|\mu|$ Increases with motion strength

Drift diffusion model



Sample paths:

$$e_i = \mathcal{N}(\mu\Delta t, \sigma^2\Delta t)$$

$$x_{i+1} = x_i + e_i$$

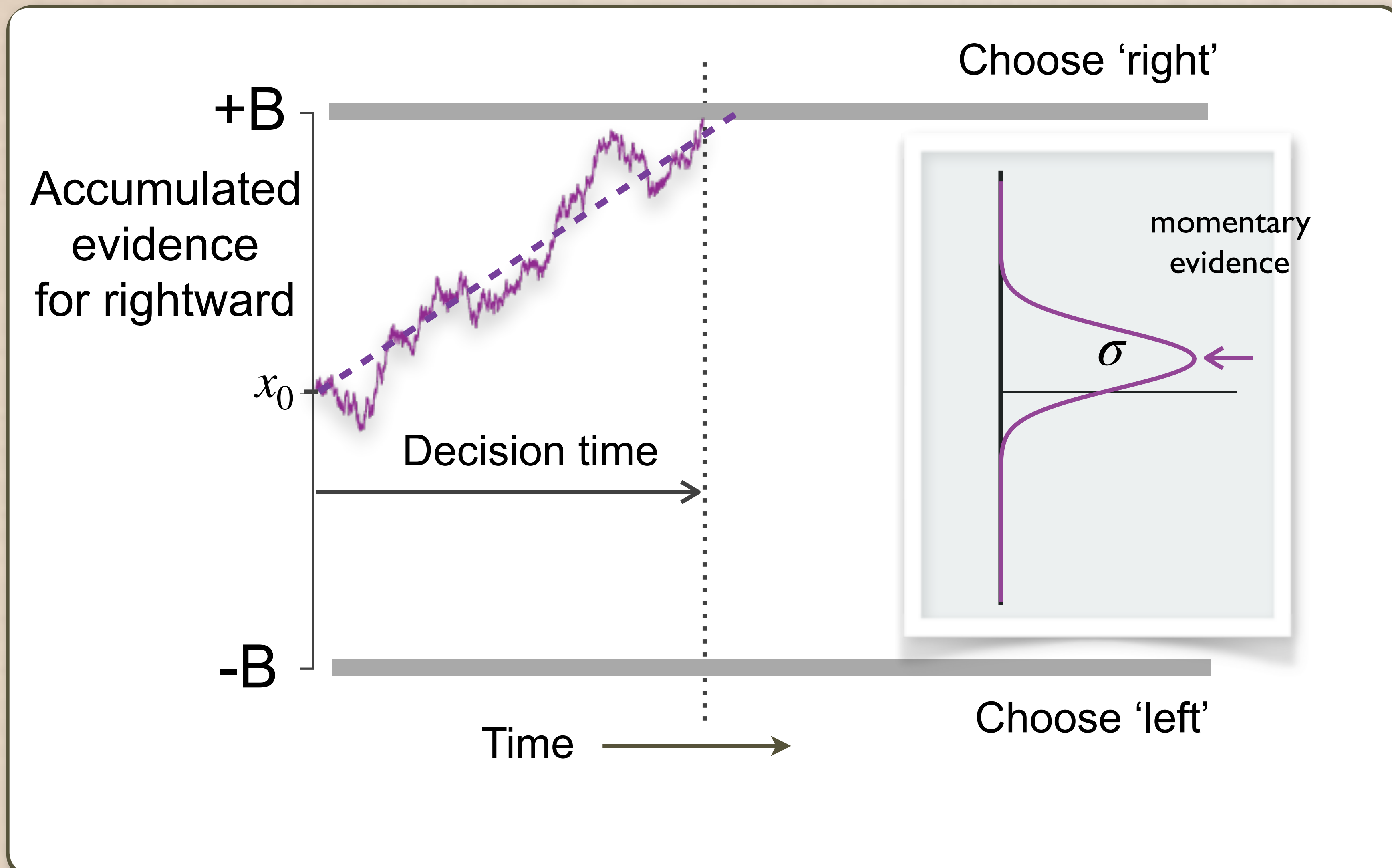
$$x_0 = 0$$

Motion task:

$\mu = \kappa \times$ Motion coherence

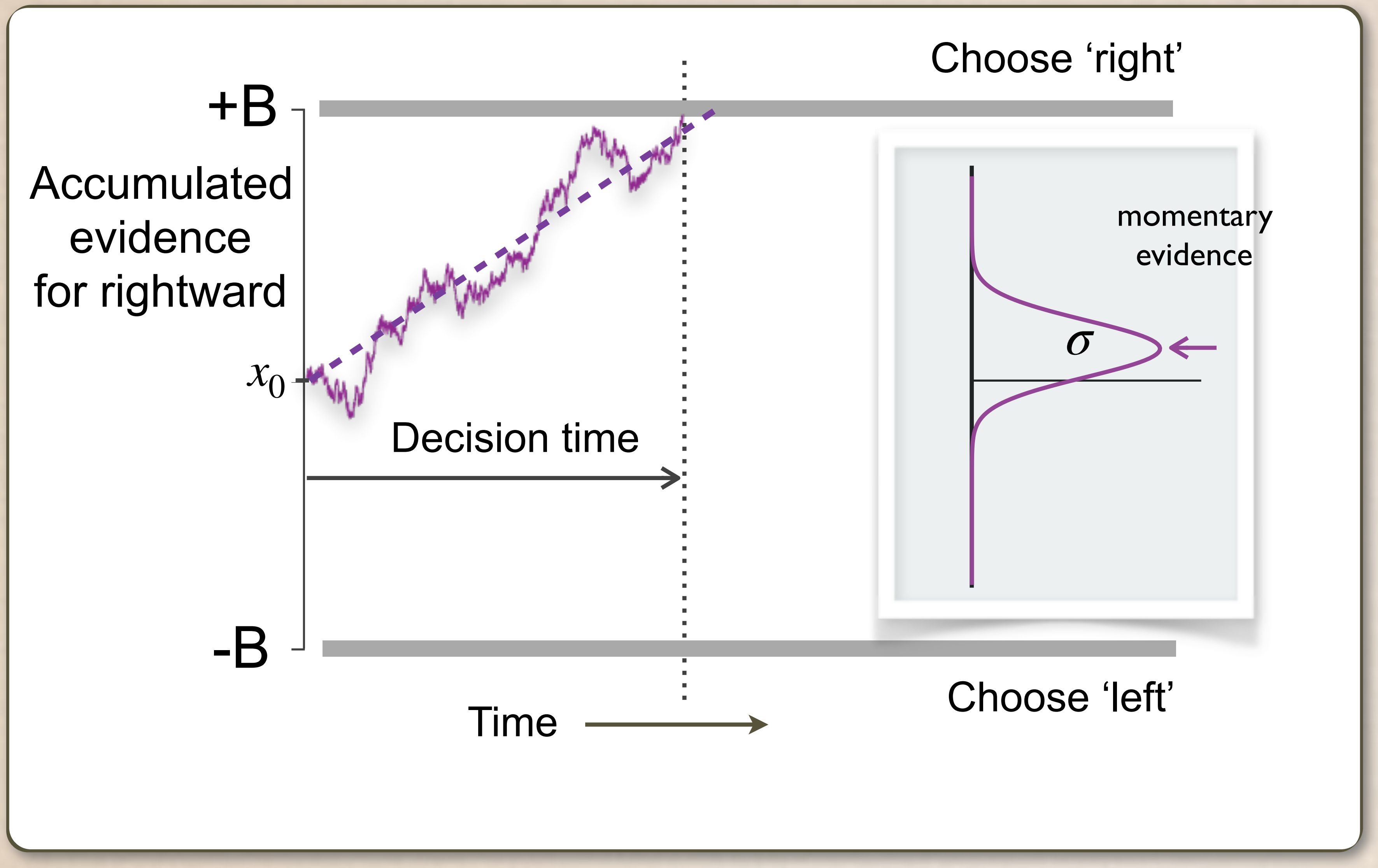
$|\mu|$ Increases with motion strength

Drift diffusion model



Parameters in the 'simple' version:

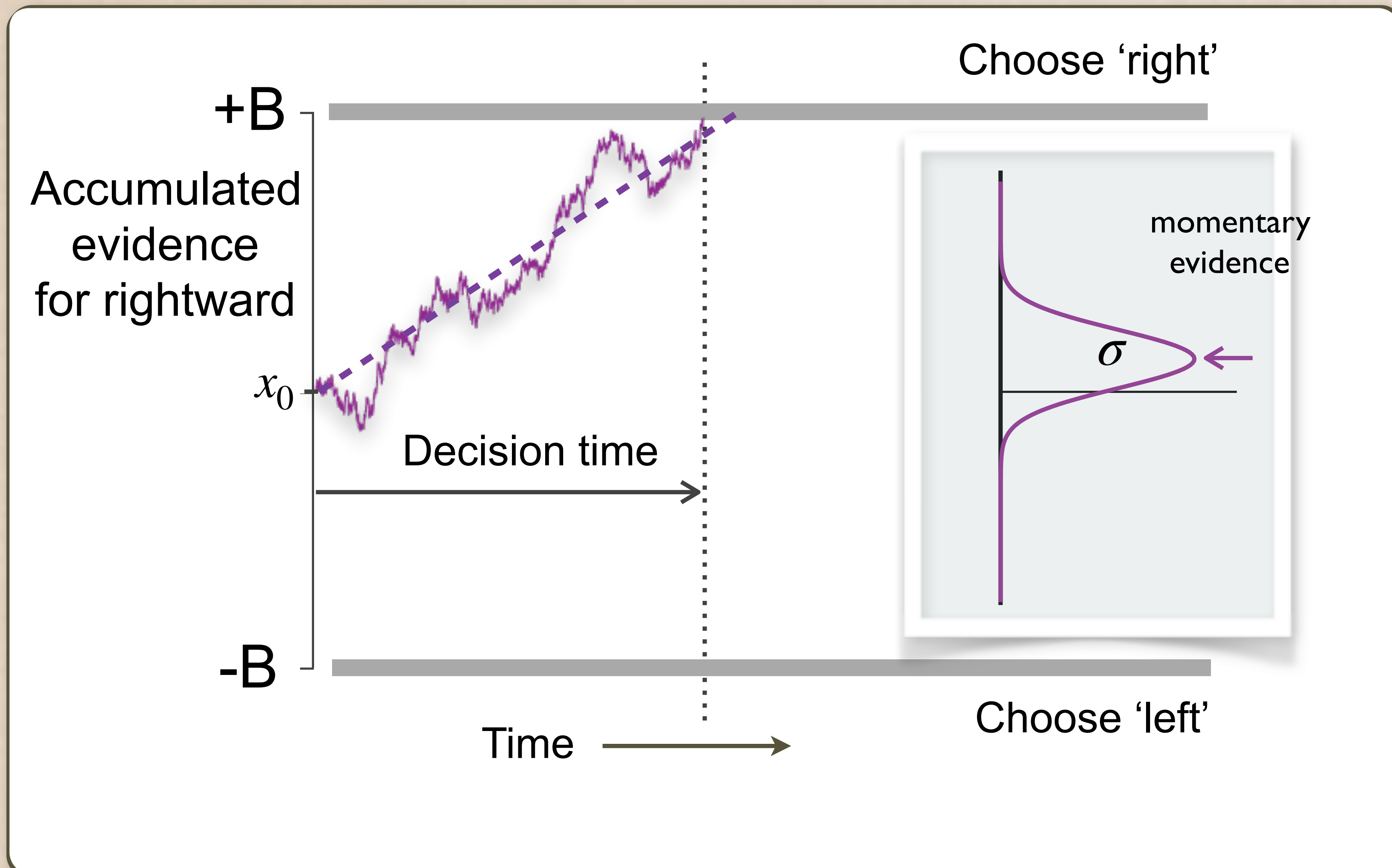
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ

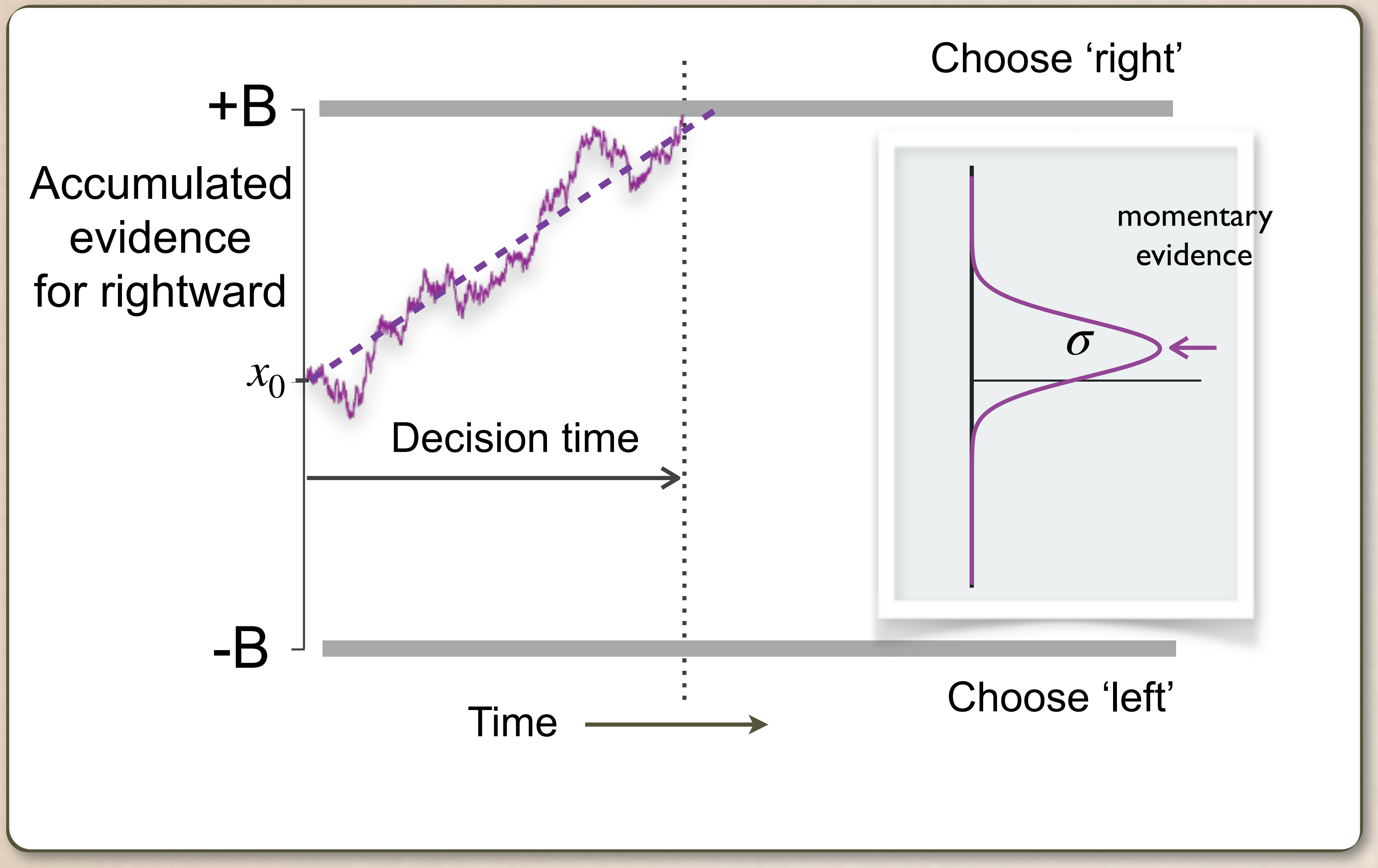
Drift diffusion model



Parameters in the 'simple' version:

» Drift rate, μ

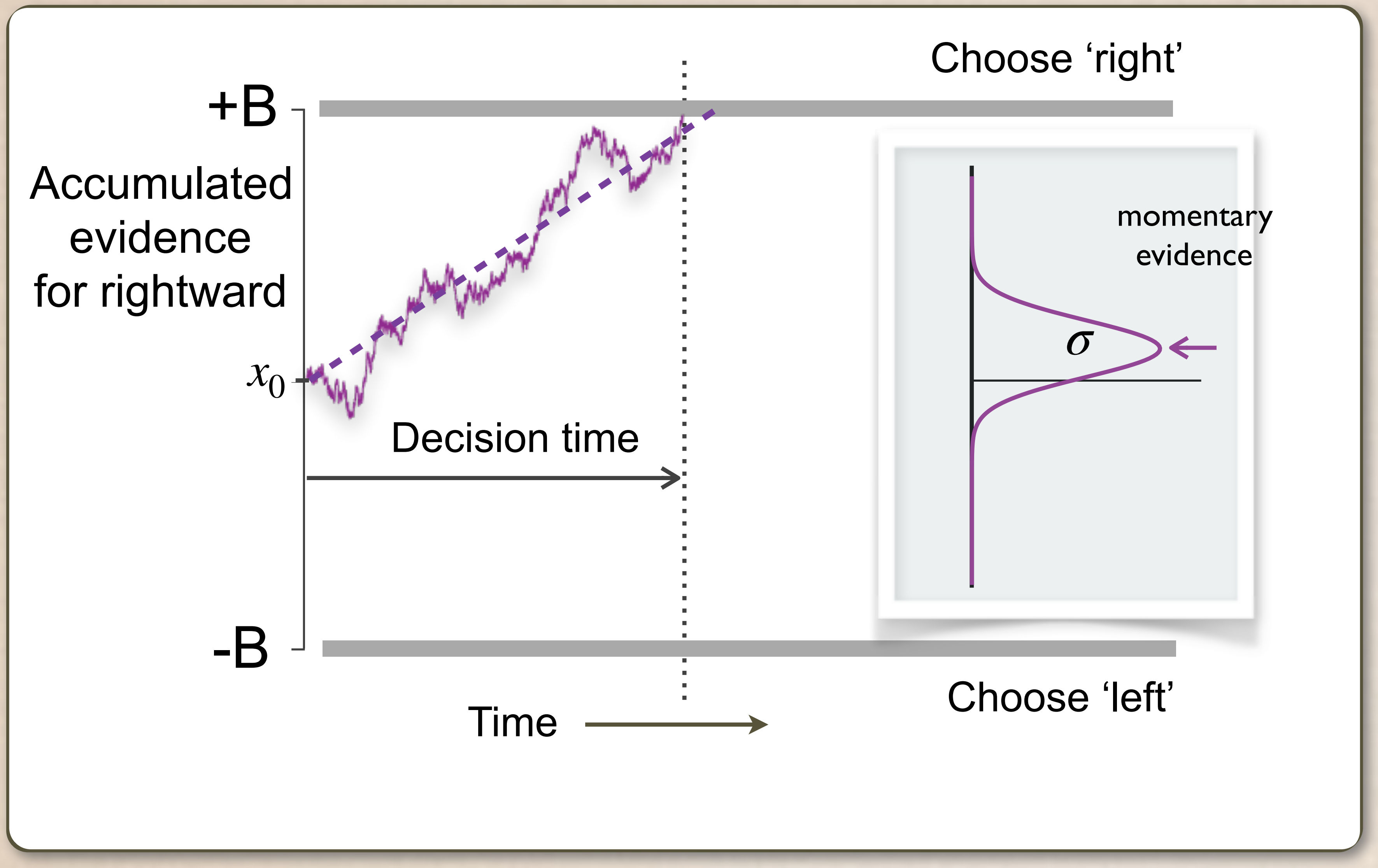
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B

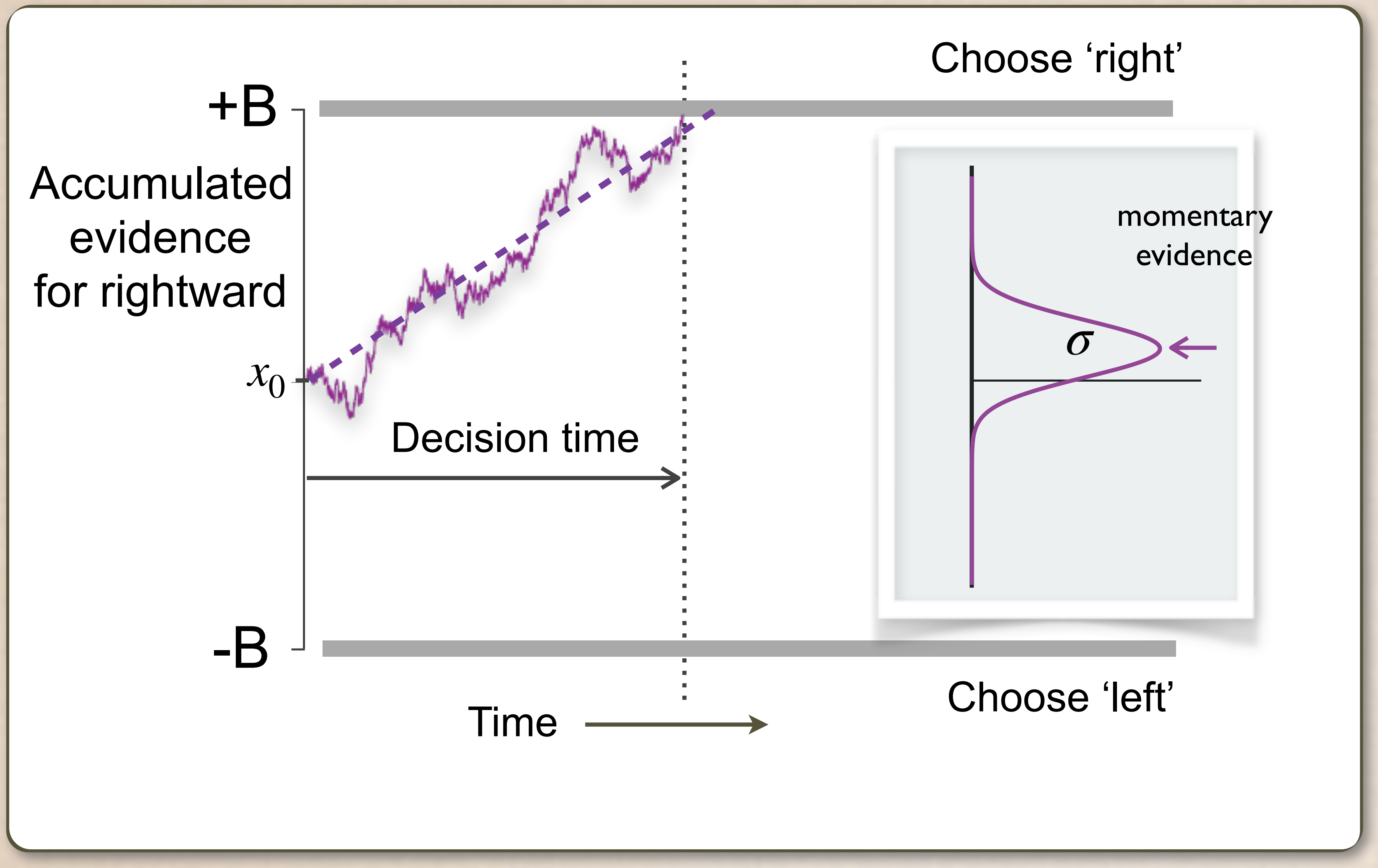
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B

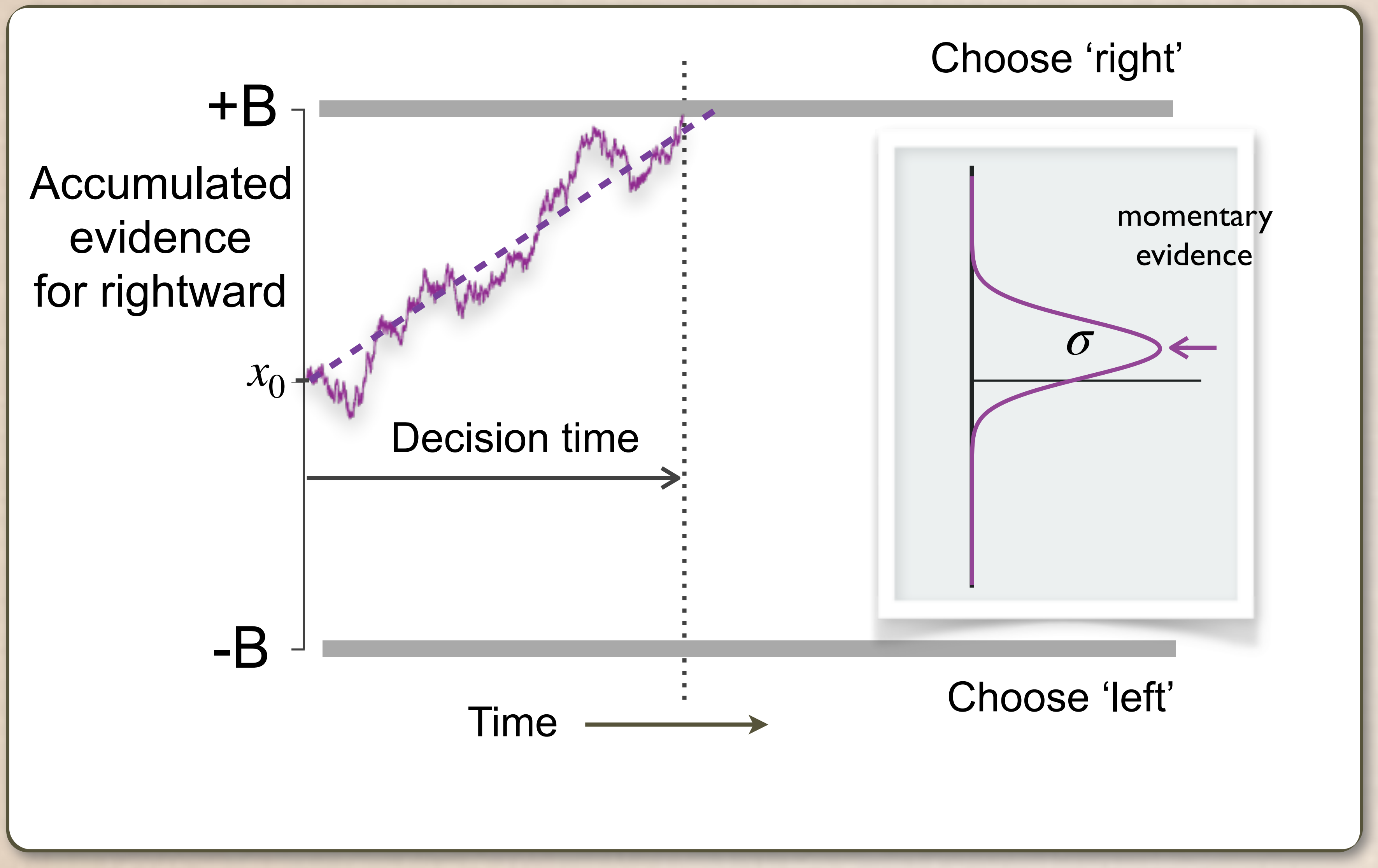
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0

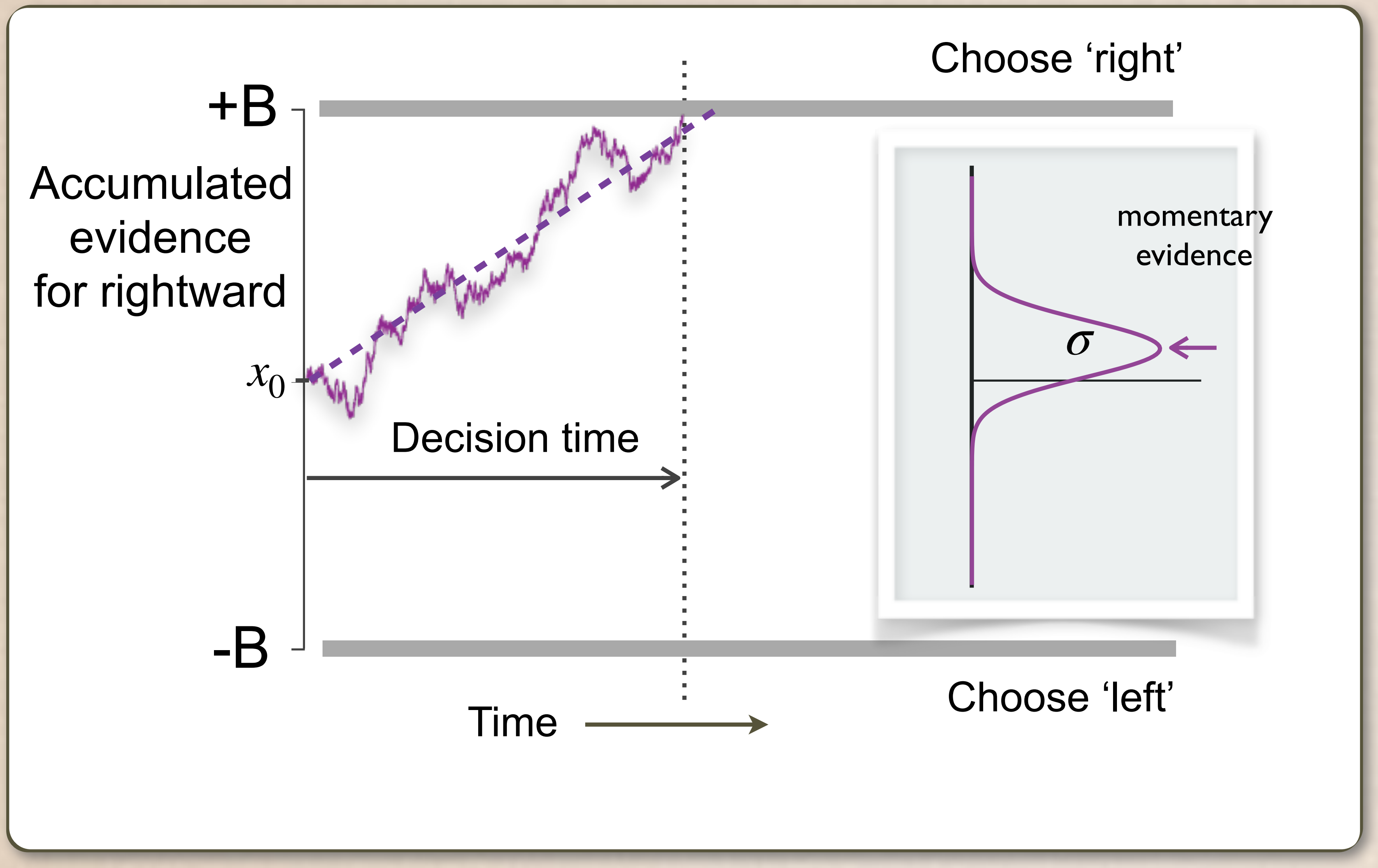
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0

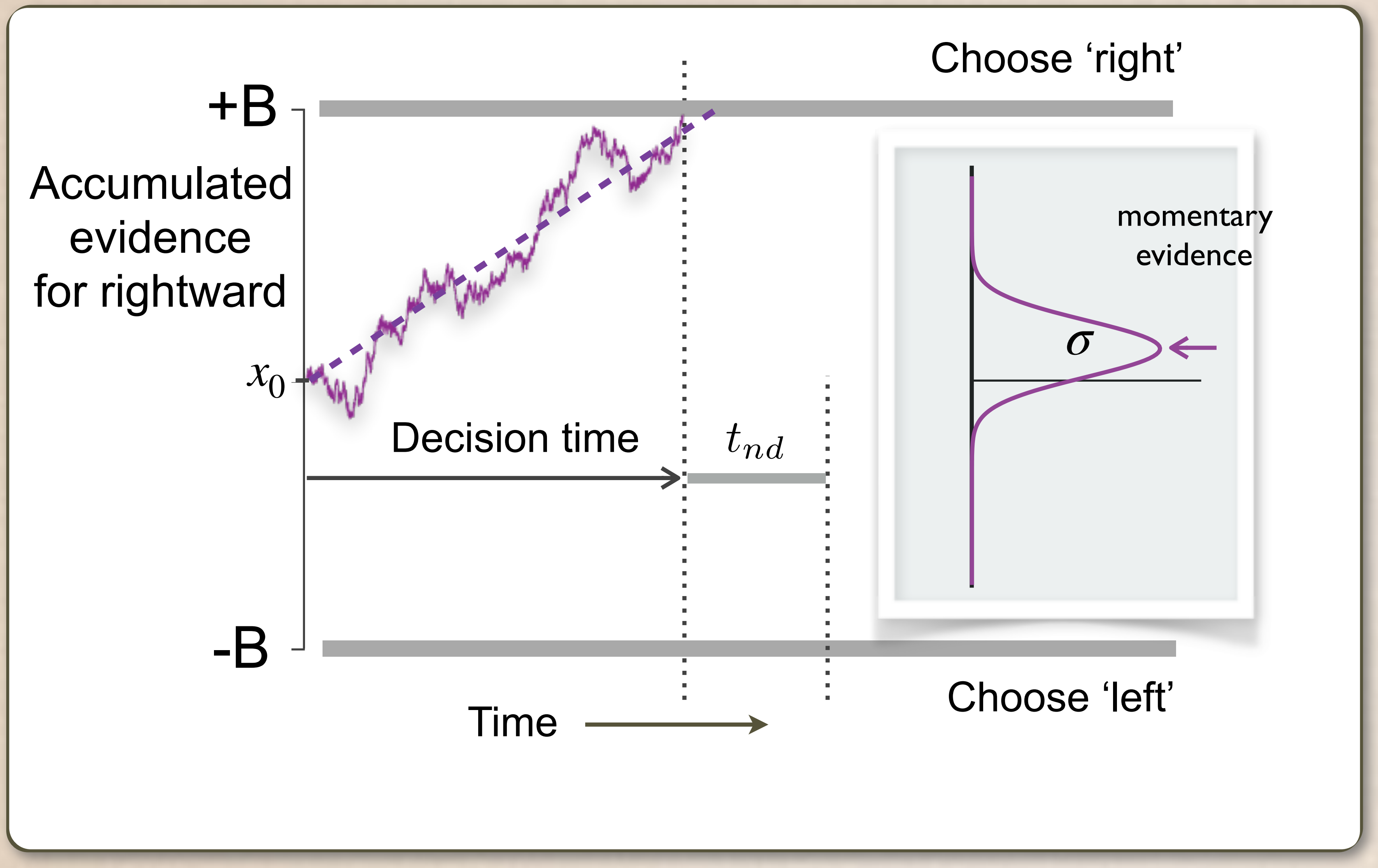
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}

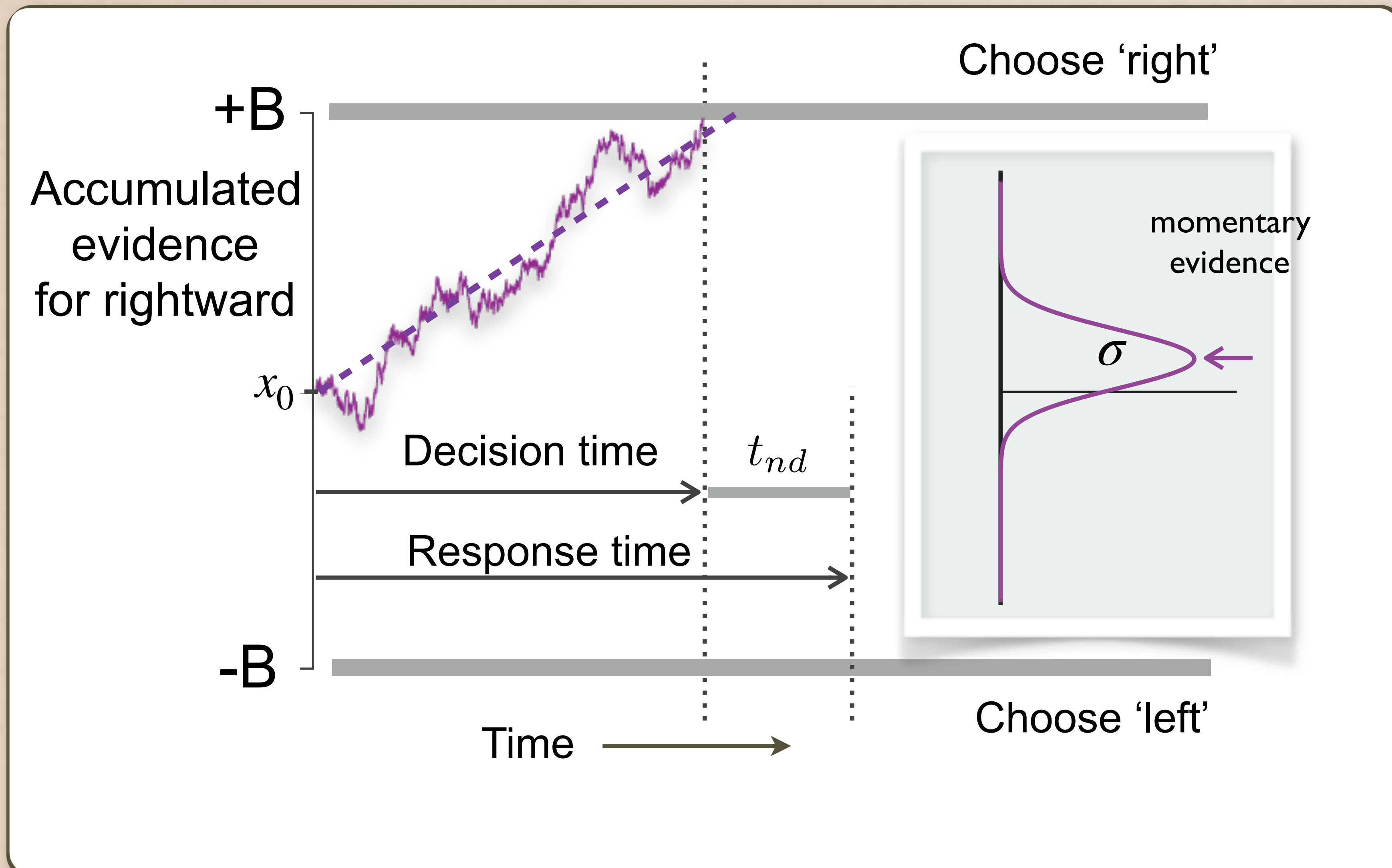
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}

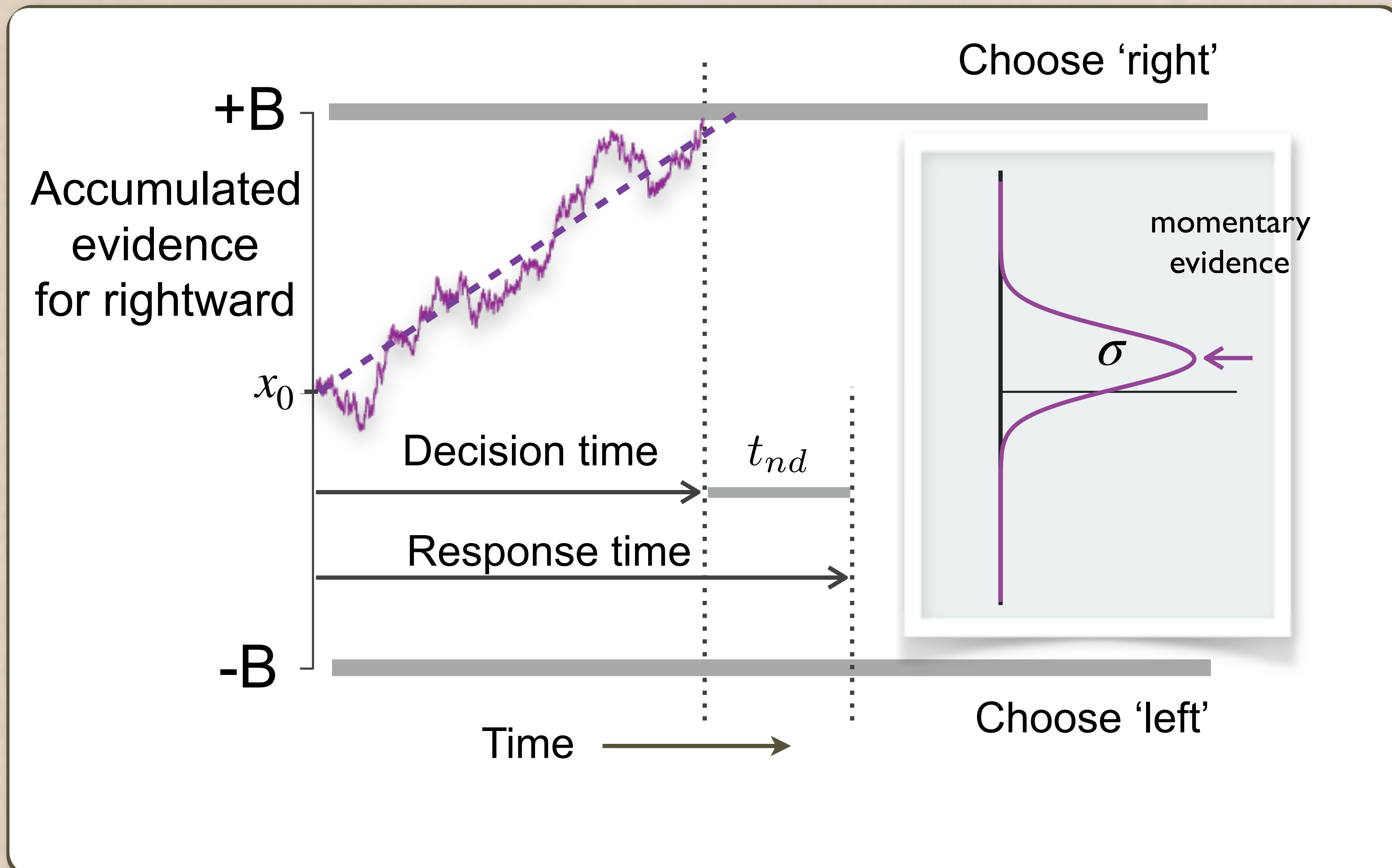
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}

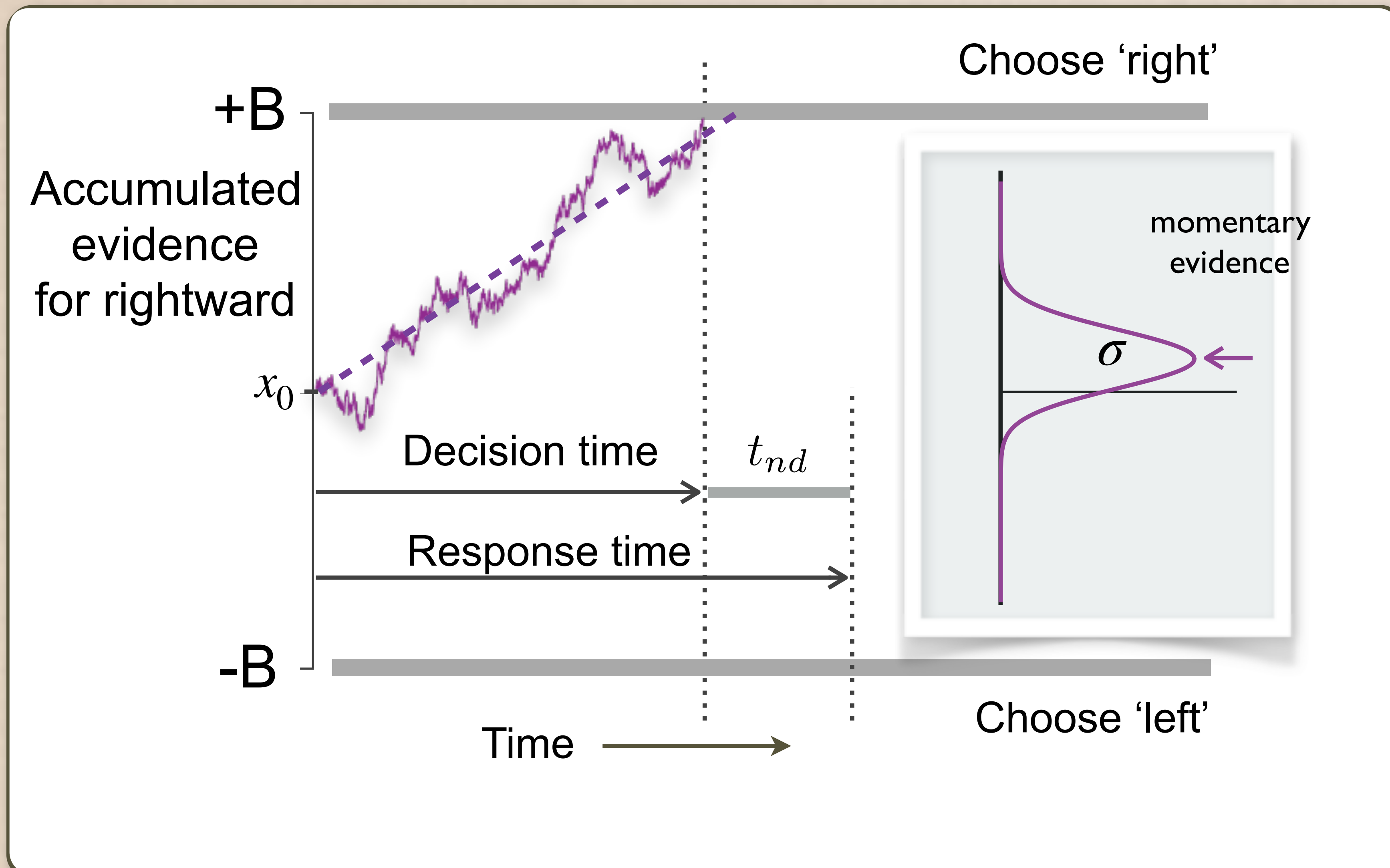
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}
- Diffusion coefficient, σ

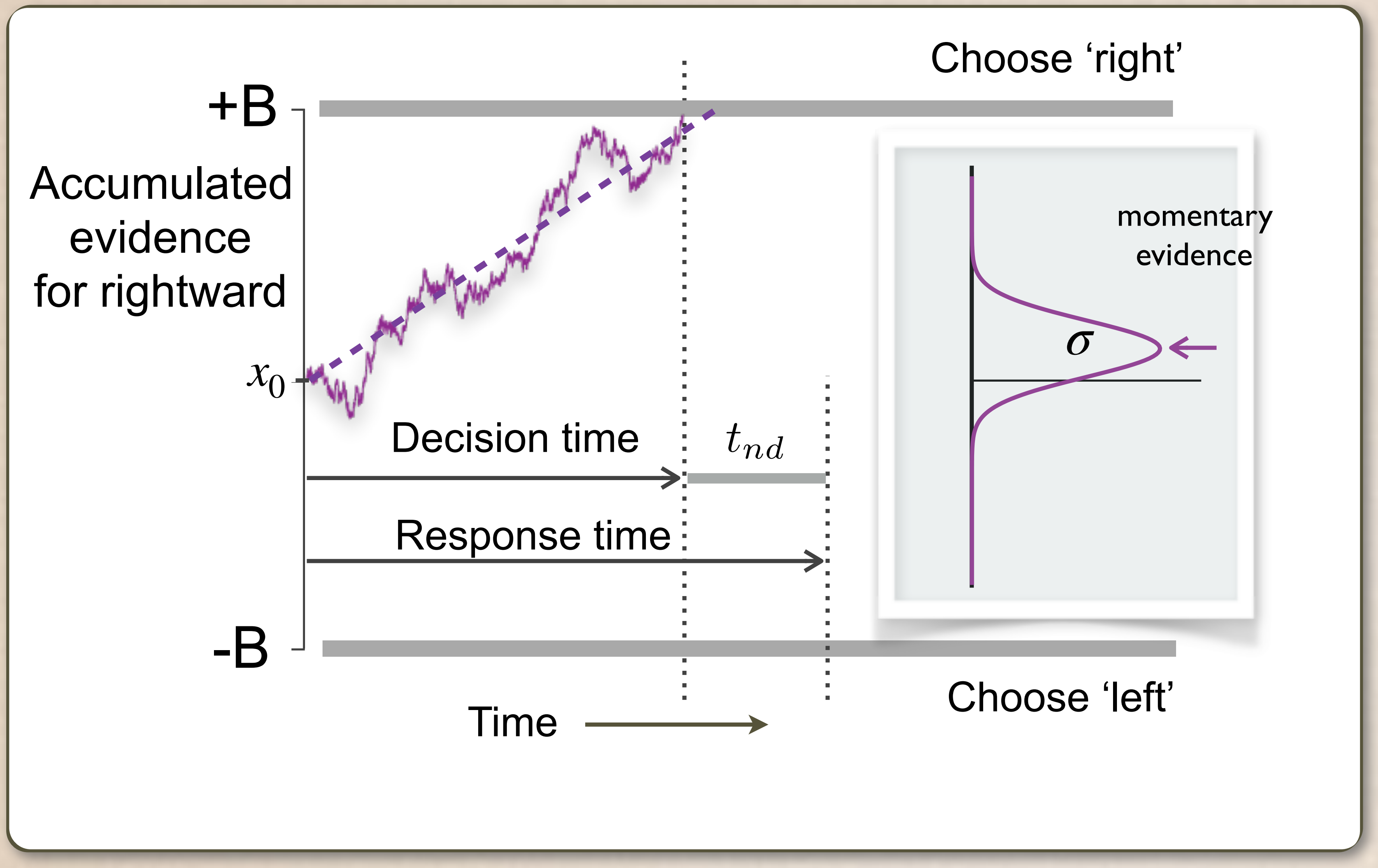
Drift diffusion model



Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}
- Diffusion coefficient, σ

Drift diffusion model

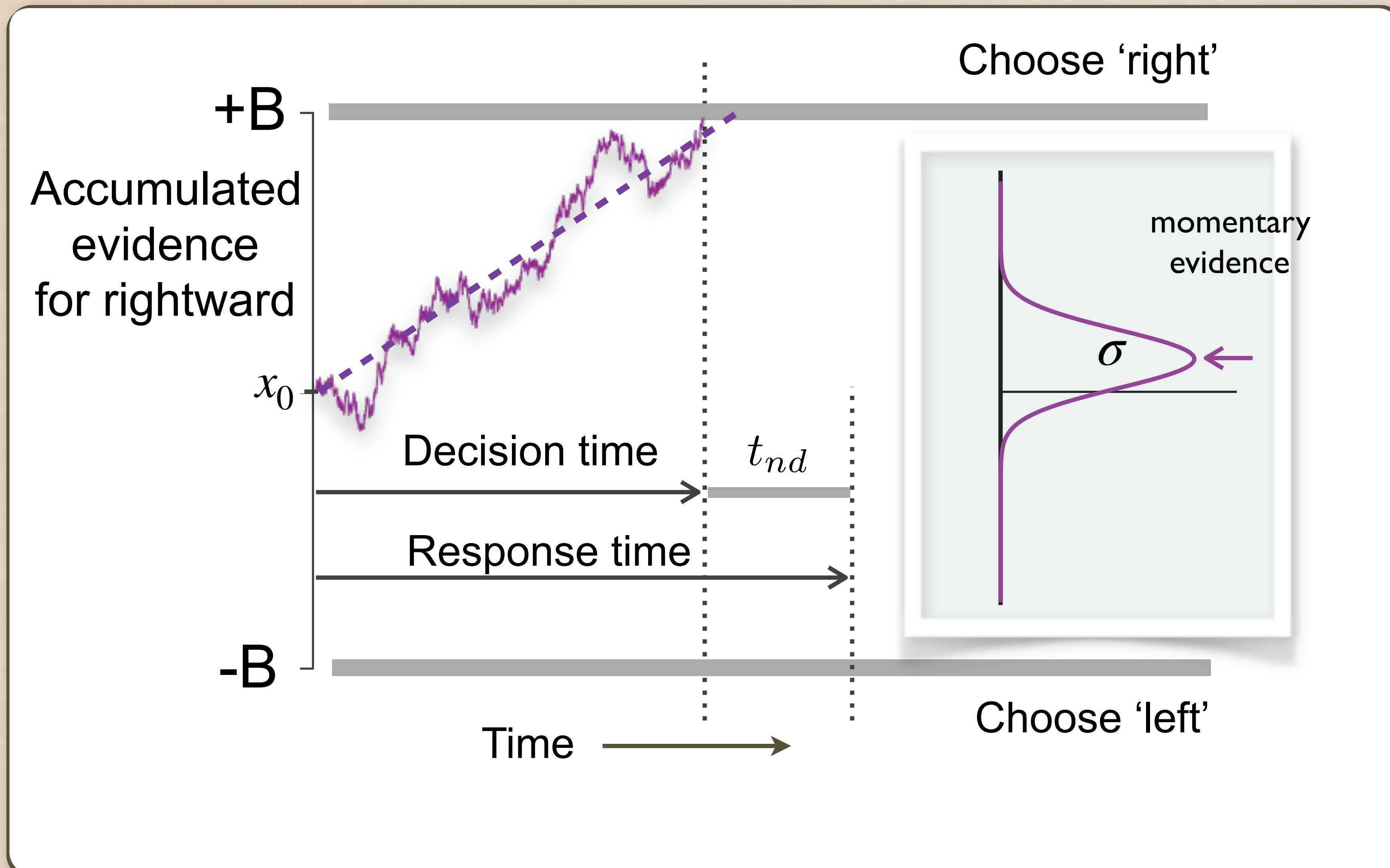


Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}
- ~~Diffusion coefficient, σ~~

$$\sigma = 1$$

Drift diffusion model

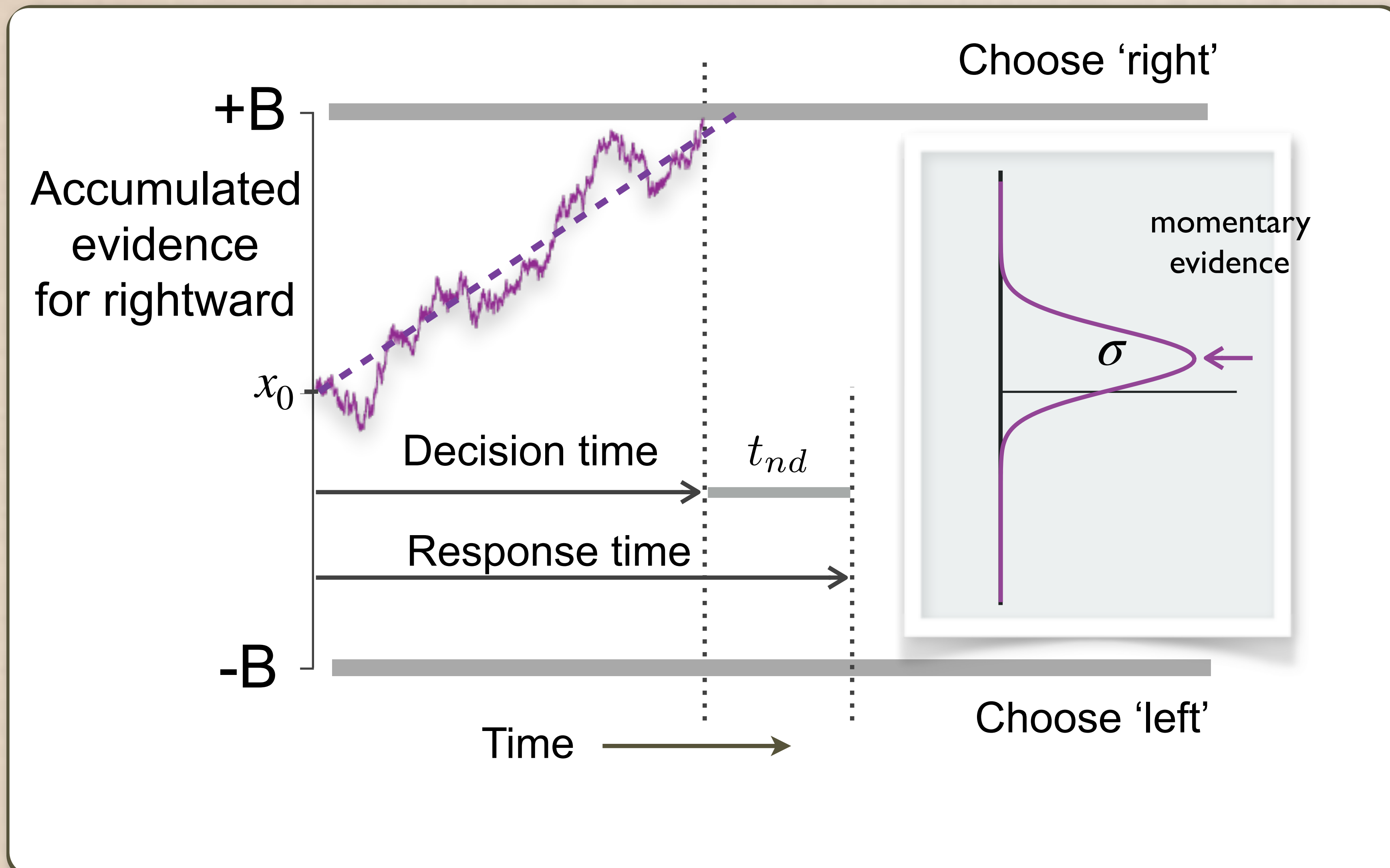


Parameters in the 'simple' version:

- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}
- ~~Diffusion coefficient, σ~~

$$\sigma = 1 \times \mathbf{2}$$

Drift diffusion model

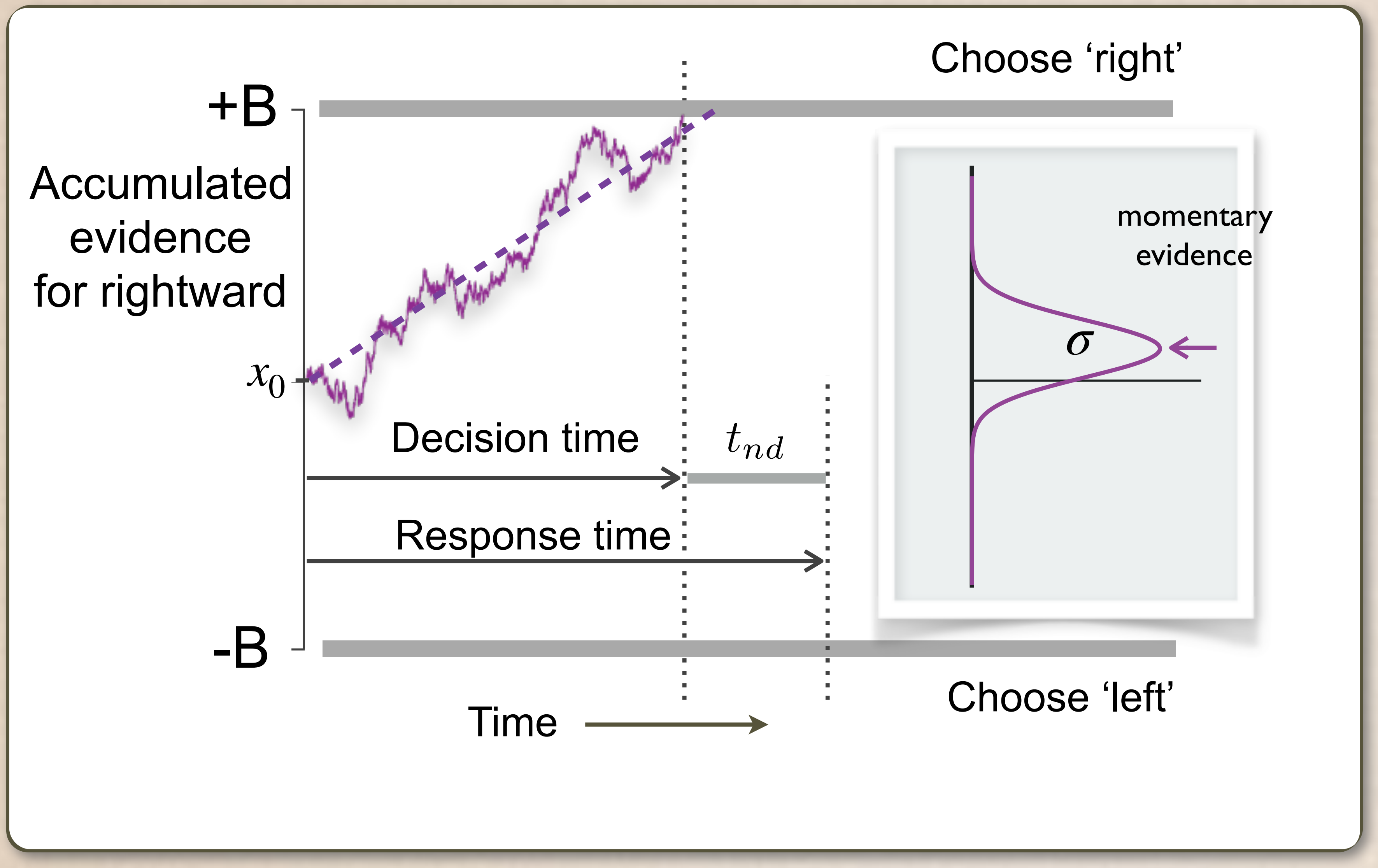


Parameters in the 'simple' version:

- Drift rate, μ $\times$ **2**
- Bound height, B $\times$ **2**
- Starting point, x_0 $\times$ **2**
- Non-decision time, t_{nd}
- ~~Diffusion coefficient, σ~~

$\sigma = 1$ $\times$ **2**

Drift diffusion model



Parameters in the 'simple' version:

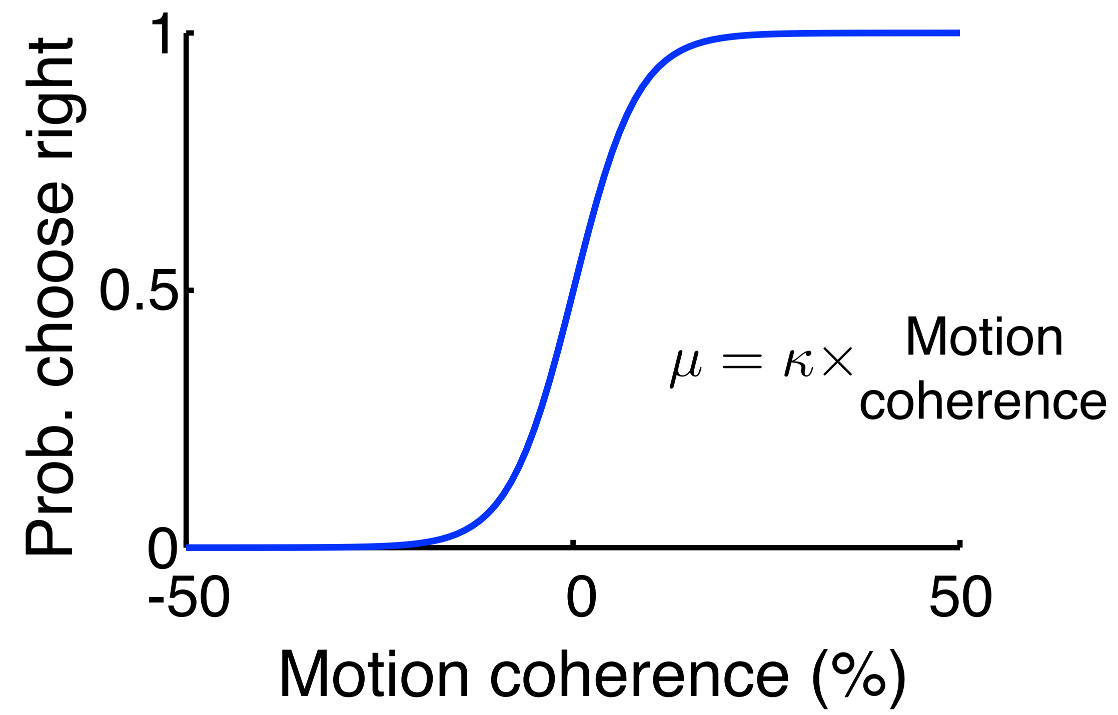
- Drift rate, μ
- Bound height, B
- Starting point, x_0
- Non-decision time, t_{nd}
- ~~Diffusion coefficient, σ~~

$$\sigma = 1$$

Analytic solution to the choice and response times

Probability of terminating at the upper bound:

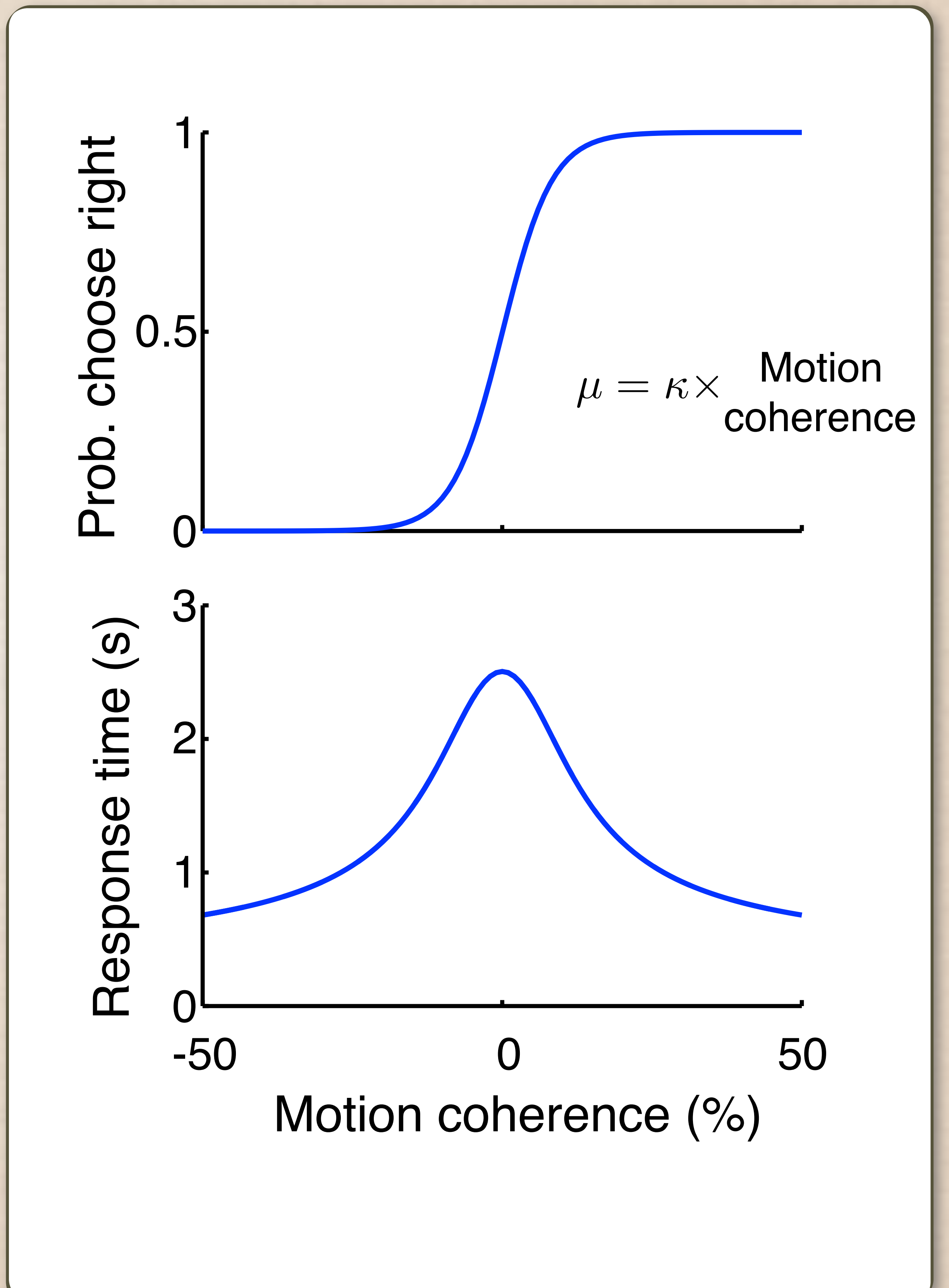
$$p_+ = \frac{1}{1 + e^{-2\mu B}}$$



Analytic solution to the choice and response times

Probability of terminating at the upper bound:

$$p_+ = \frac{1}{1 + e^{-2\mu B}}$$



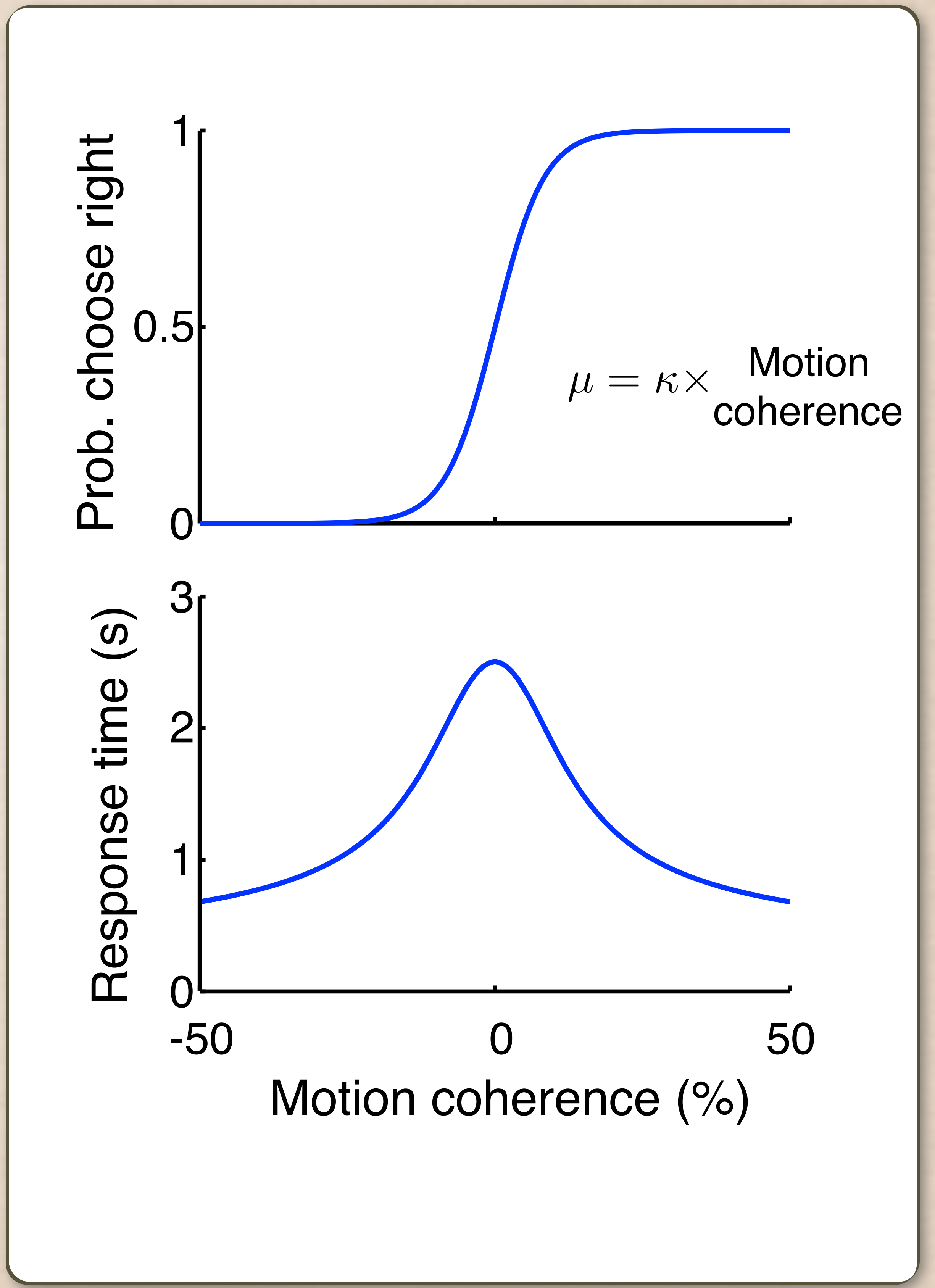
Analytic solution to the choice and response times

Probability of terminating at the upper bound:

$$p_+ = \frac{1}{1 + e^{-2\mu B}}$$

Average response time:

$$T = \frac{B}{\mu} \tanh(\mu B) + t_{nd}$$



Analytic solution to the choice and response times

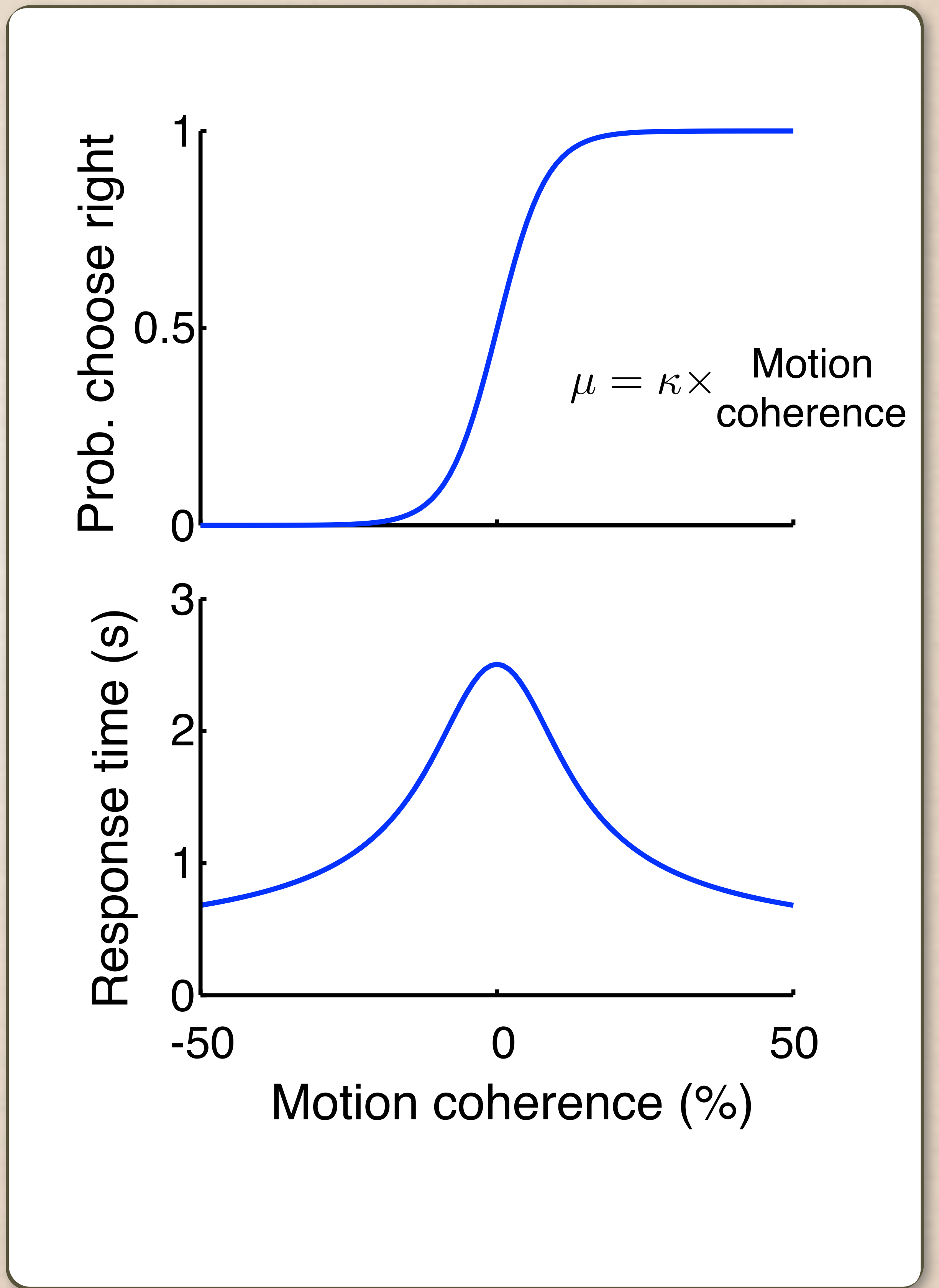
Probability of terminating at the upper bound:

$$p_+ = \frac{1}{1 + e^{-2\mu B}}$$

Average response time:

$$T = \frac{B}{\mu} \tanh(\mu B) + t_{nd}$$

When $\mu = 0$: $T = B^2 + t_{nd}$



Analytic solution to the choice and response times

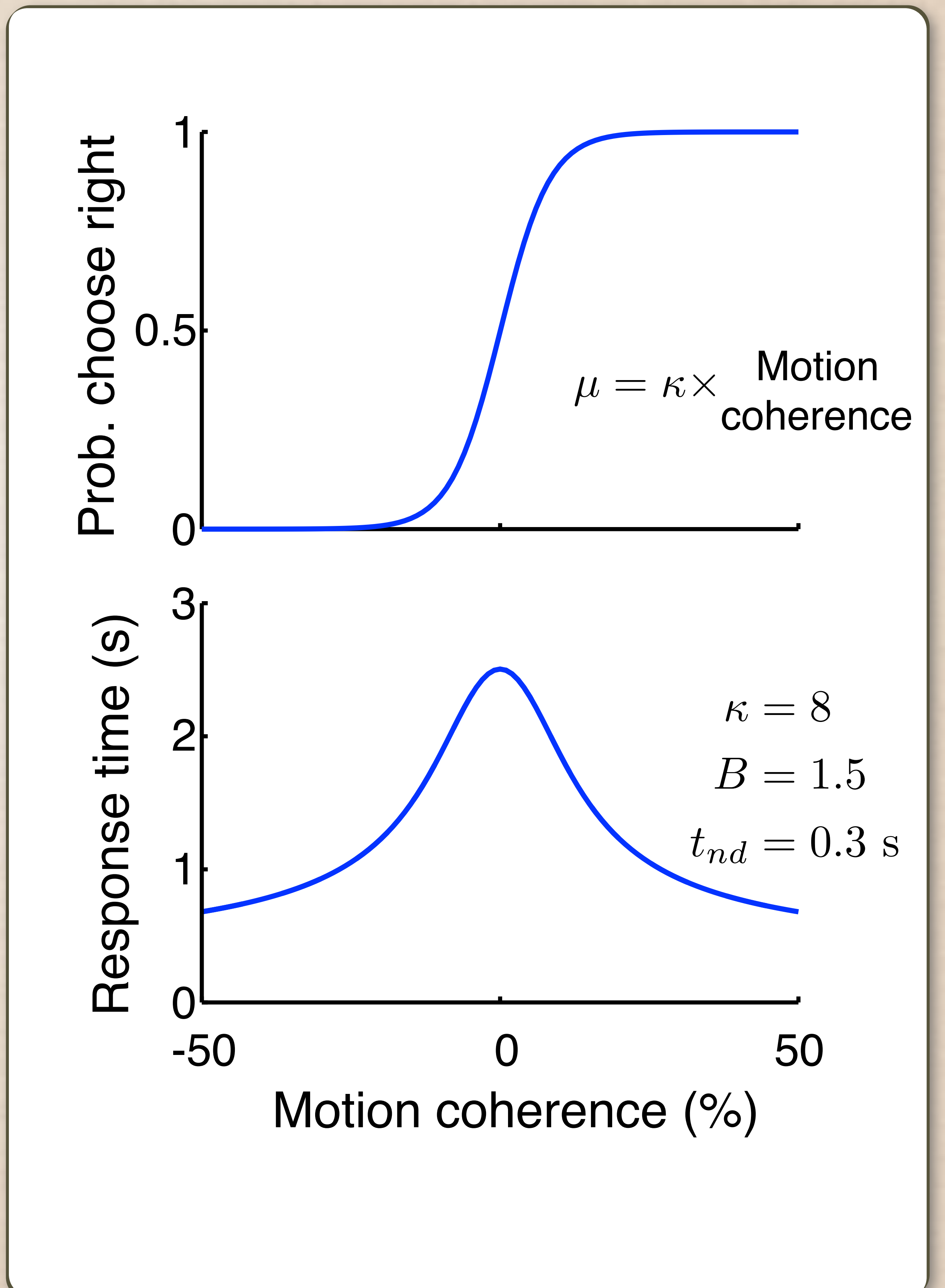
Probability of terminating at the upper bound:

$$p_+ = \frac{1}{1+e^{-2\mu B}}$$

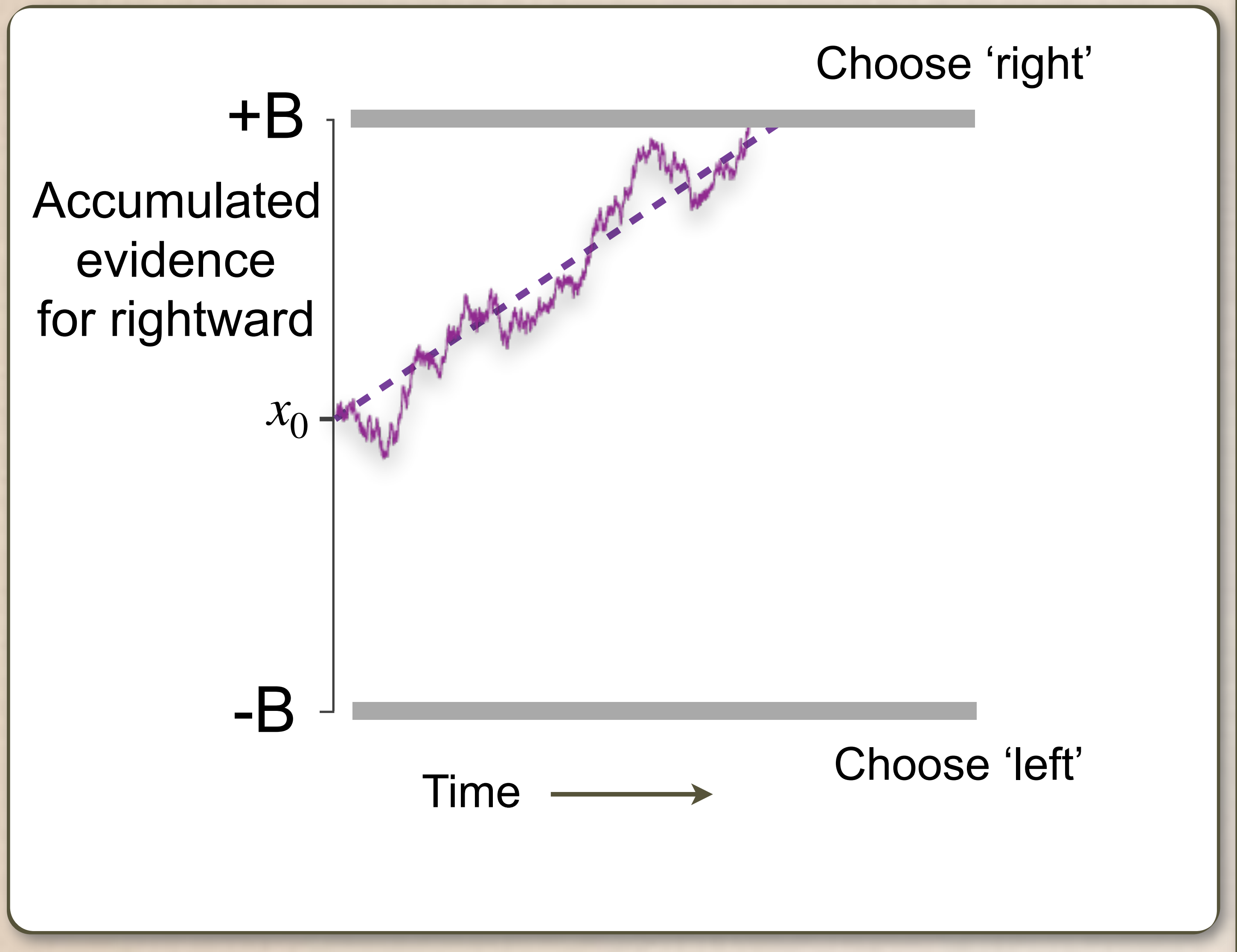
Average response time:

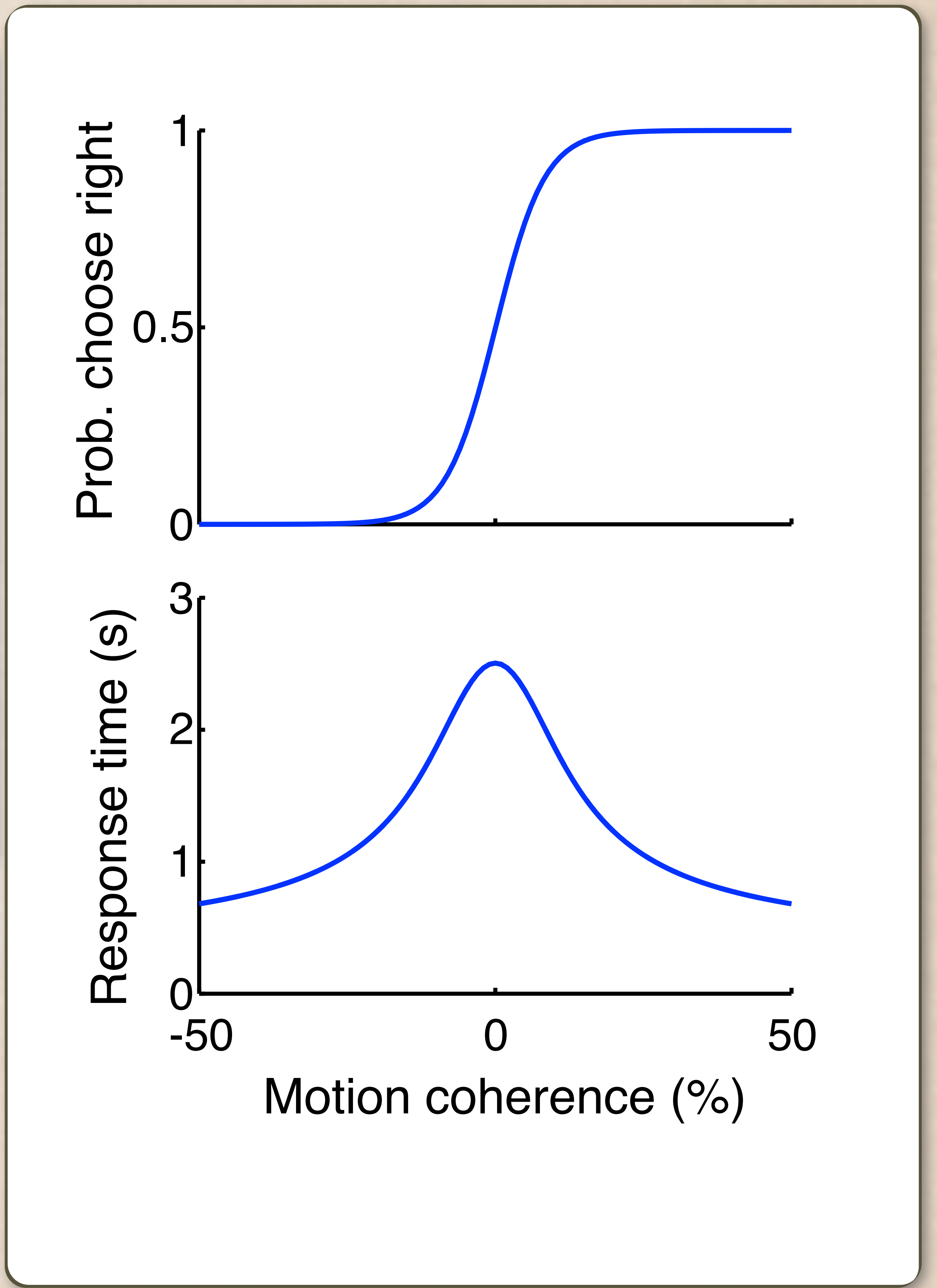
$$T = \frac{B}{\mu}\tanh(\mu B) + t_{nd}$$

When $\mu = 0$: $T = B^2 + t_{nd}$

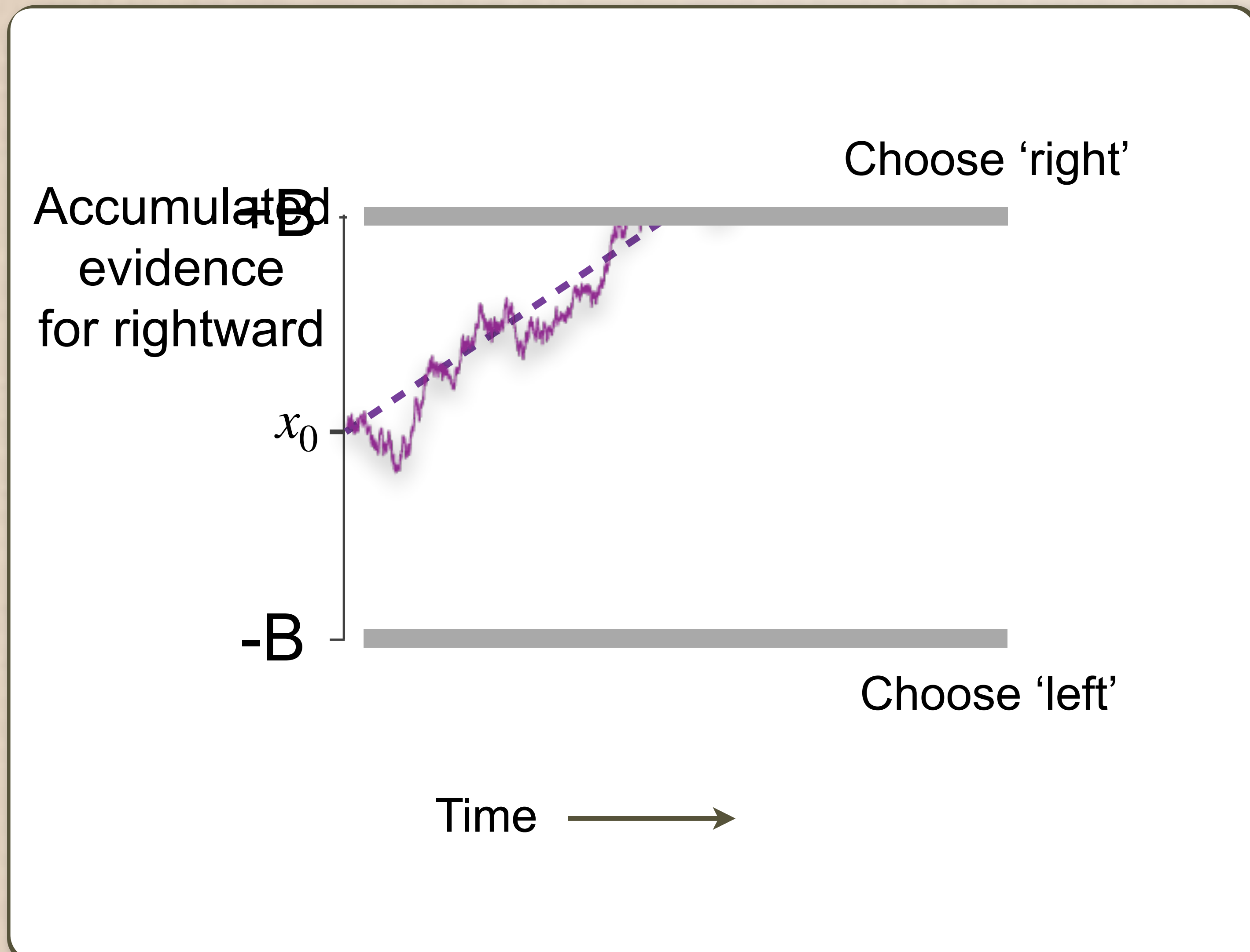


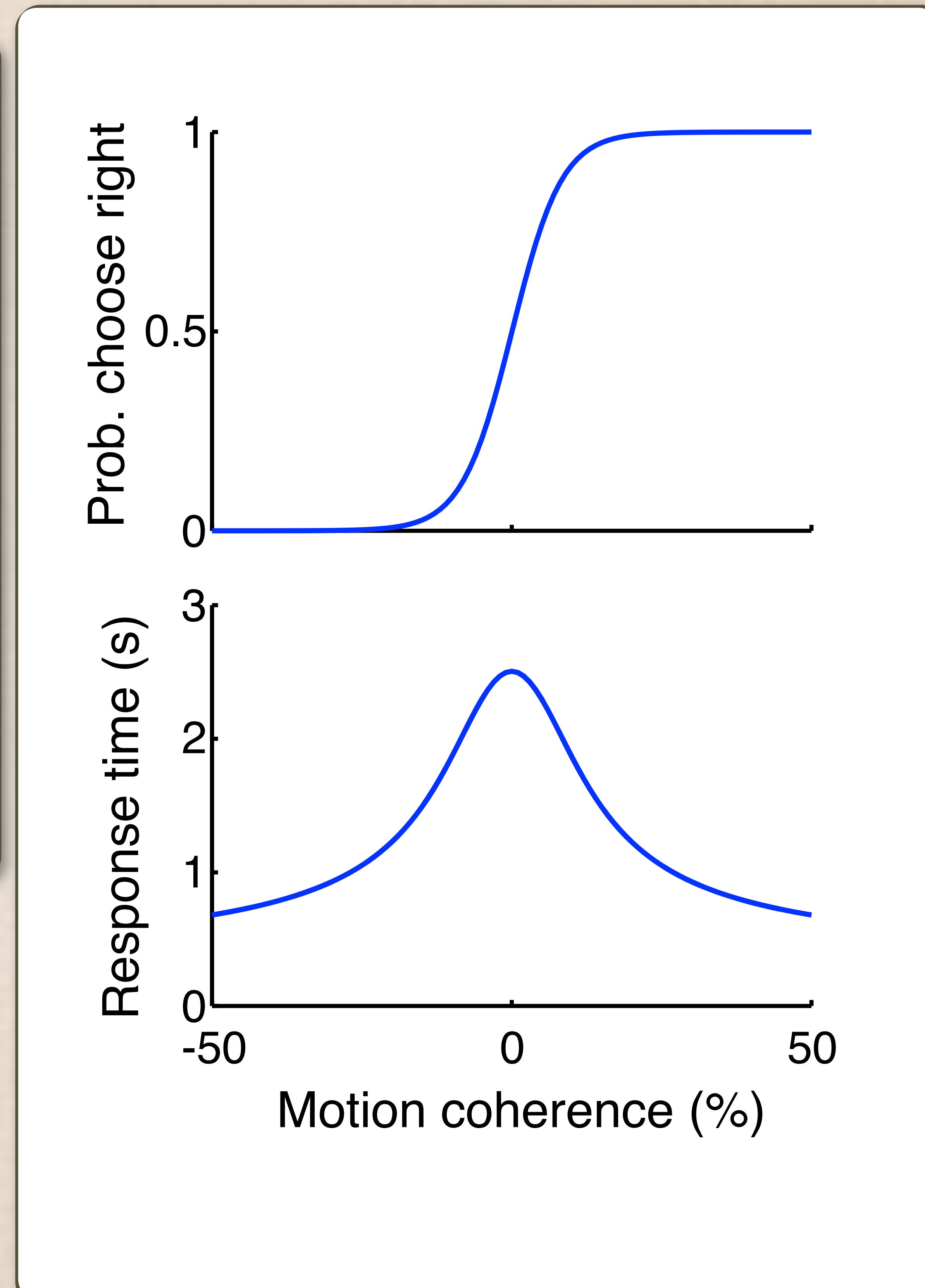
Change in bound height



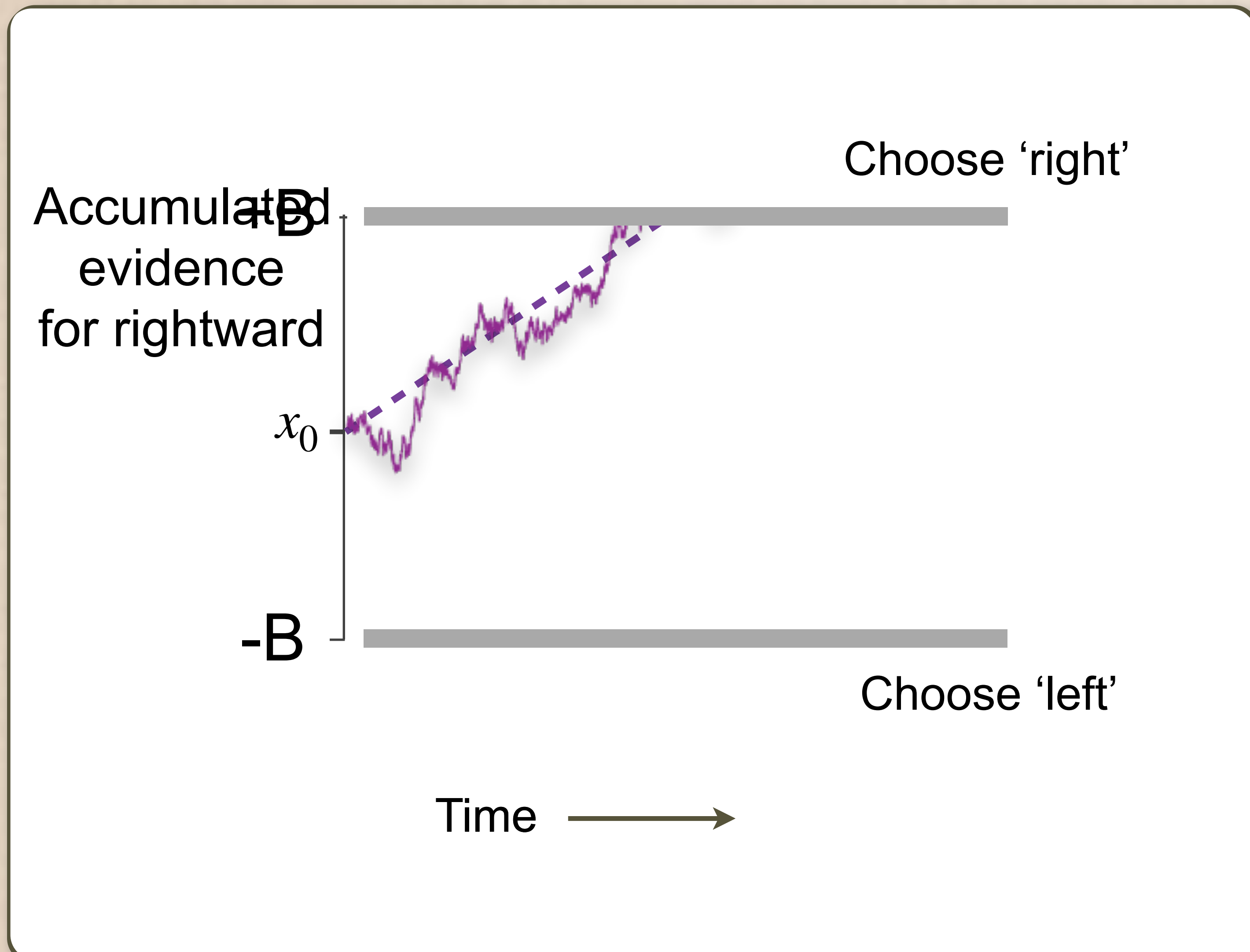


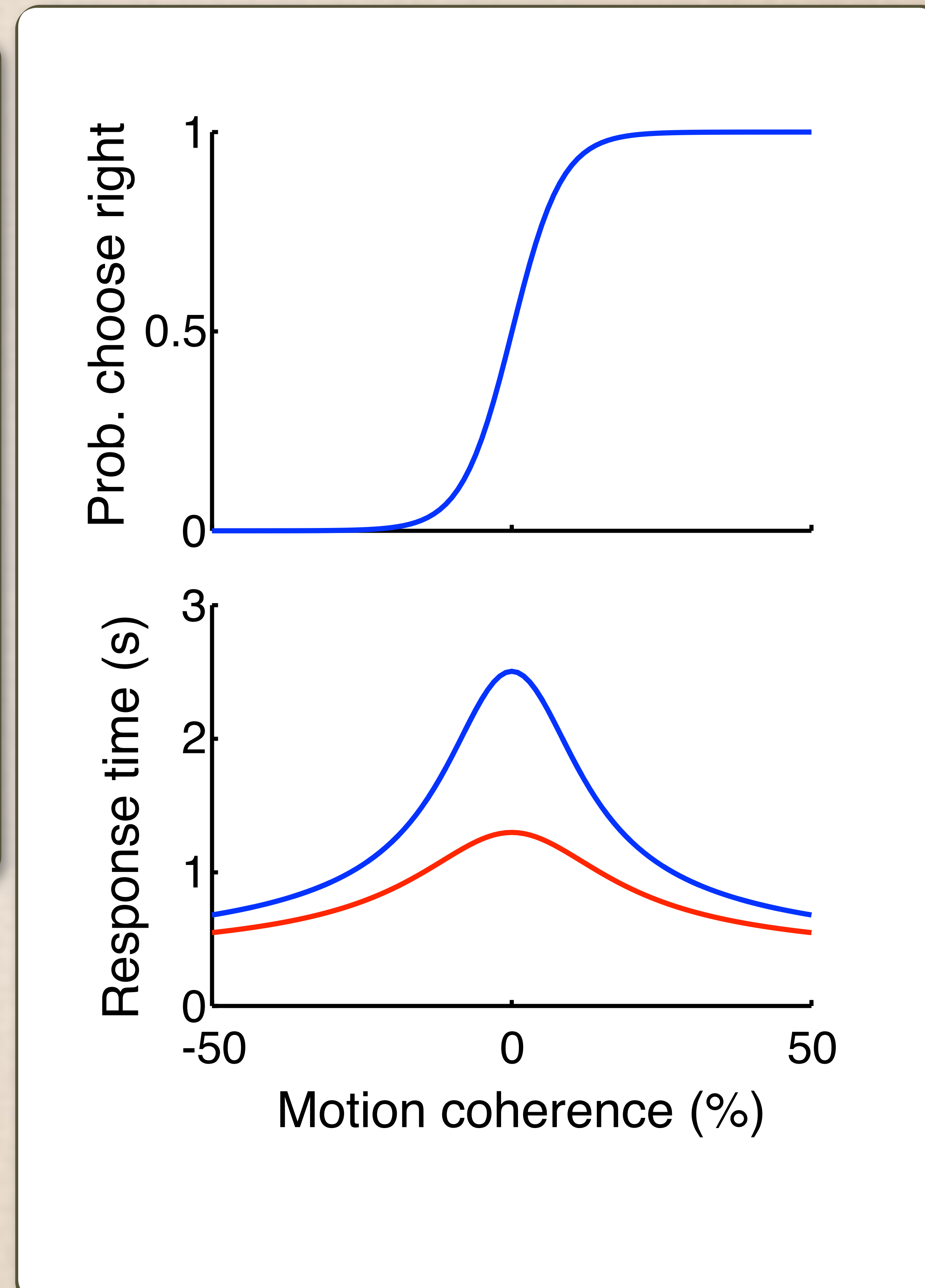
Change in bound height



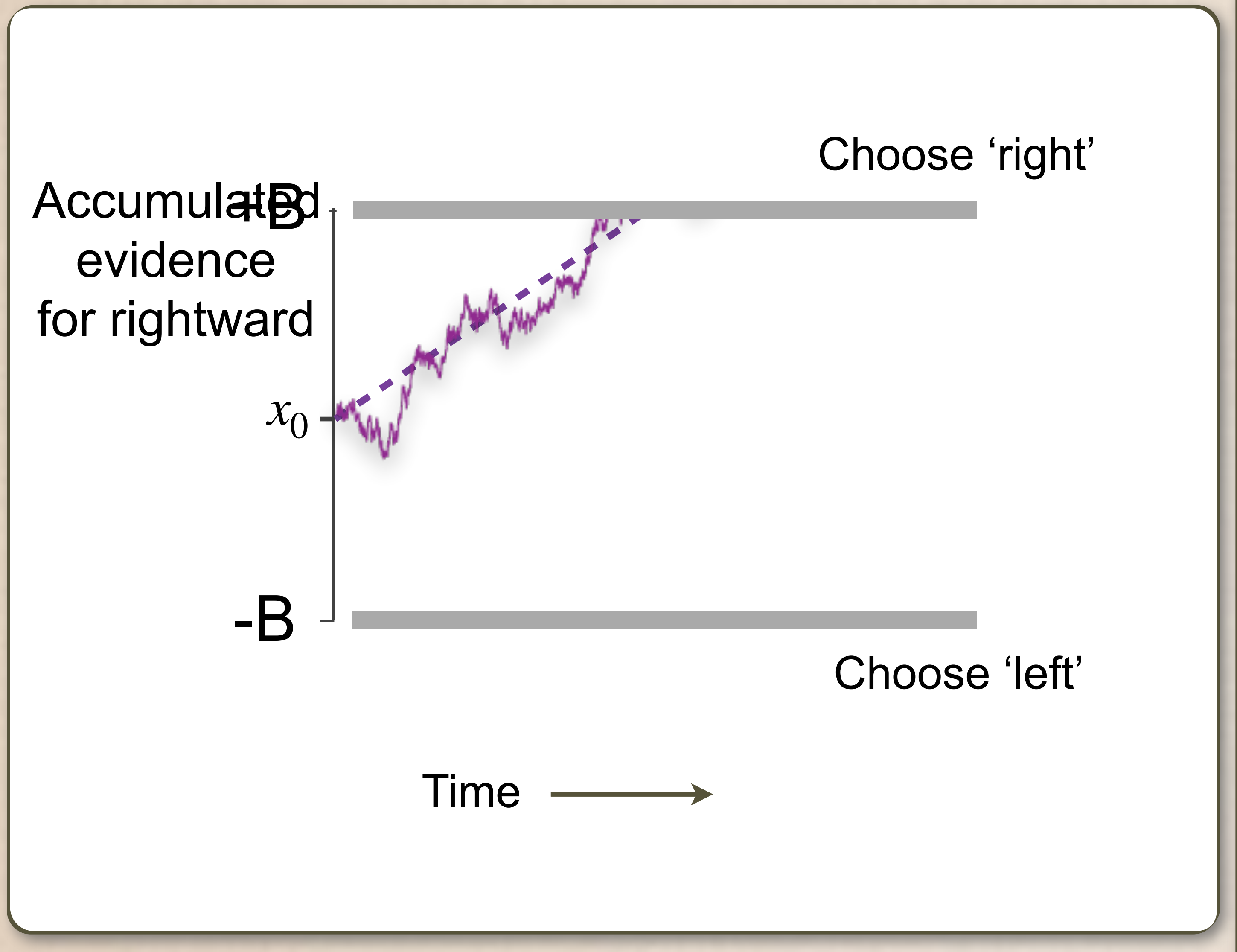


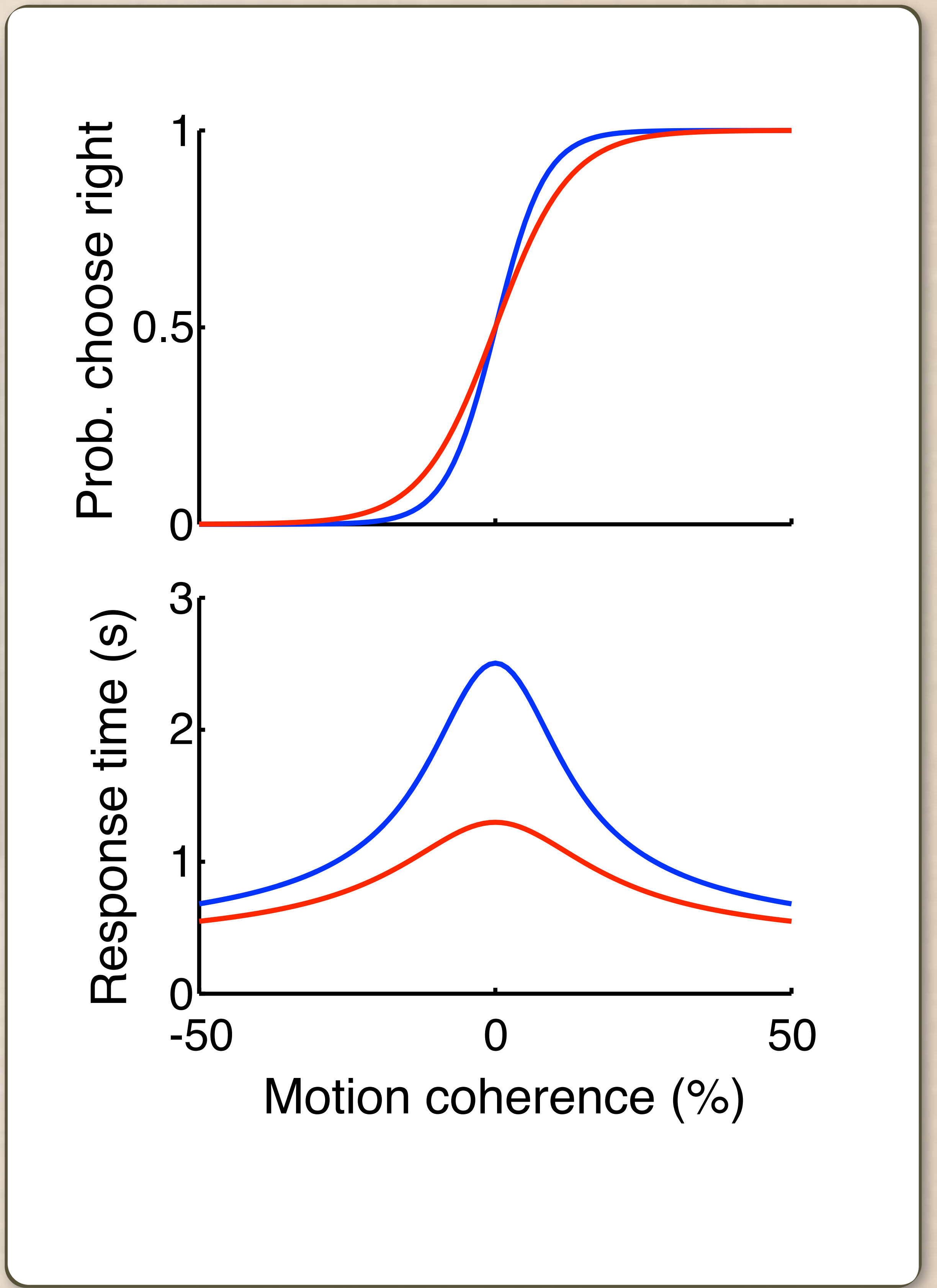
Change in bound height



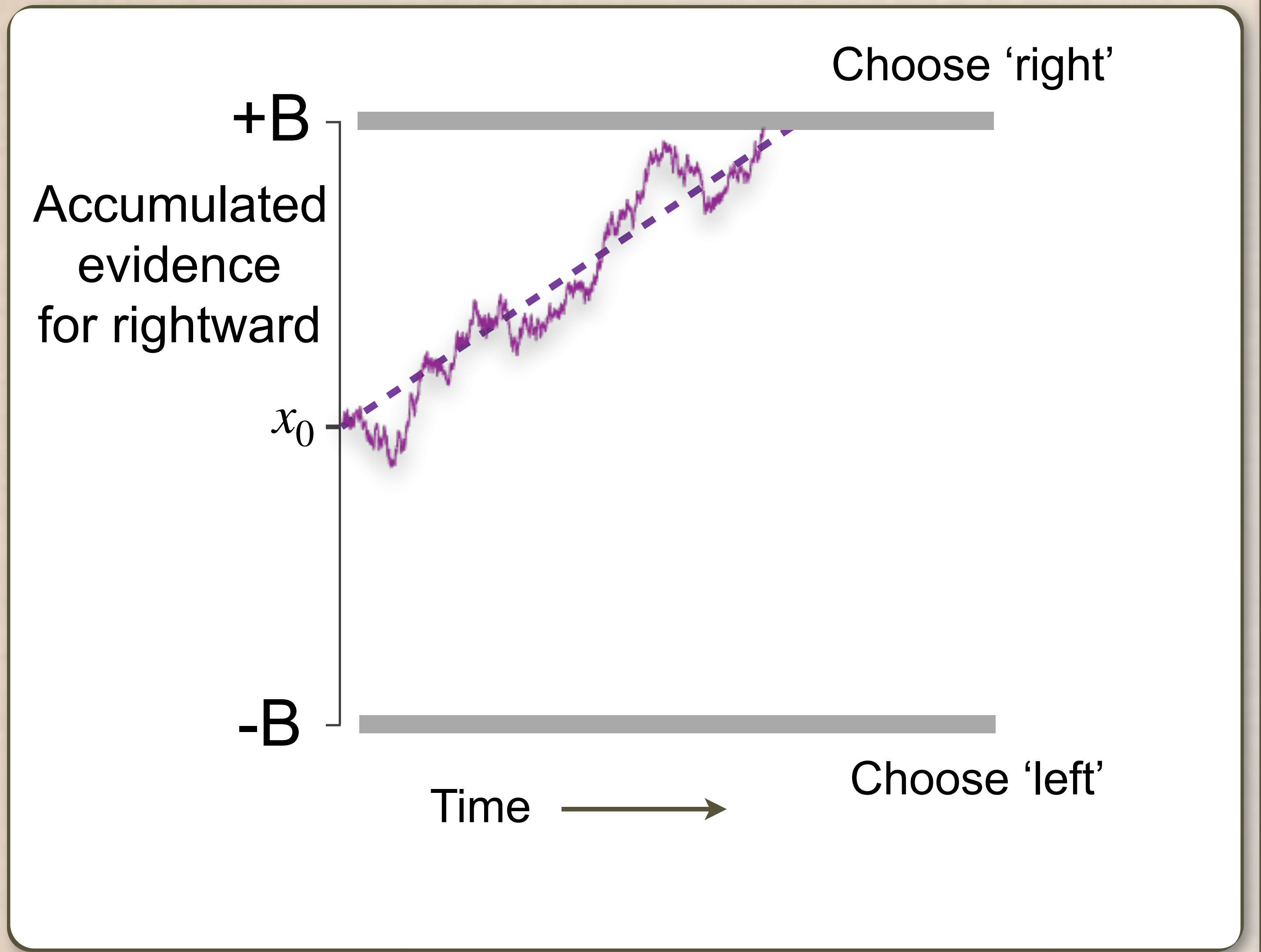


Change in bound height



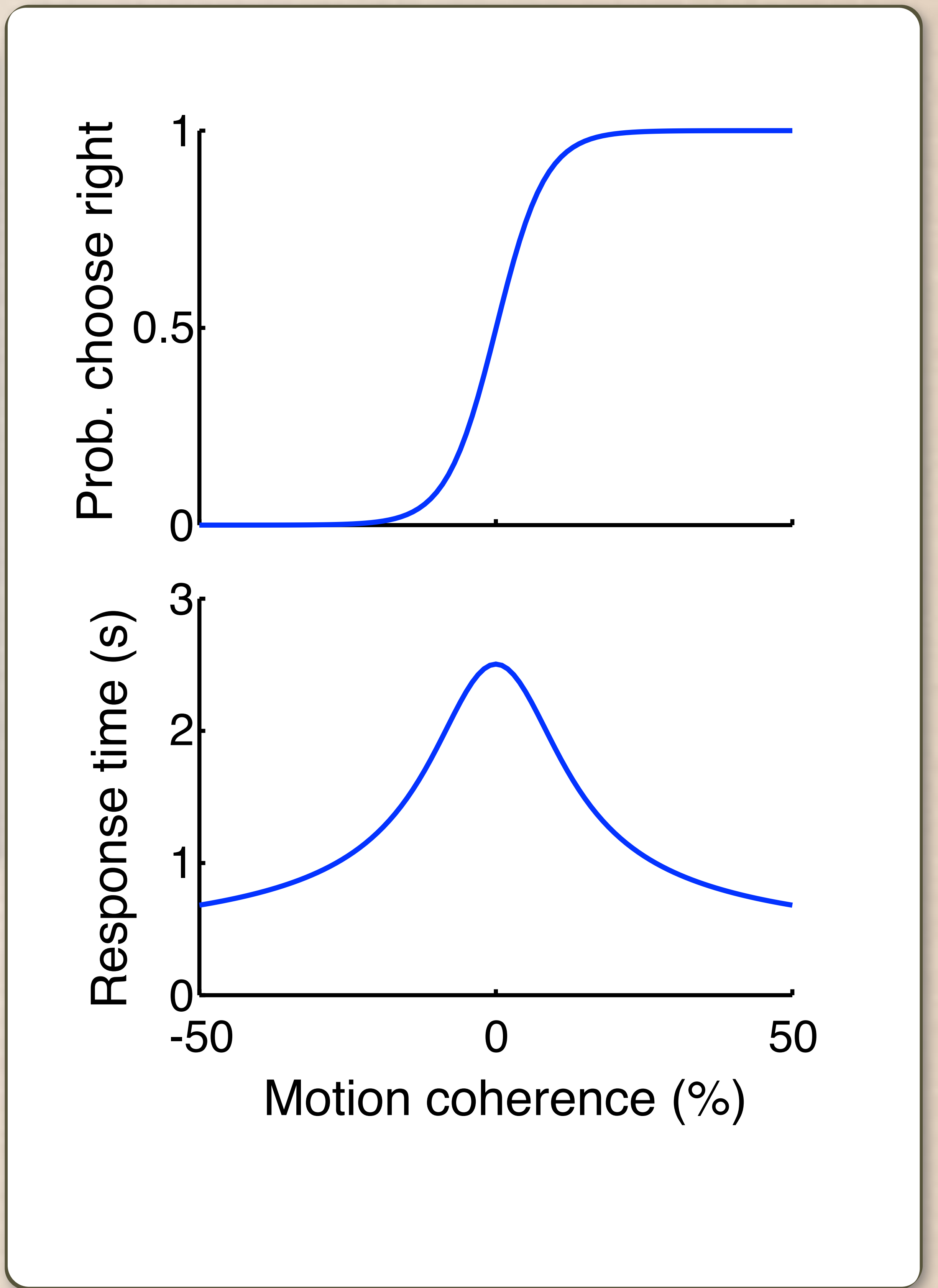


Change in drift rate

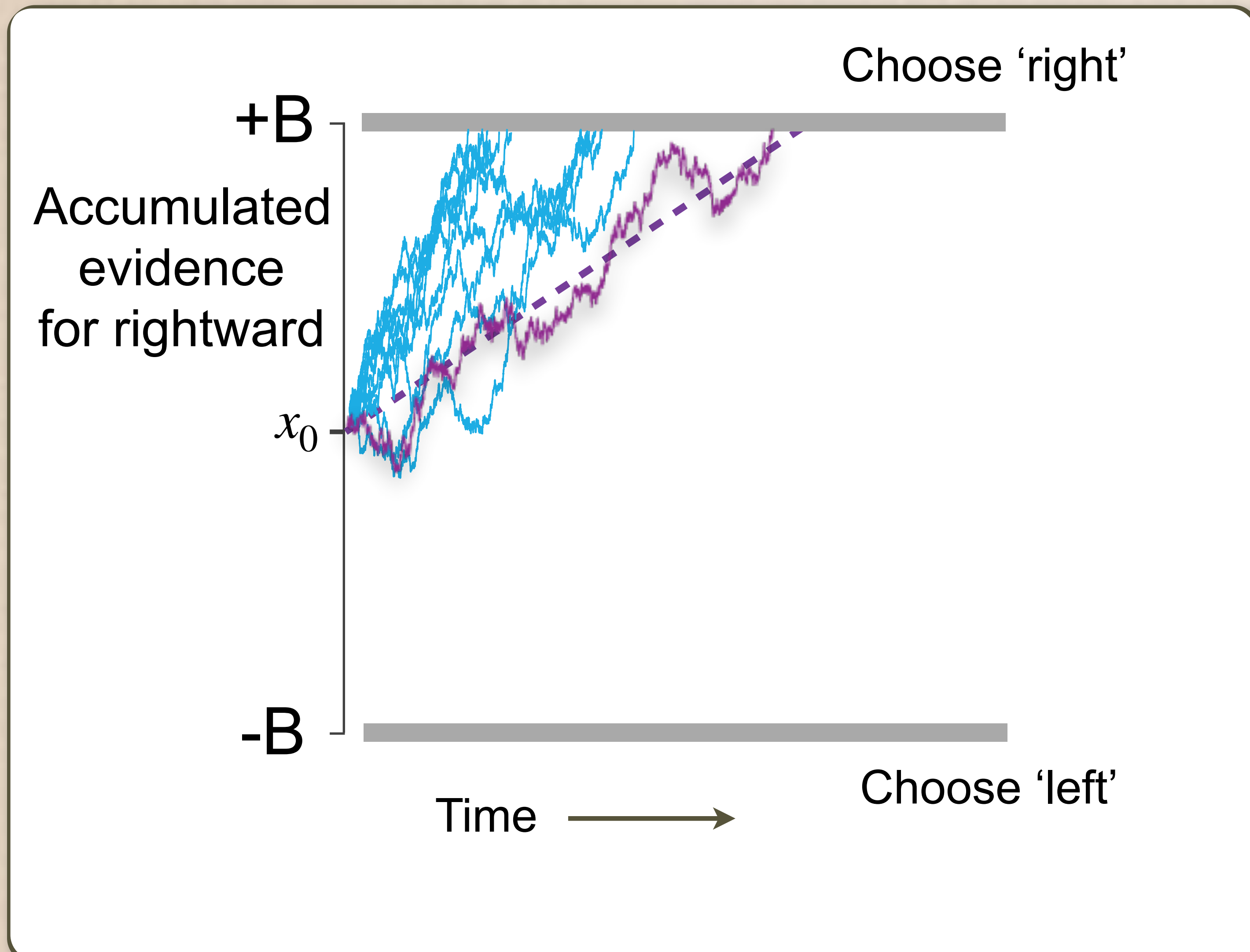


$\mu = \kappa \times$ Motion coherence

signal-to-noise

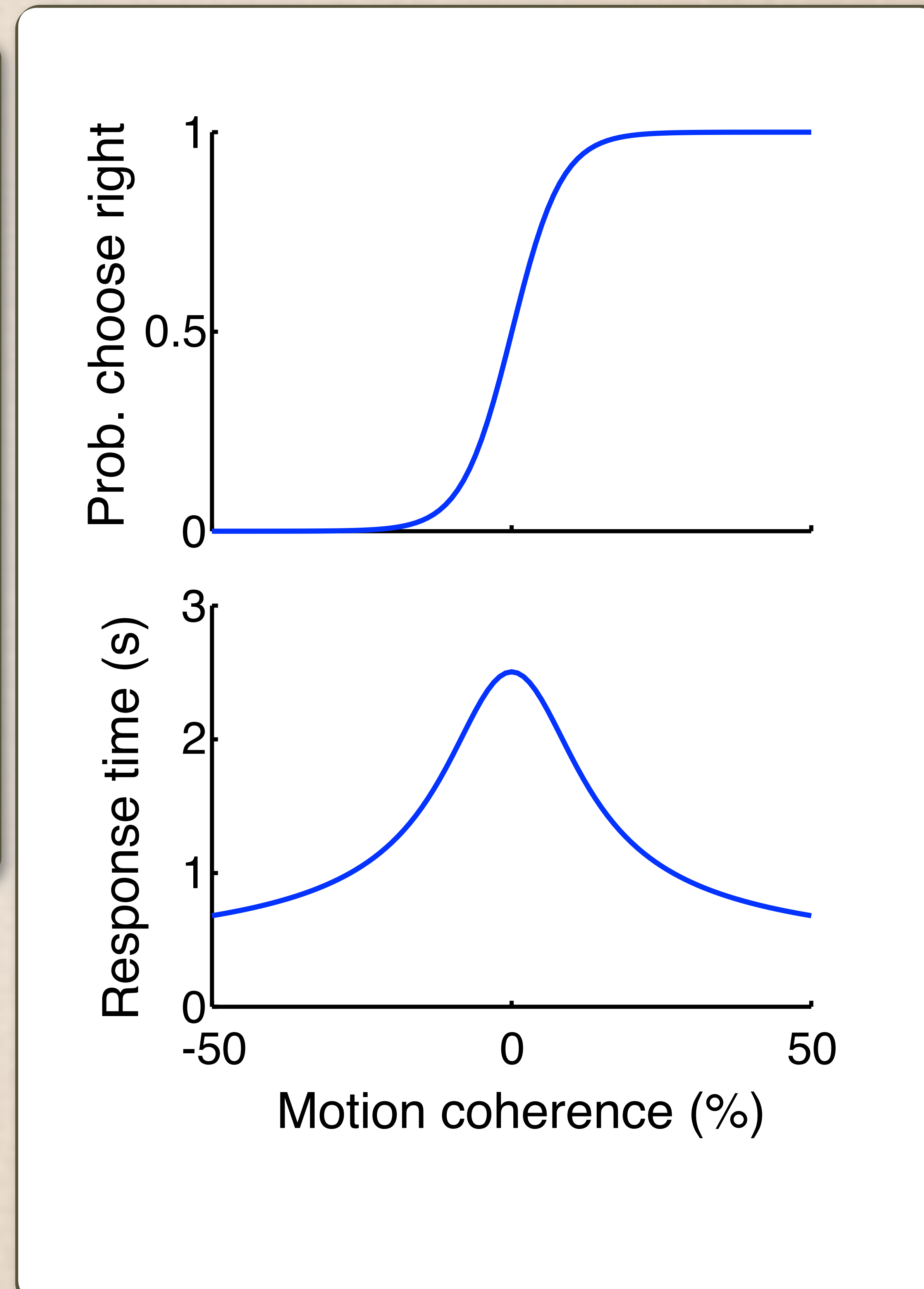


Change in drift rate

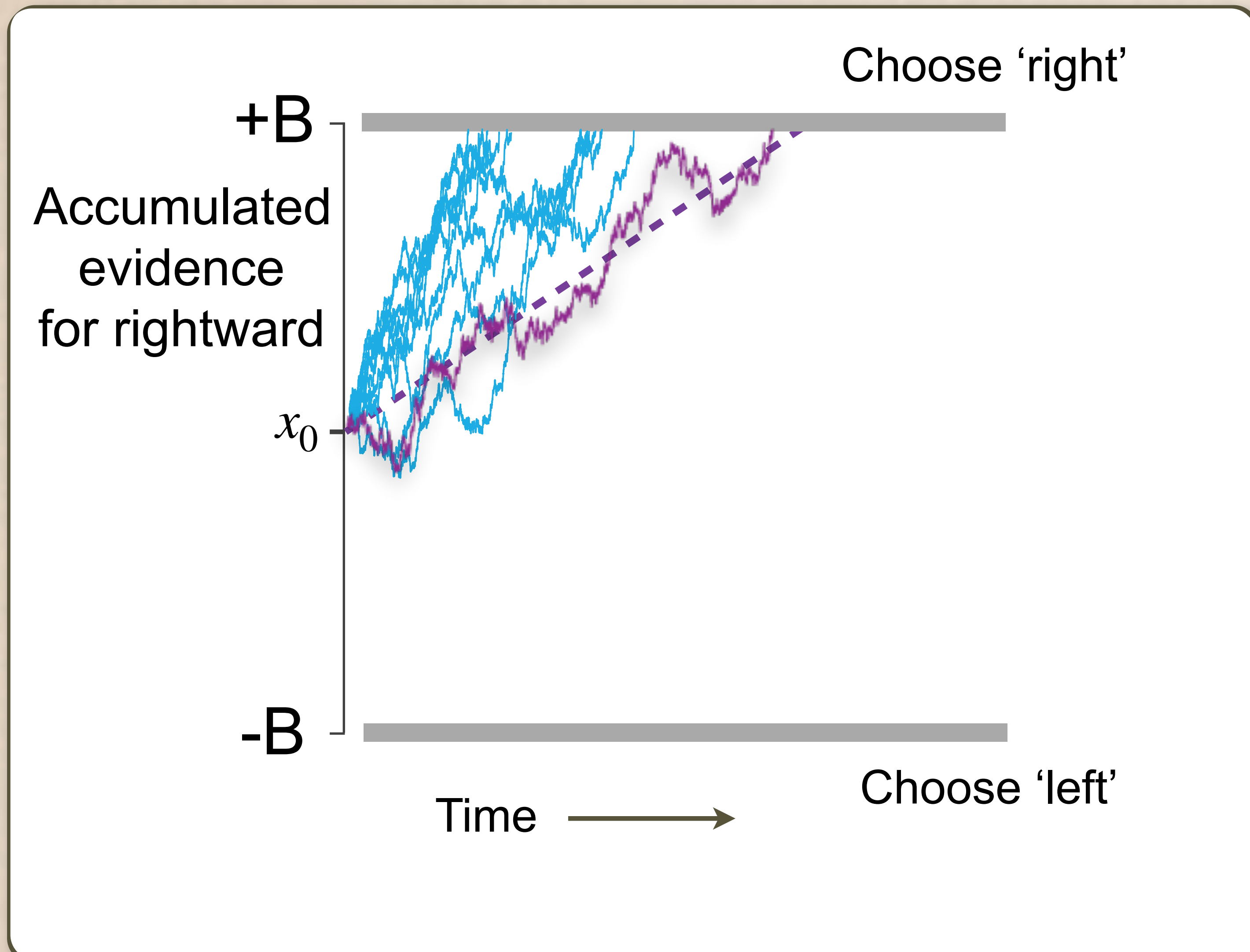


$\mu = \kappa \times$ Motion coherence

signal-to-noise

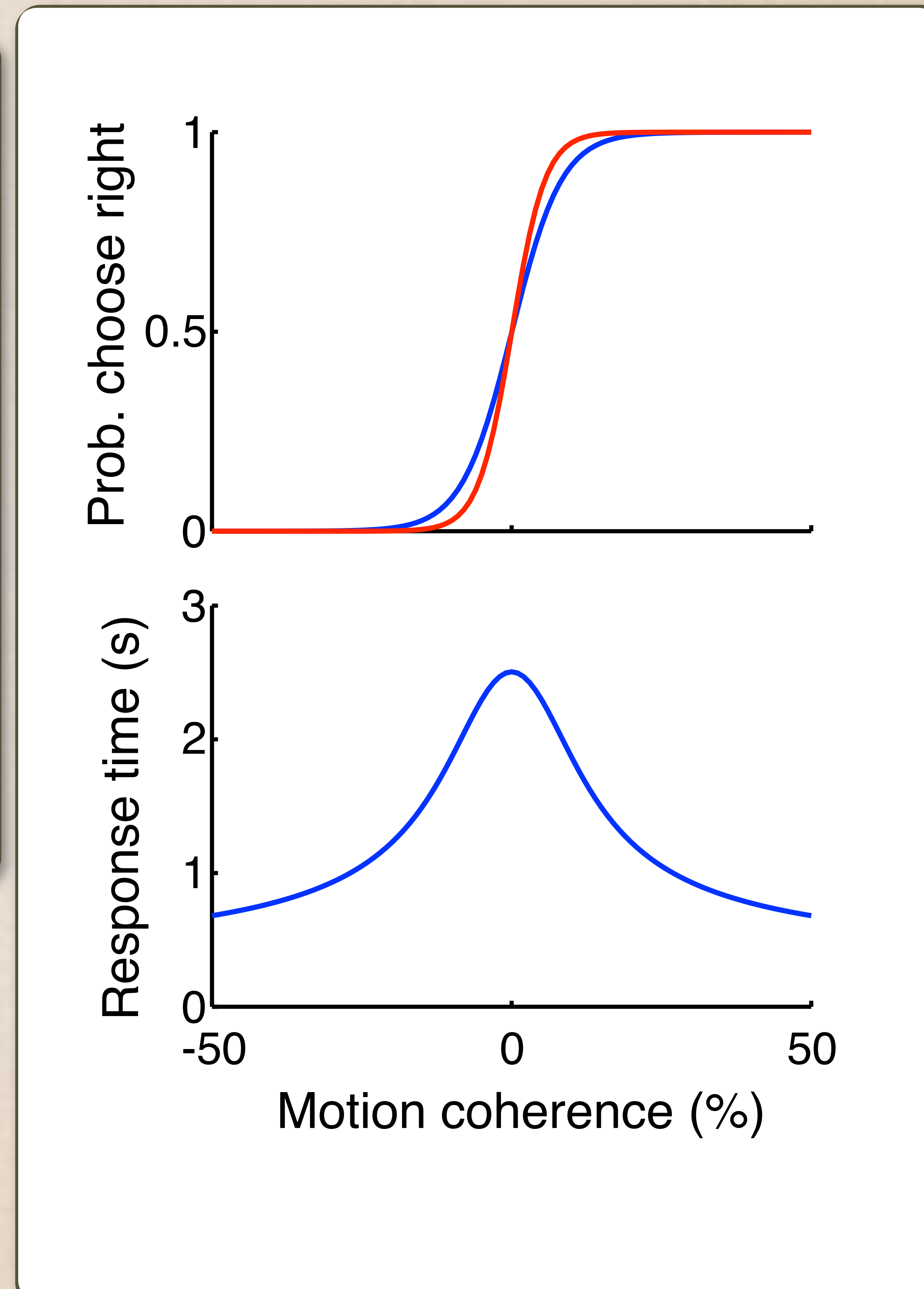


Change in drift rate

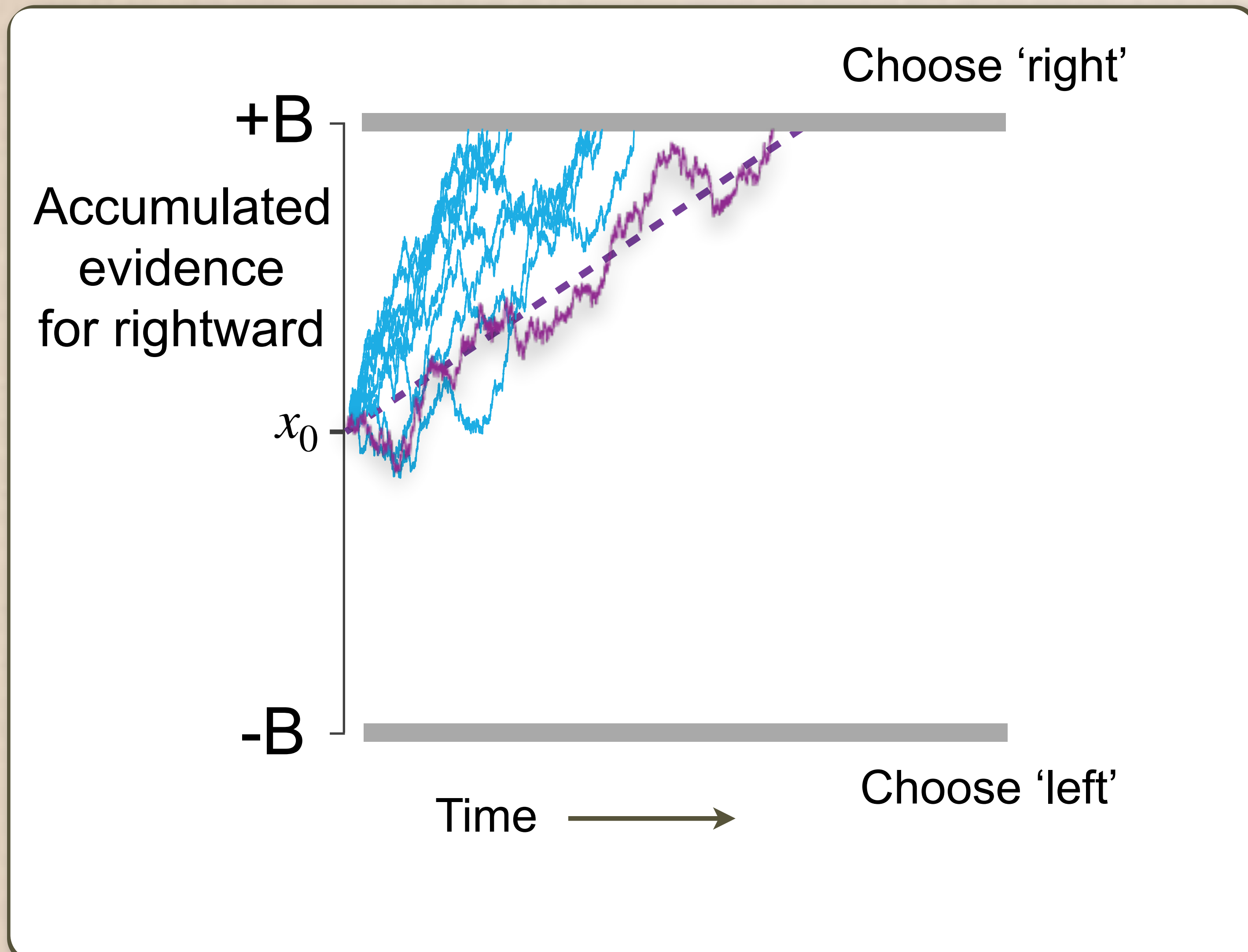


$\mu = \kappa \times$ Motion coherence

signal-to-noise

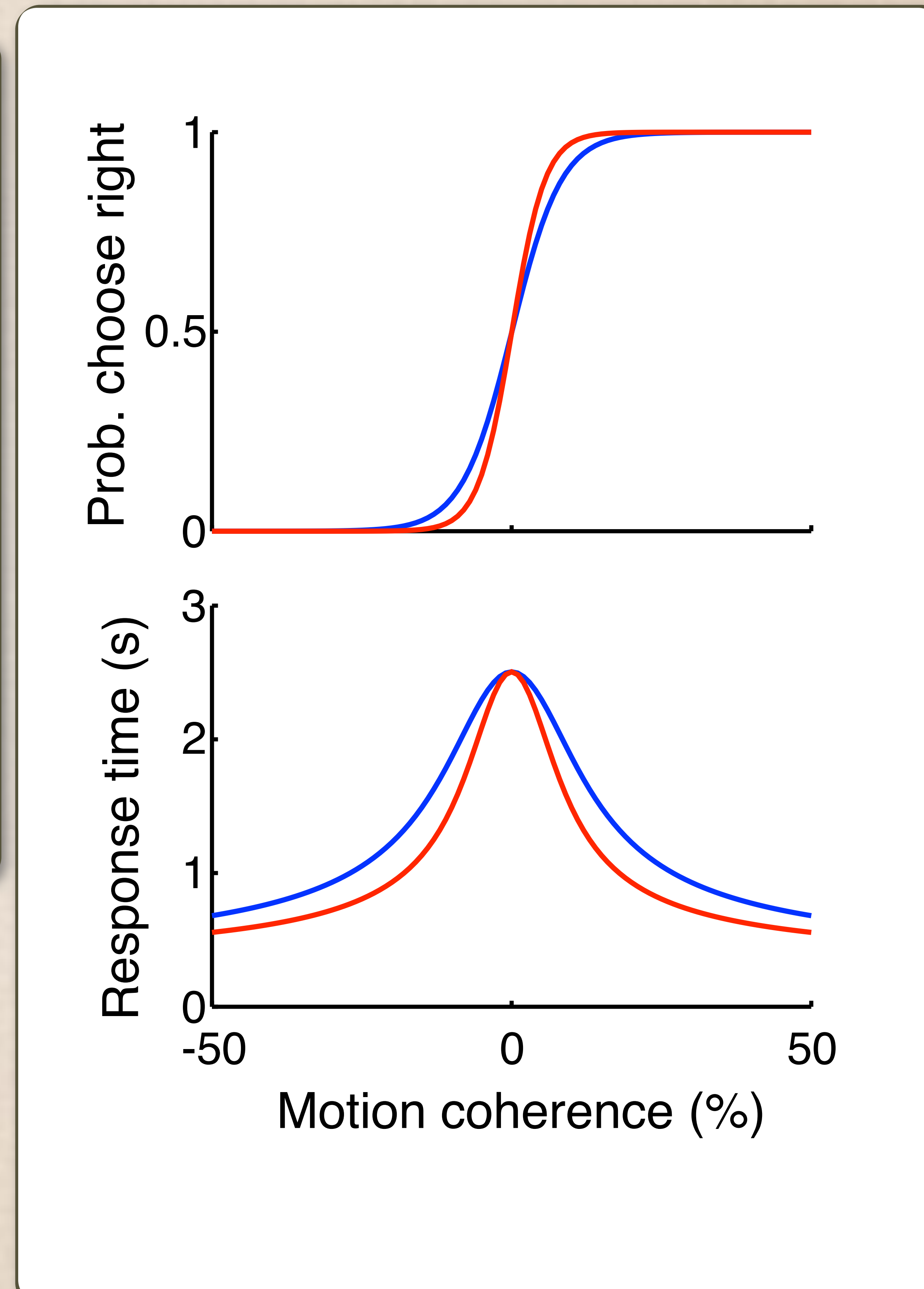


Change in drift rate

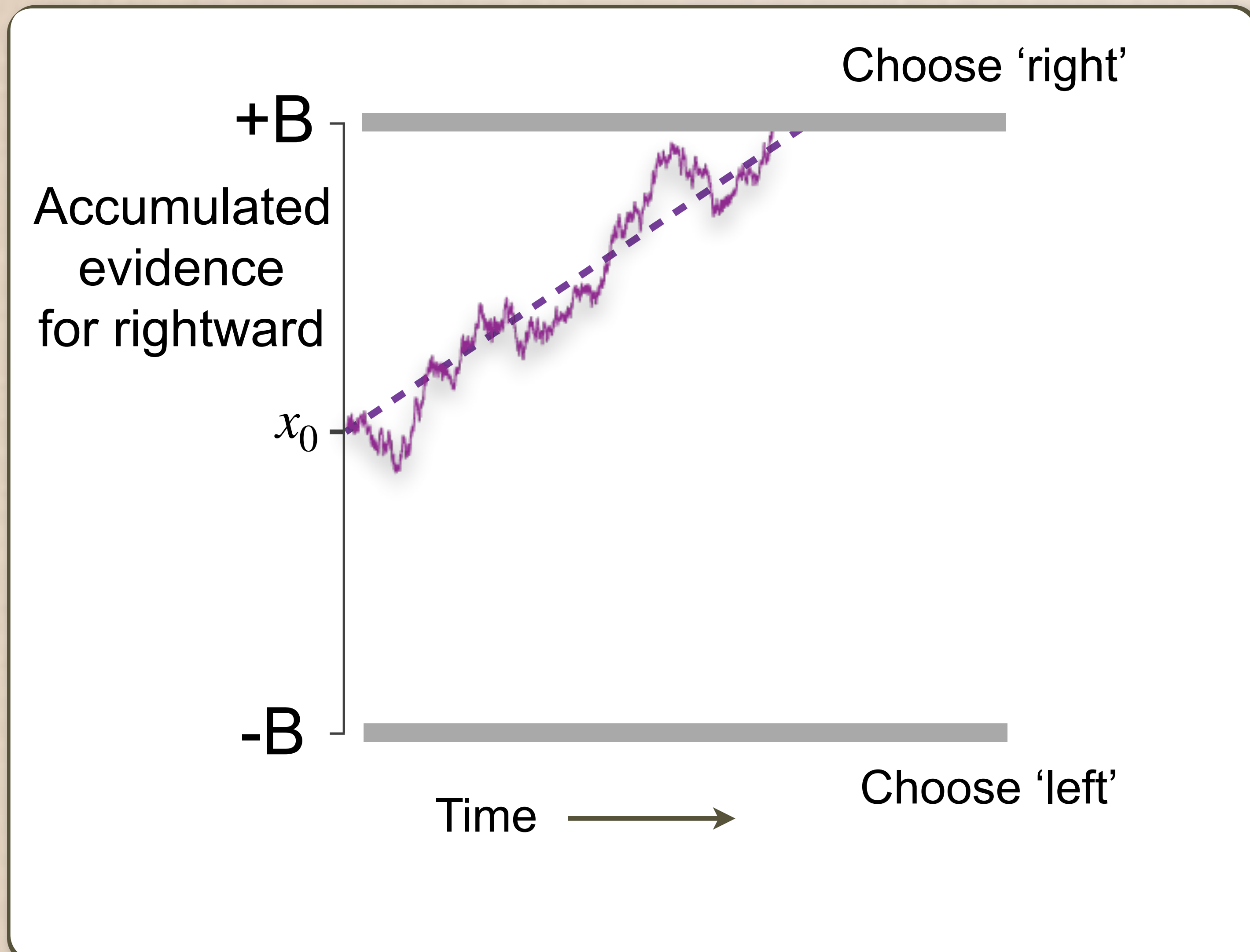


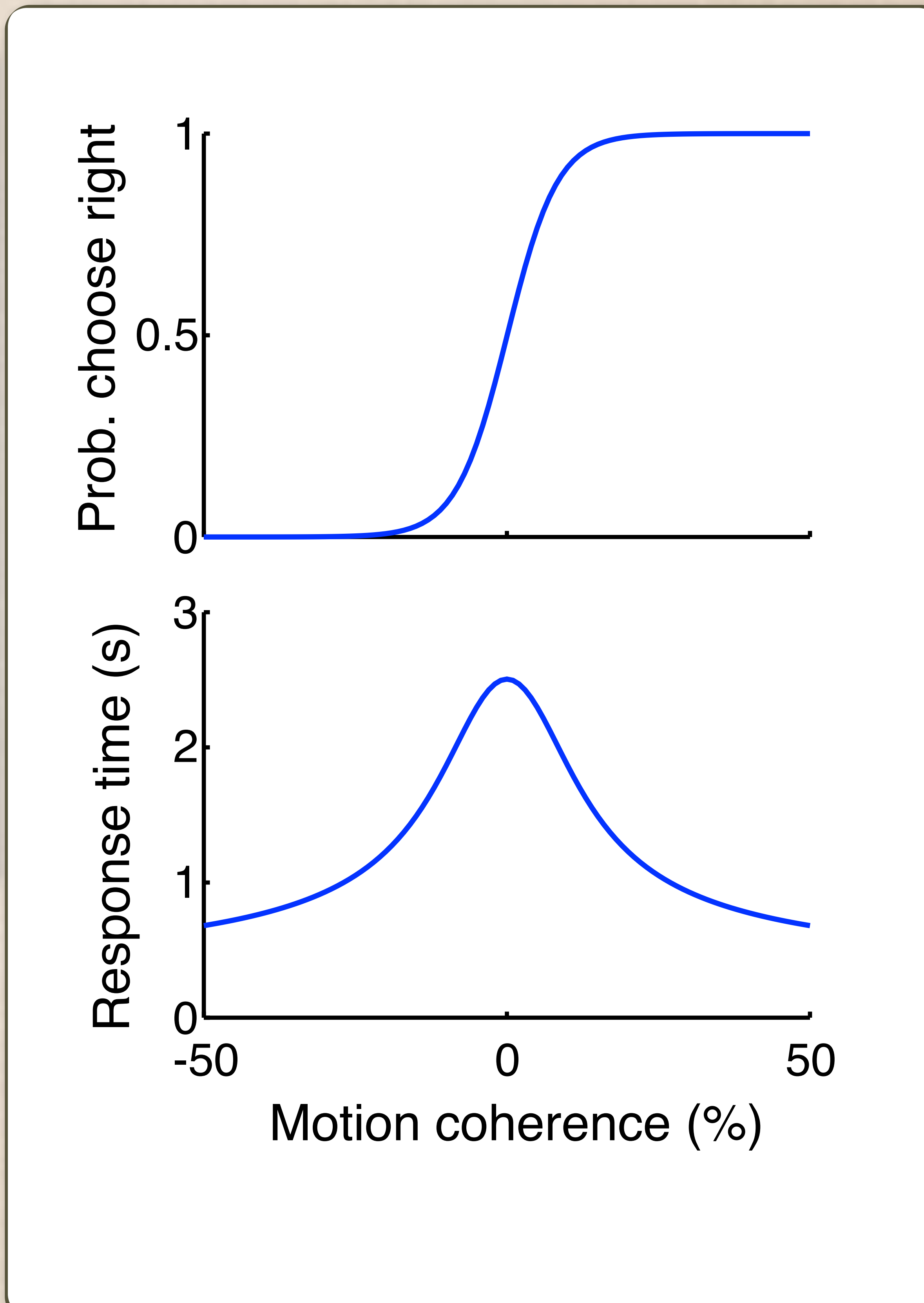
$\mu = \kappa \times$ Motion coherence

signal-to-noise

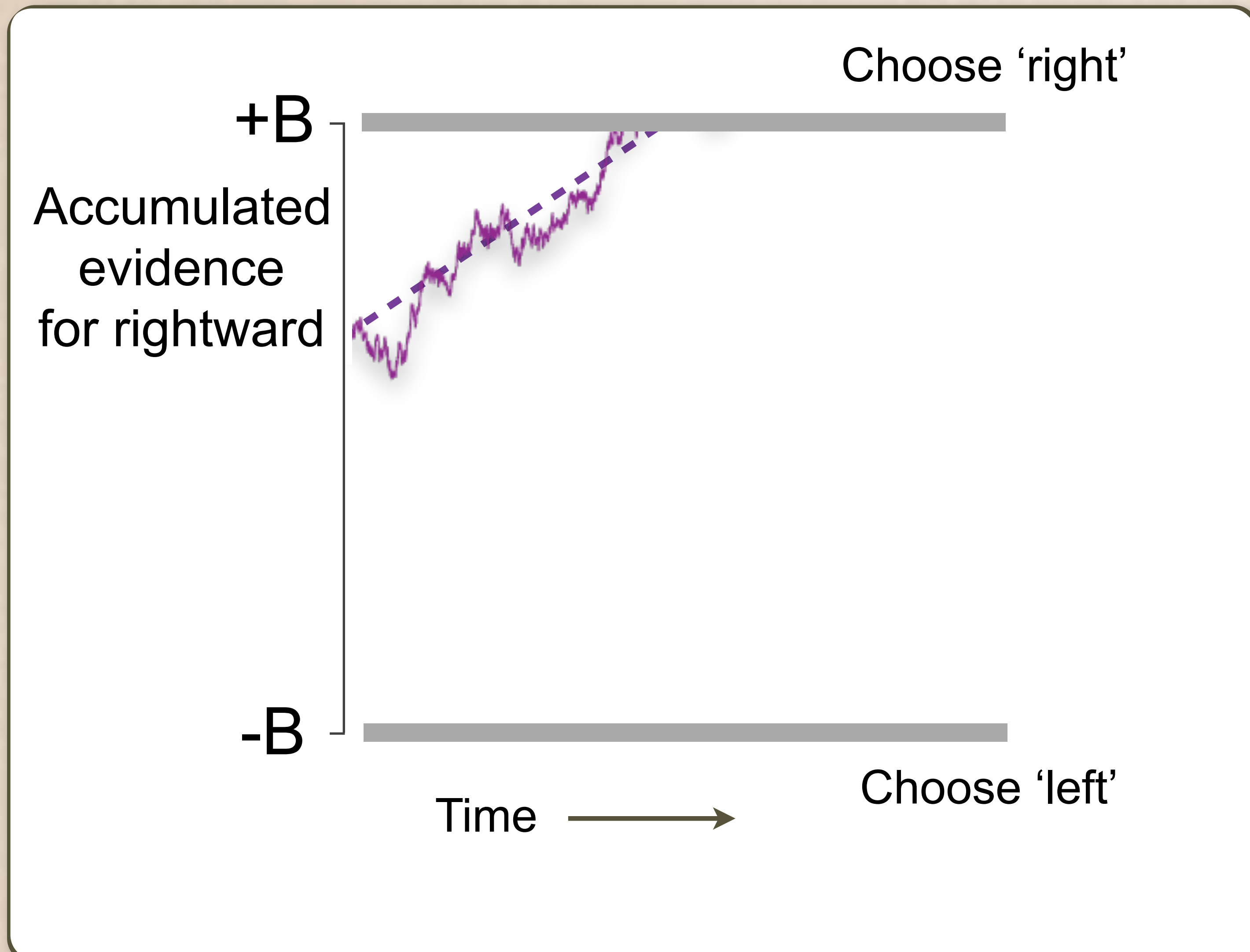


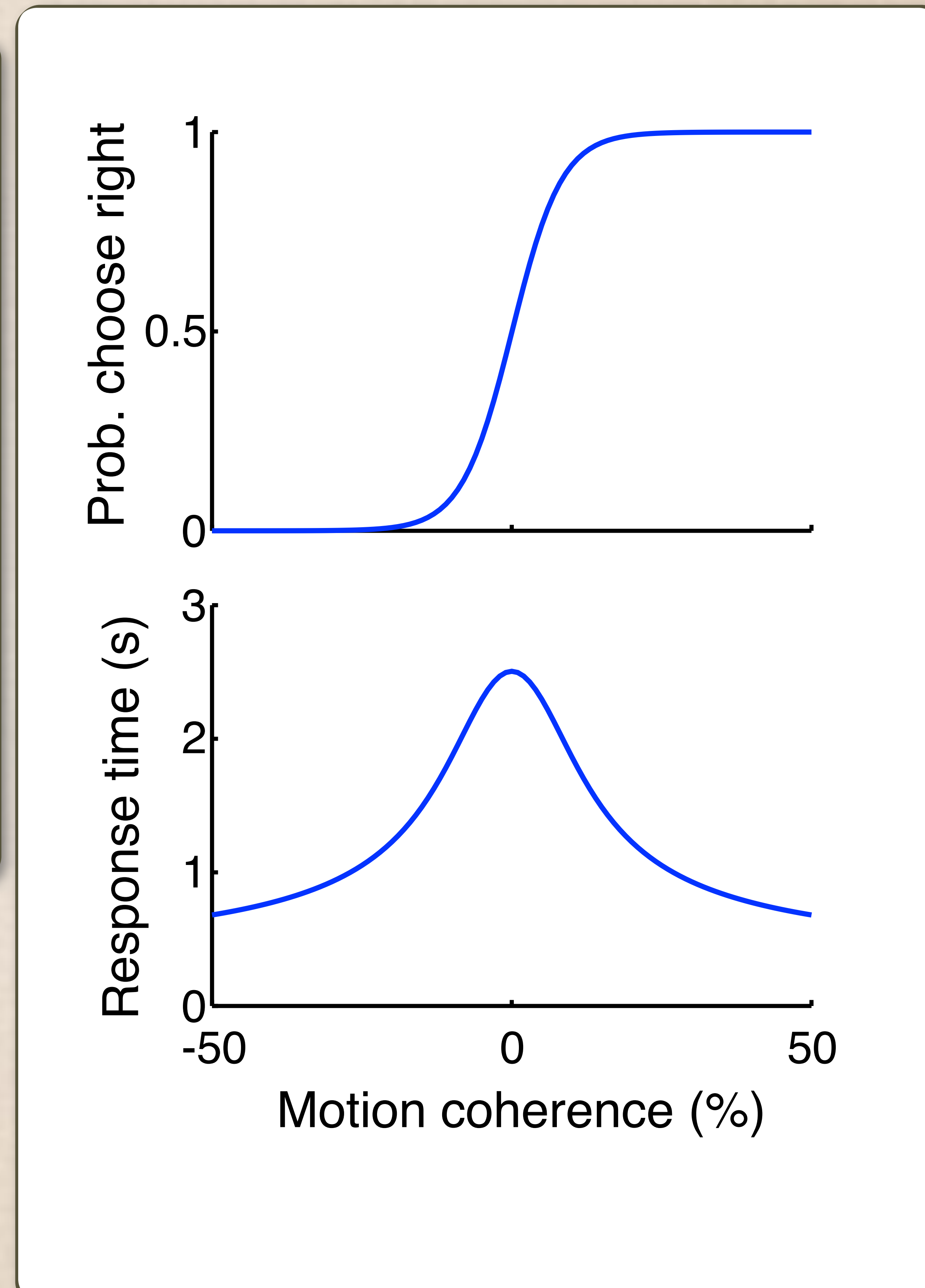
Change in starting point



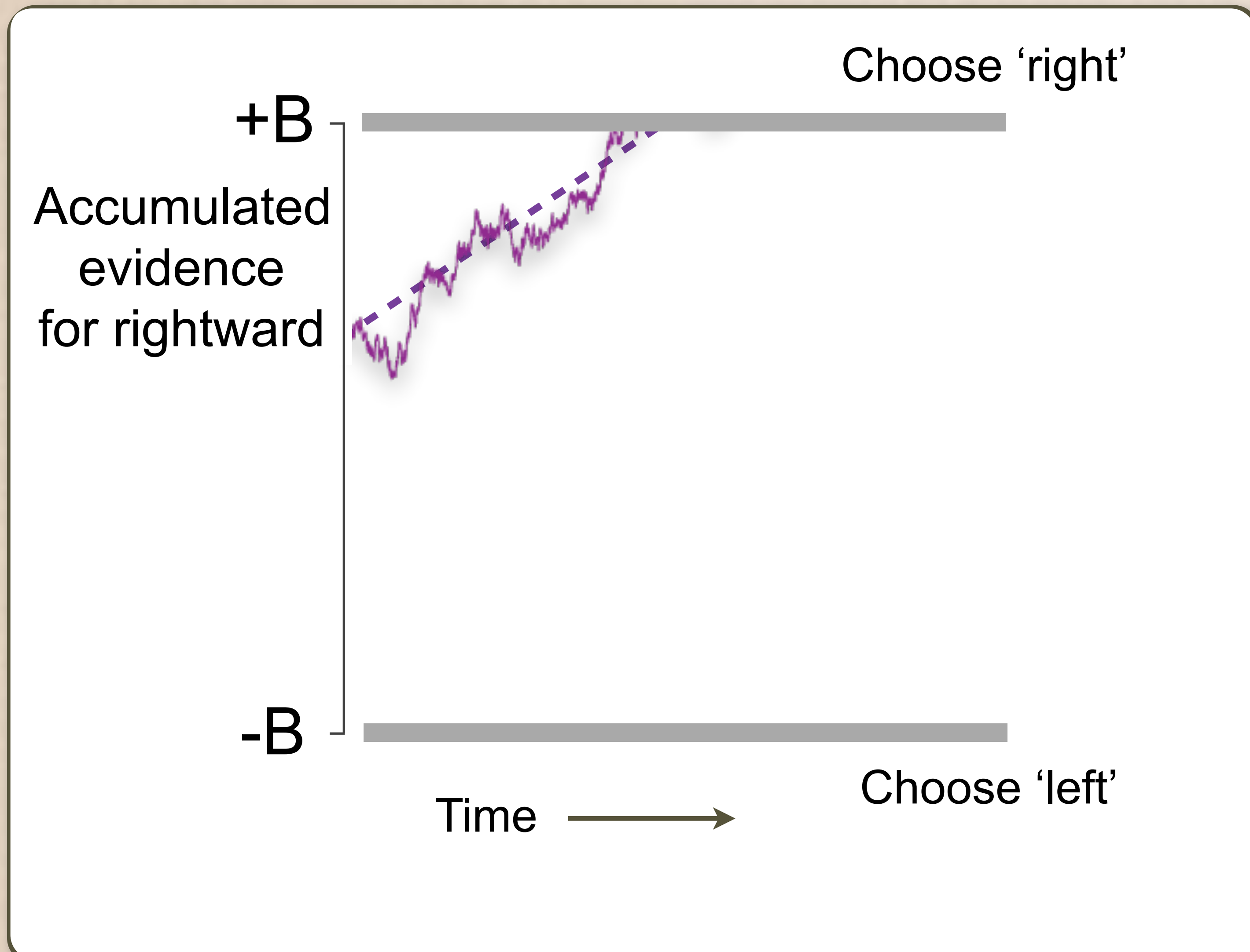


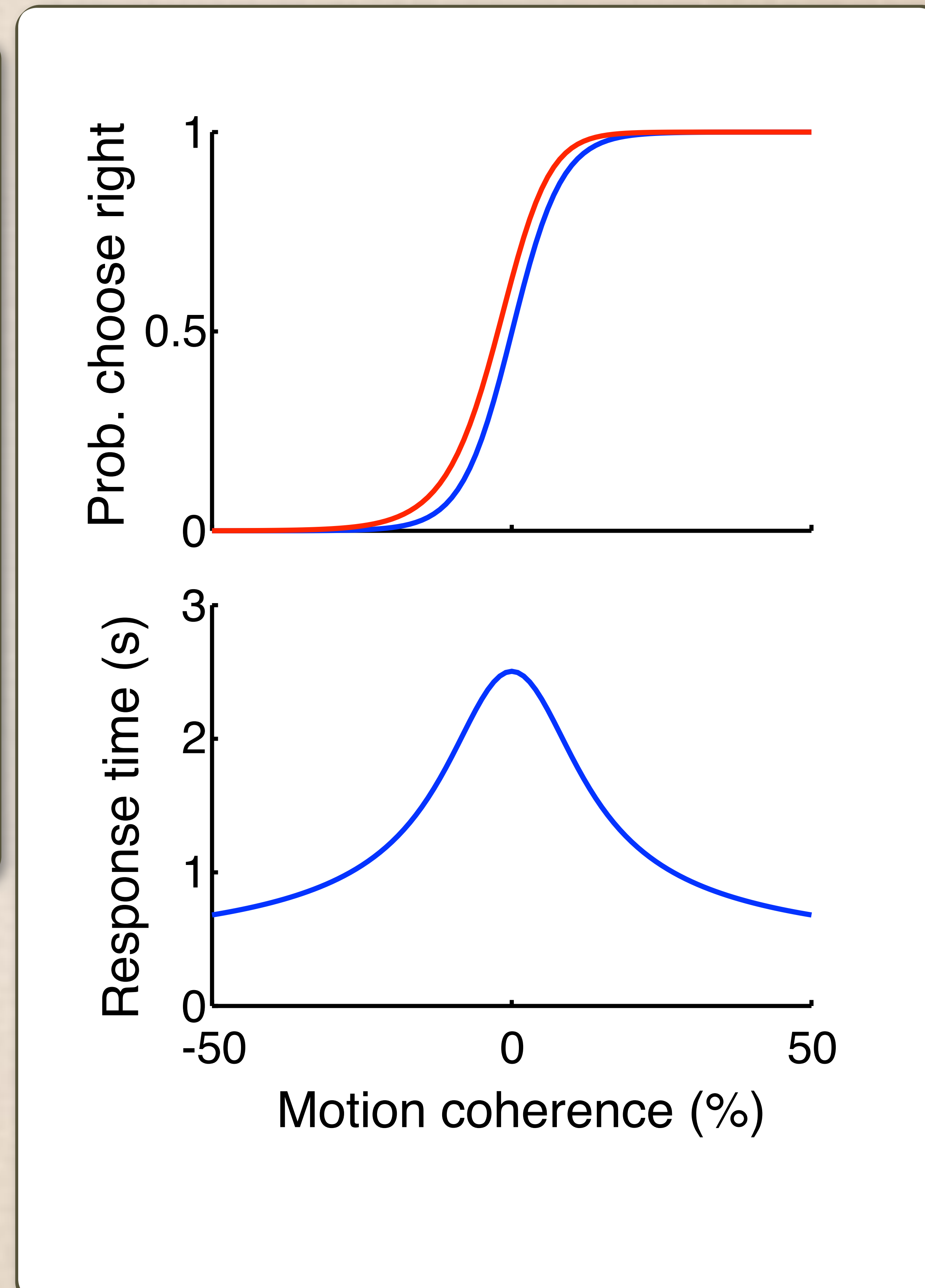
Change in starting point



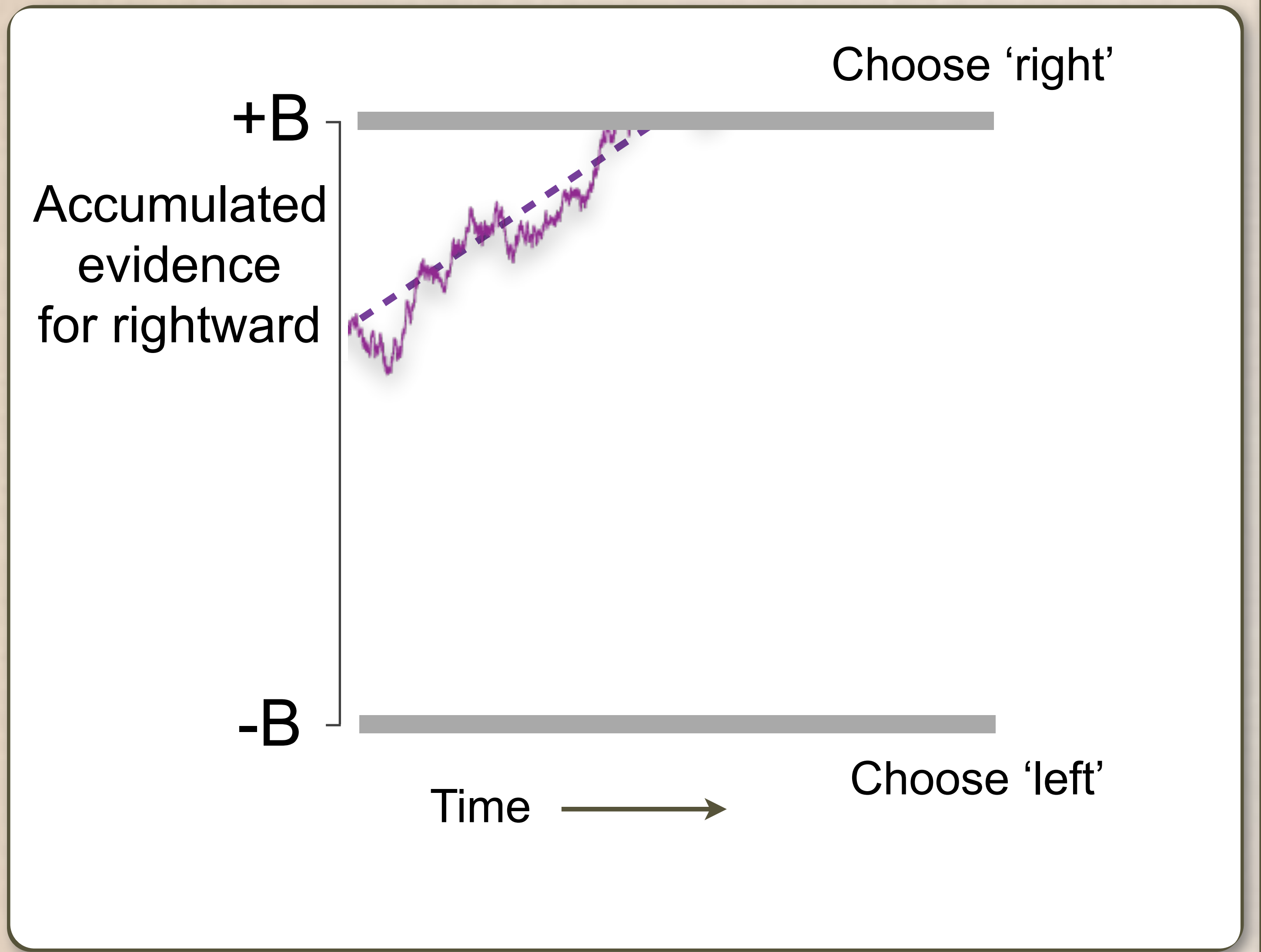


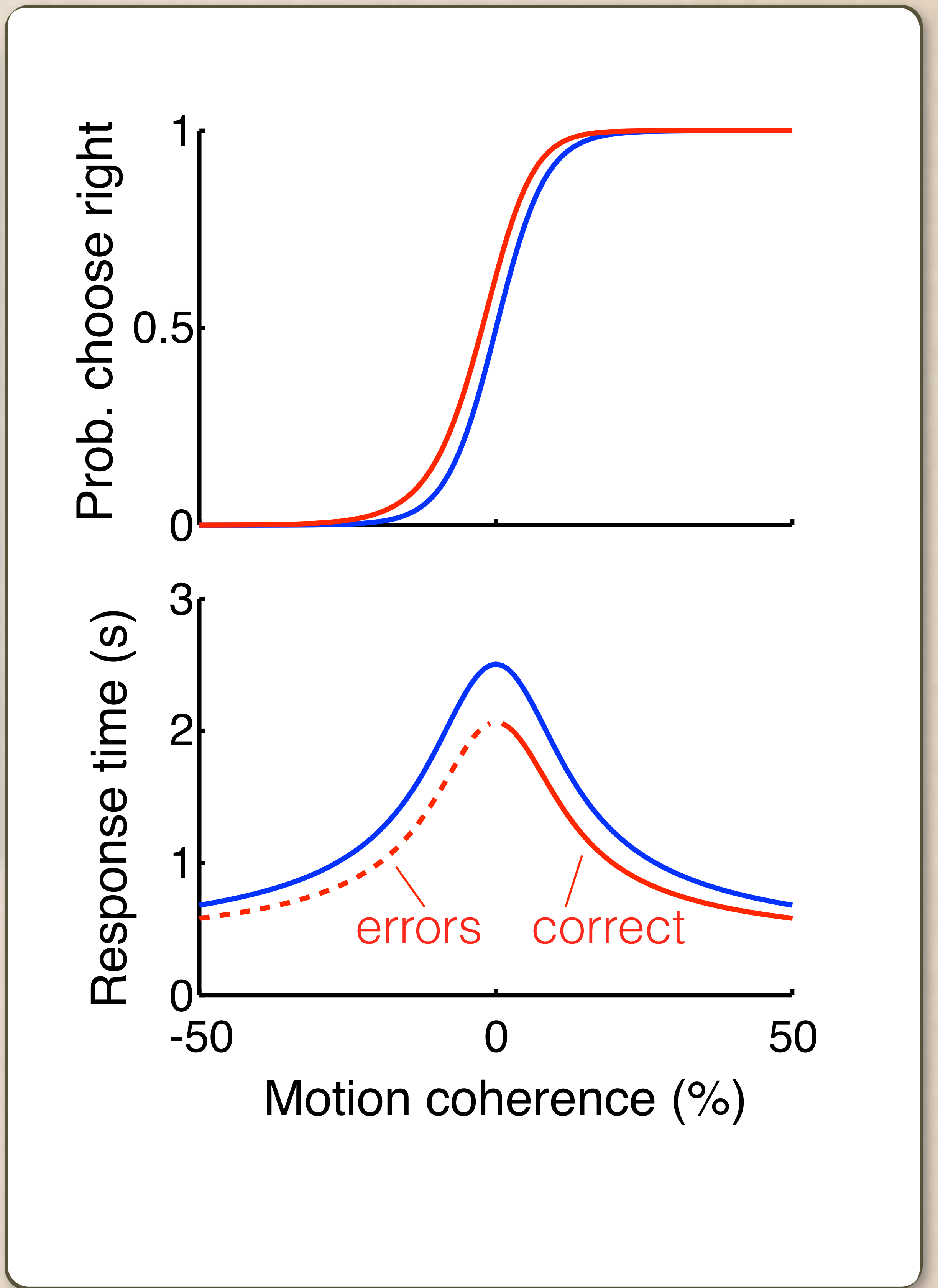
Change in starting point



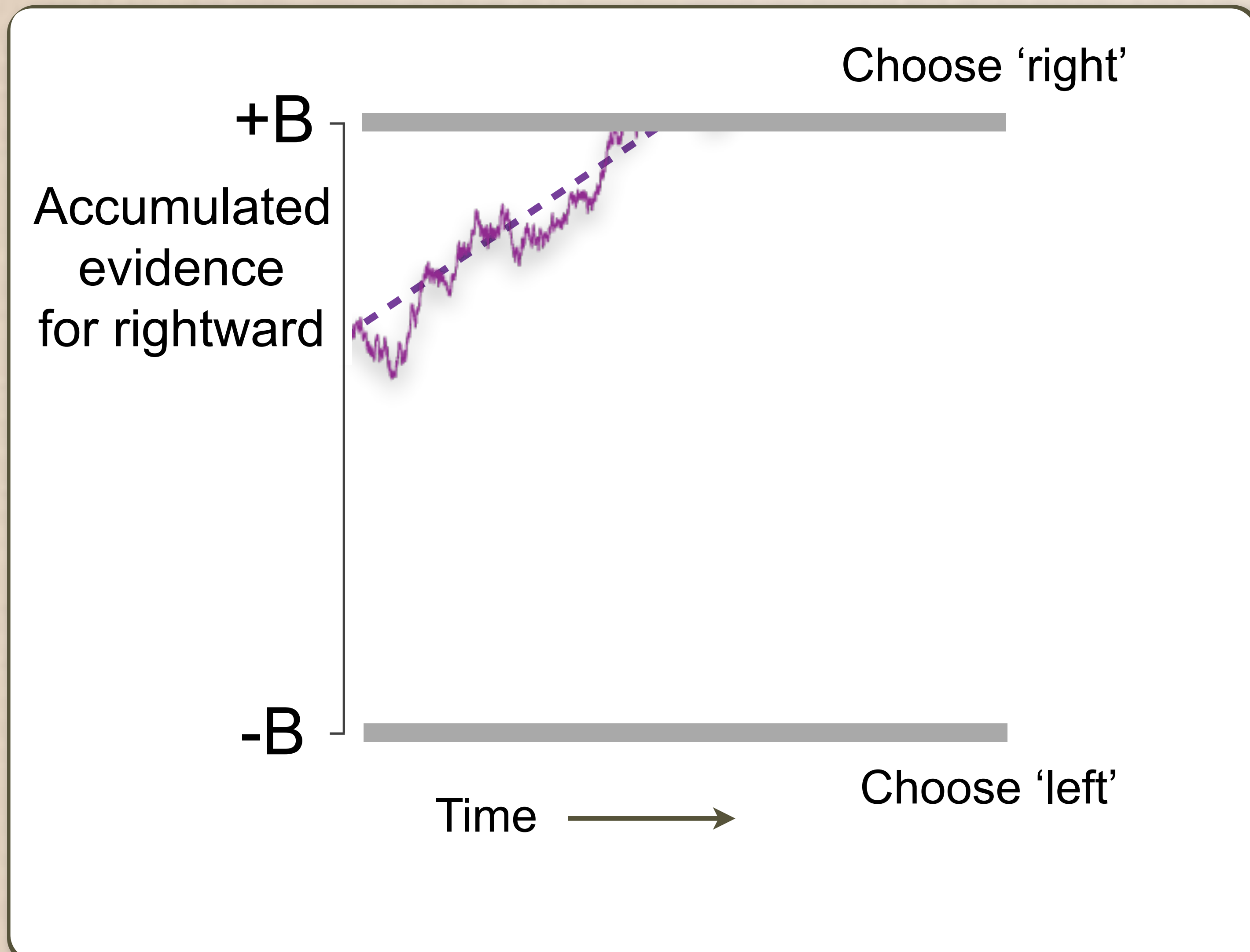


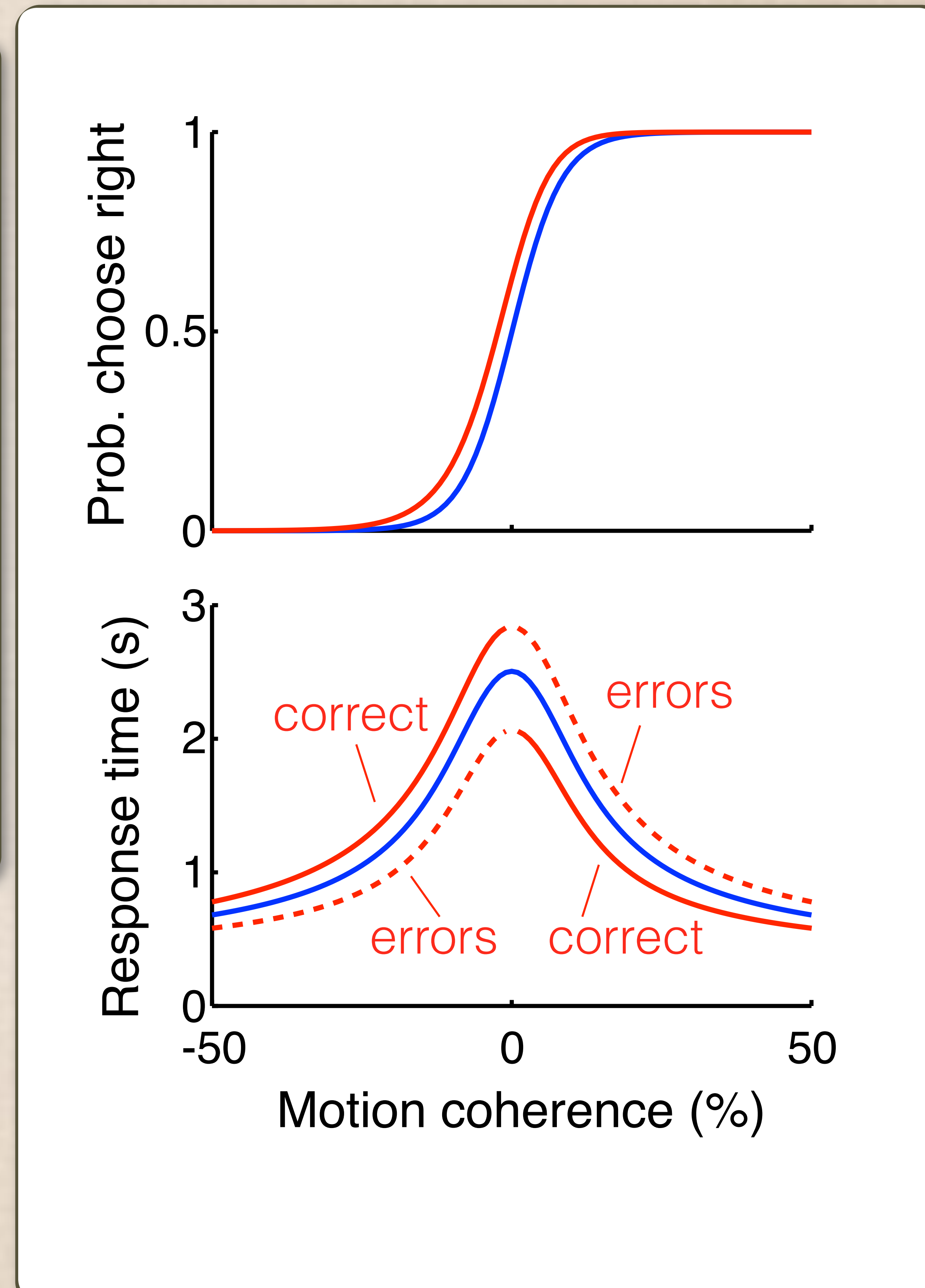
Change in starting point



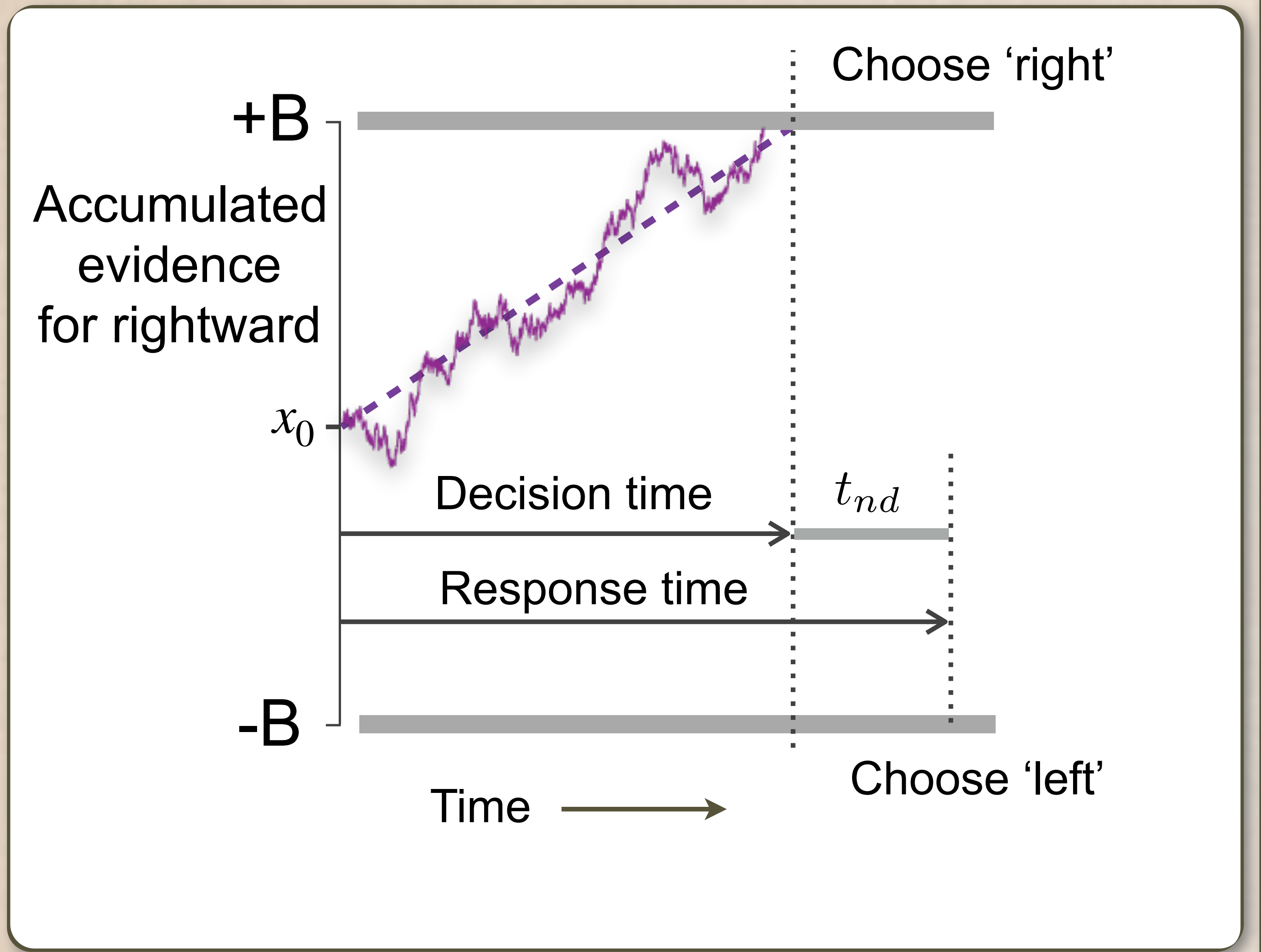


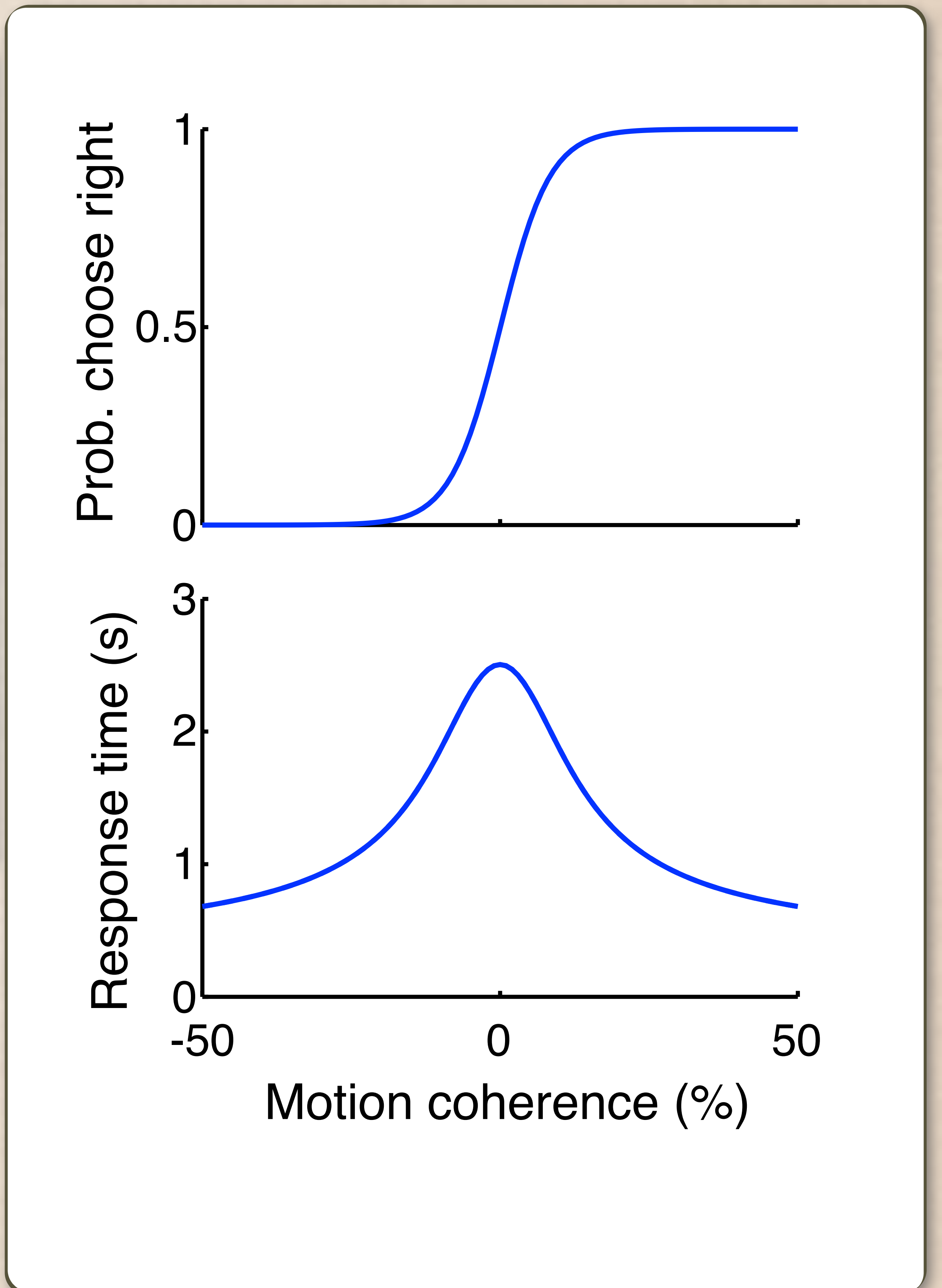
Change in starting point



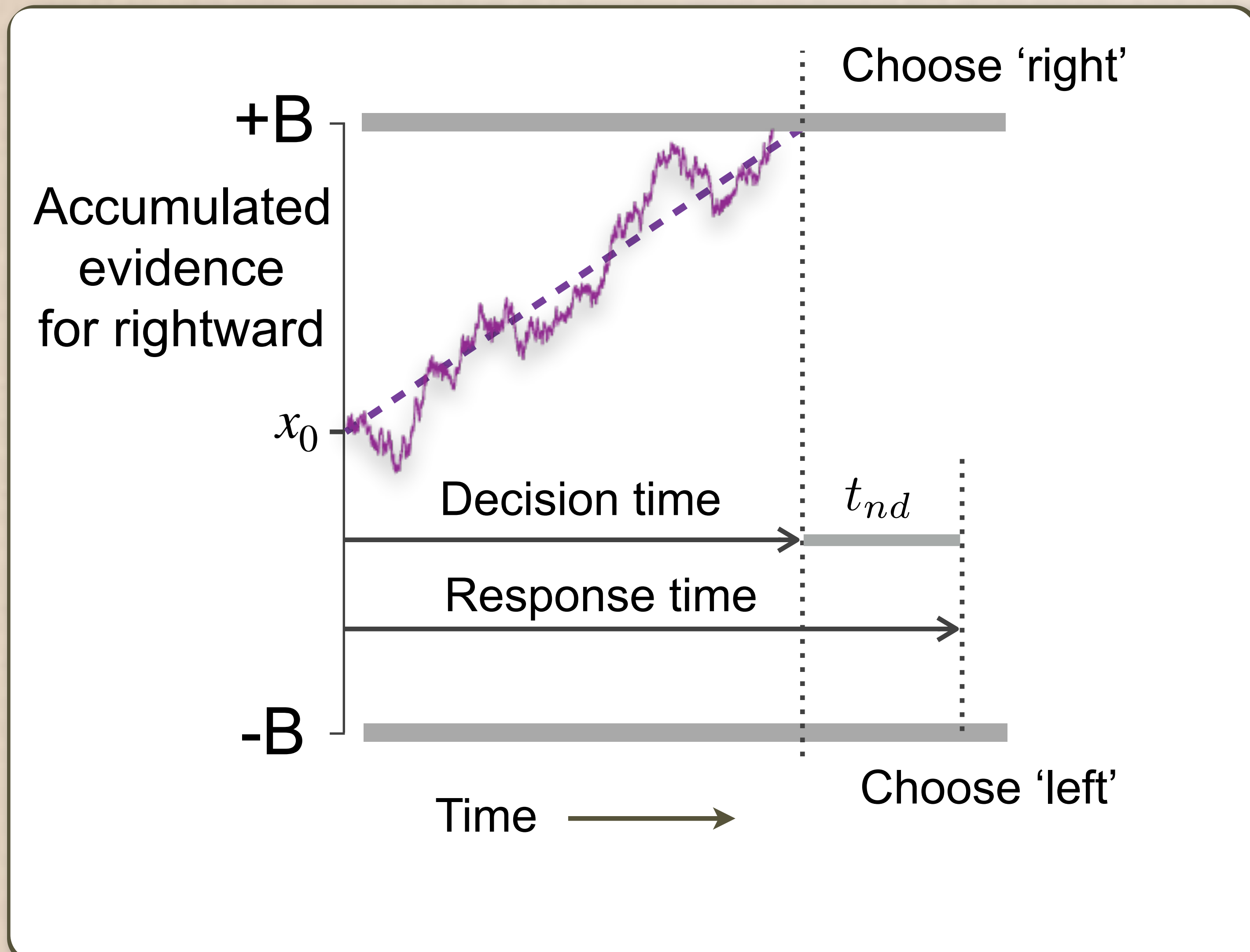


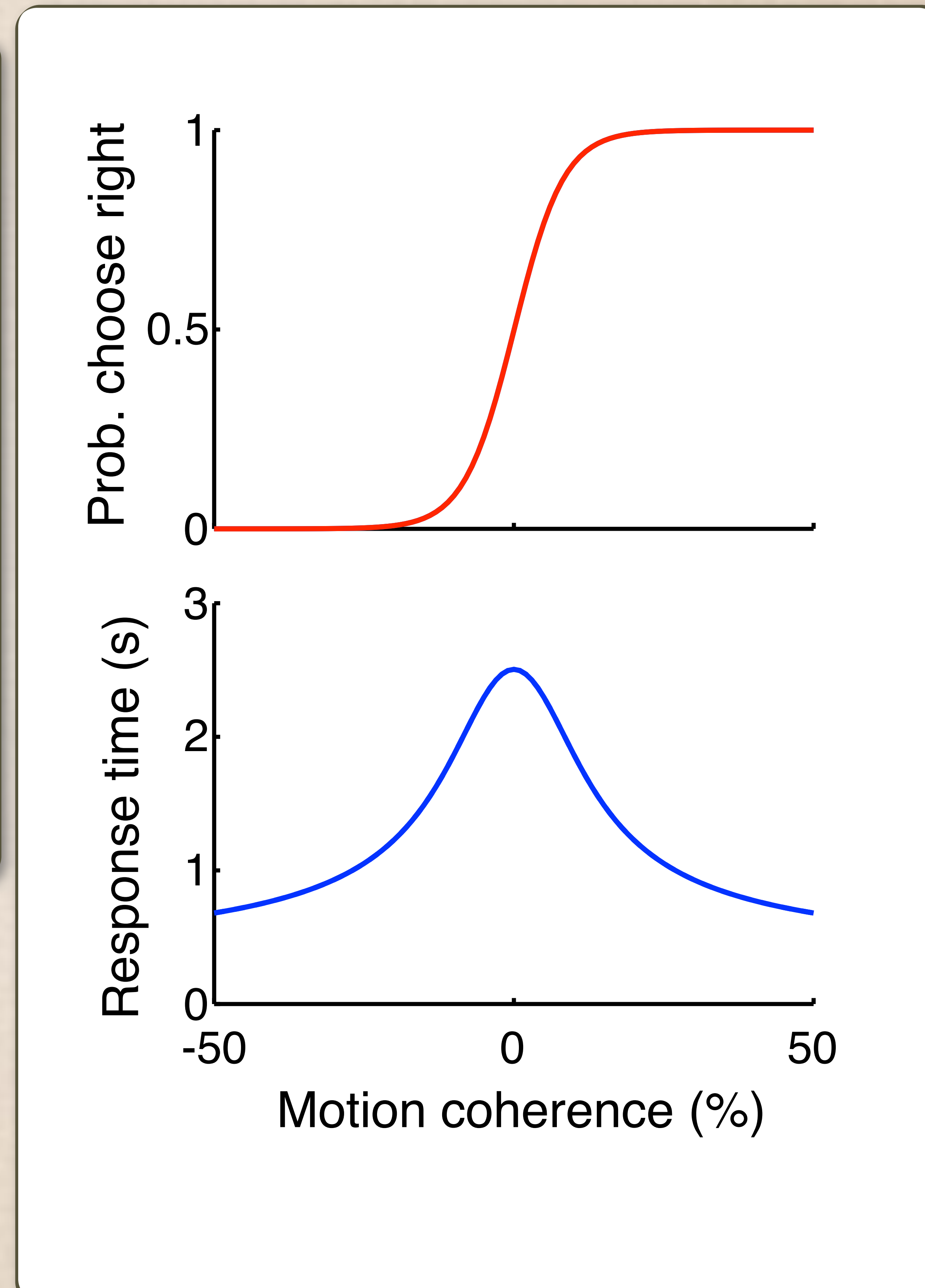
Change in non-decision time



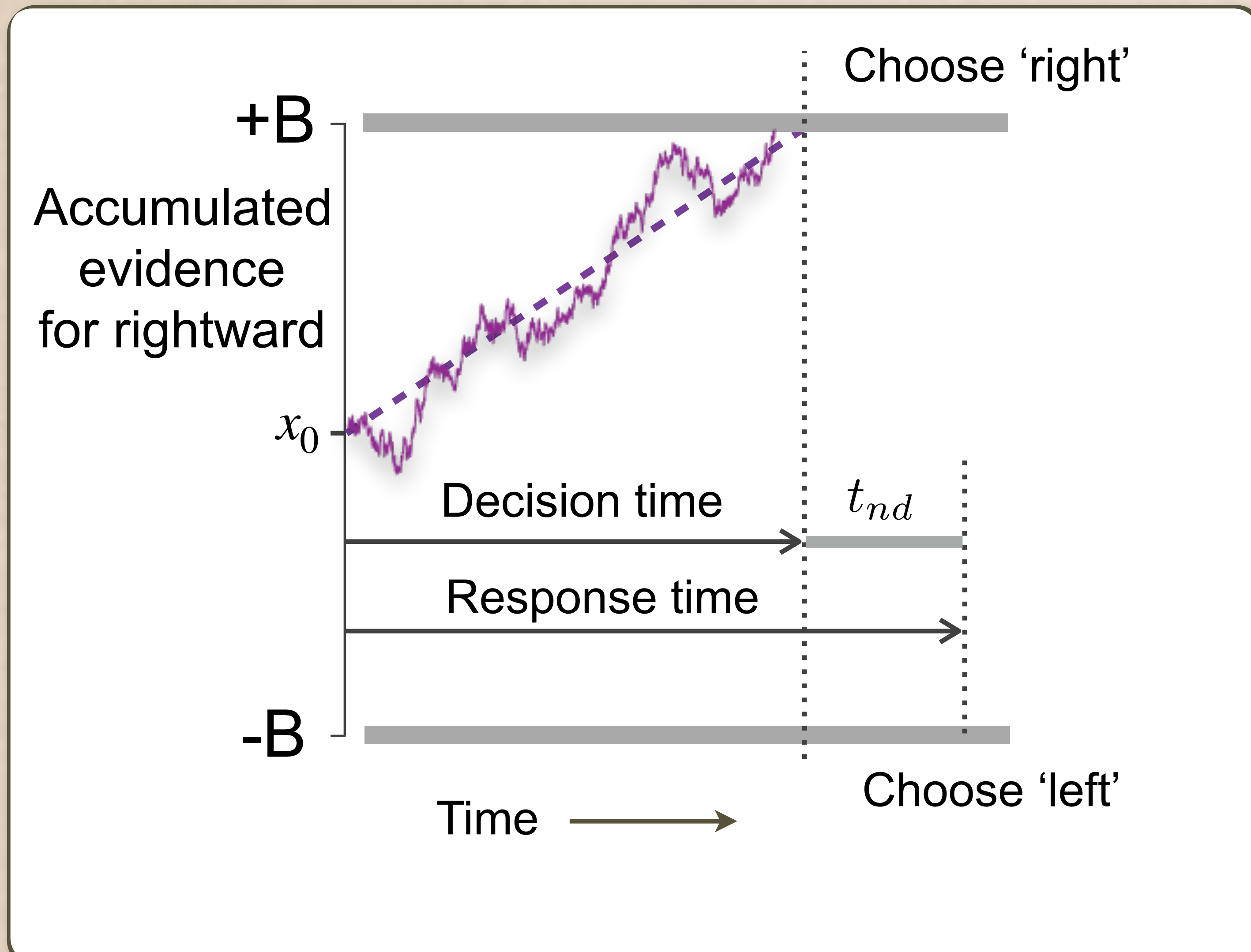


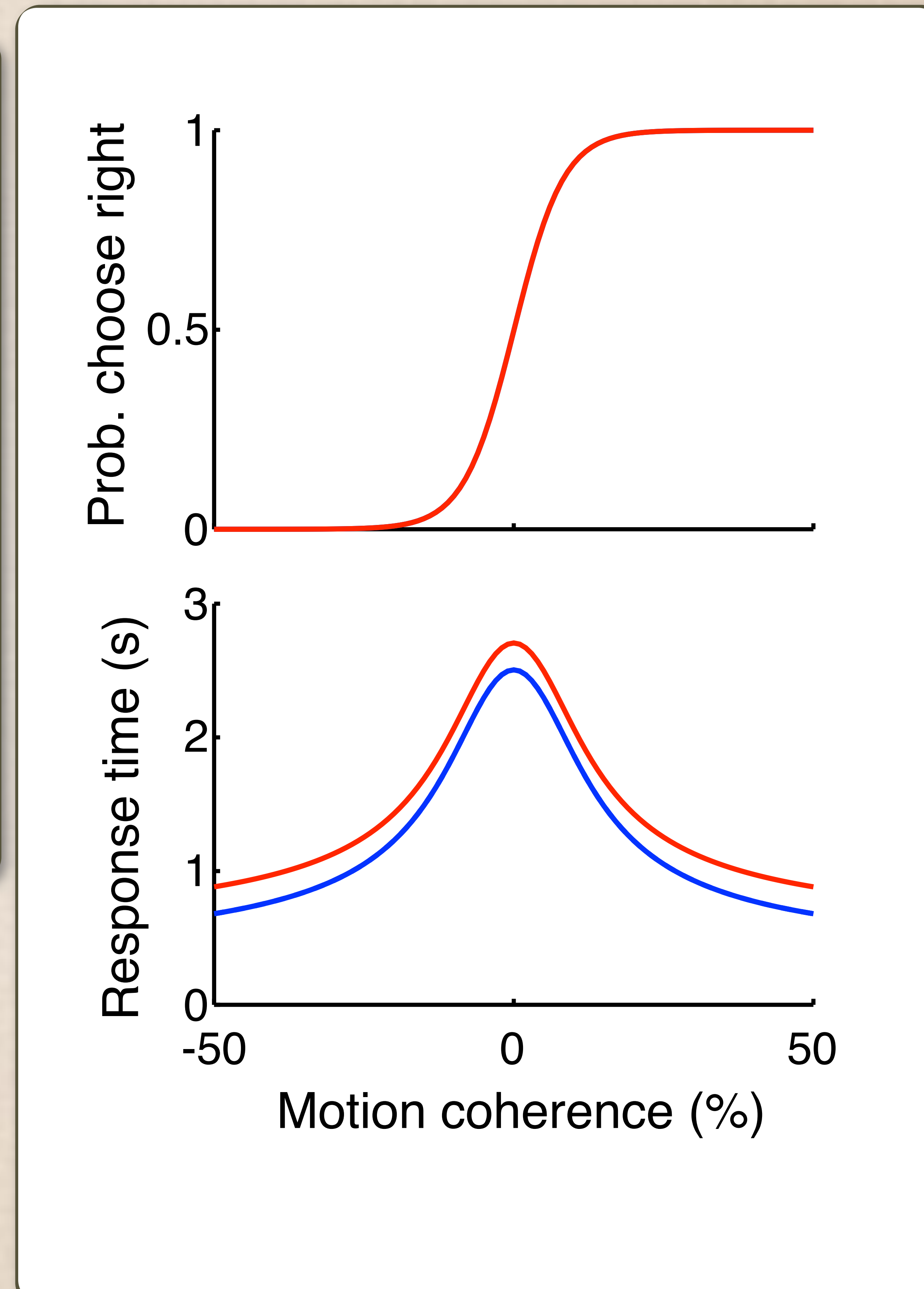
Change in non-decision time





Change in non-decision time

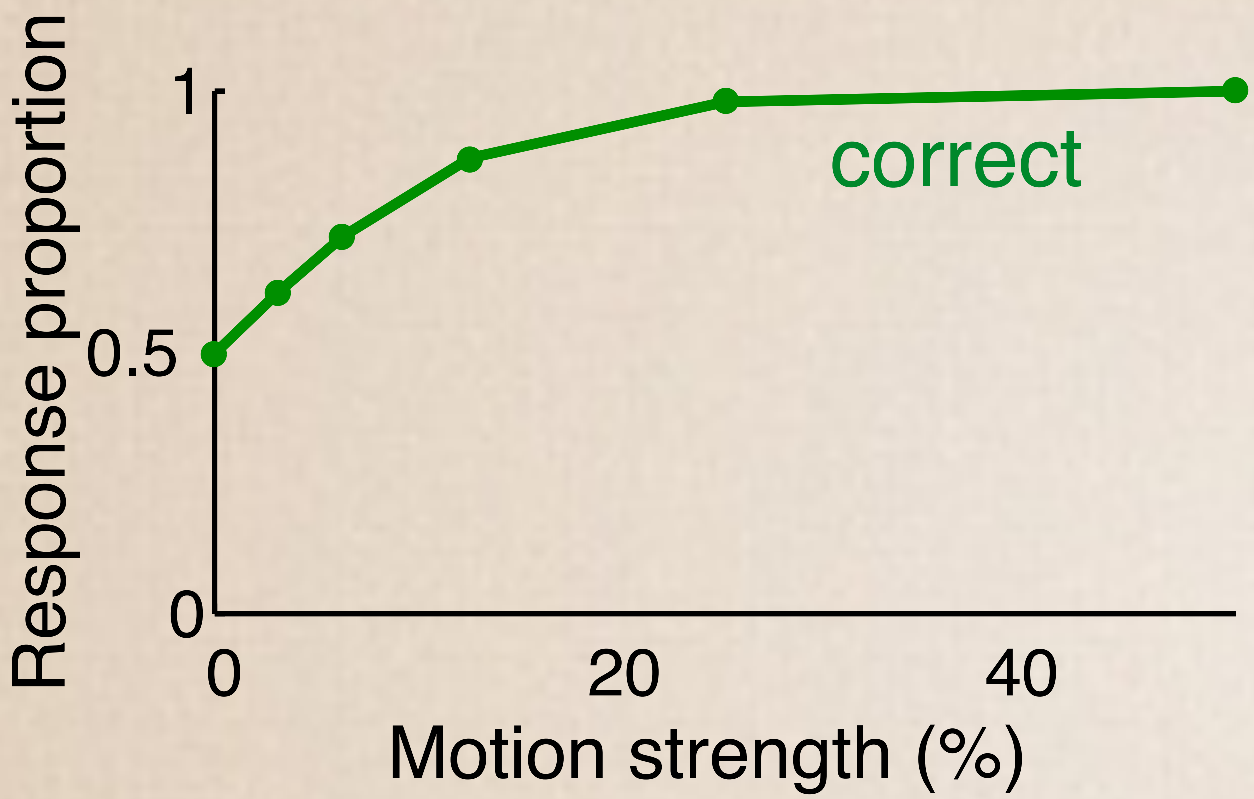




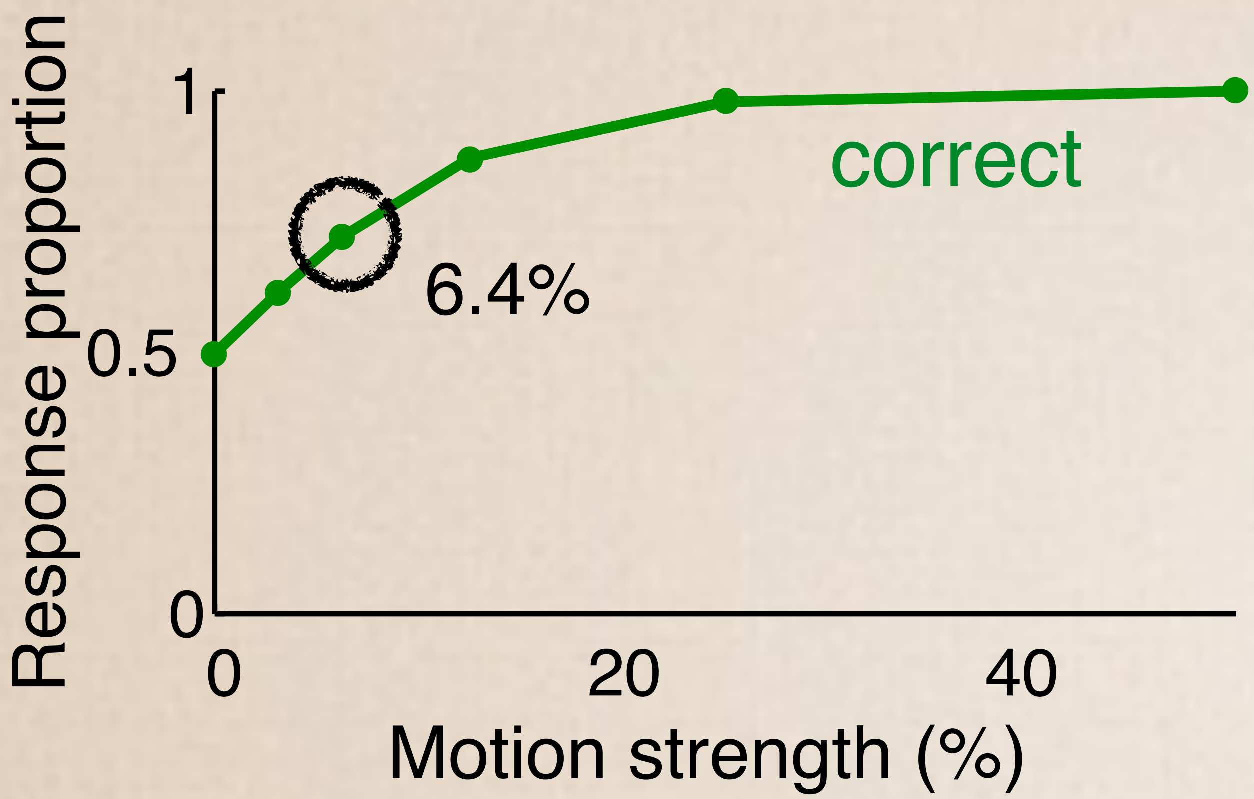
Quantile probability plot

Ratcliff et al., *Psychology and aging* (2001)

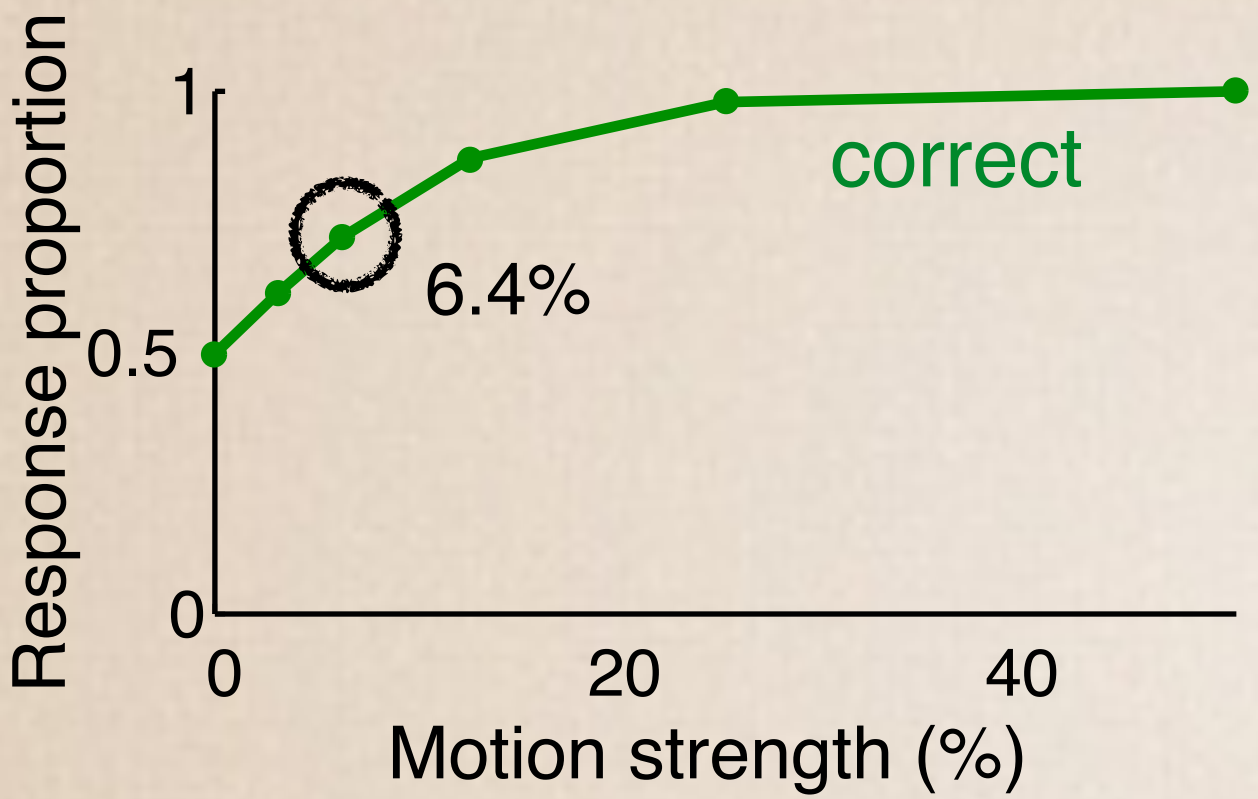
Quantile probability plot

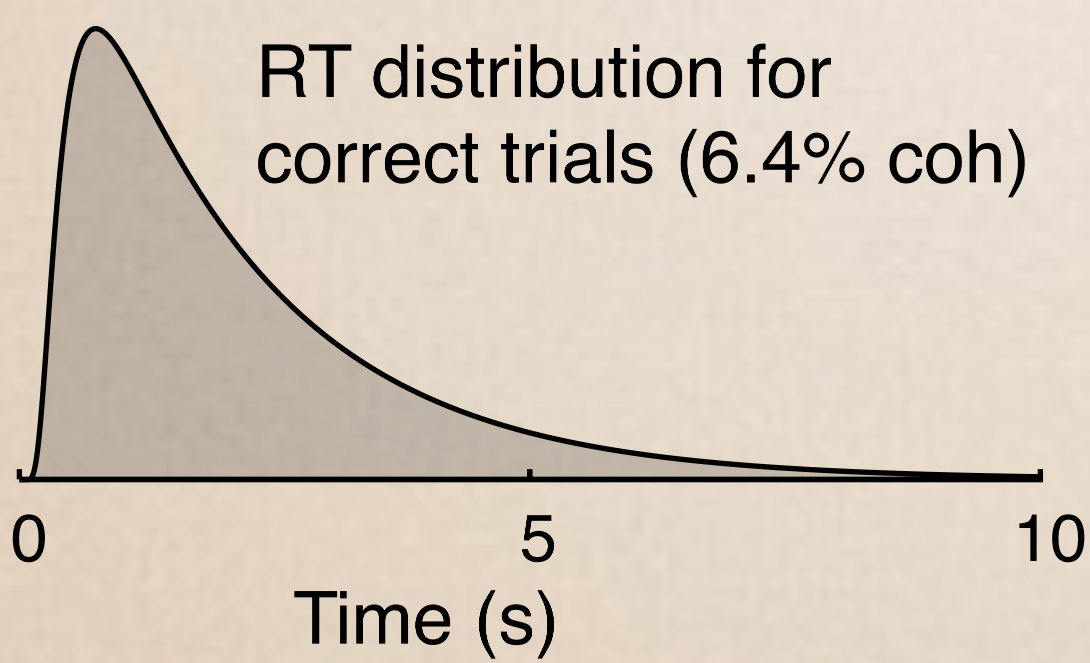


Quantile probability plot

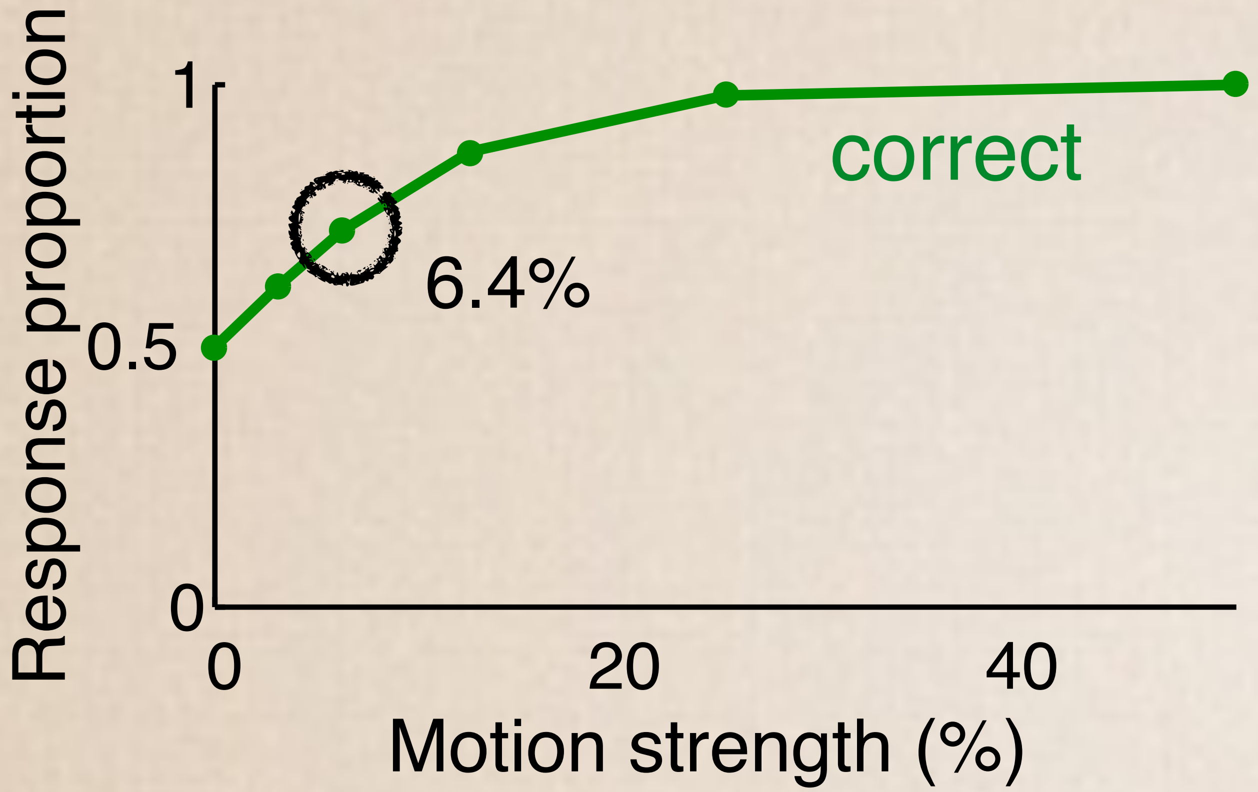


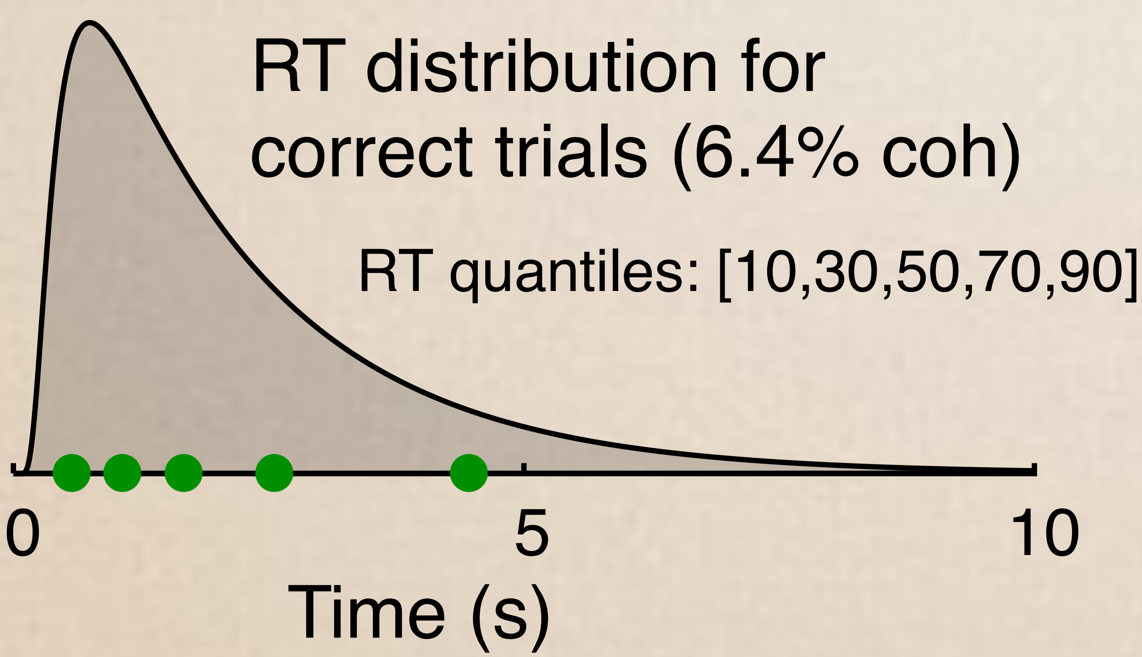
Quantile probability plot



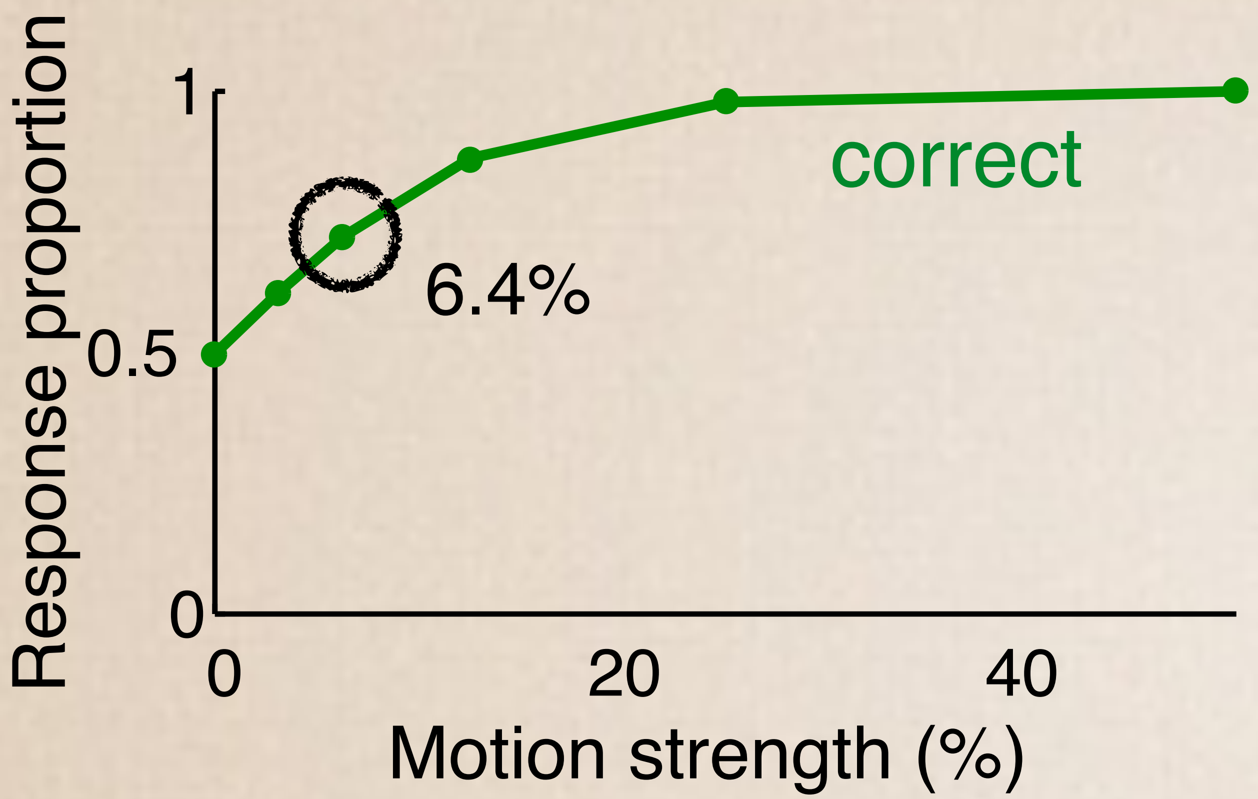


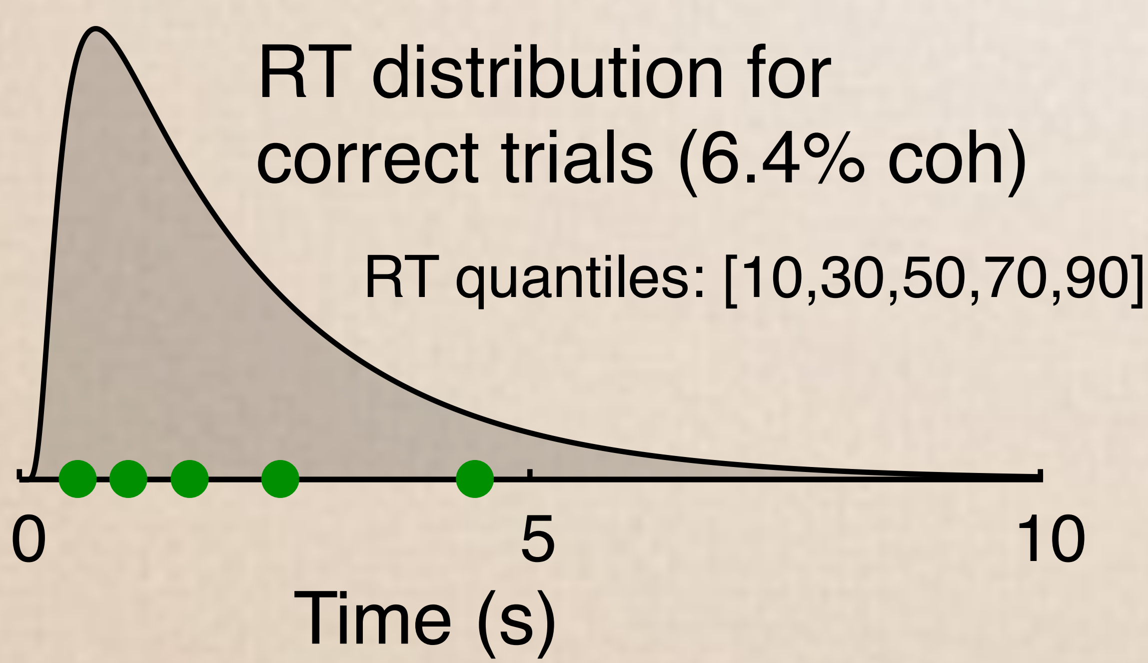
Quantile probability plot

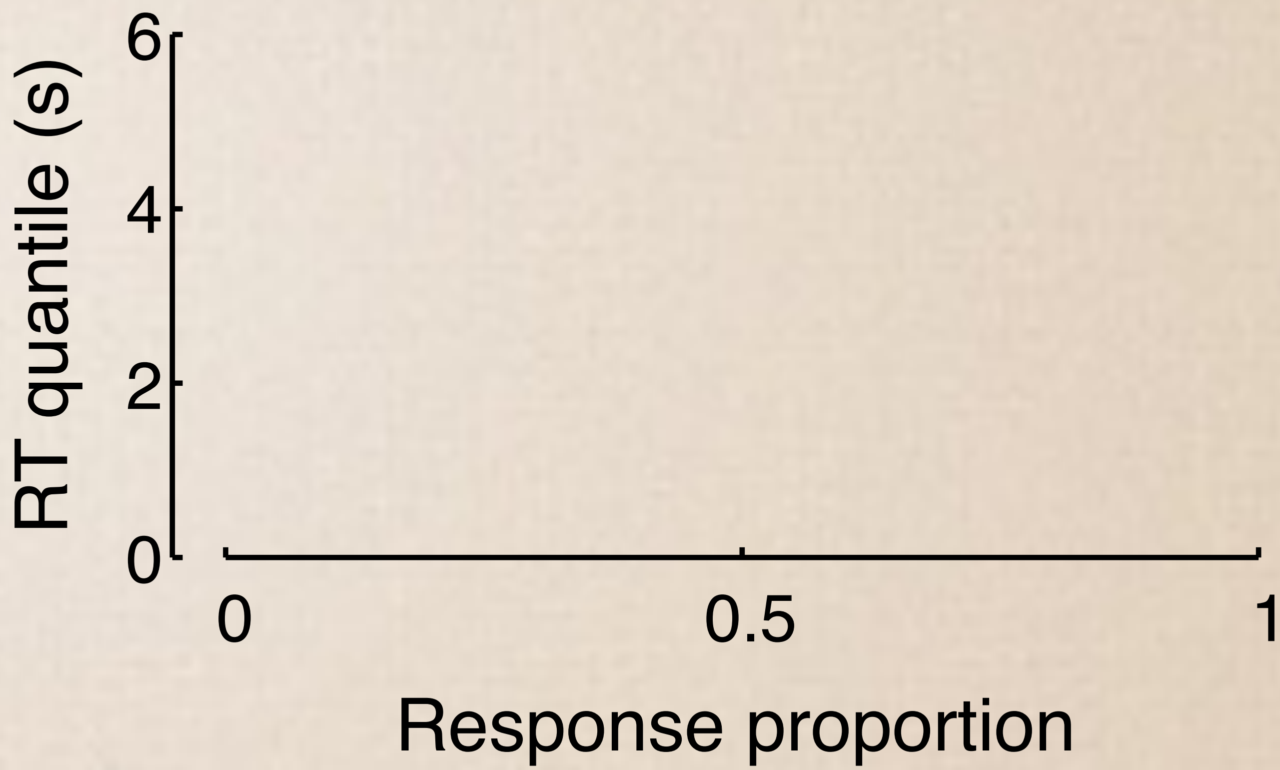




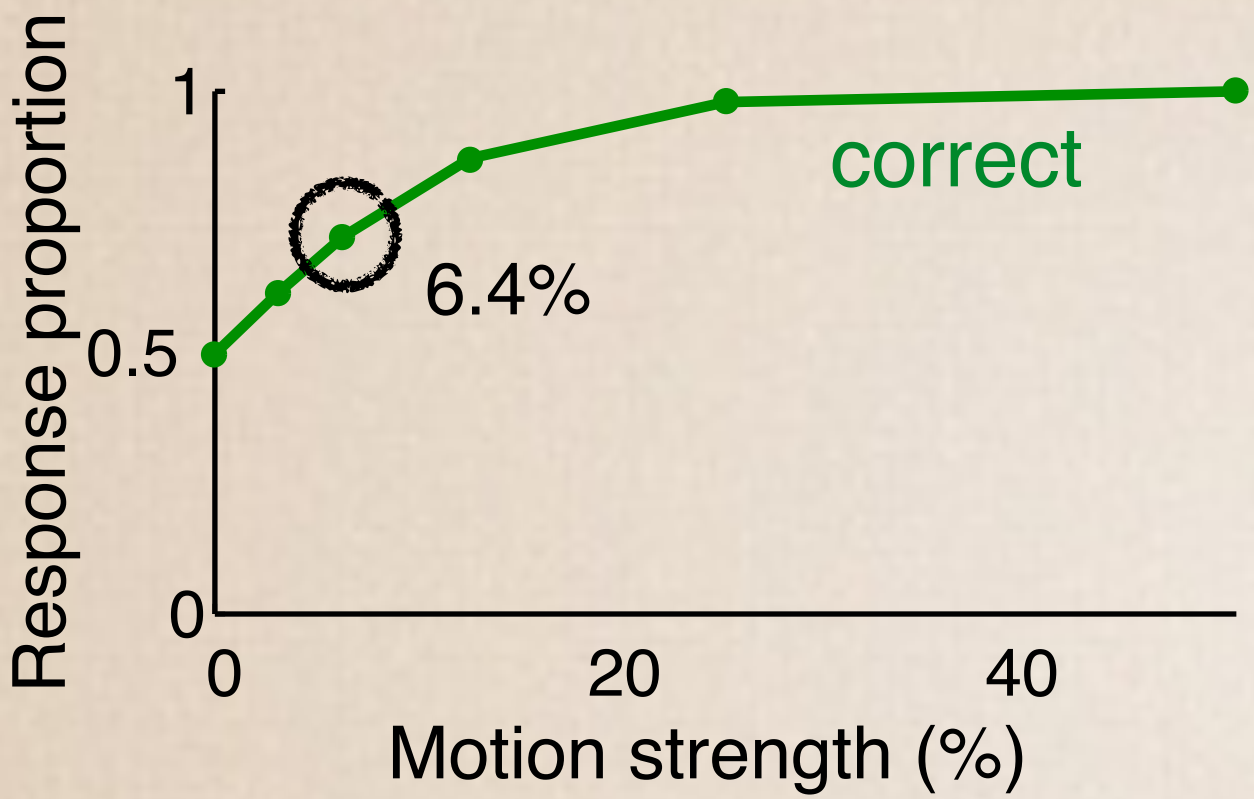
Quantile probability plot

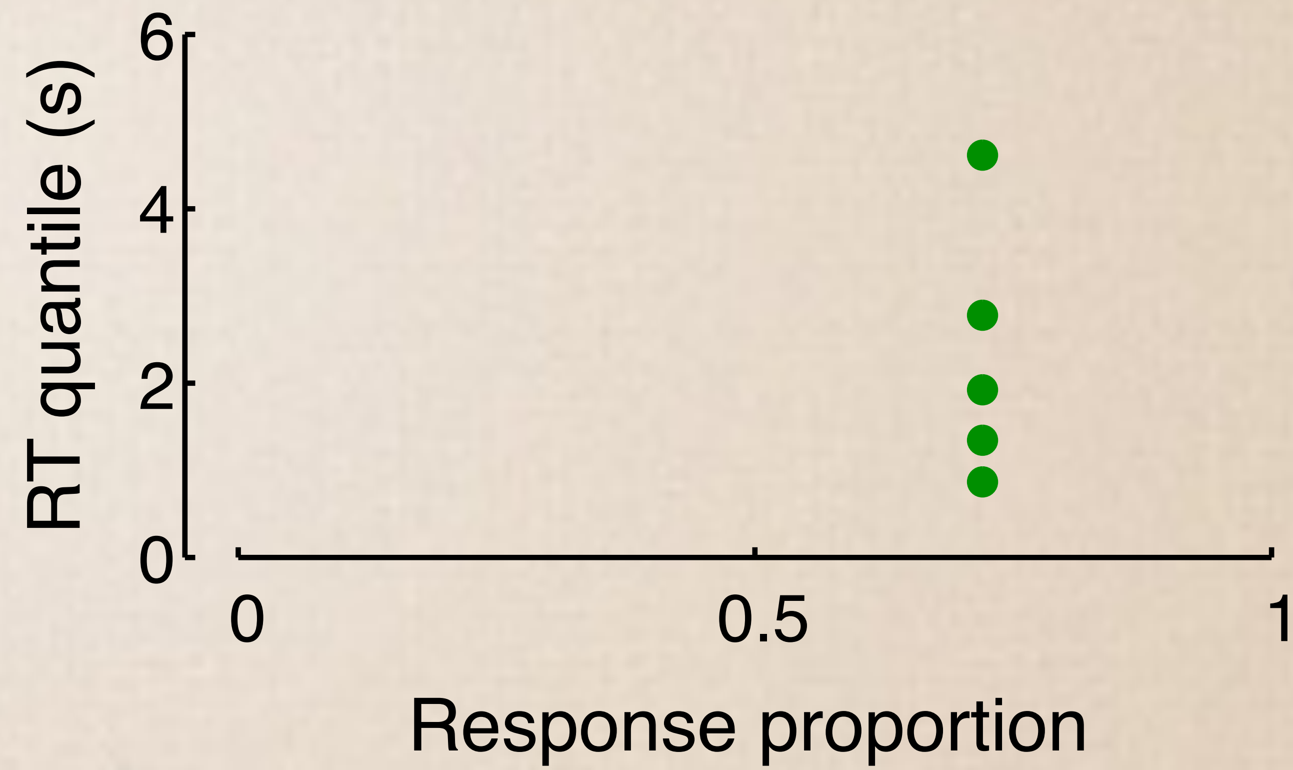


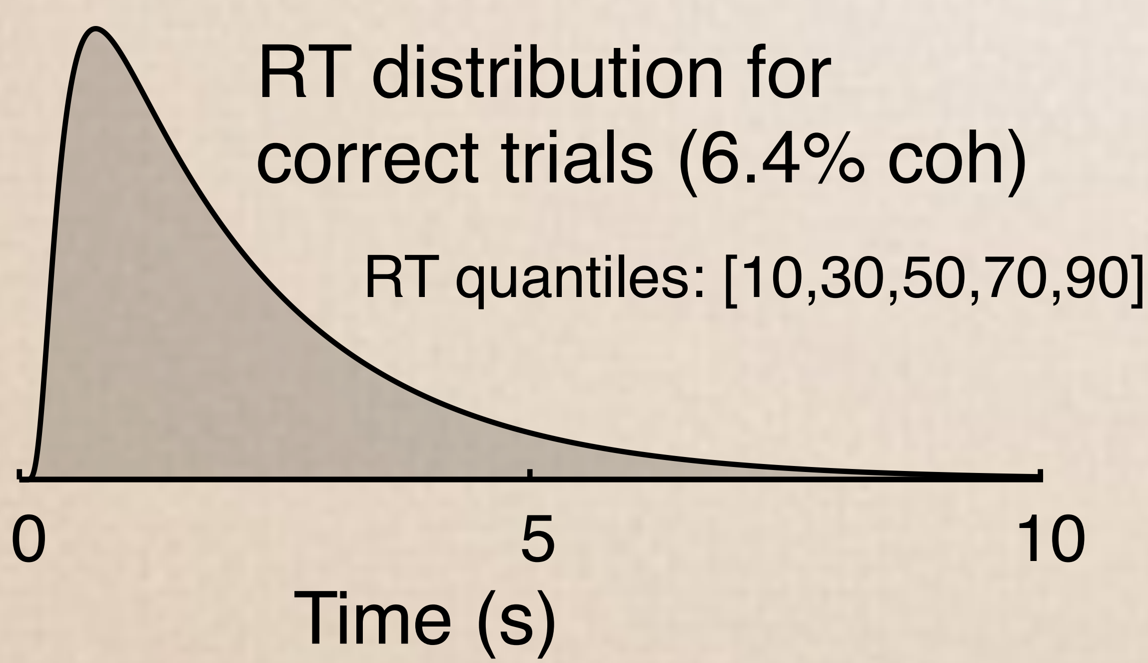




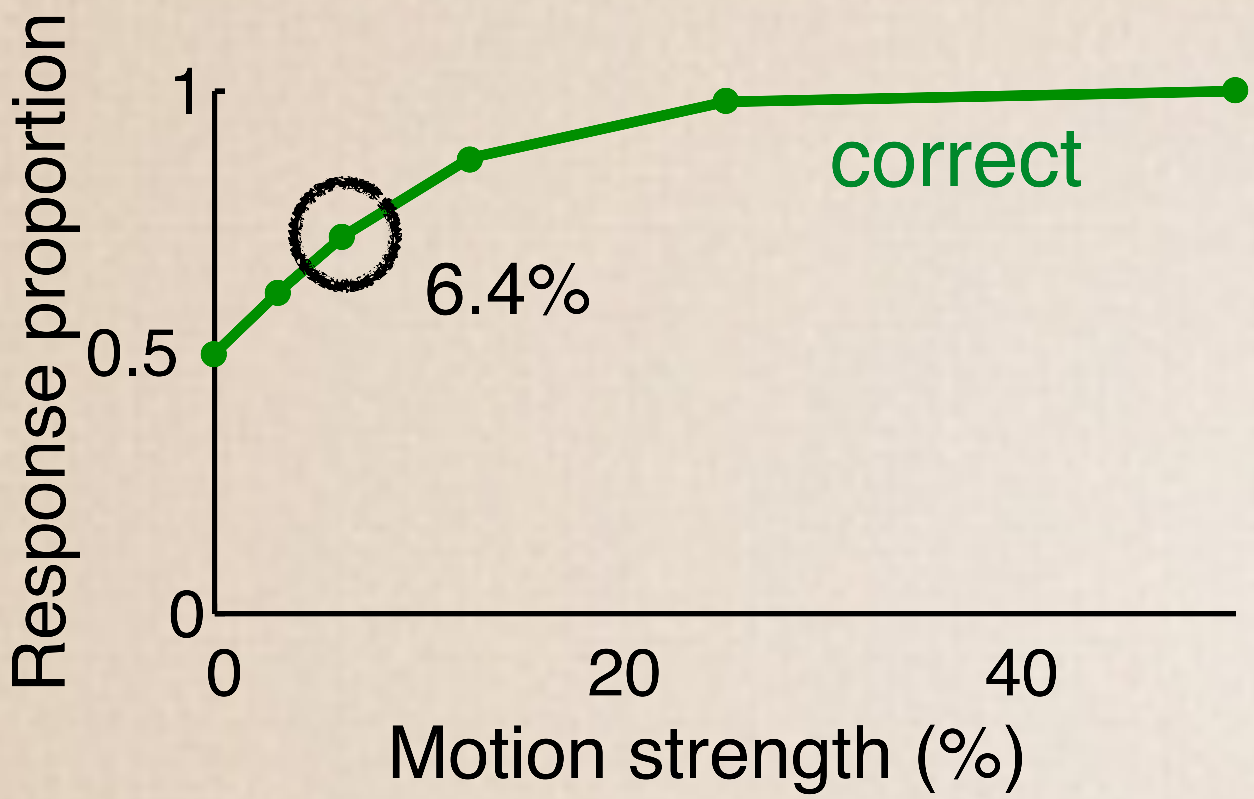
Quantile probability plot

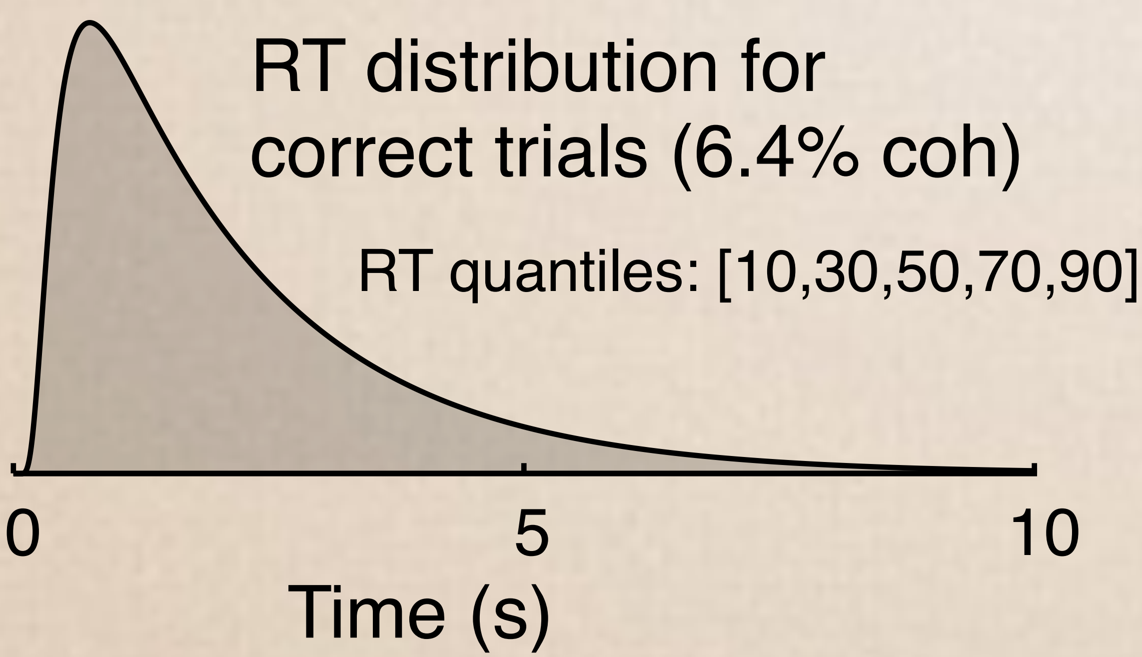


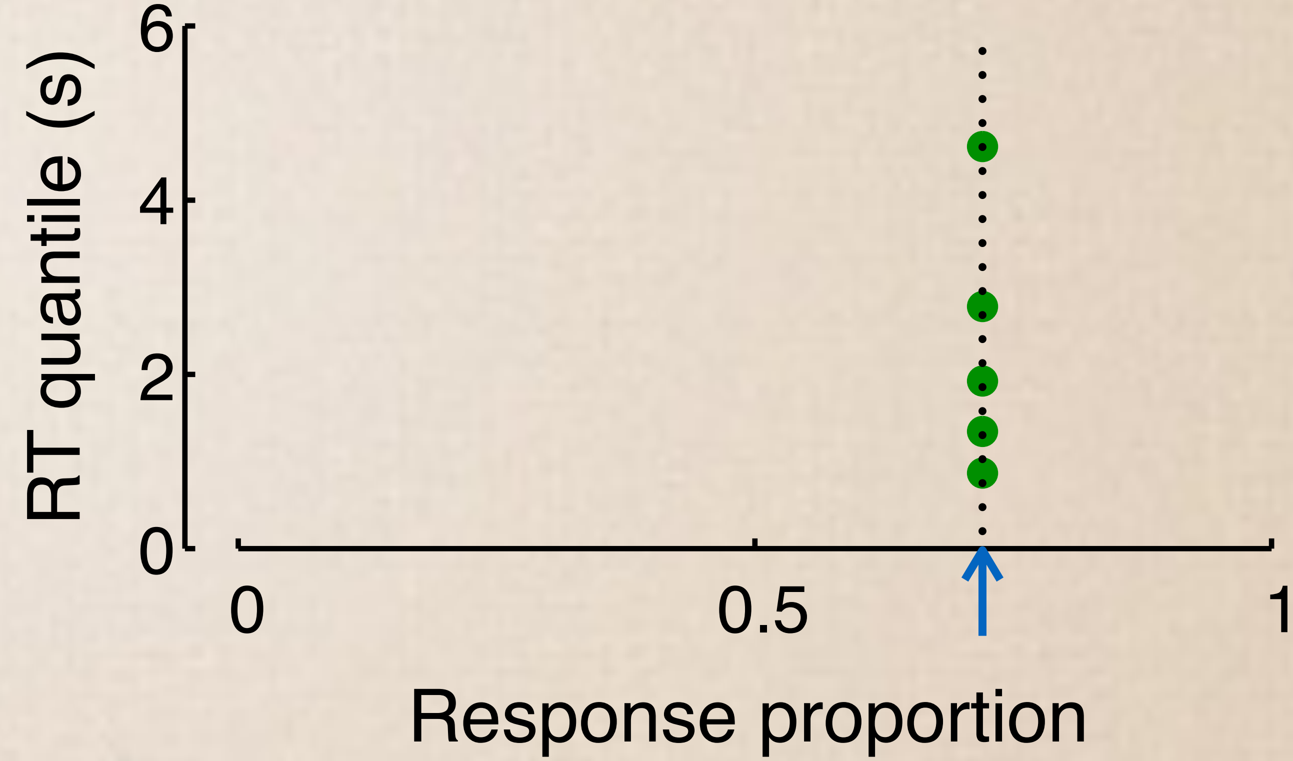




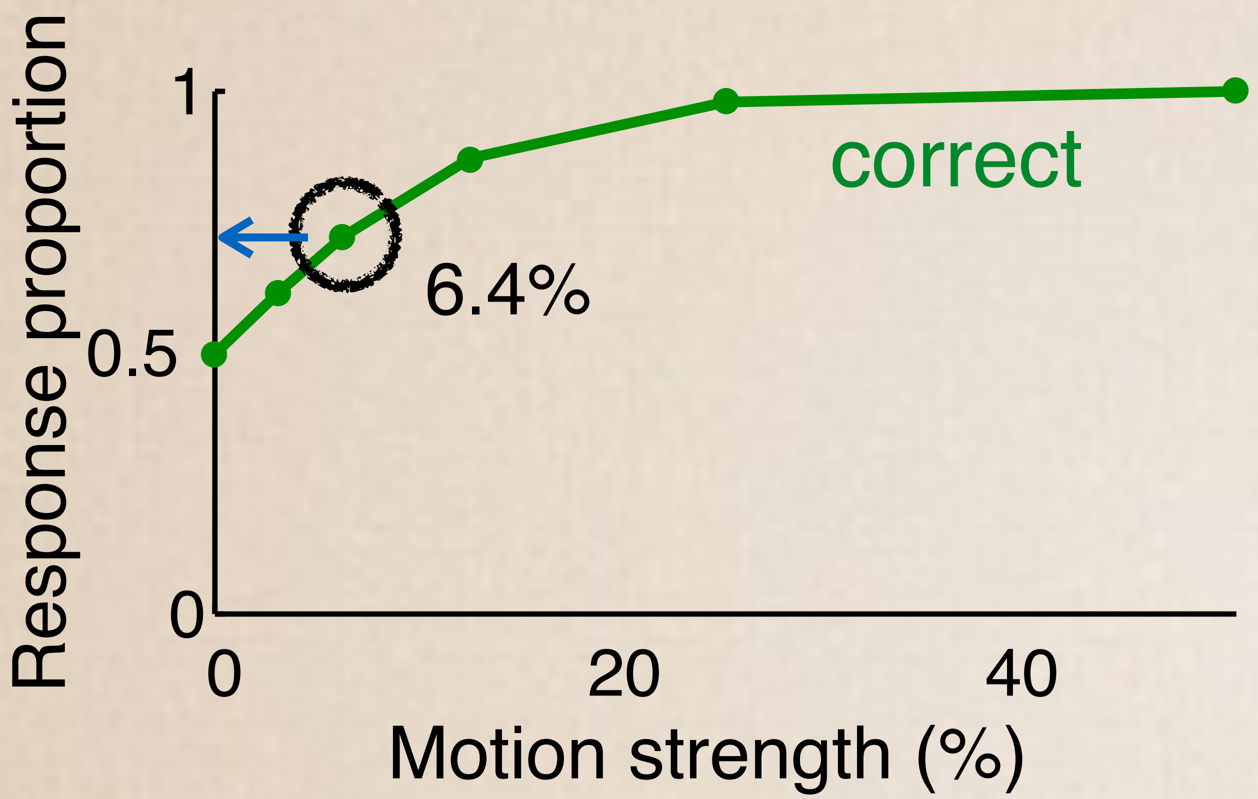
Quantile probability plot

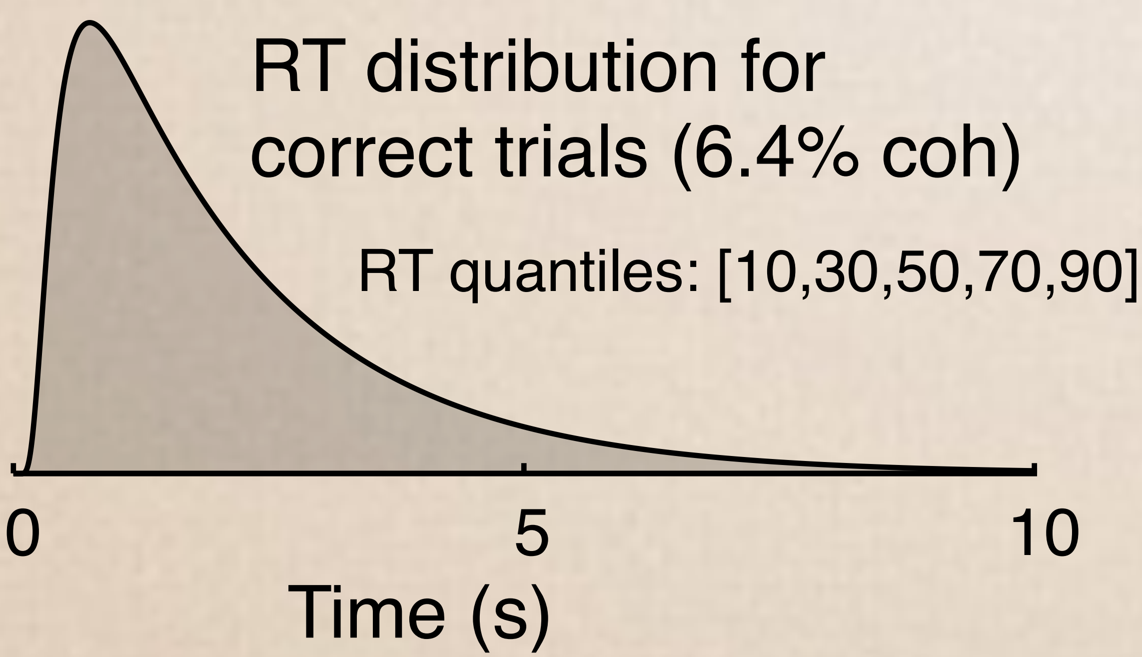


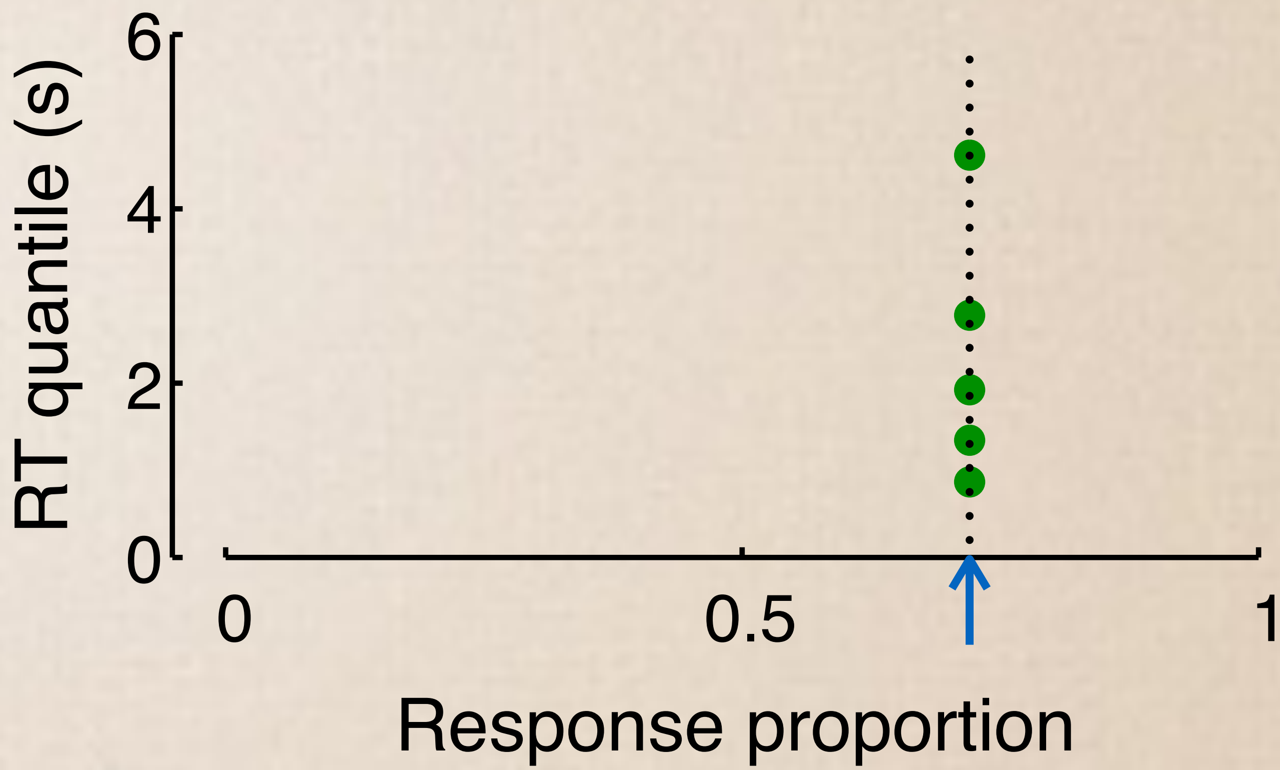




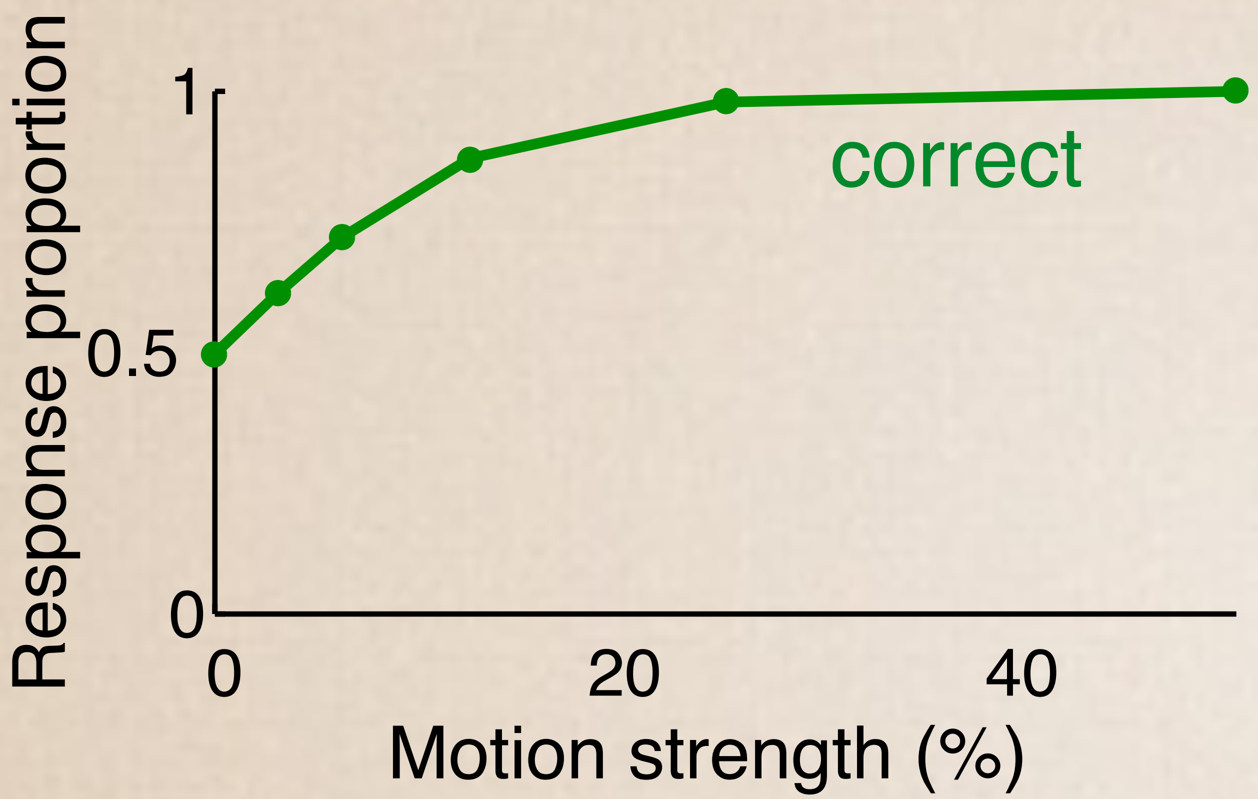
Quantile probability plot

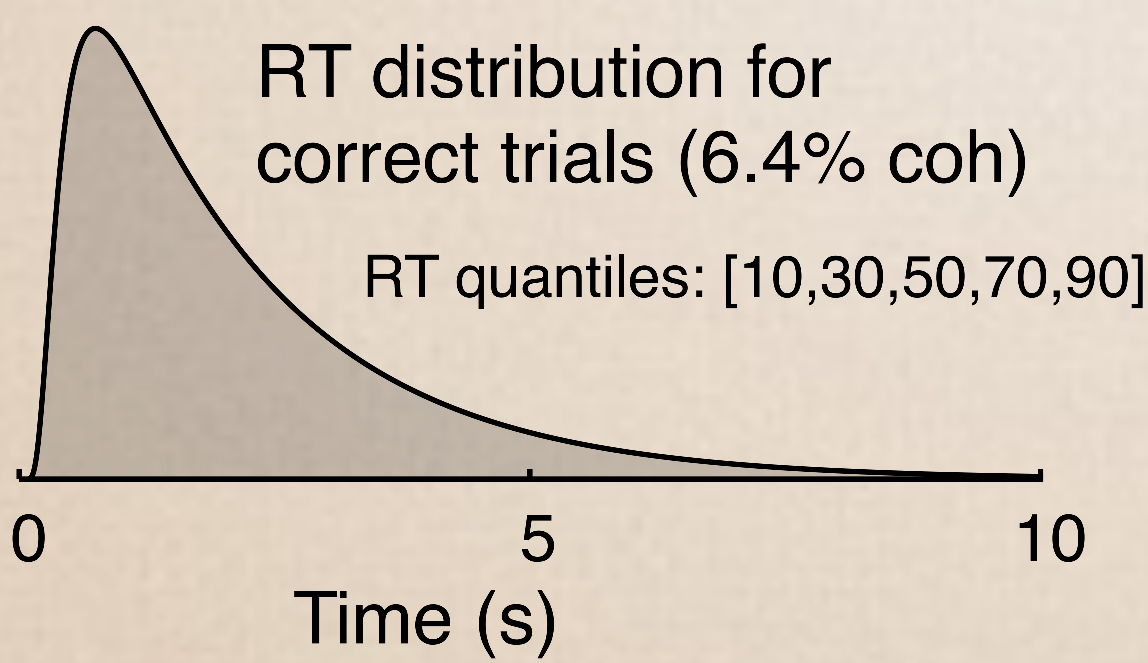




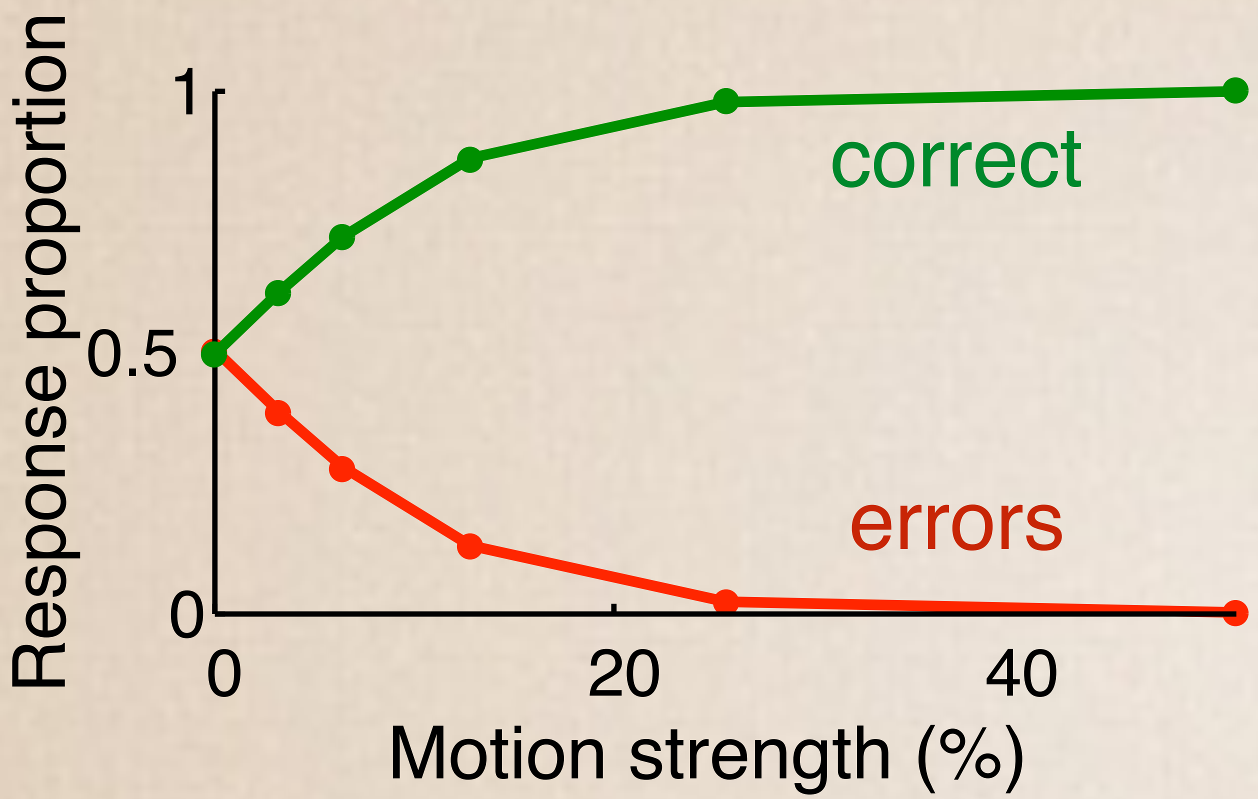


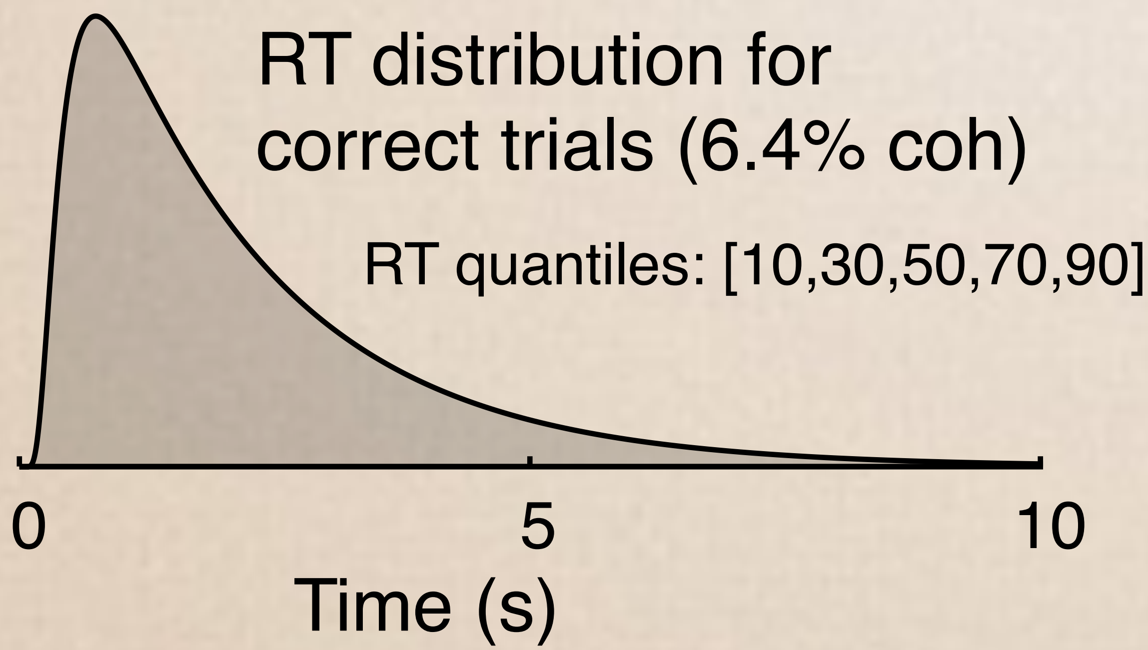
Quantile probability plot



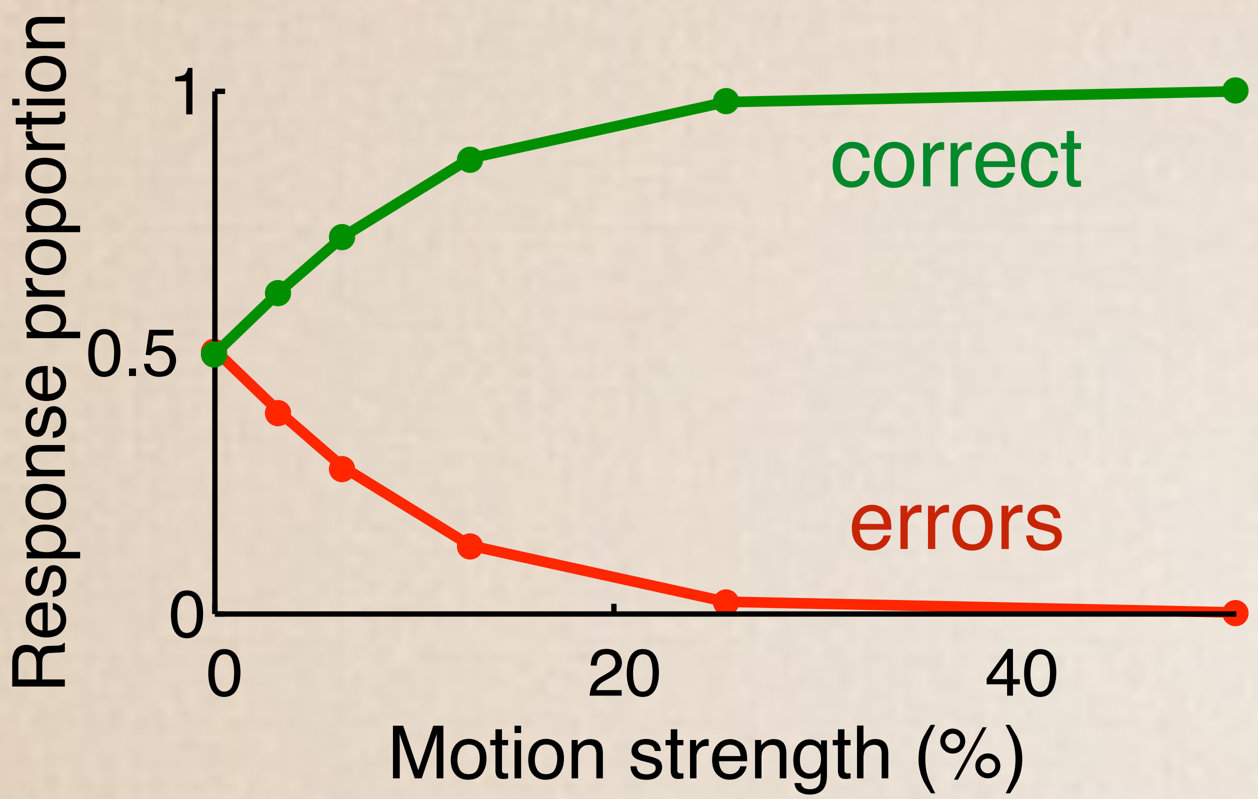


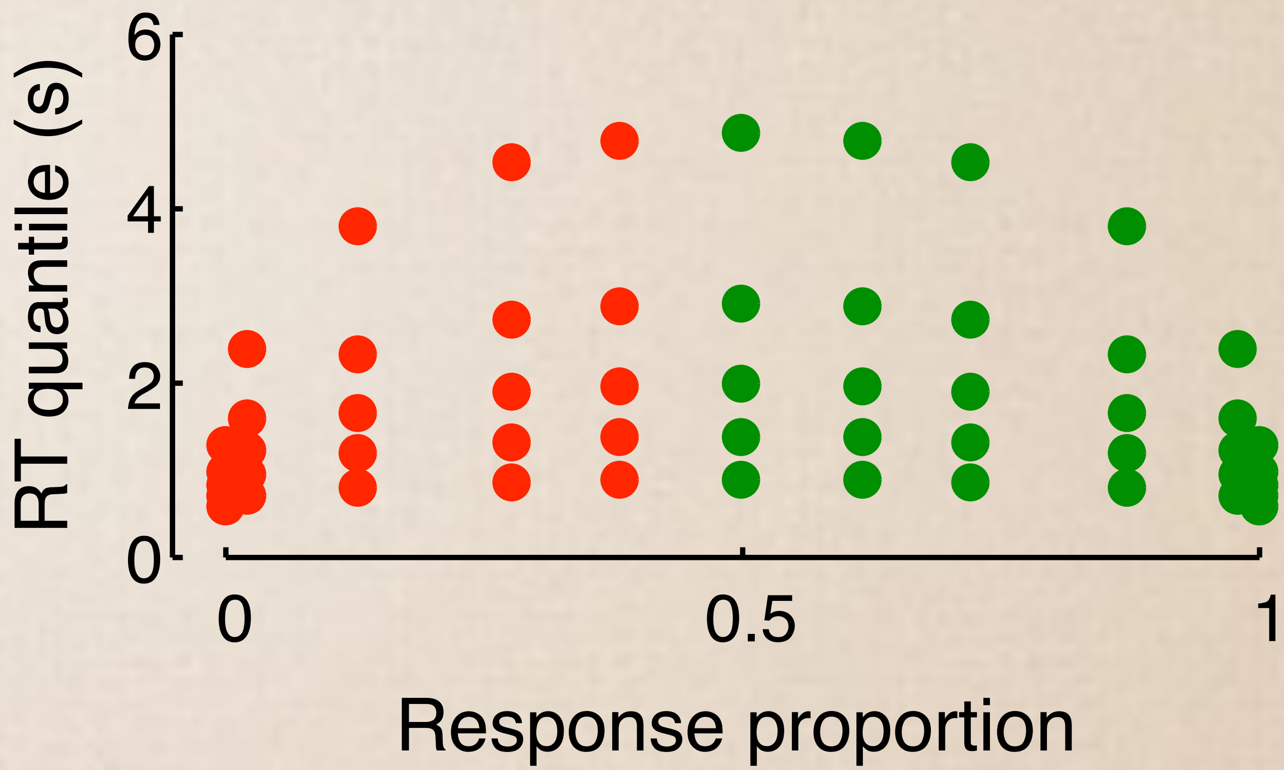
Quantile probability plot

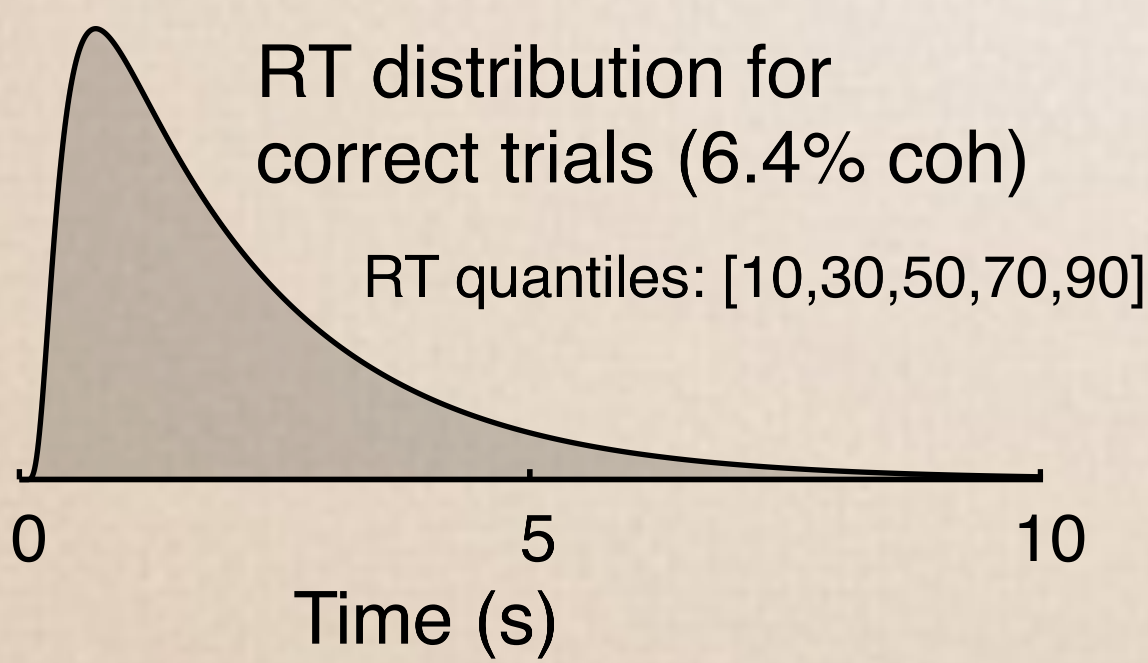




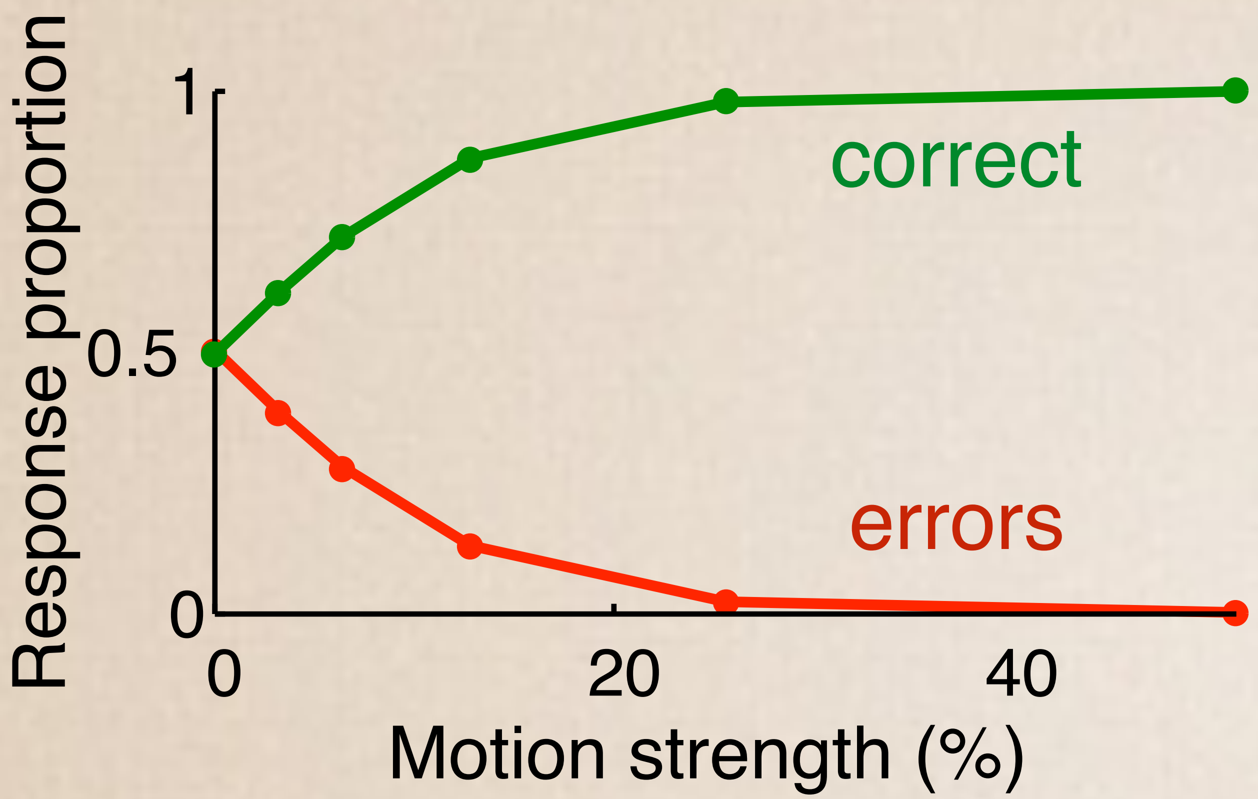
Quantile probability plot

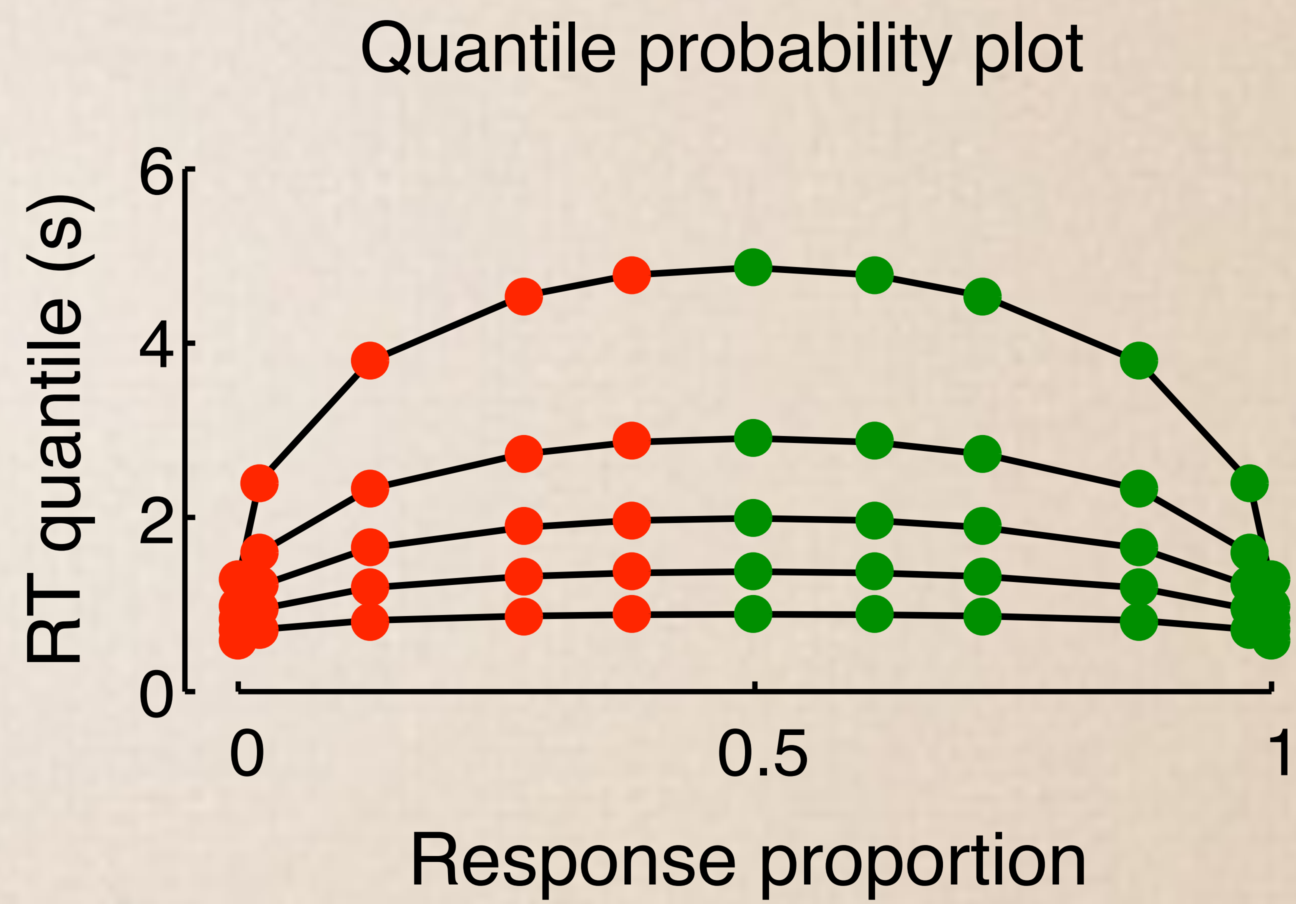


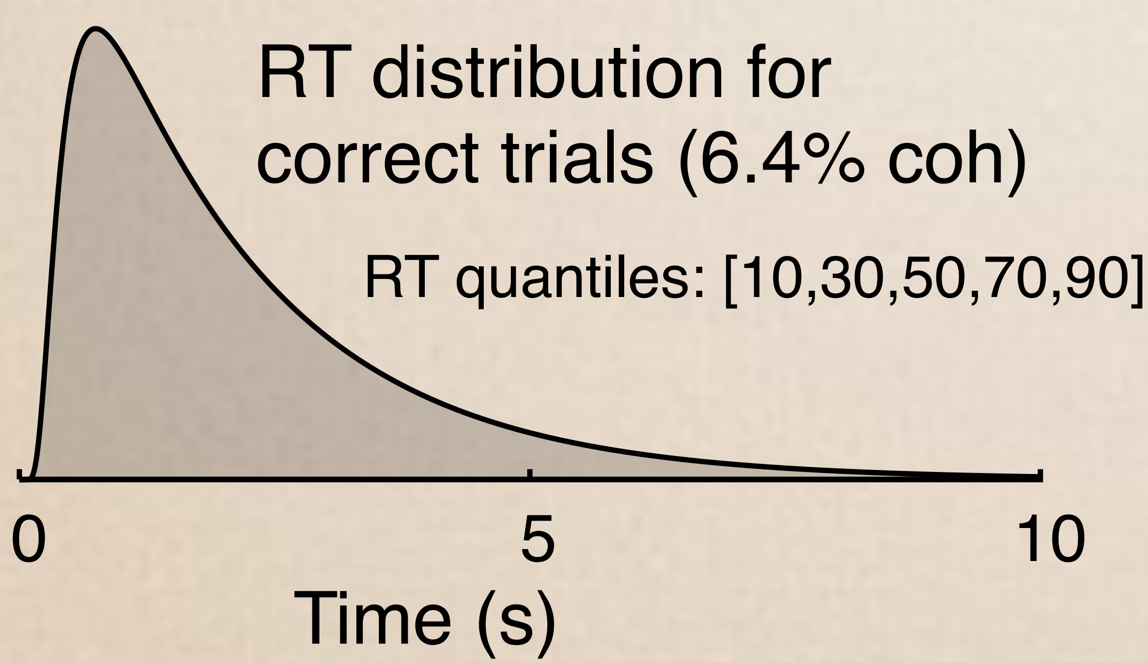




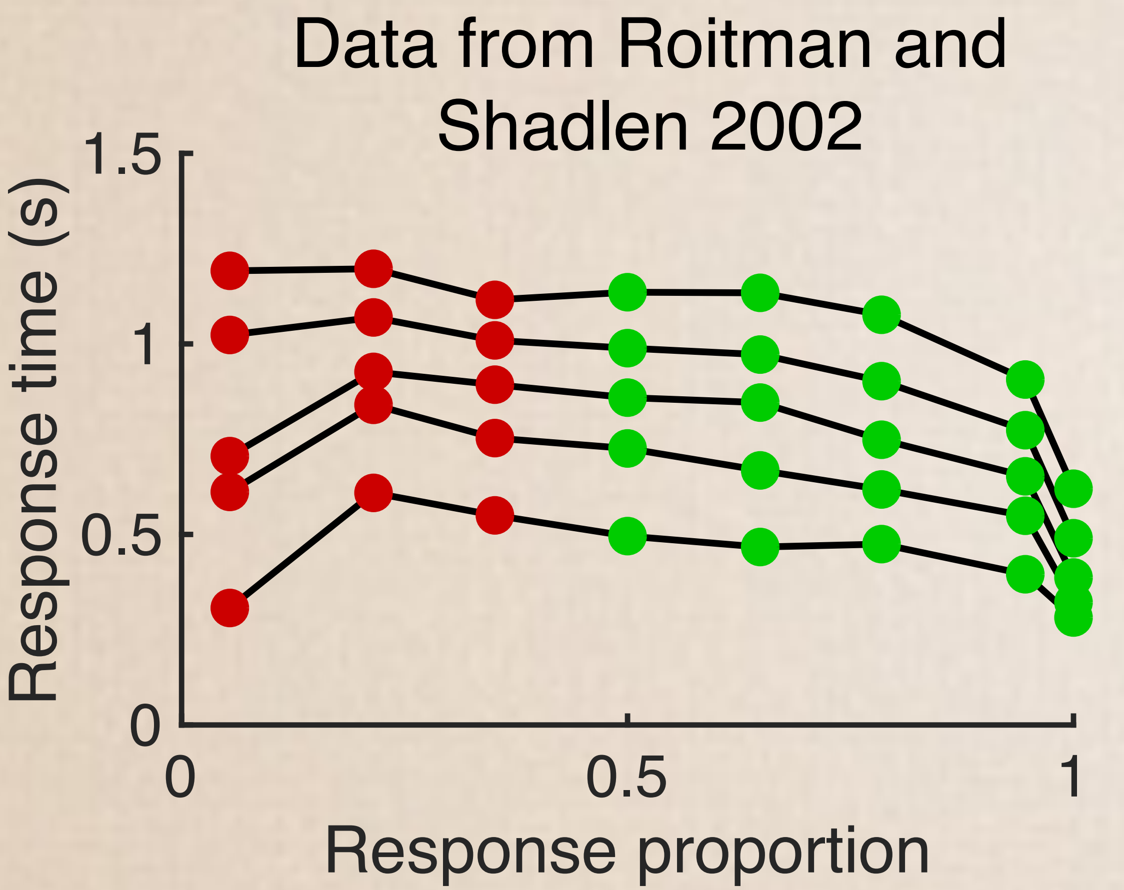
Quantile probability plot

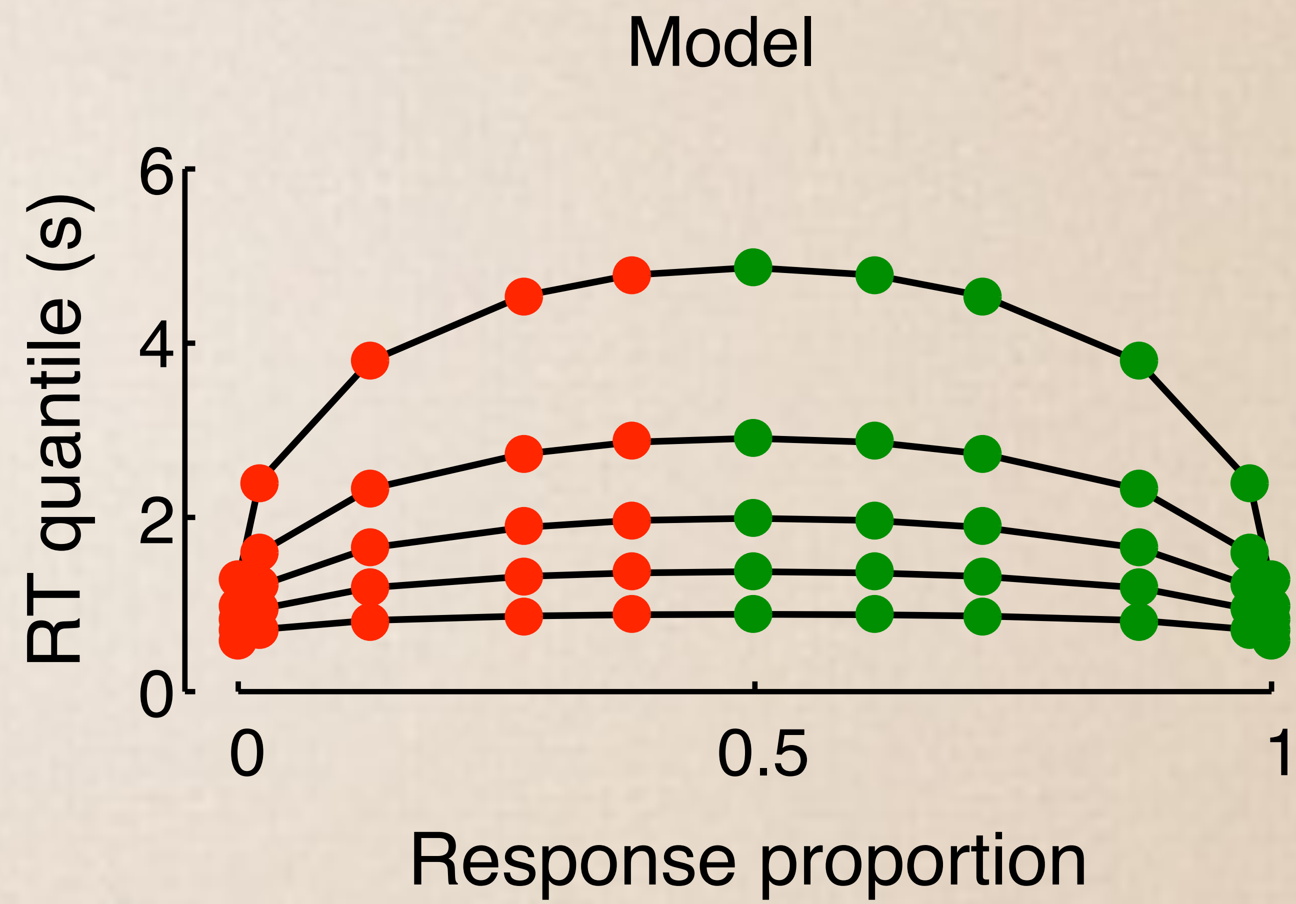




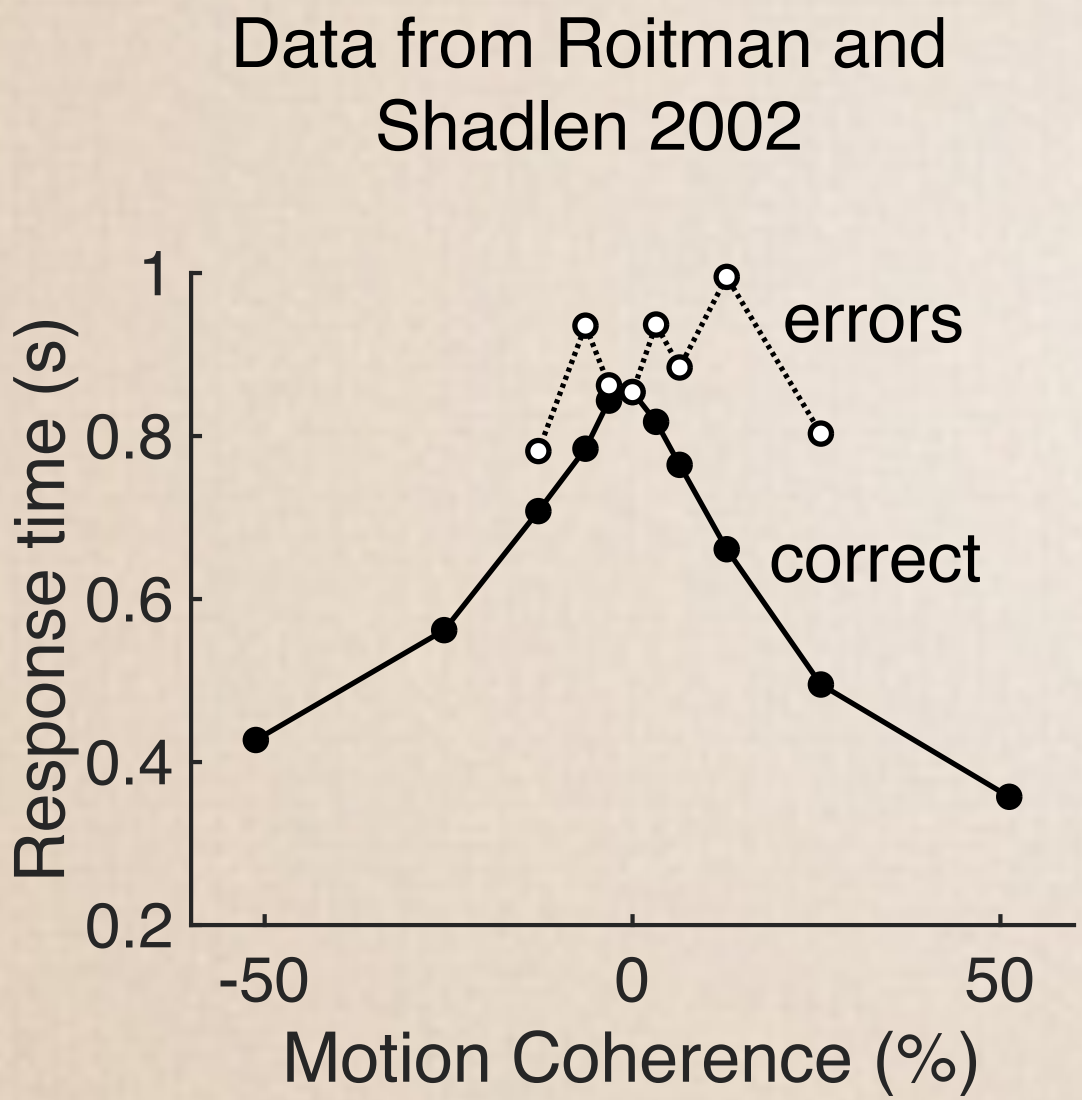


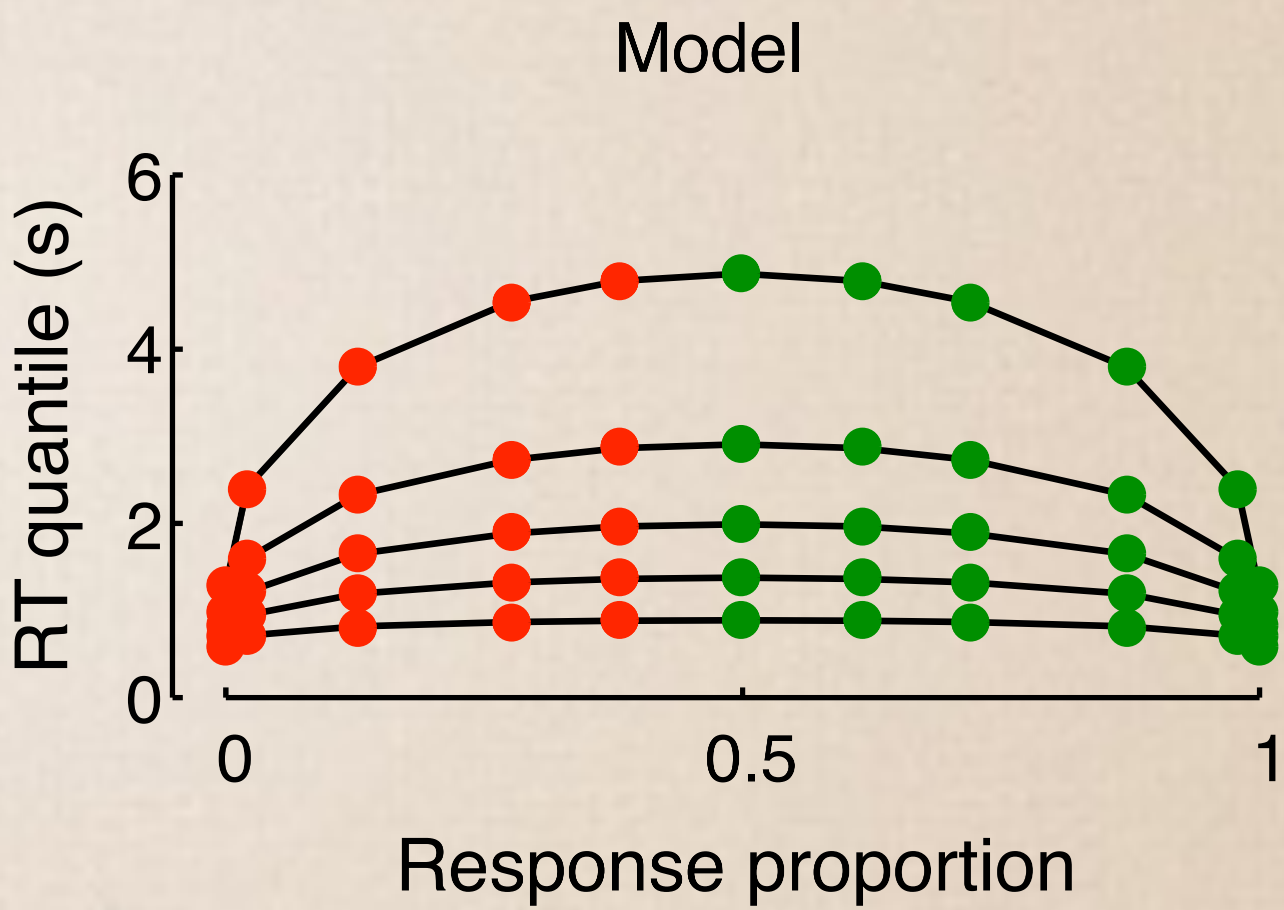
Problem with simple DDM: symmetric RT distributions



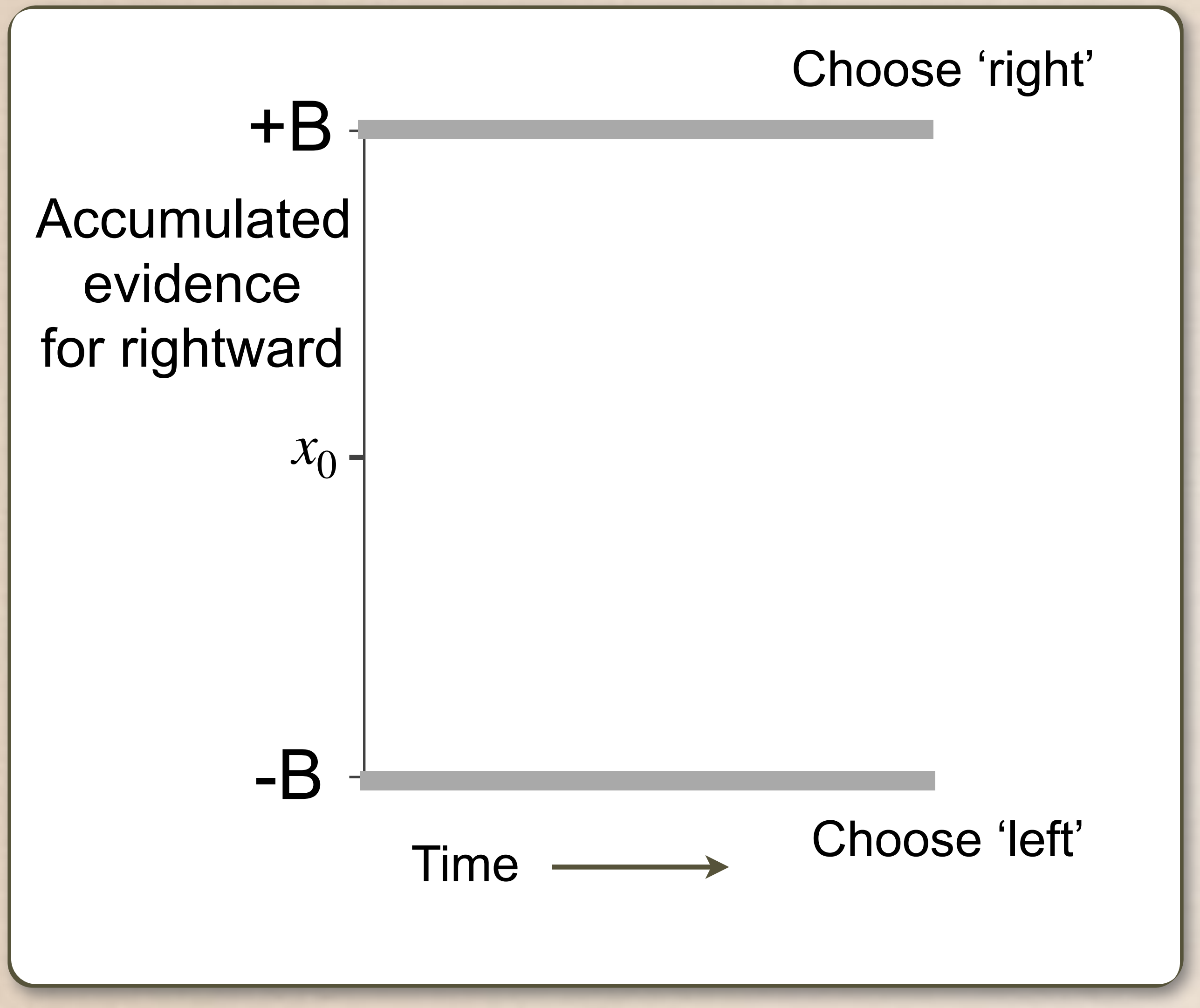


Problem with simple DDM: symmetric RT distributions

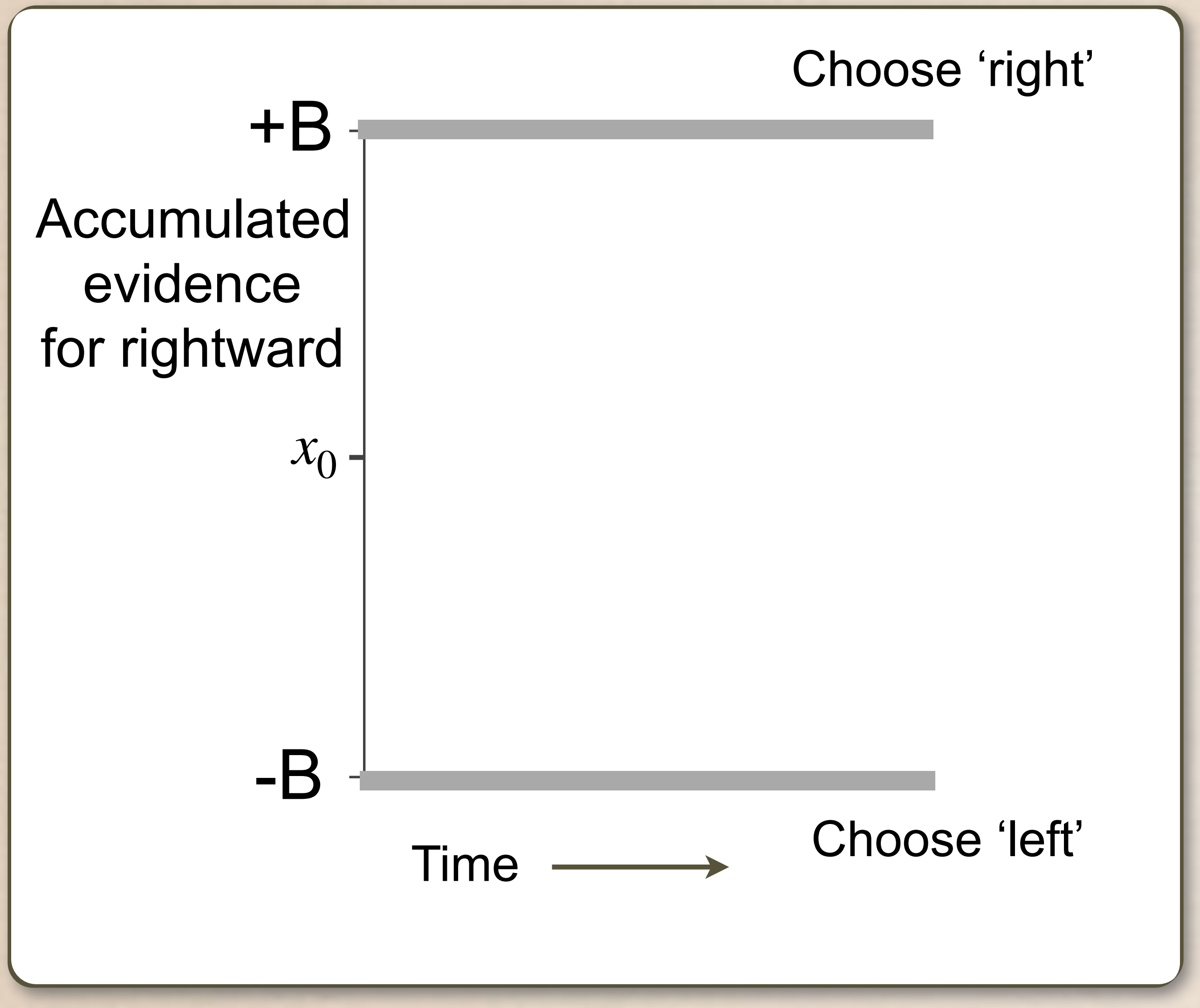




Why flat bounds lead to identical RT distributions for correct and incorrect decisions

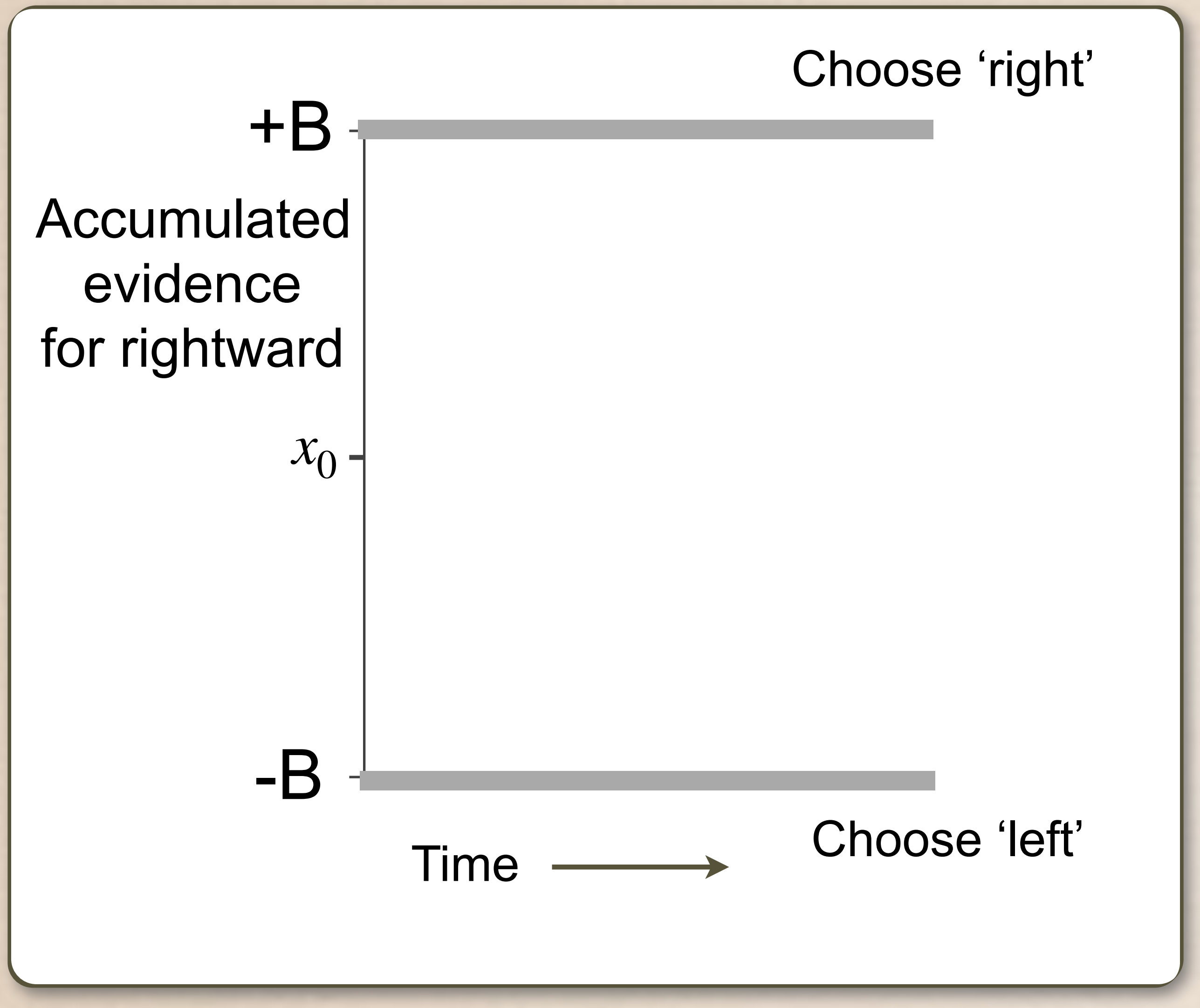


Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\uparrow$ with probability p

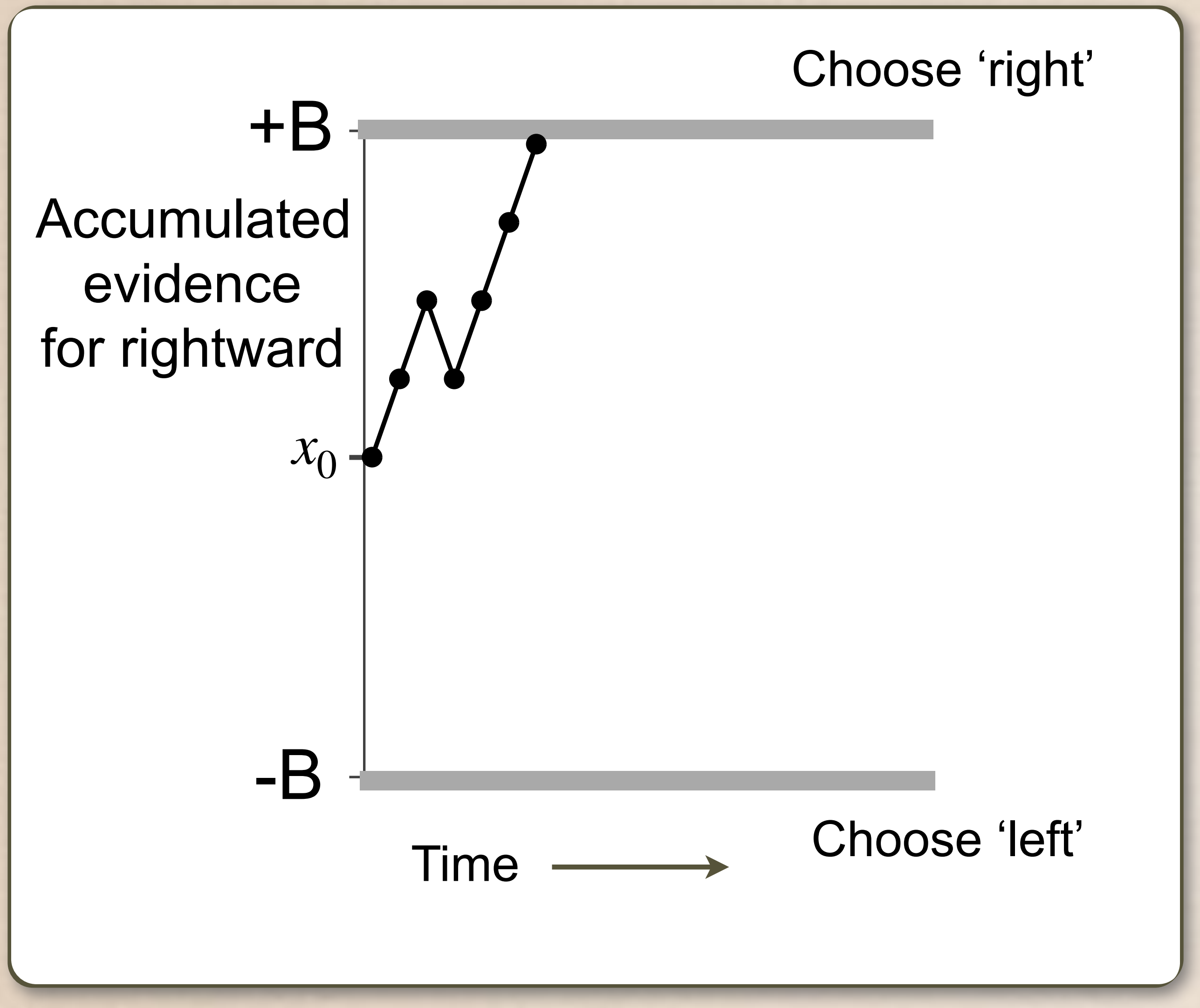
Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

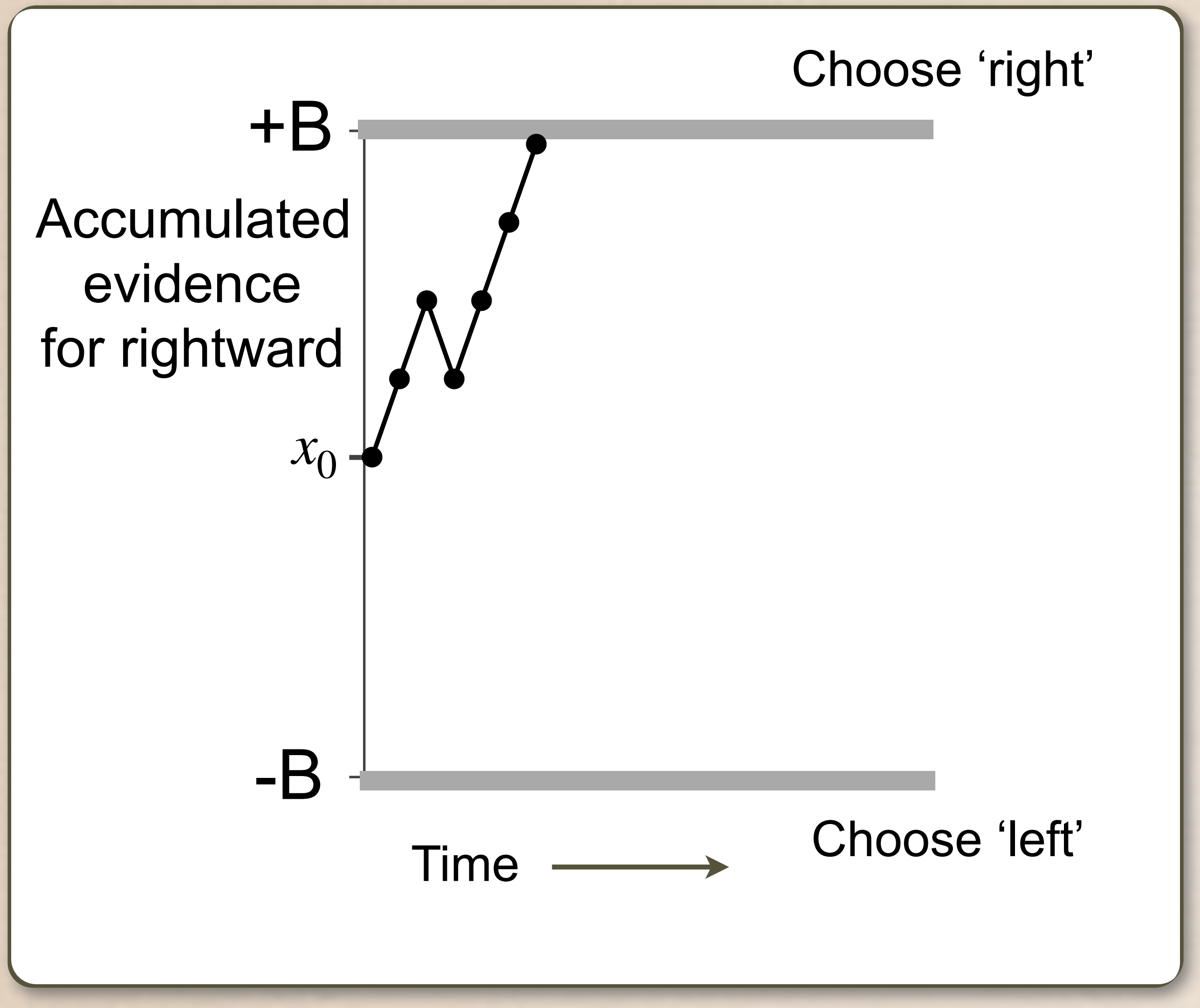
Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions

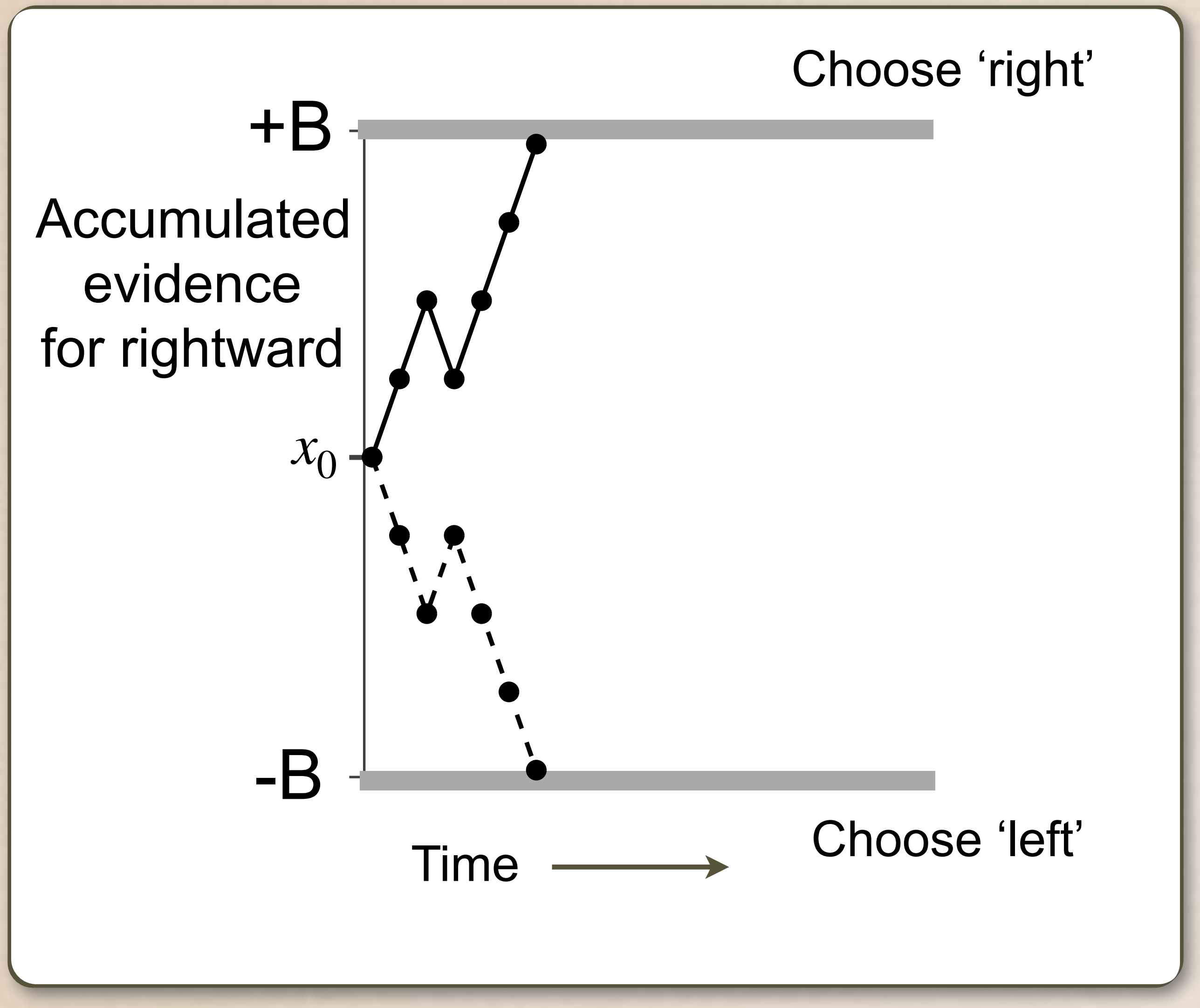


$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions

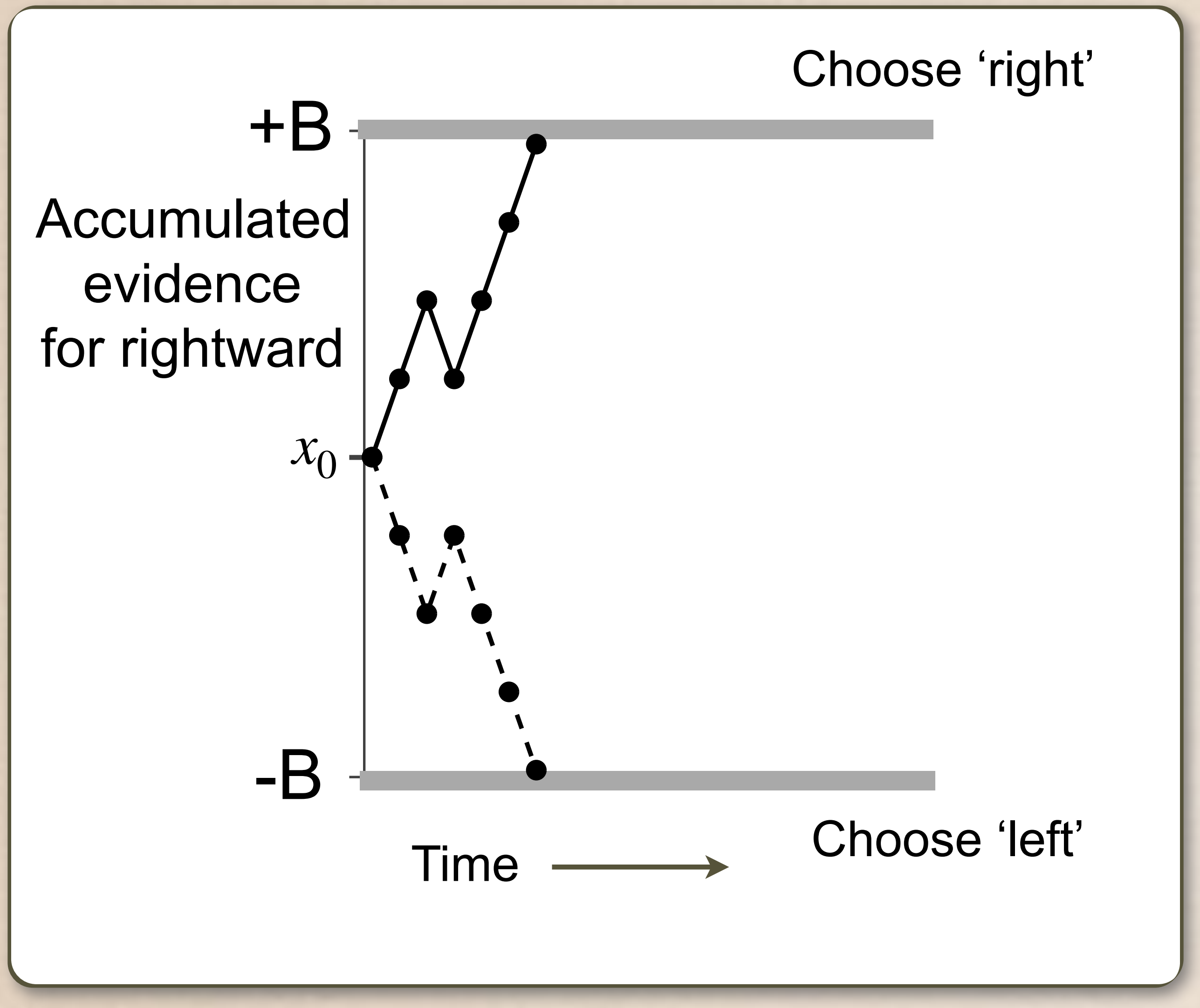


$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



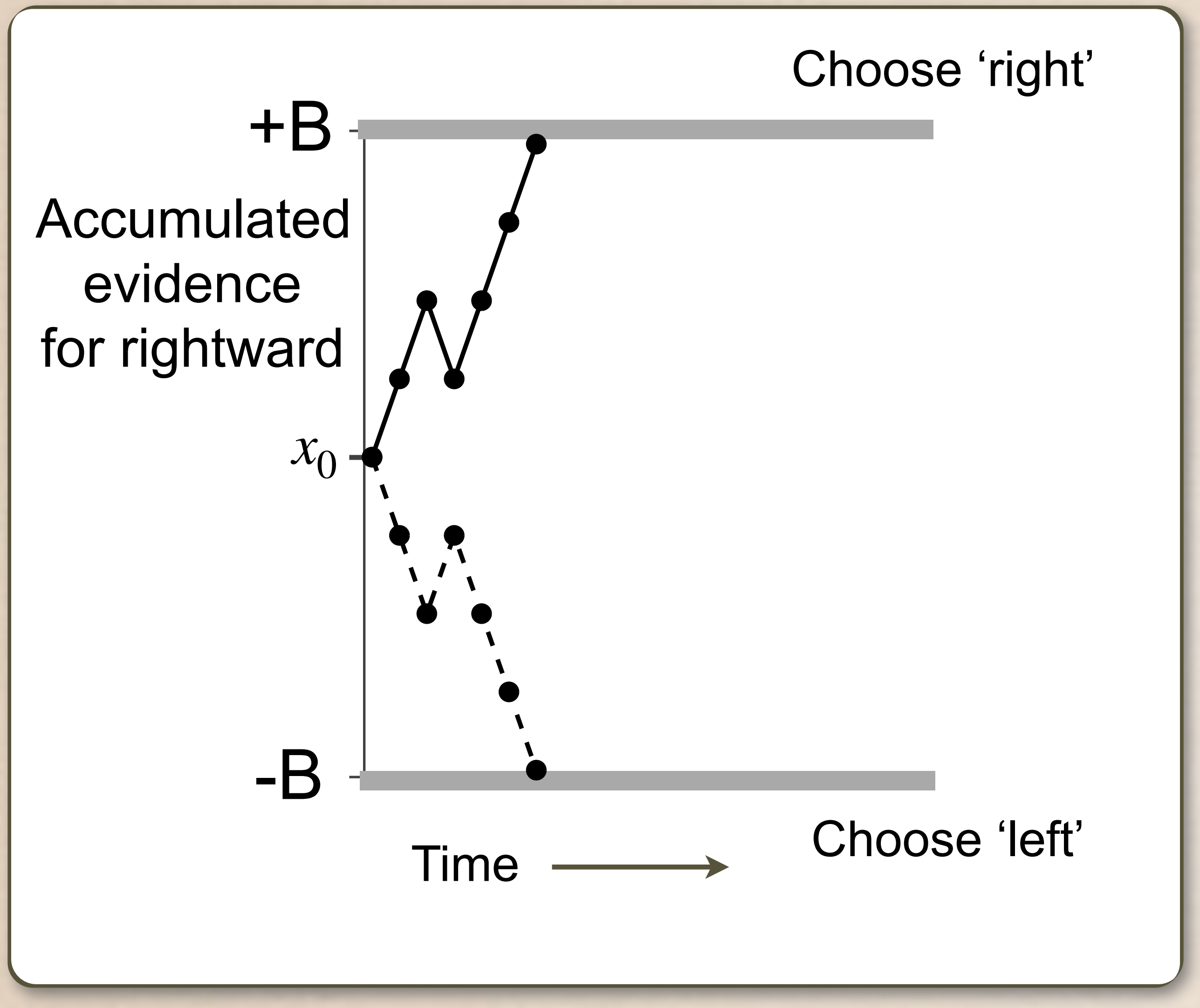
$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Pr(mirror) = $qqpqqq$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\uparrow$ with probability p

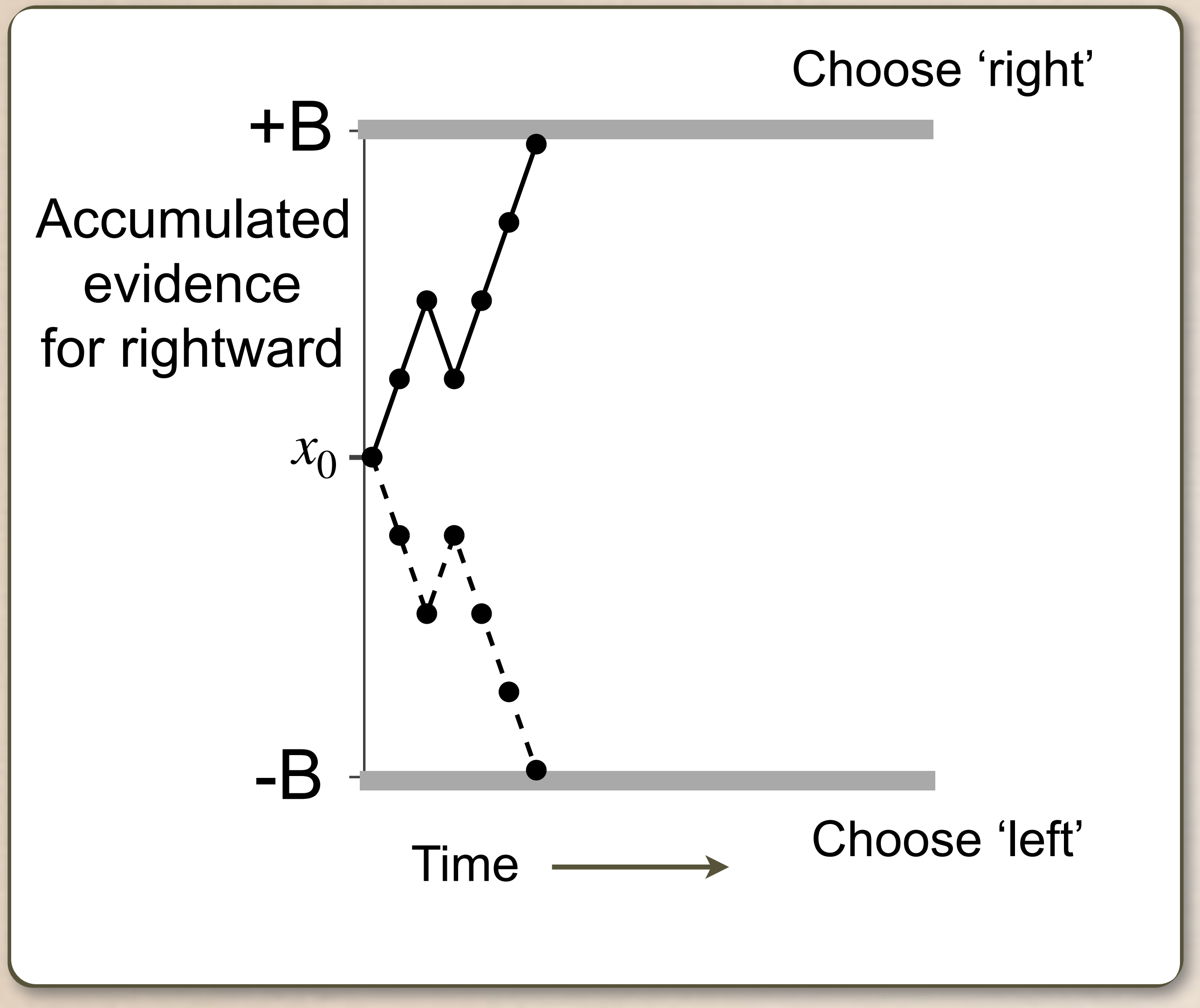
$\downarrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Pr(mirror) = $qqpqqq$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma}$$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

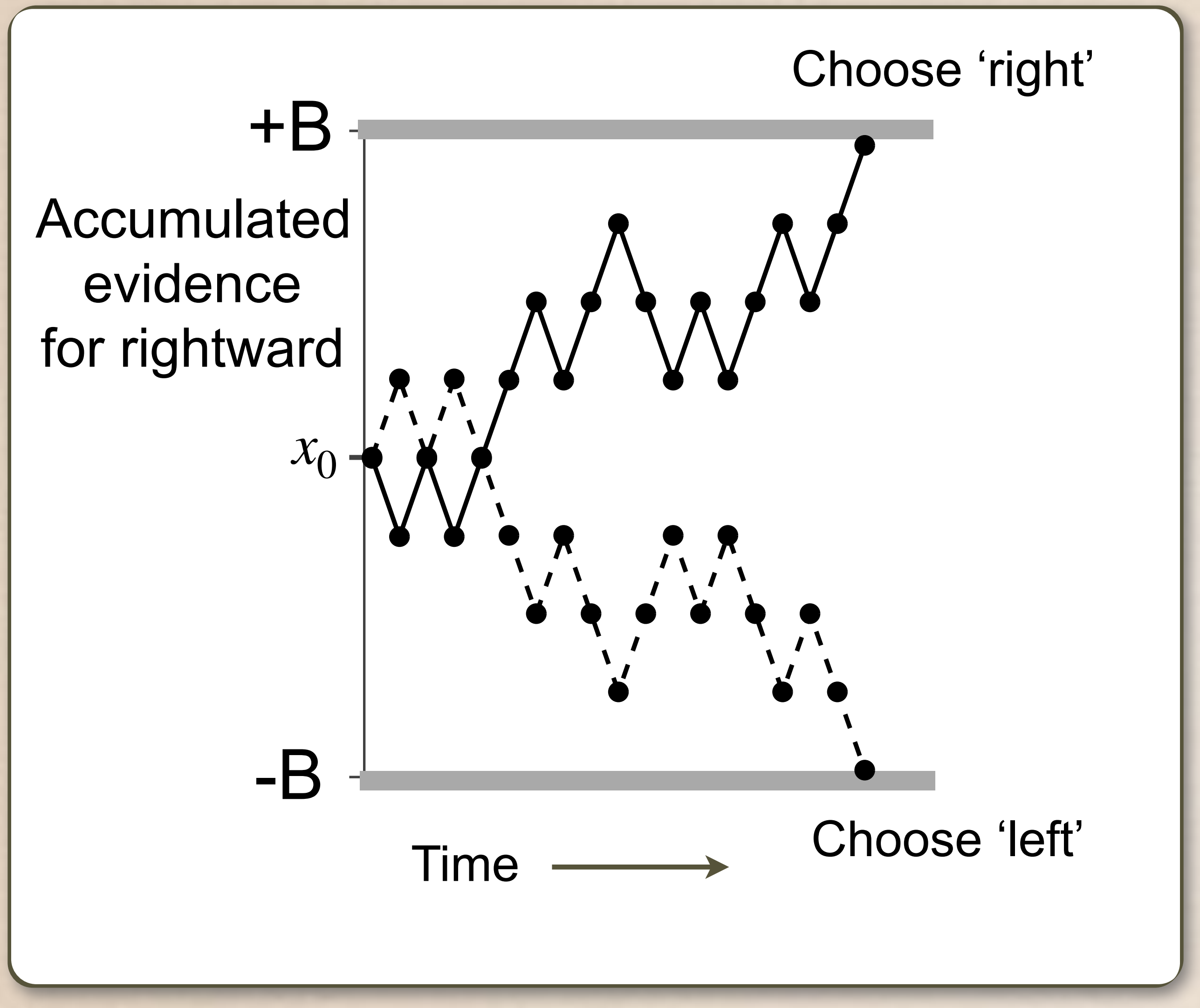
$\searrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Pr(mirror) = $qqpqqq$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma = 4}$$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

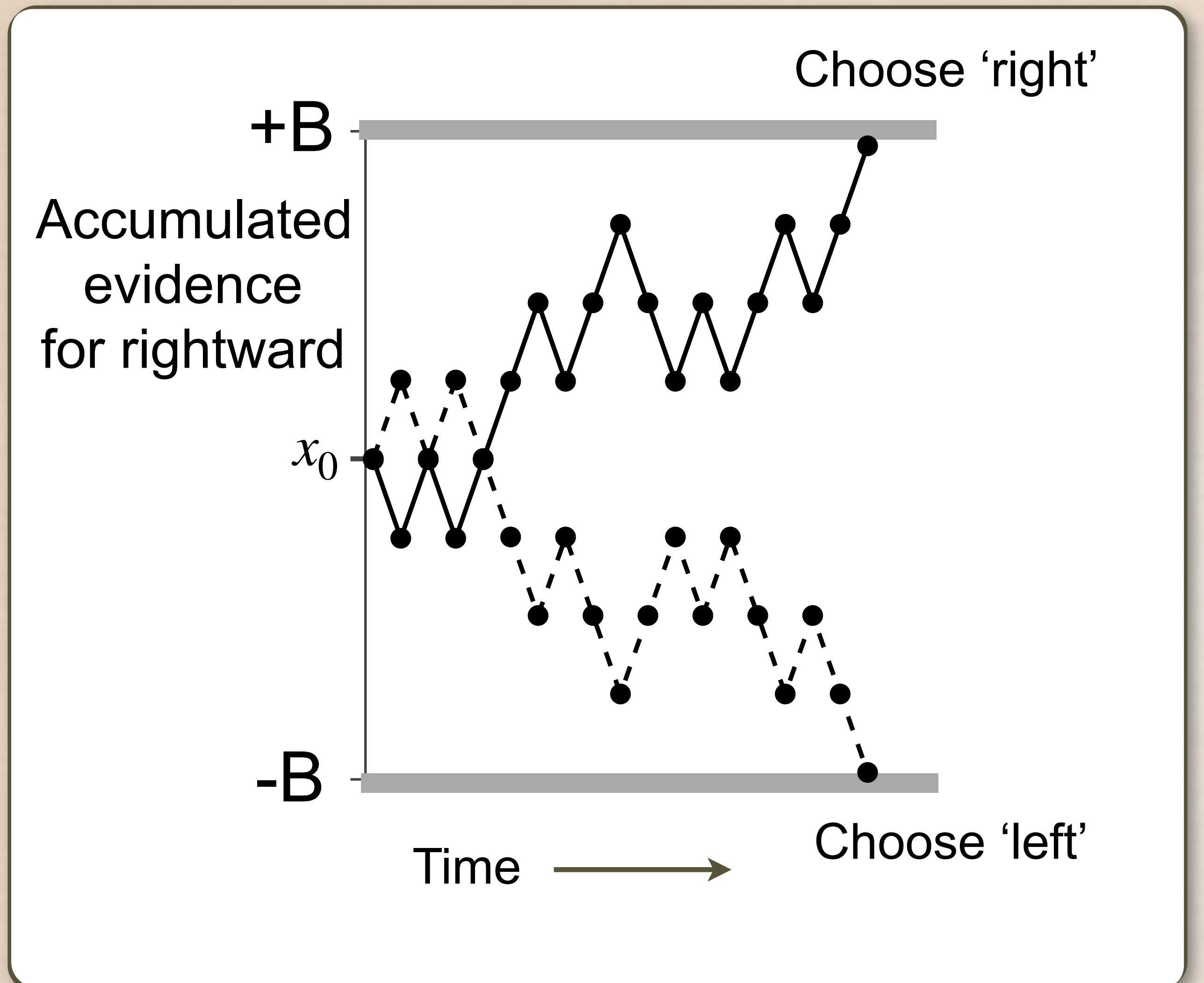
$\searrow$ with probability $q = 1 - p$

Pr = $ppqppp$

Pr(mirror) = $qqpqqq$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma \, =4}$$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\uparrow$ with probability p

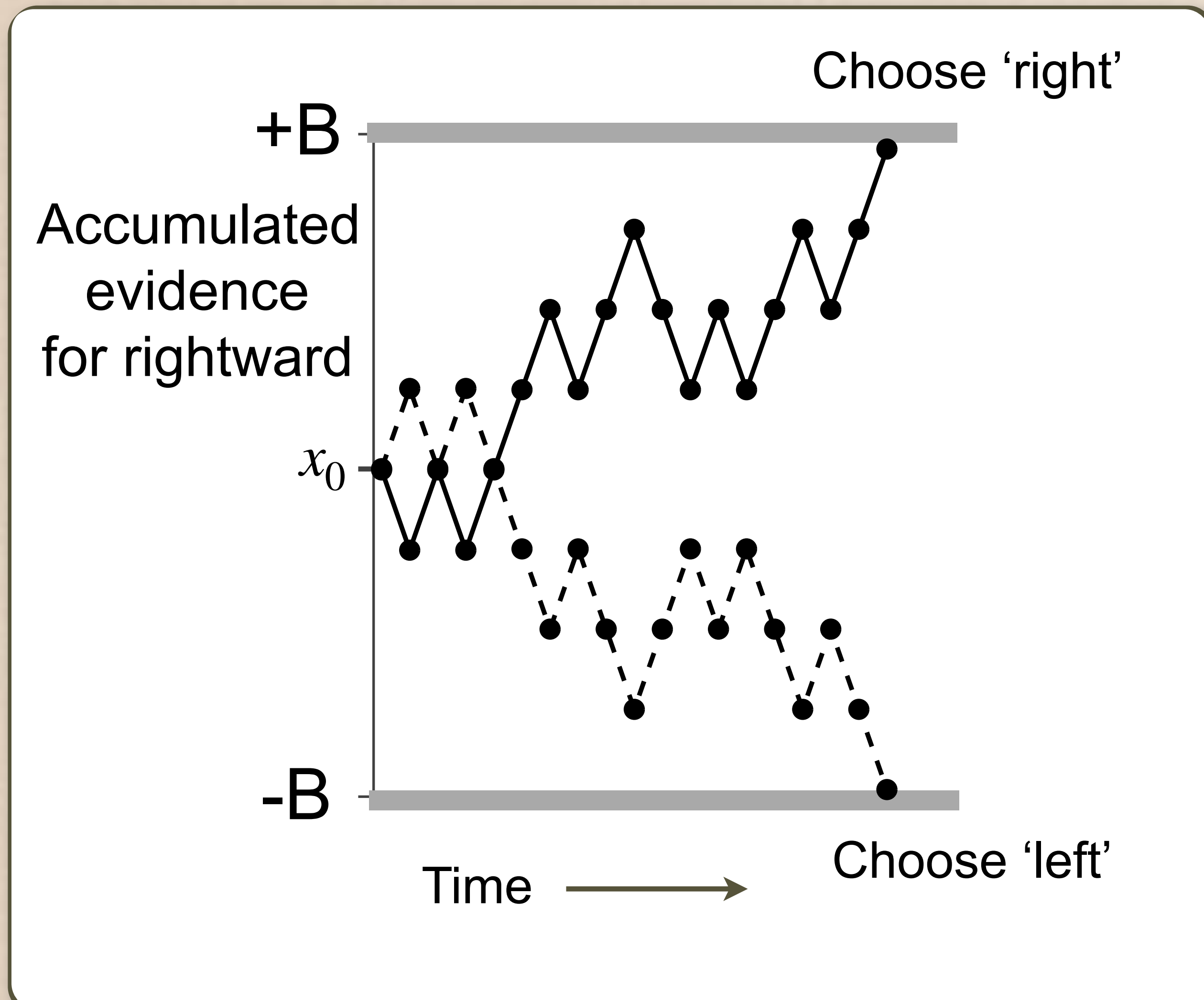
$\downarrow$ with probability $q = 1 - p$

Pr = $qpqpppqppqqpqppqpp$

Pr(mirror) = $pqpqqqpqqppqpqqpqq$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma \,=4}$$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

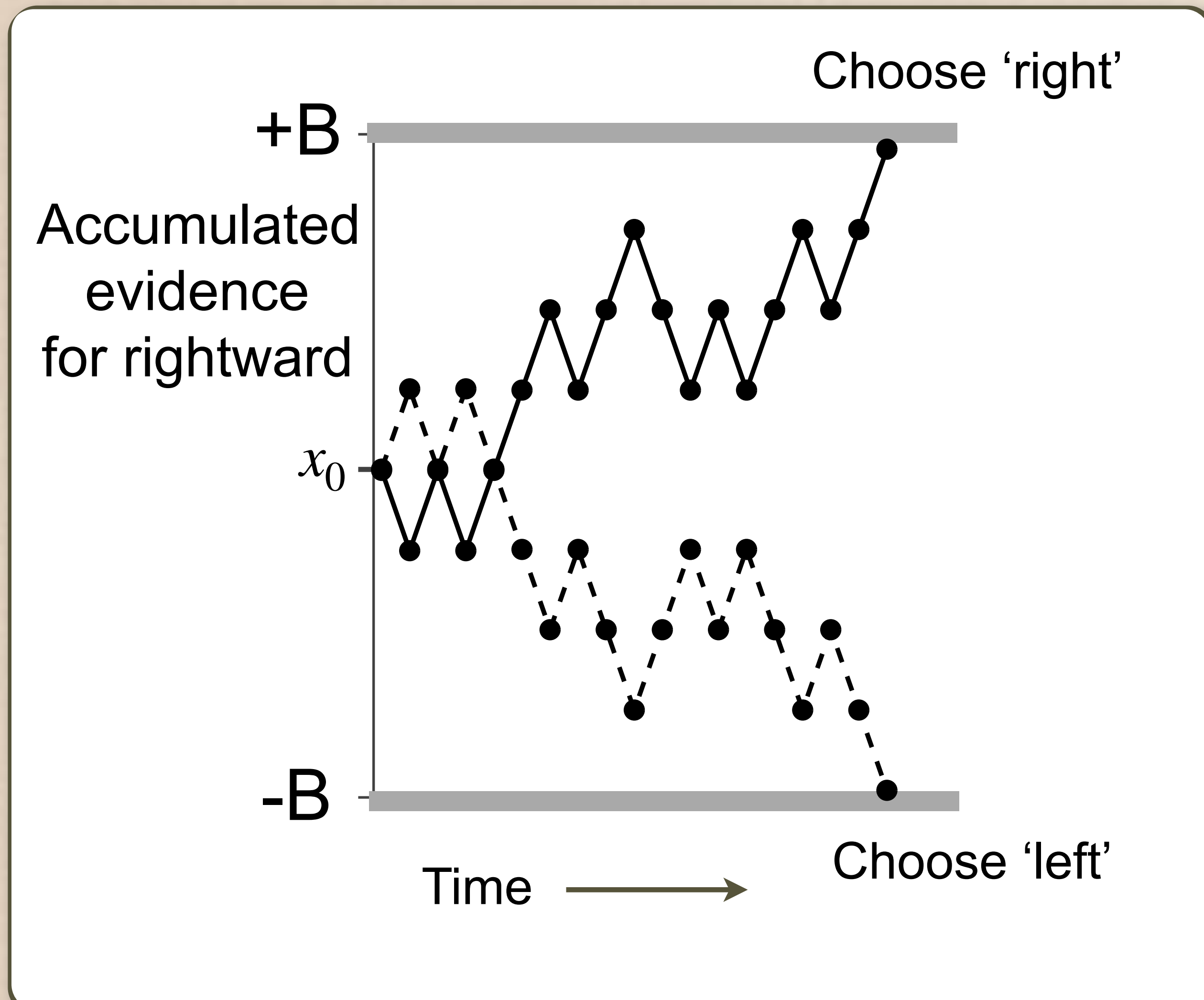
$\text{Pr} = qpqpppqppqqpqppqpp$

$\text{Pr(mirror)} = pqpqqqpqqppqpqqpqq$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma \, =4}$$

$$\boxed{P(t, \text{right}) = aP(t, \text{left})}$$

Why flat bounds lead to identical RT distributions for correct and incorrect decisions



$\nearrow$ with probability p

$\searrow$ with probability $q = 1 - p$

$$\text{Pr} = qpqpppqppqqpqppqpp$$

$$\text{Pr(mirror)} = pqpqqqpqqppqpqqpqq$$

$$\frac{\text{Pr}}{\text{Pr(mirror)}} = \frac{pppp}{qqqq} = \left(\frac{p}{1-p}\right)^{\gamma}$$

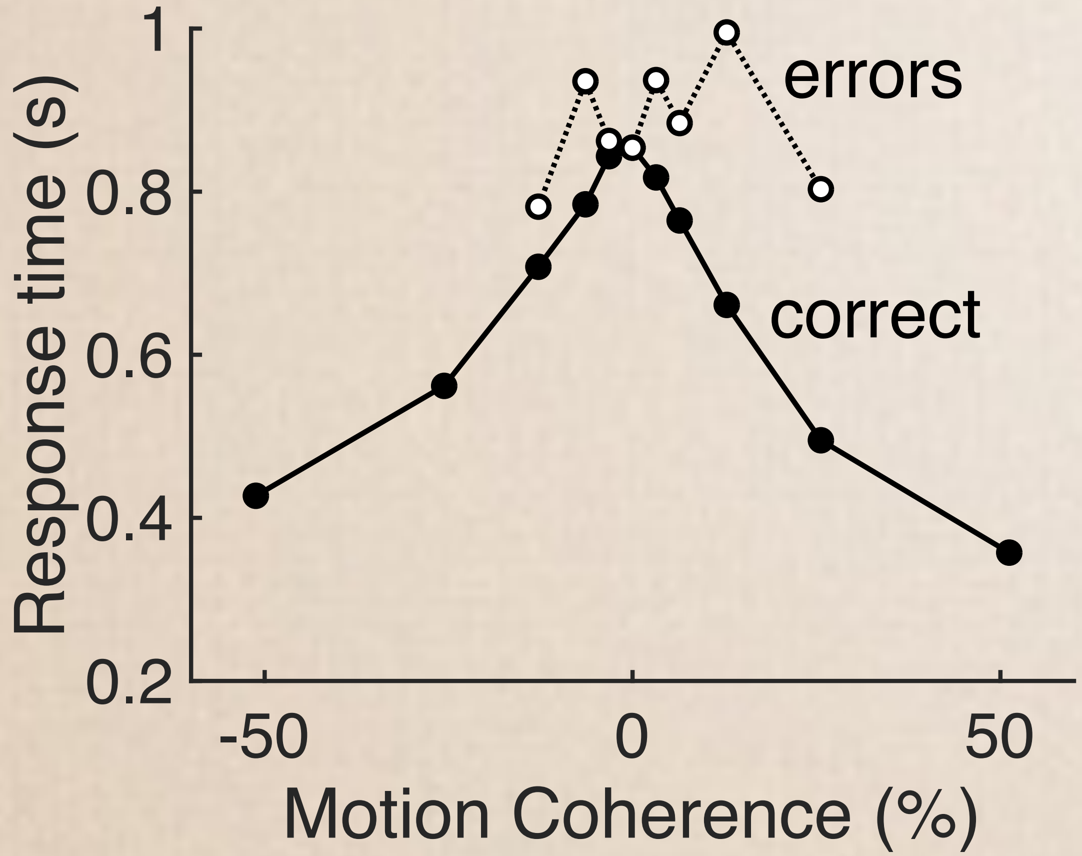
=4

After normalization:

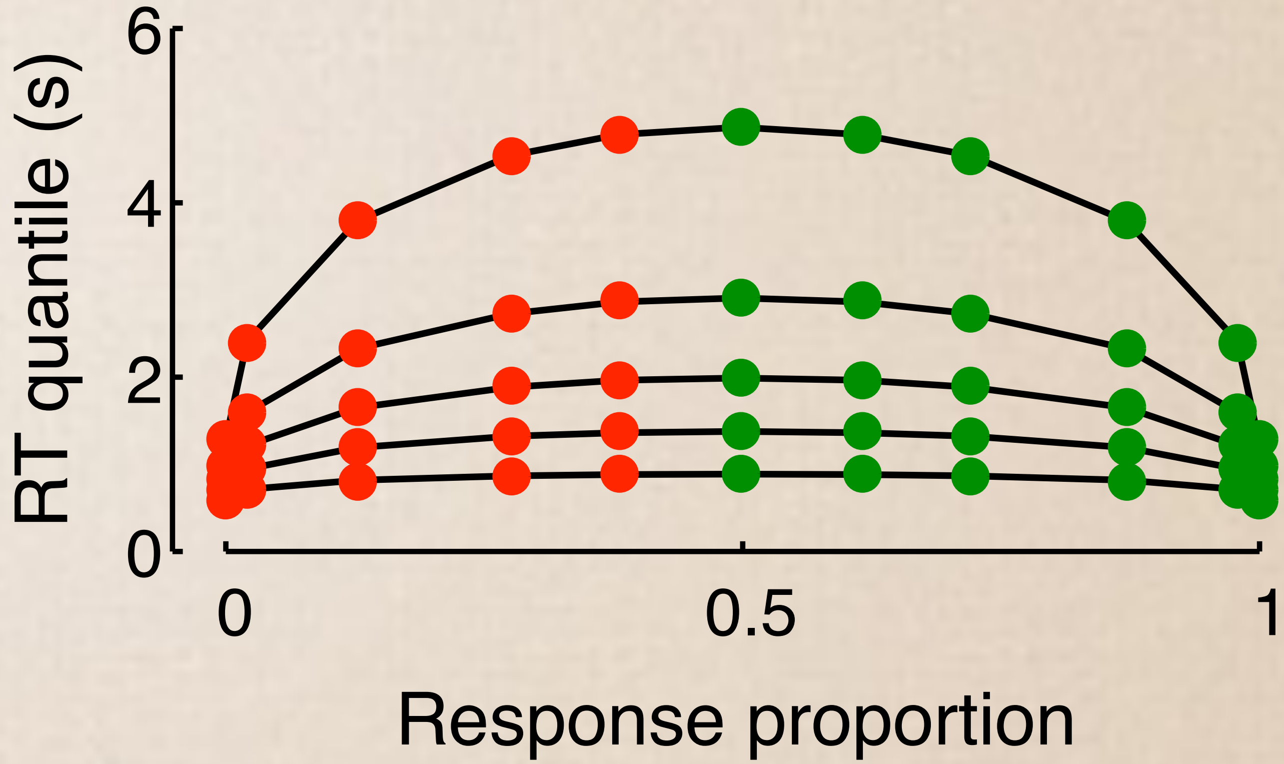
$$\boxed{P(t\,|\,\text{right}) = P(t\,|\,\text{left})}$$

Problem with simple DDM: symmetric RT distributions

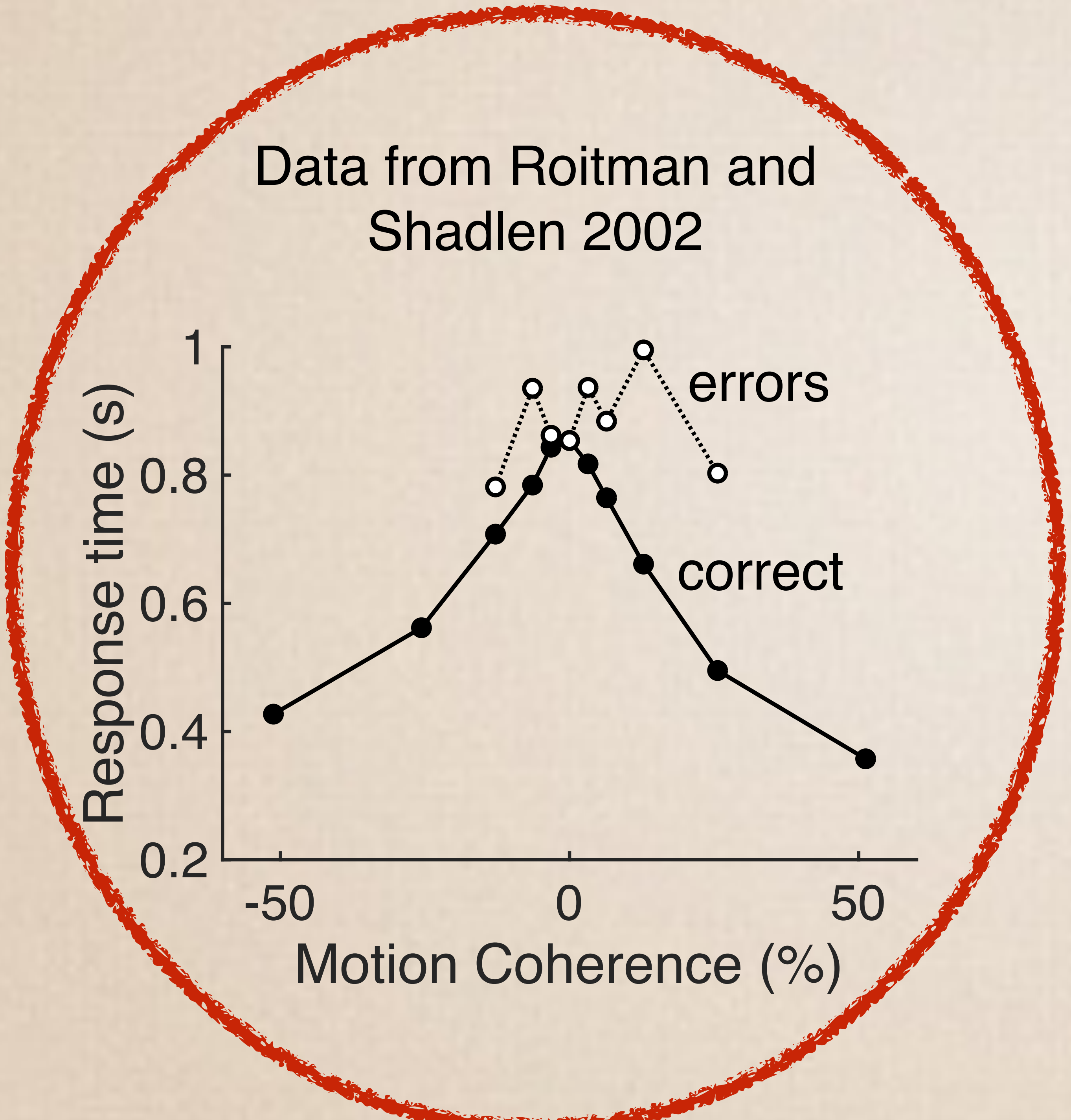
Data from Roitman and Shadlen 2002

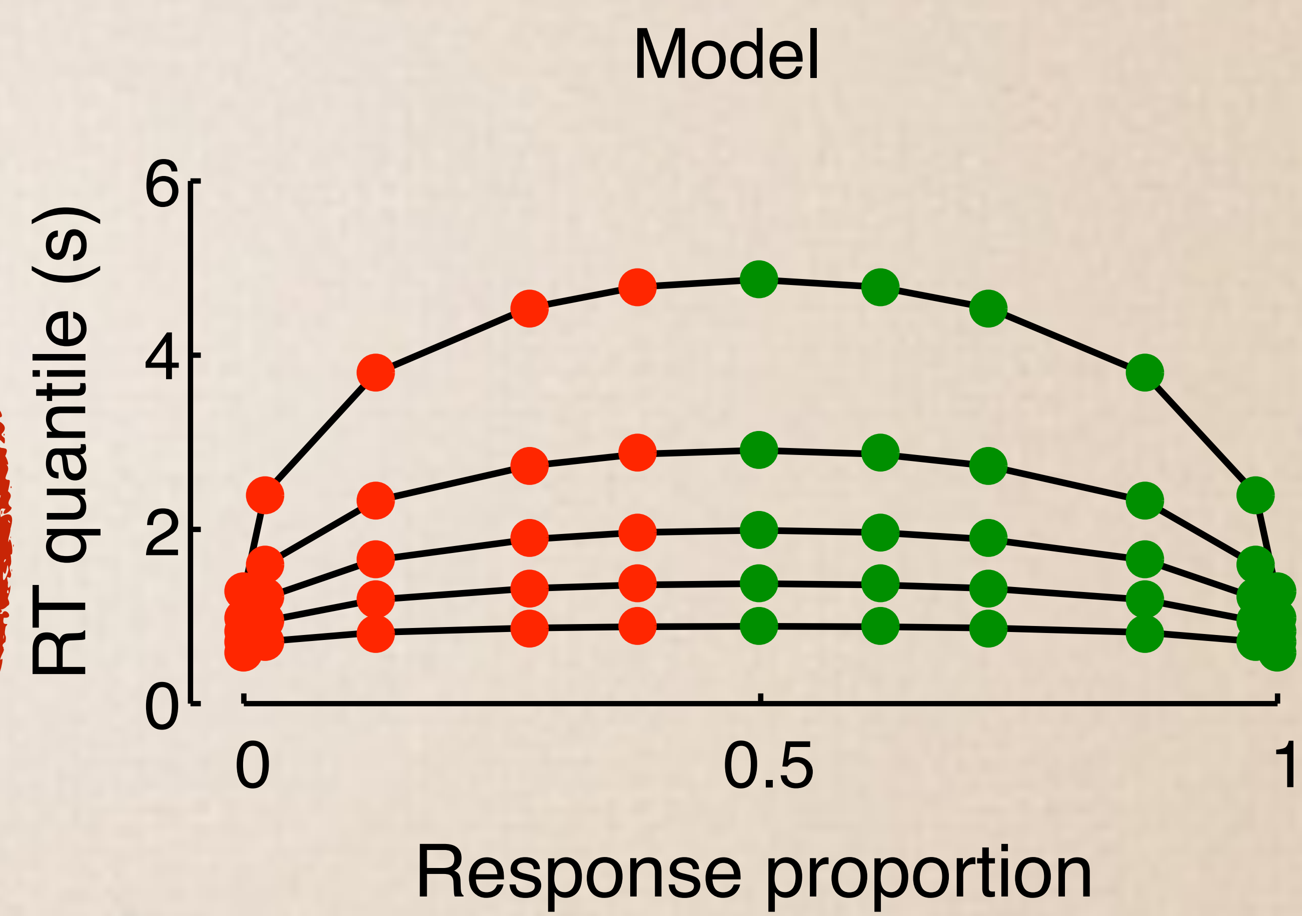


Model

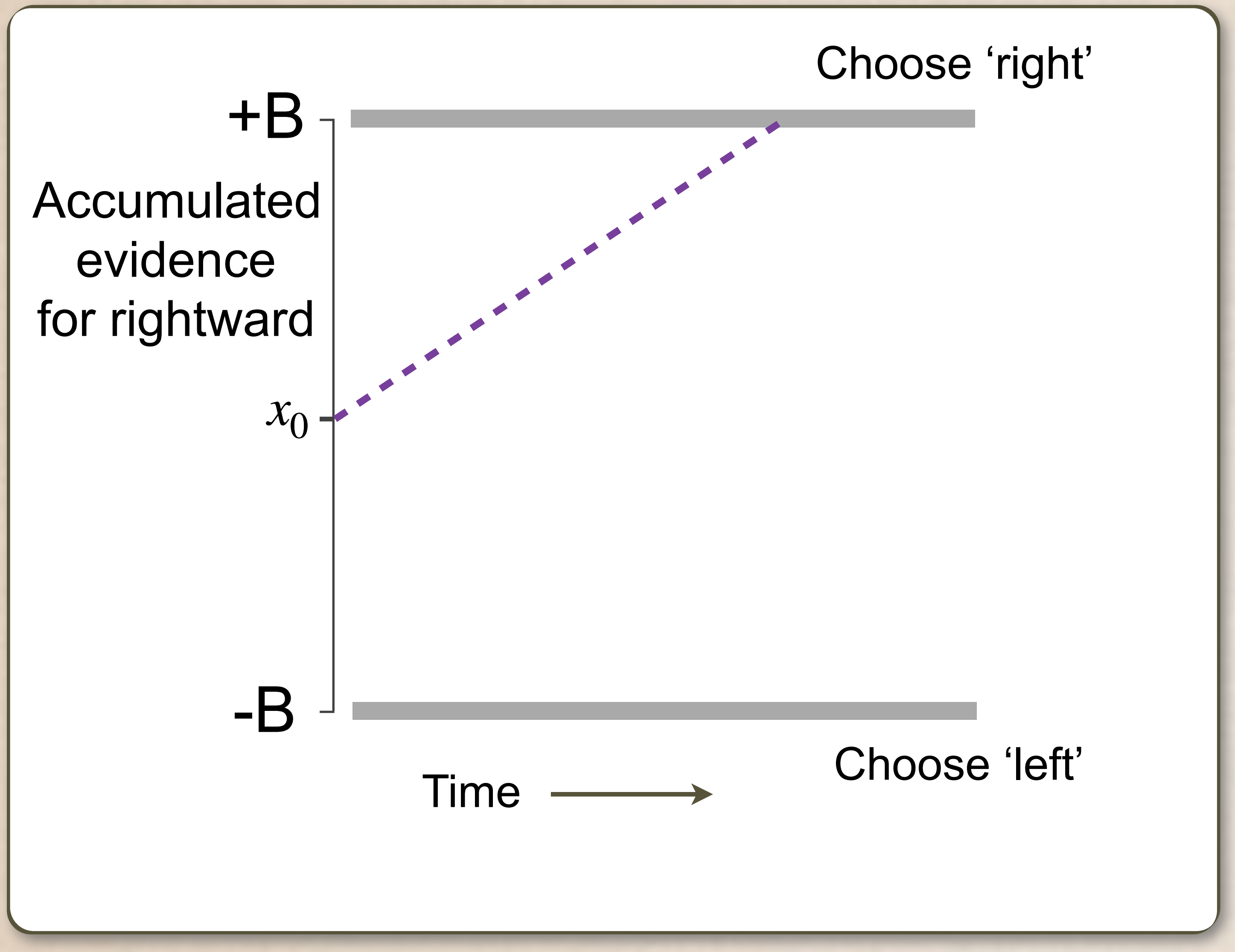


Problem with simple DDM: symmetric RT distributions



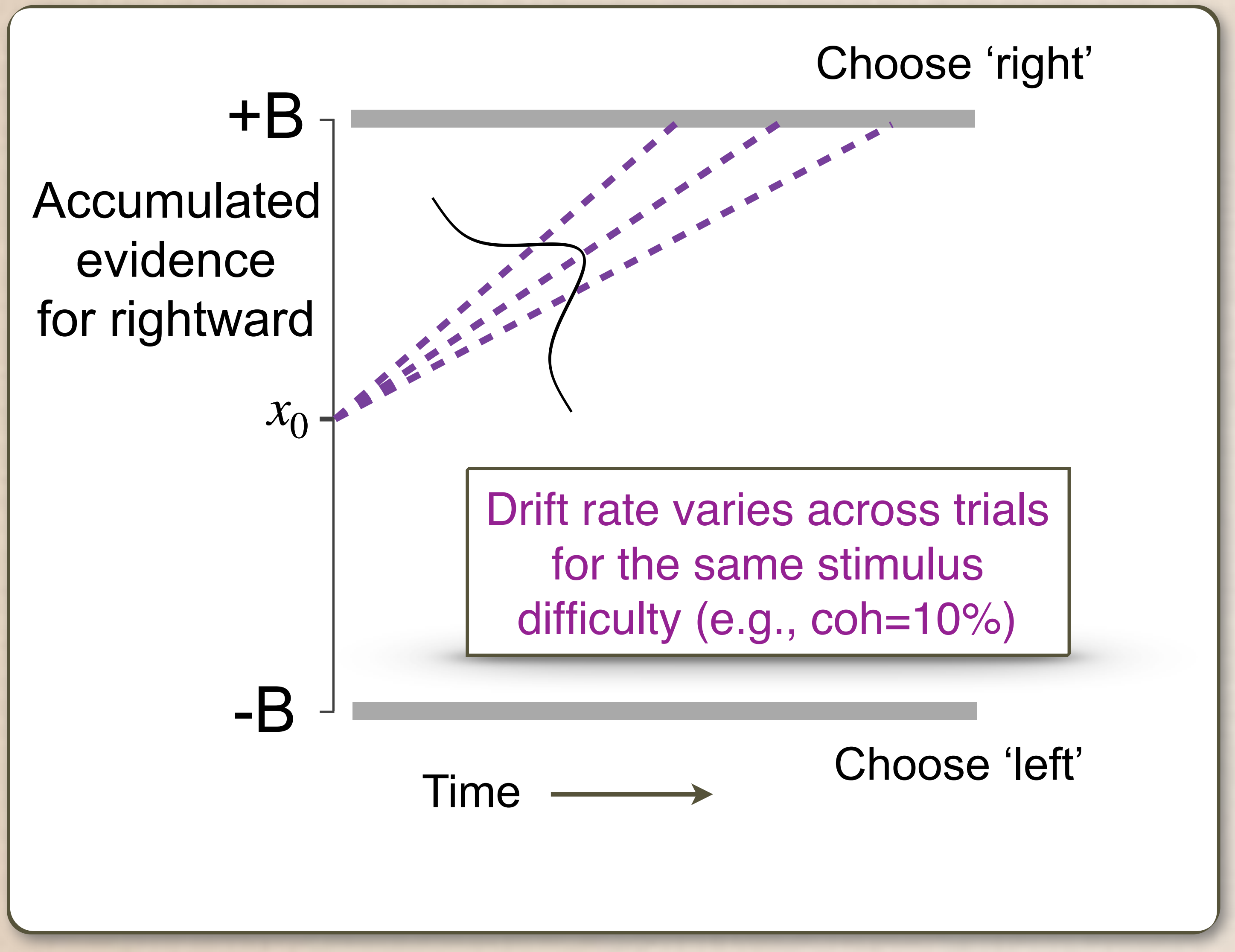


Heuristic solution: variability in drift-rate across trials



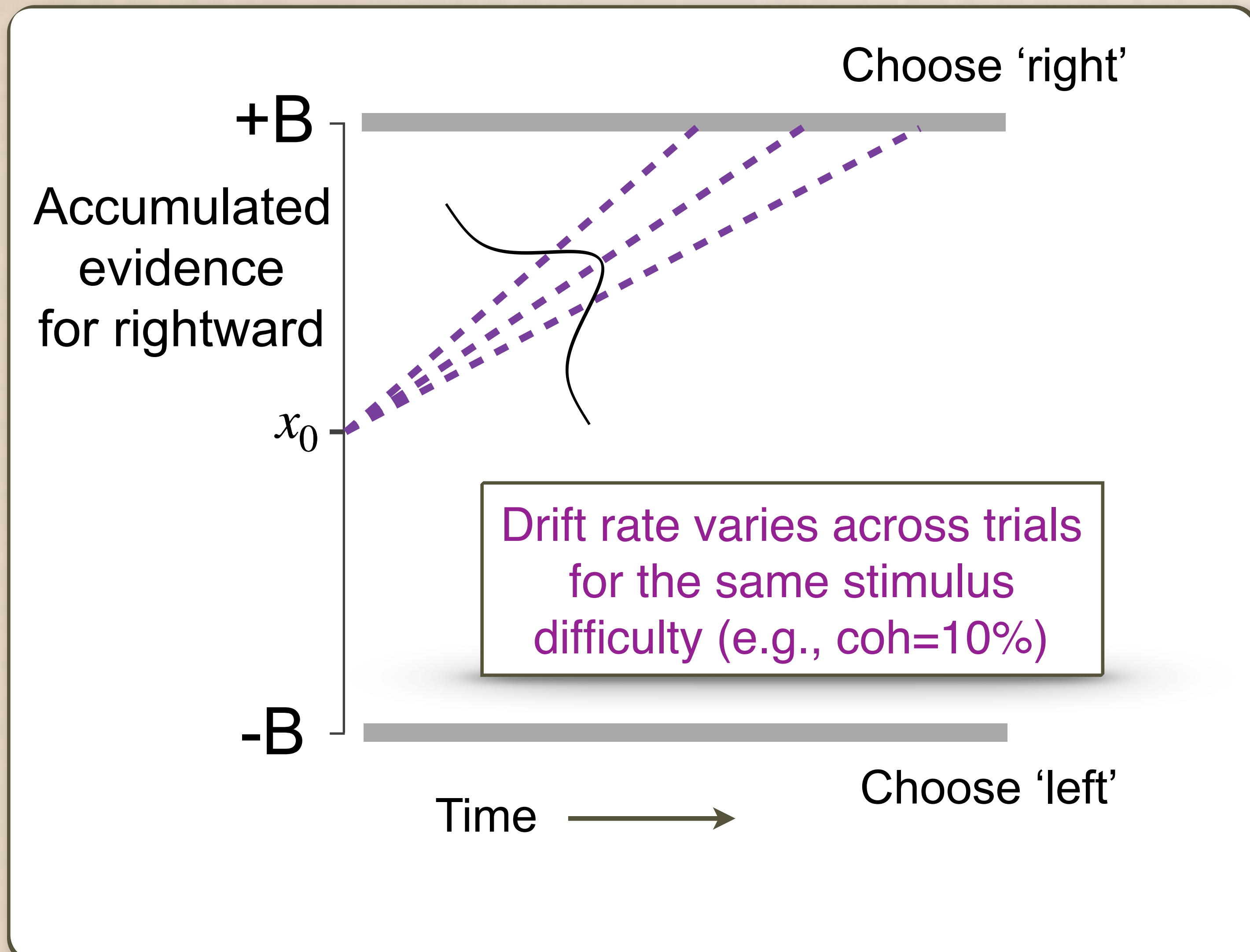
Ratcliff & McKoon (2008)

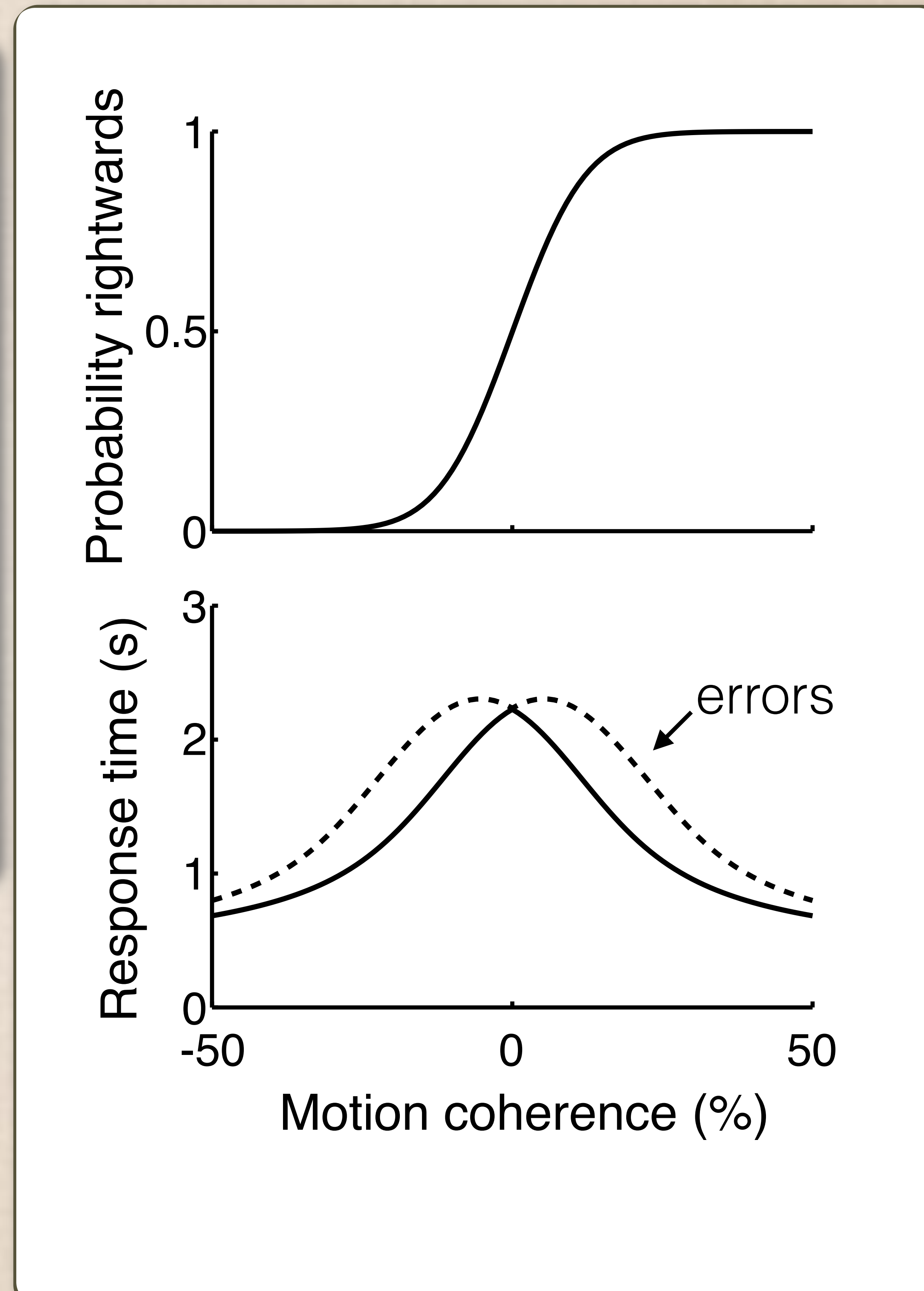
Heuristic solution: variability in drift-rate across trials



Ratcliff & McKoon (2008)

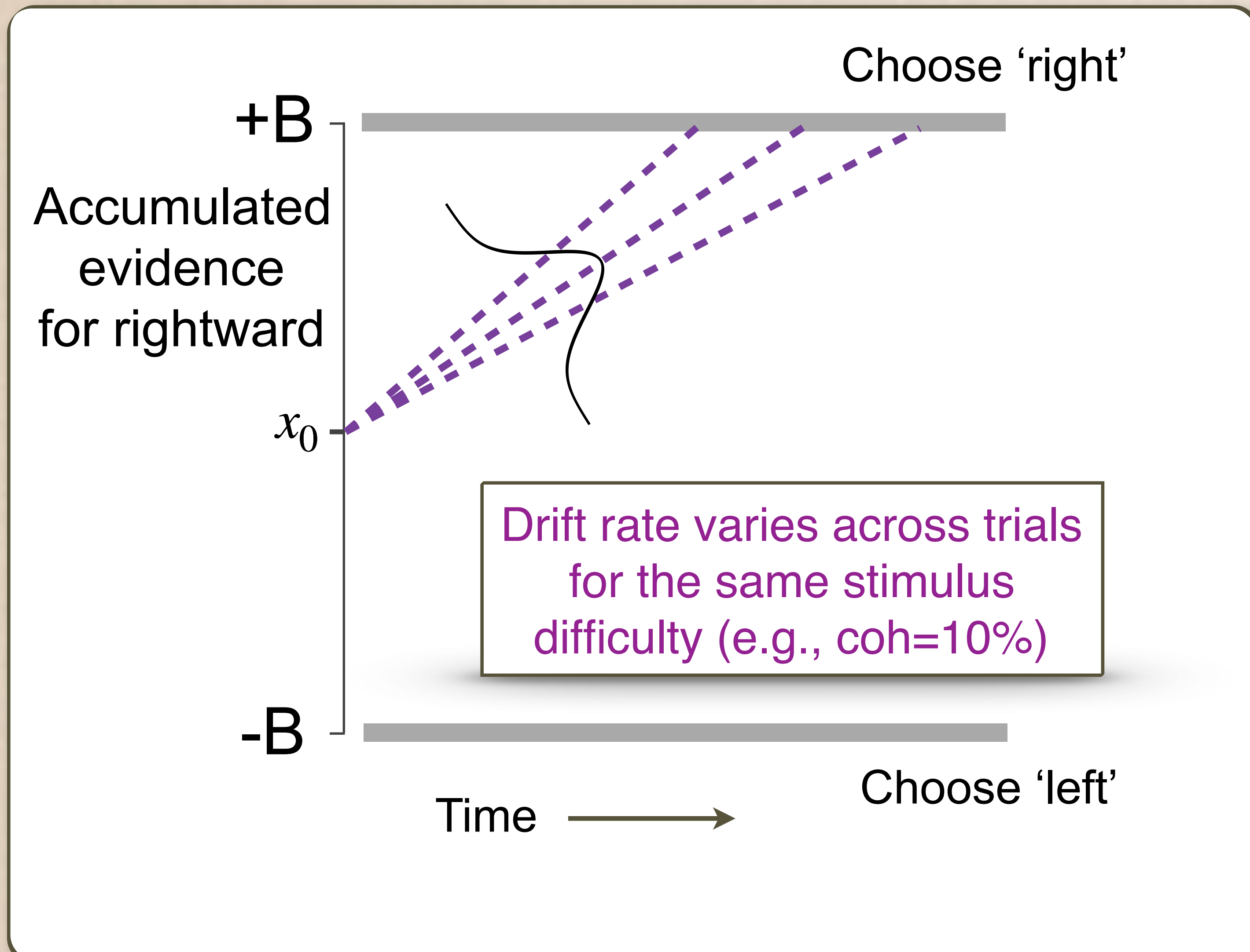
Heuristic solution: variability in drift-rate across trials

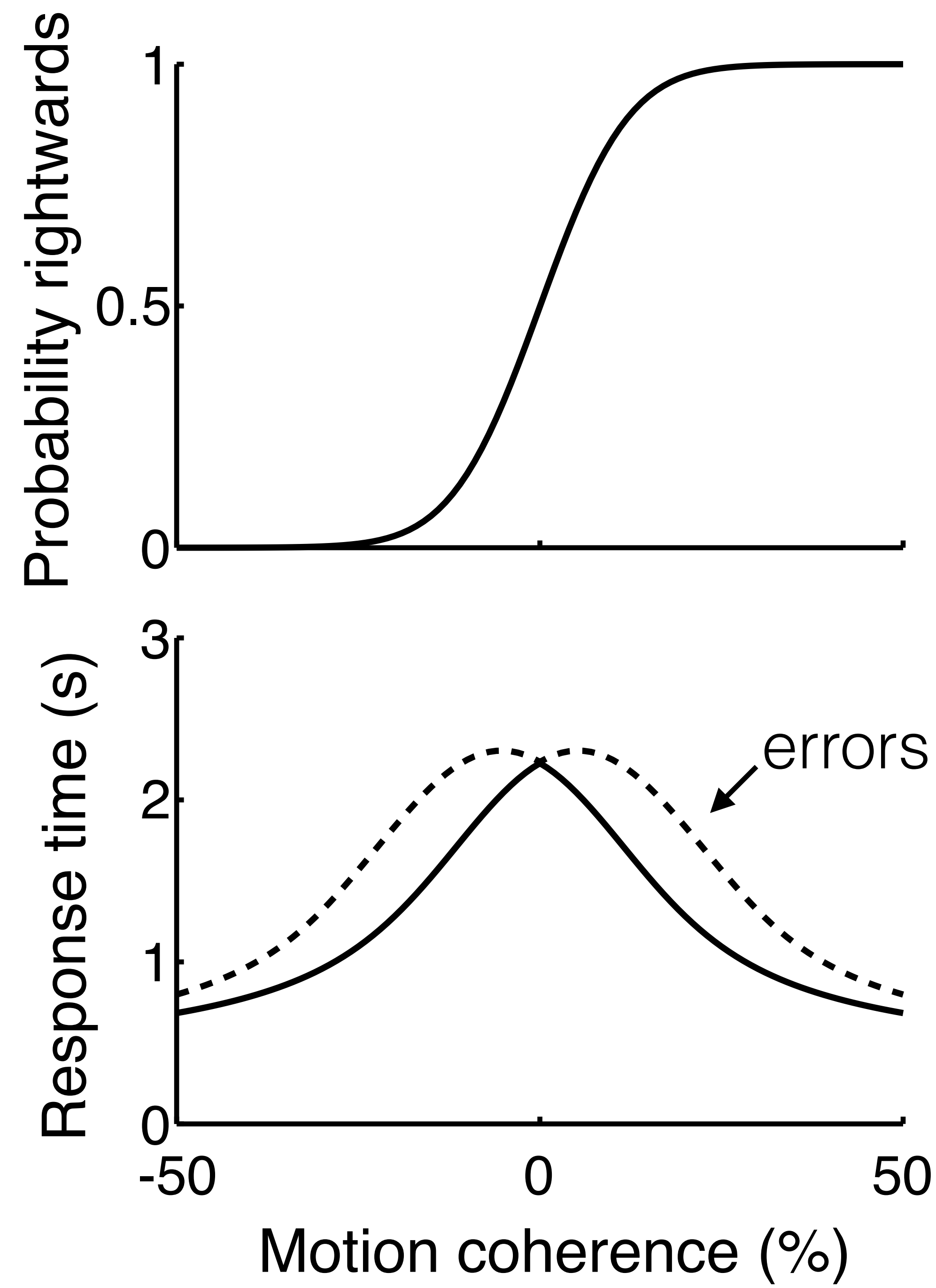




Ratcliff & McKoon (2008)

Heuristic solution: variability in drift-rate across trials



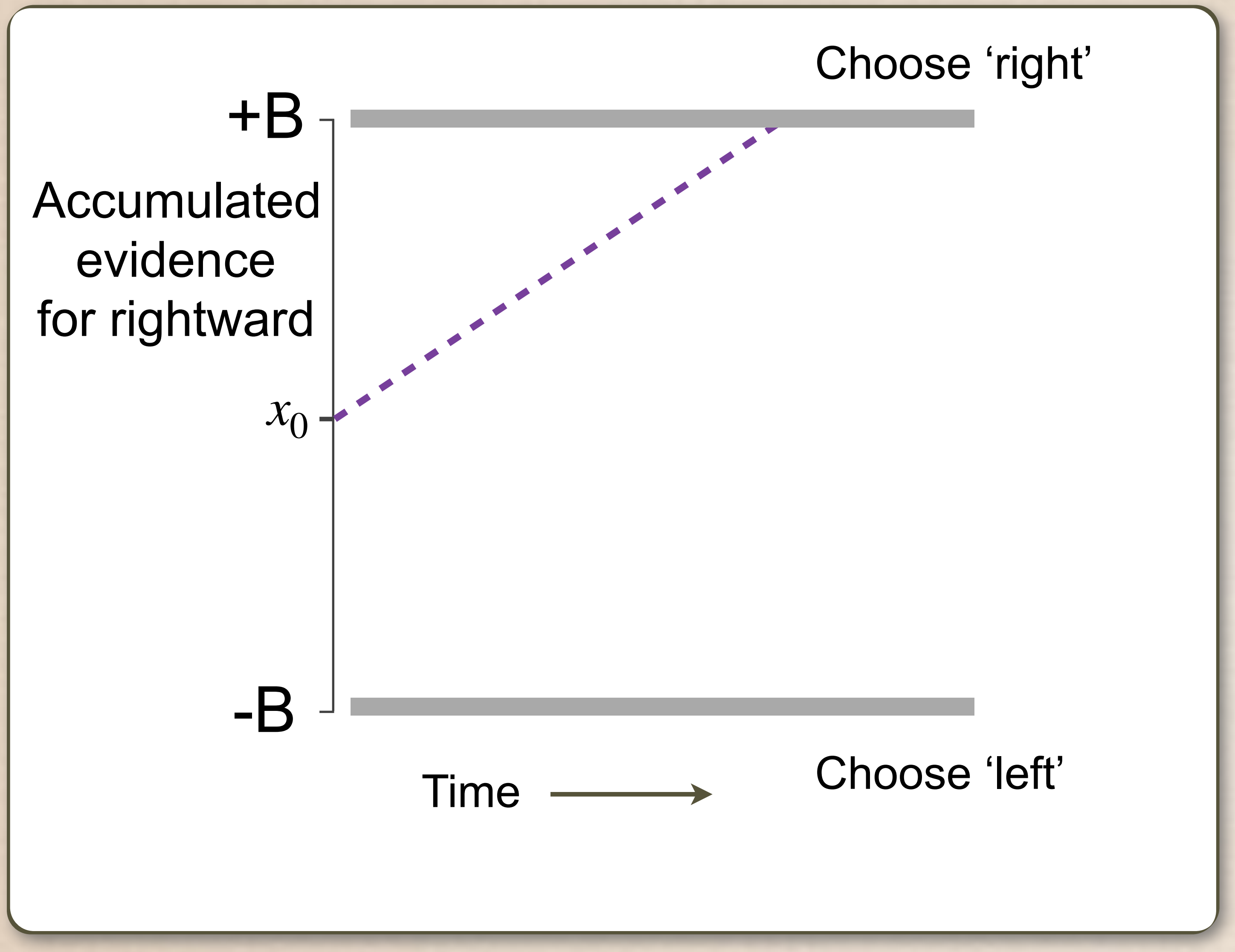


Drift-rates closer to zero → Slower responses

Drift-rates closer to zero → Higher error rates

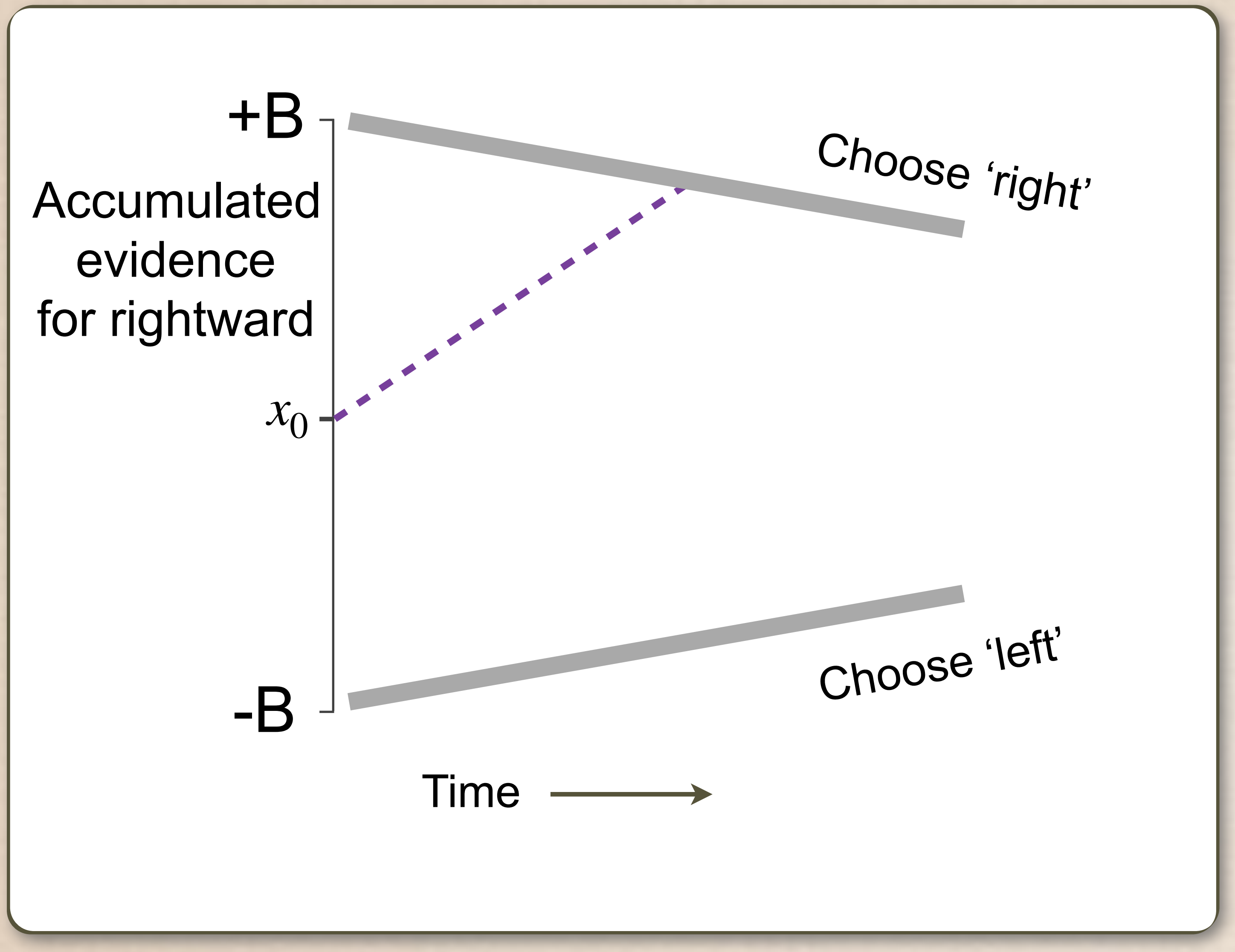
Ratcliff & McKoon (2008)

Collapsing bounds



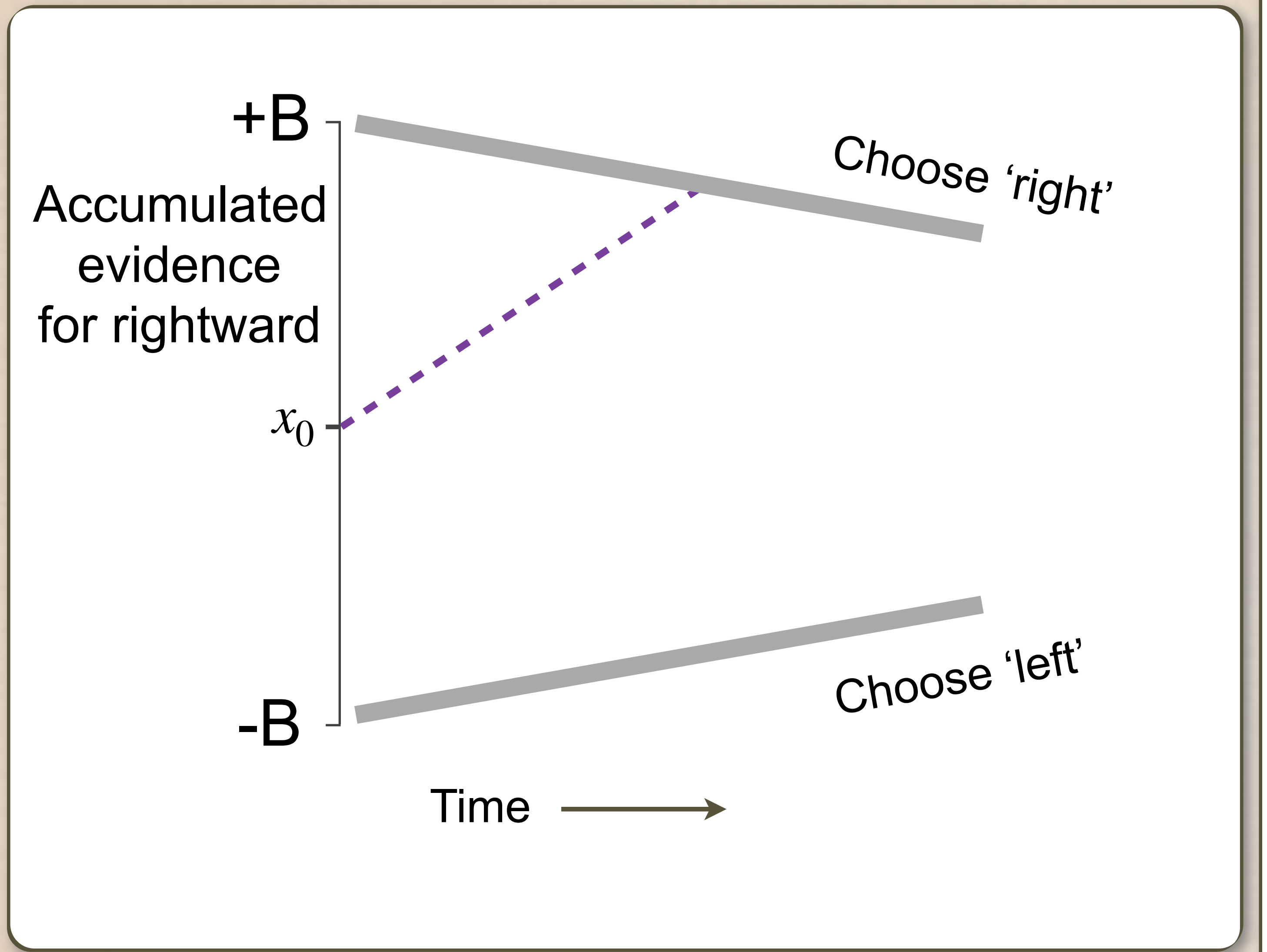
MN Shadlen et al. (2006)

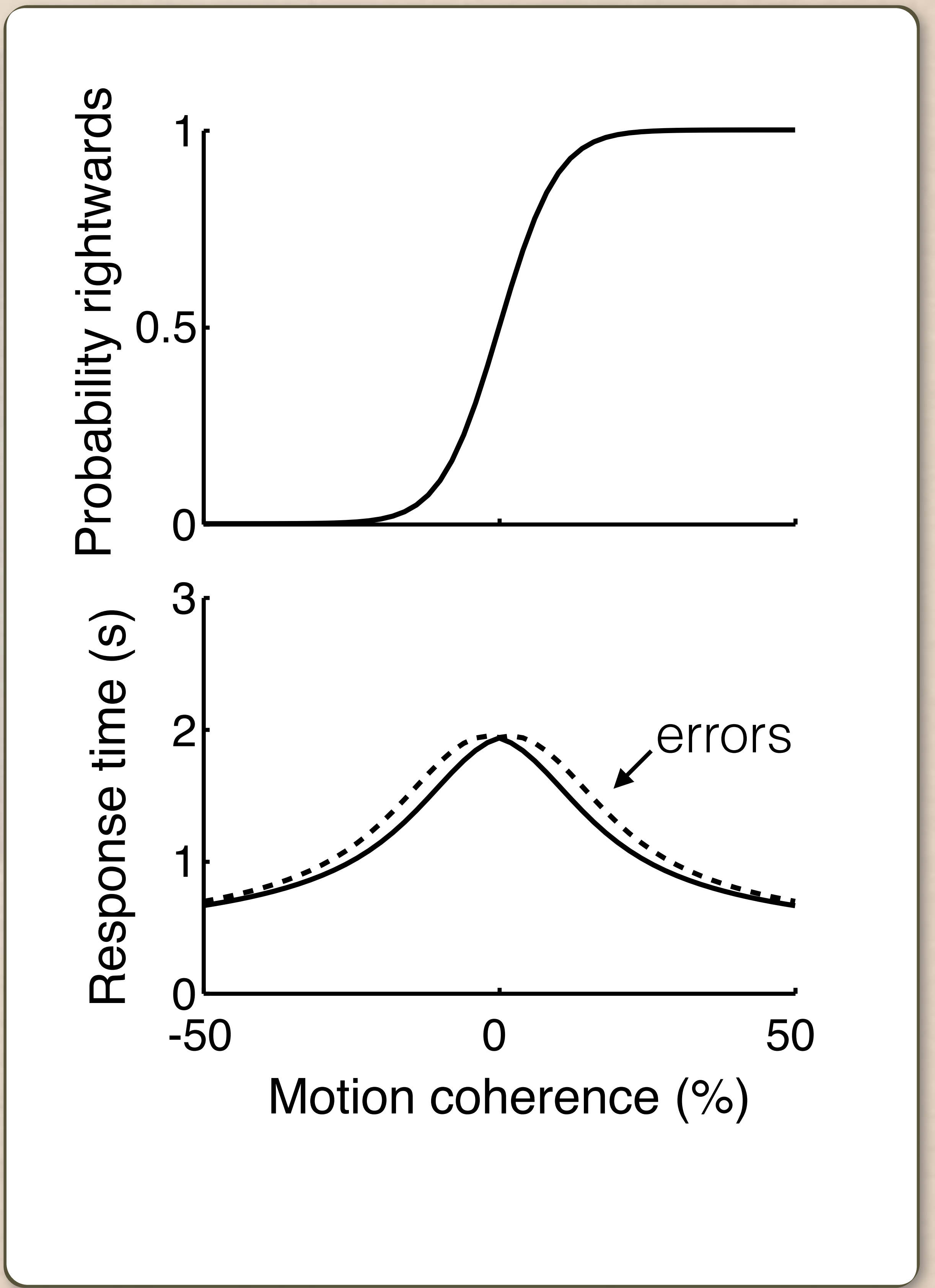
Collapsing bounds



MN Shadlen et al. (2006)

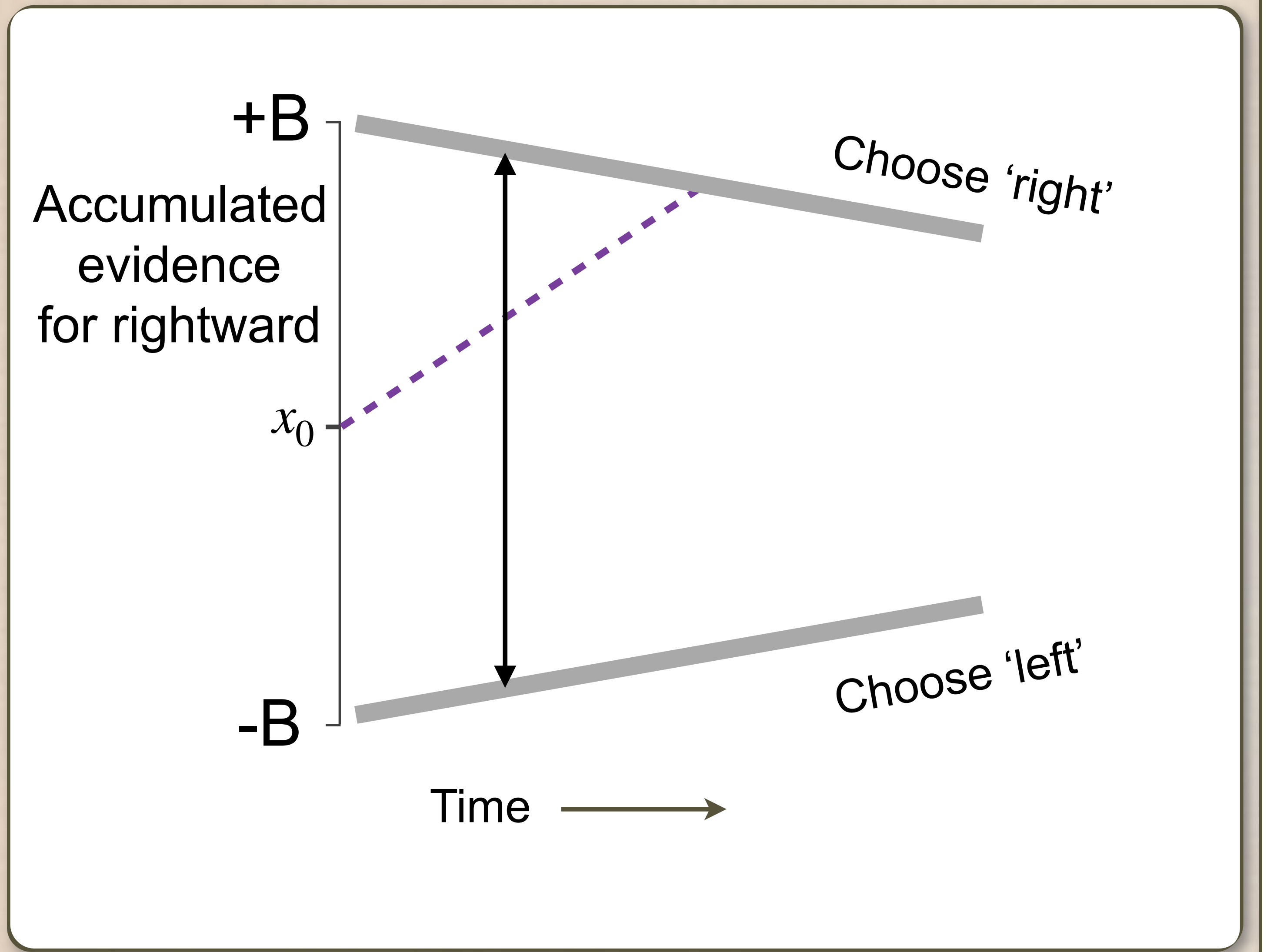
Collapsing bounds

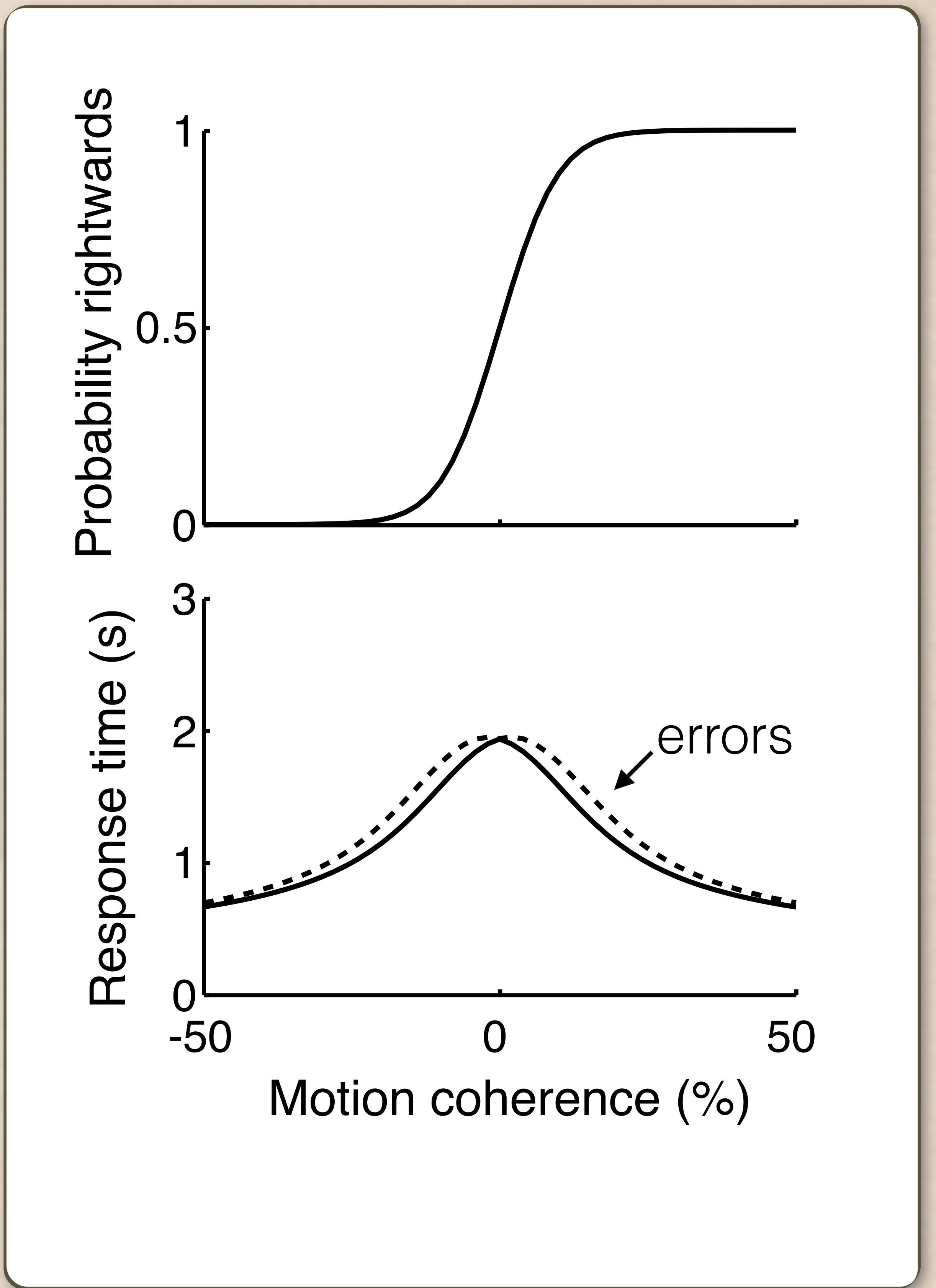




MN Shadlen et al. (2006)

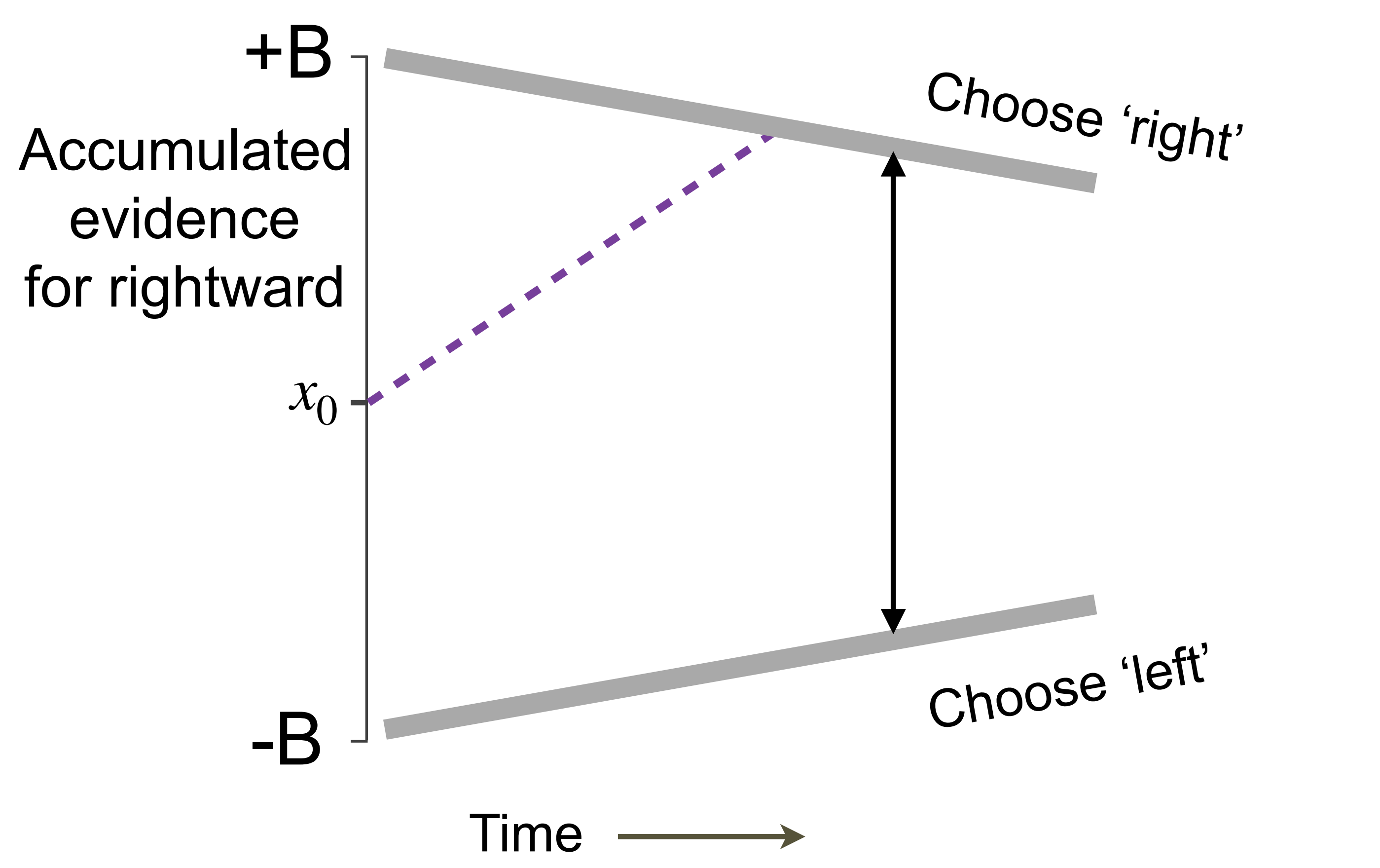
Collapsing bounds

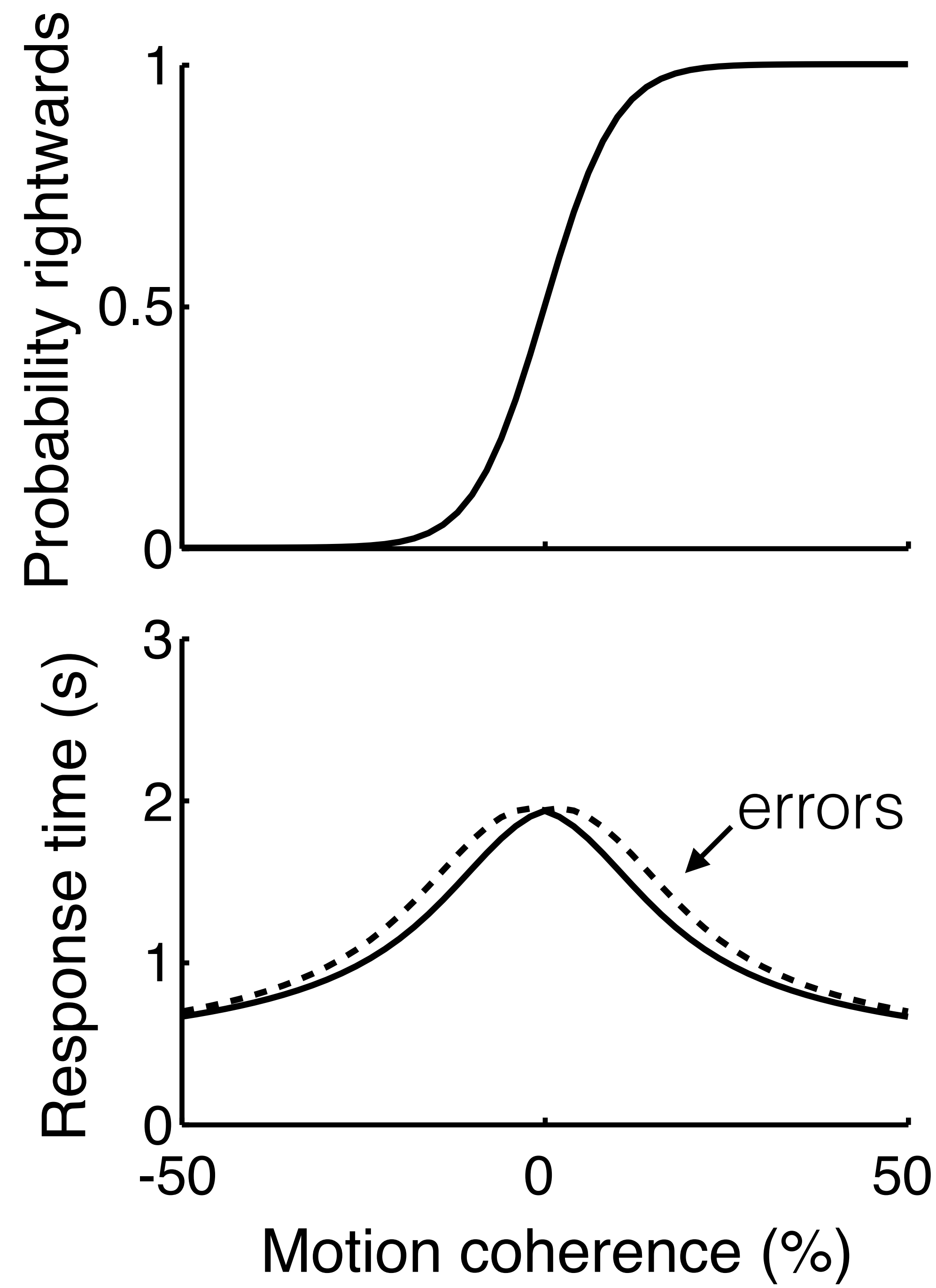




MN Shadlen et al. (2006)

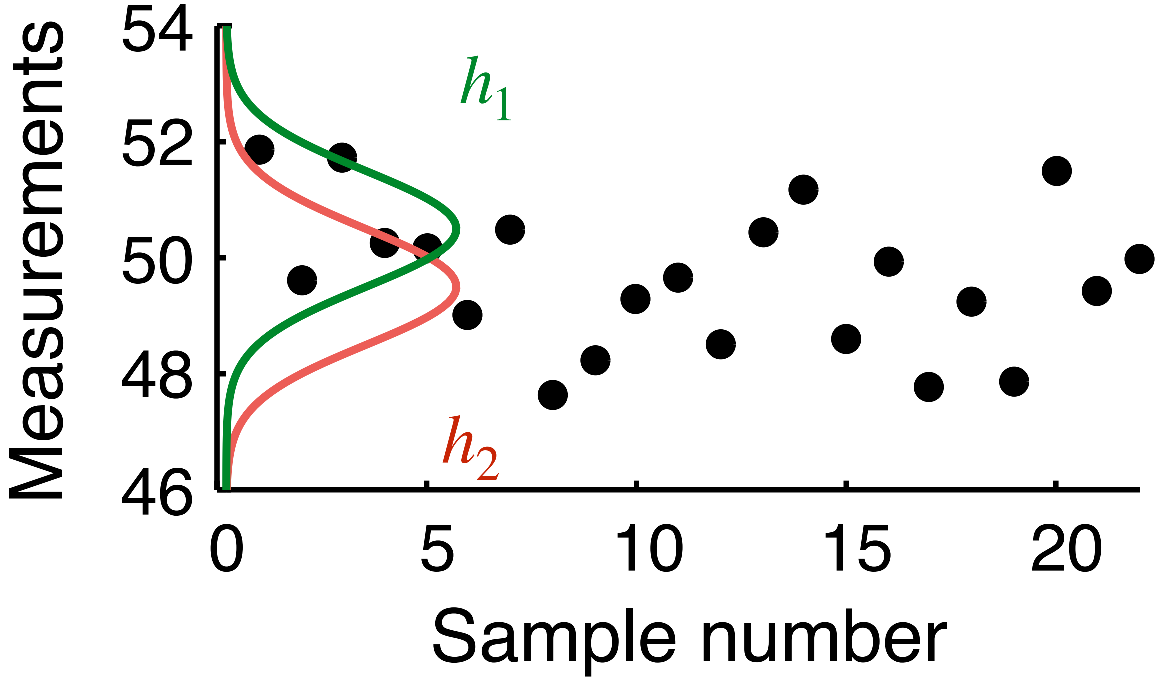
Collapsing bounds





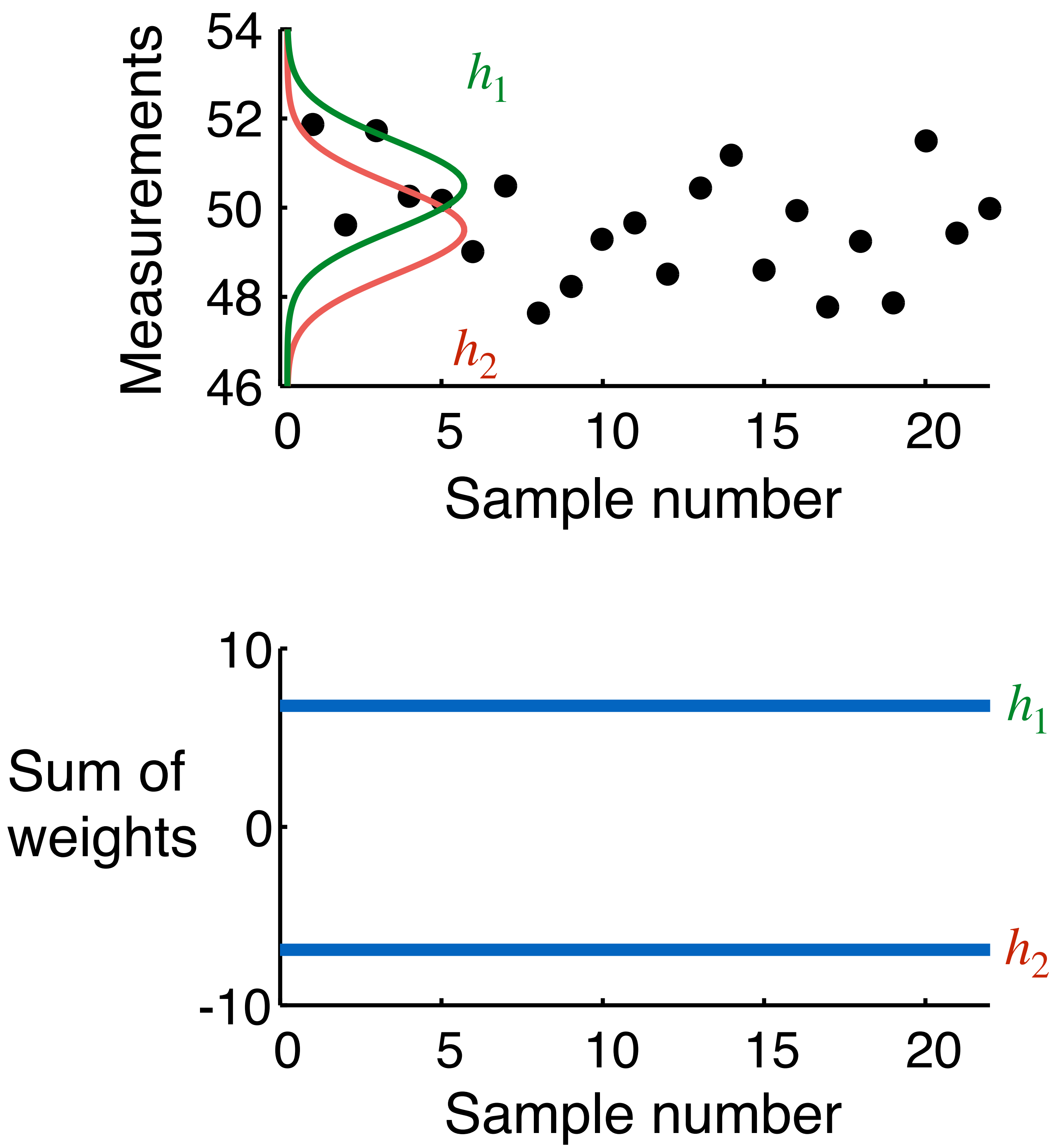
MN Shadlen et al. (2006)

Collapsing bounds maximize reward-rate when difficulty varies across trials



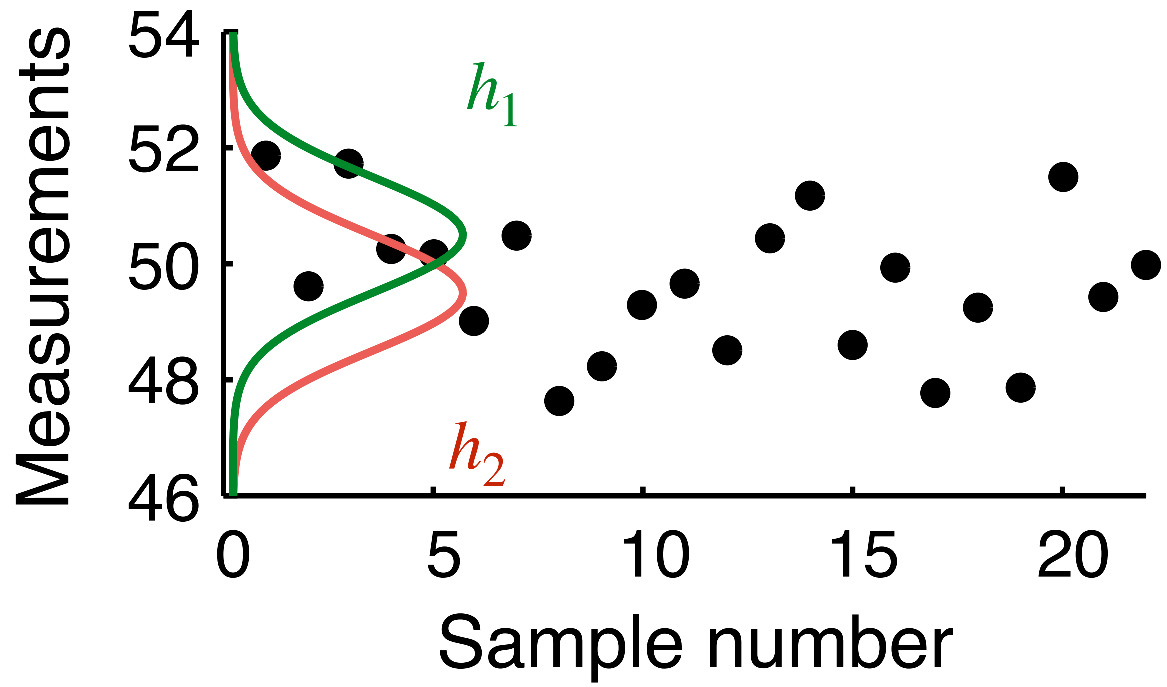
Frazier & Yu, NiPS 2008
Drugowitsch et al., JNeuro 2012

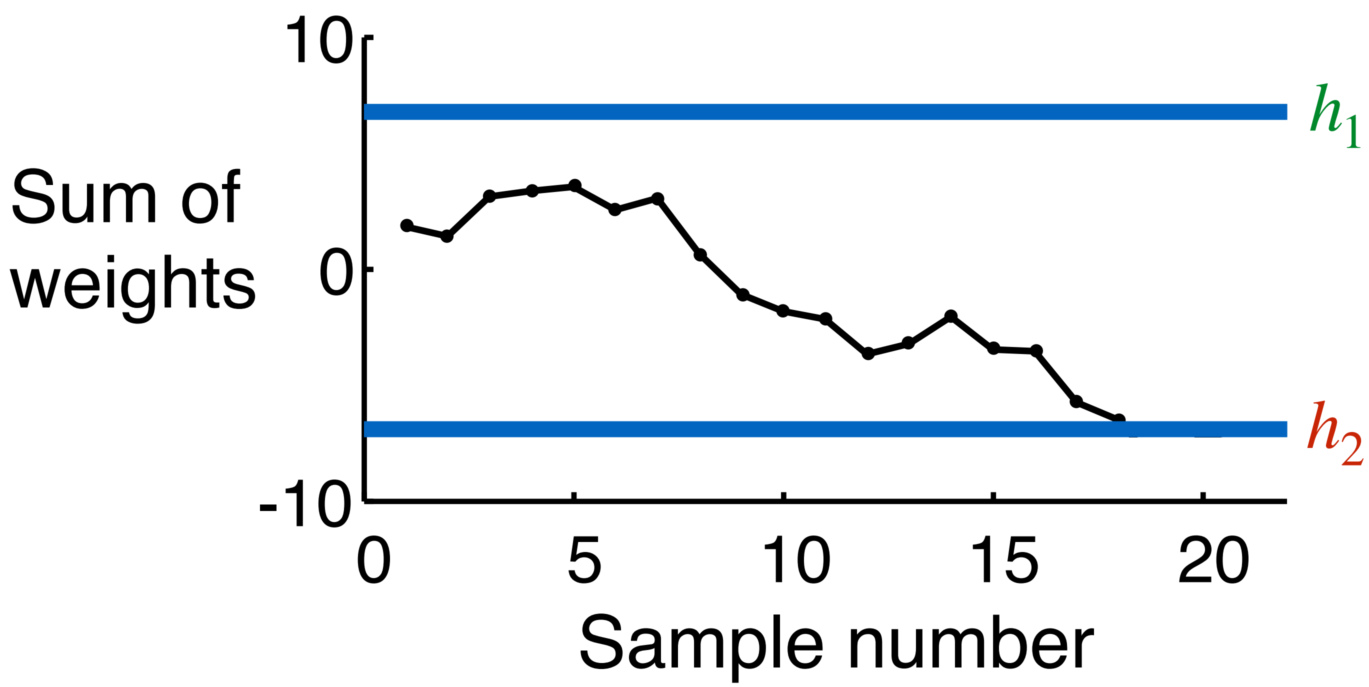
Collapsing bounds maximize reward-rate when difficulty varies across trials



Frazier & Yu, NiPS 2008
Drugowitsch et al., JNeuro 2012

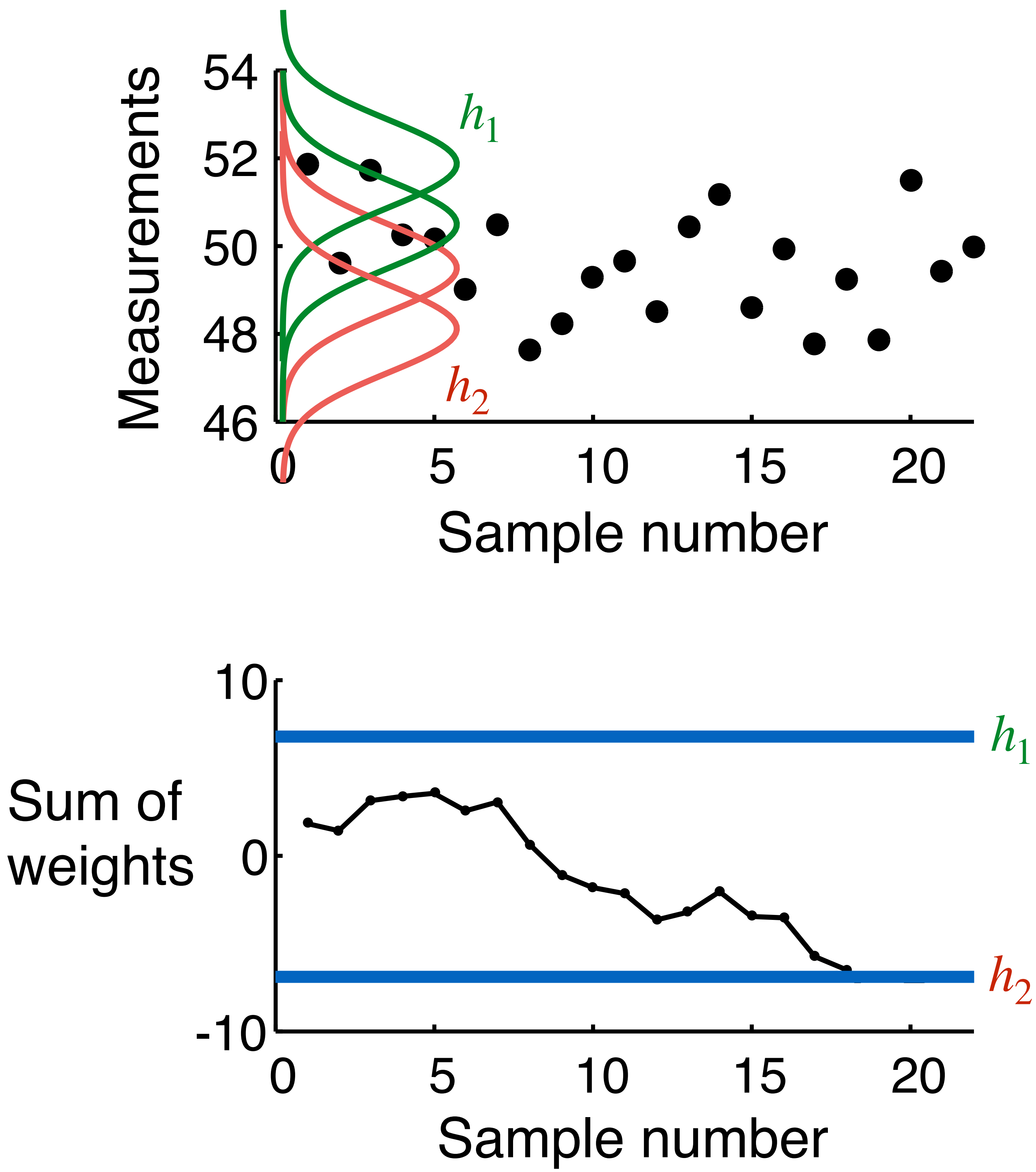
Collapsing bounds maximize reward-rate when difficulty varies across trials





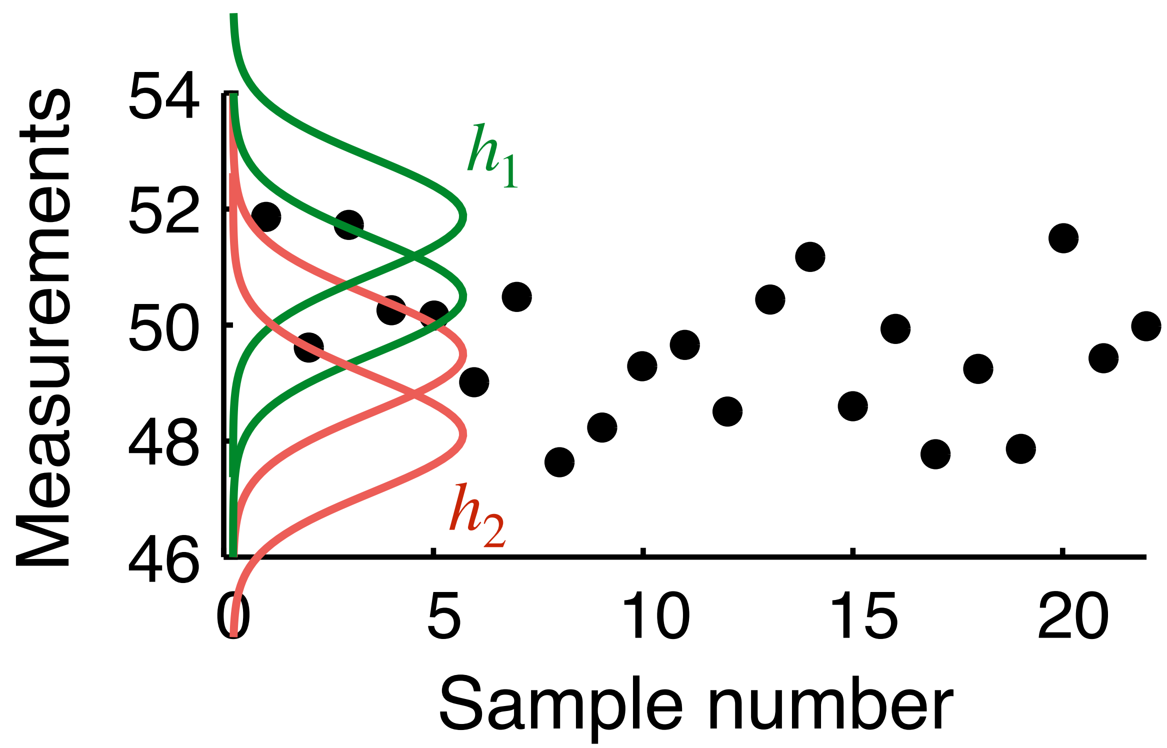
Frazier & Yu, NiPS 2008
Drugowitsch et al., JNeuro 2012

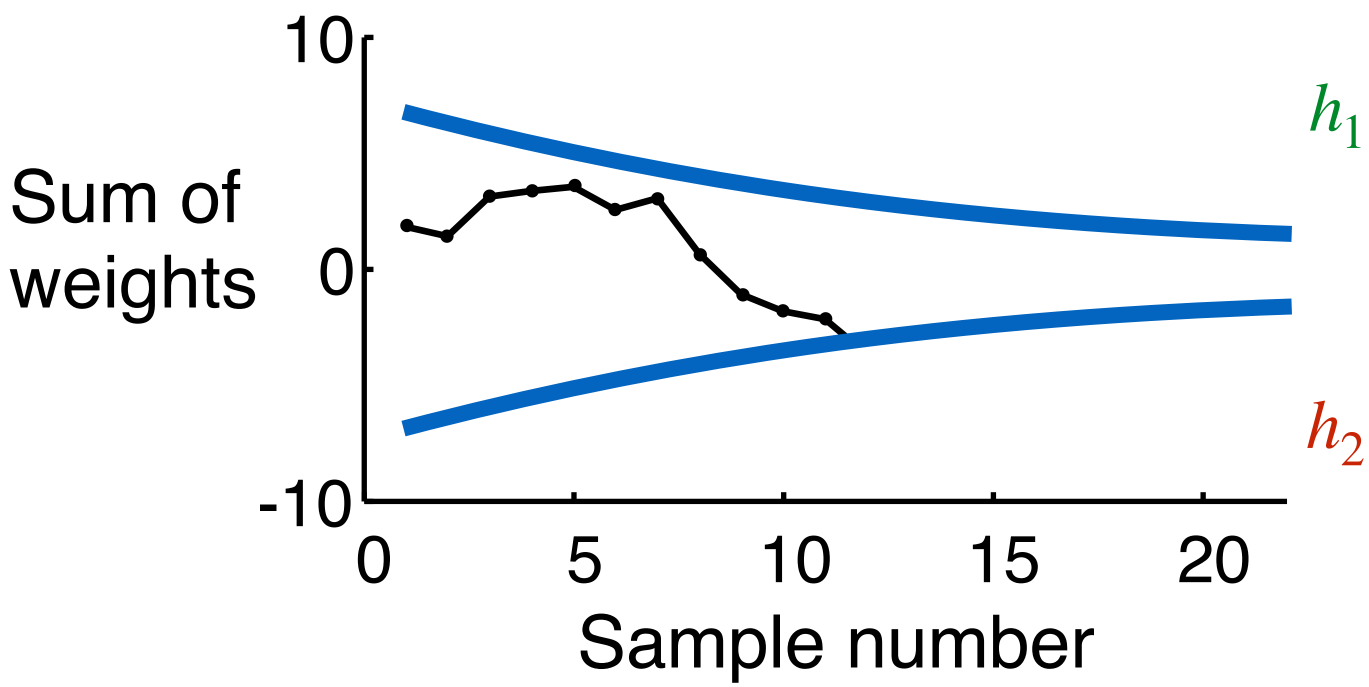
Collapsing bounds maximize reward-rate when difficulty varies across trials



Frazier & Yu, NiPS 2008
Drugowitsch et al., JNeuro 2012

Collapsing bounds maximize reward-rate when difficulty varies across trials





Frazier & Yu, NiPS 2008
Drugowitsch et al., JNeuro 2012

Model fitting

Model fitting

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019

test_data.mat

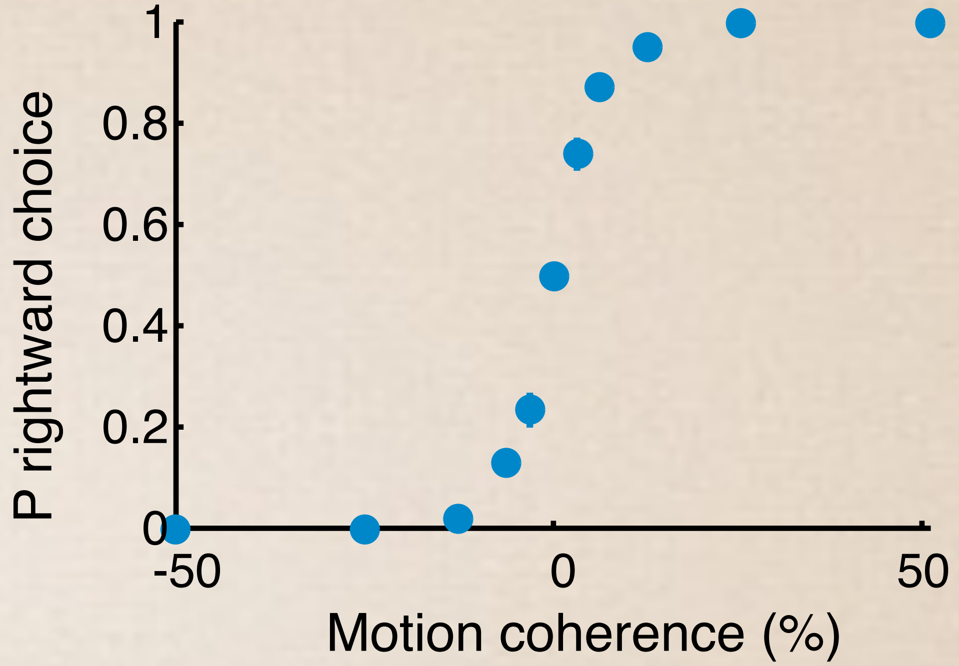
Motion Coherence (%)	Choice	RT (s)	Correct
-3.2	0	2.32	1
-12.8	0	0.76	1
0.0	0	1.31	0
51.2	1	0.39	1
-51.2	0	0.38	1
-25.6	0	0.55	1
-12.8	0	0.52	1
0.0	0	1.40	0
51.2	1	0.31	1
-3.2	0	2.05	1

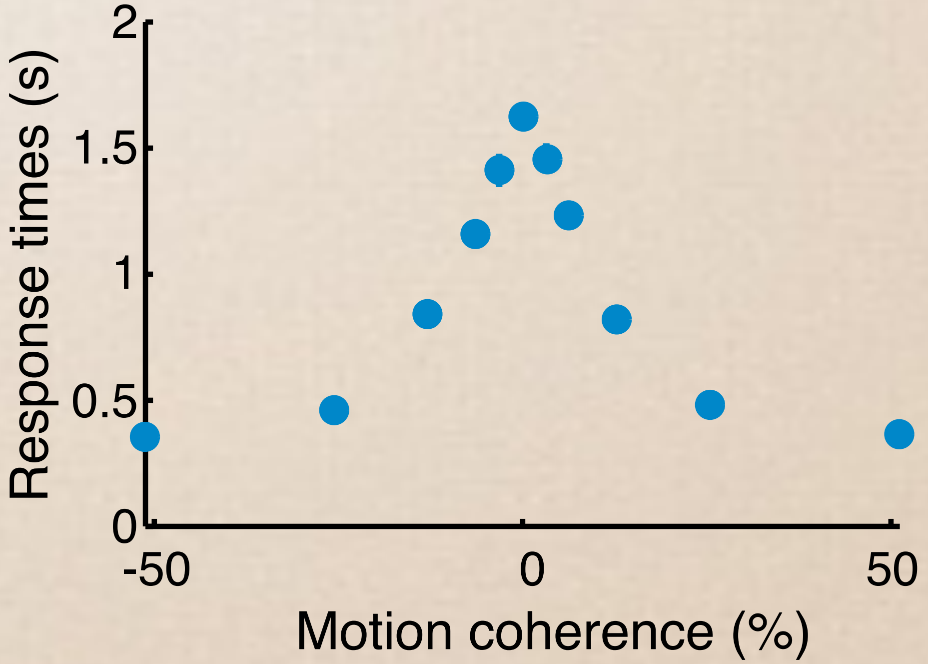
Model fitting

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019

test_data.mat

Motion Coherence (%)	Choice	RT (s)	Correct
-3.2	0	2.32	1
-12.8	0	0.76	1
0.0	0	1.31	0
51.2	1	0.39	1
-51.2	0	0.38	1
-25.6	0	0.55	1
-12.8	0	0.52	1
0.0	0	1.40	0
51.2	1	0.31	1
-3.2	0	2.05	1



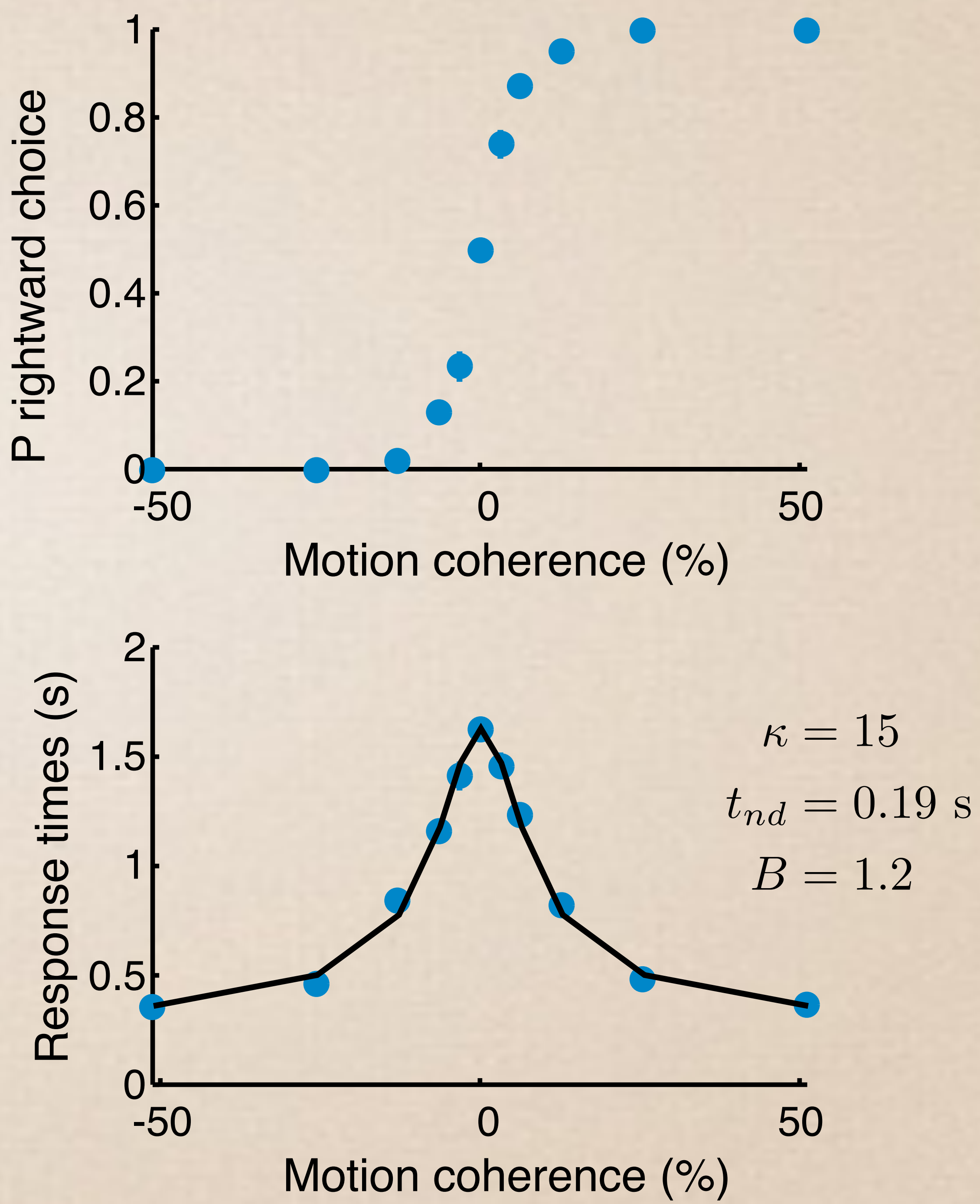


Model fitting

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019

test_data.mat

Motion Coherence (%)	Choice	RT (s)	Correct
-3.2	0	2.32	1
-12.8	0	0.76	1
0.0	0	1.31	0
51.2	1	0.39	1
-51.2	0	0.38	1
-25.6	0	0.55	1
-12.8	0	0.52	1
0.0	0	1.40	0
51.2	1	0.31	1
-3.2	0	2.05	1

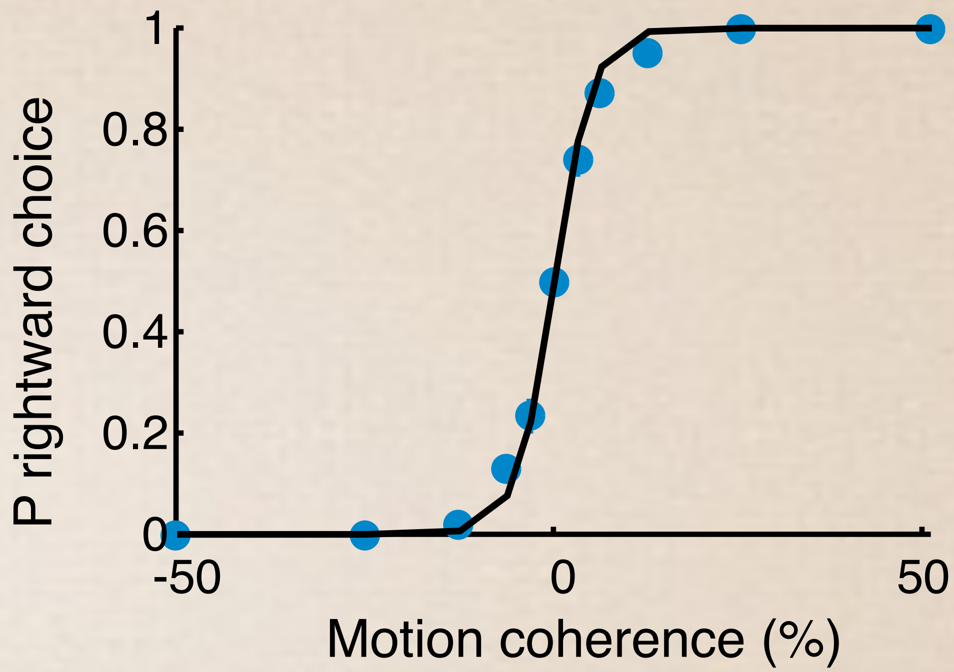


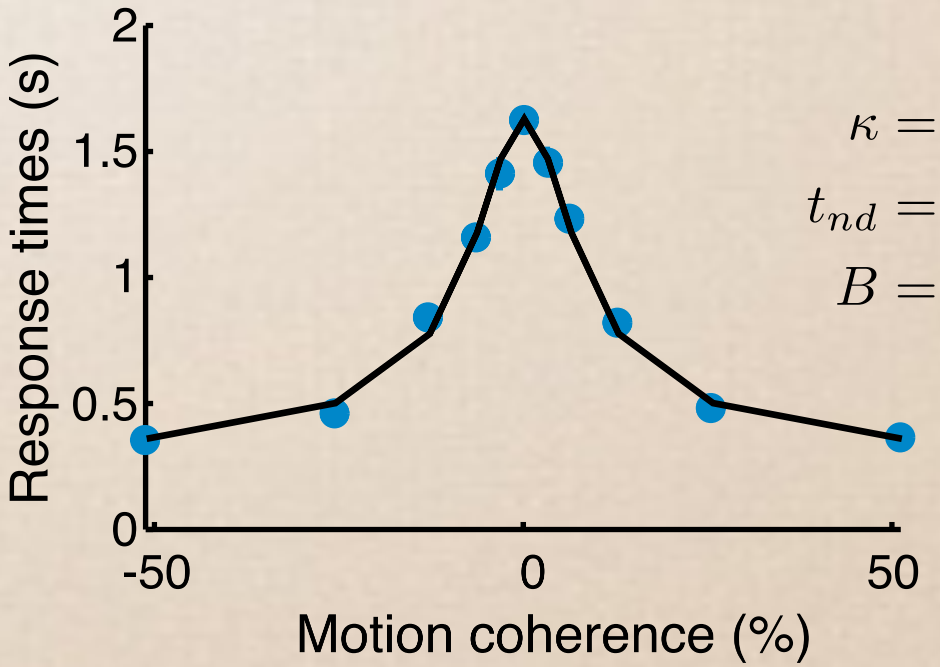
Model fitting

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019

test_data.mat

Motion Coherence (%)	Choice	RT (s)	Correct
-3.2	0	2.32	1
-12.8	0	0.76	1
0.0	0	1.31	0
51.2	1	0.39	1
-51.2	0	0.38	1
-25.6	0	0.55	1
-12.8	0	0.52	1
0.0	0	1.40	0
51.2	1	0.31	1
-3.2	0	2.05	1



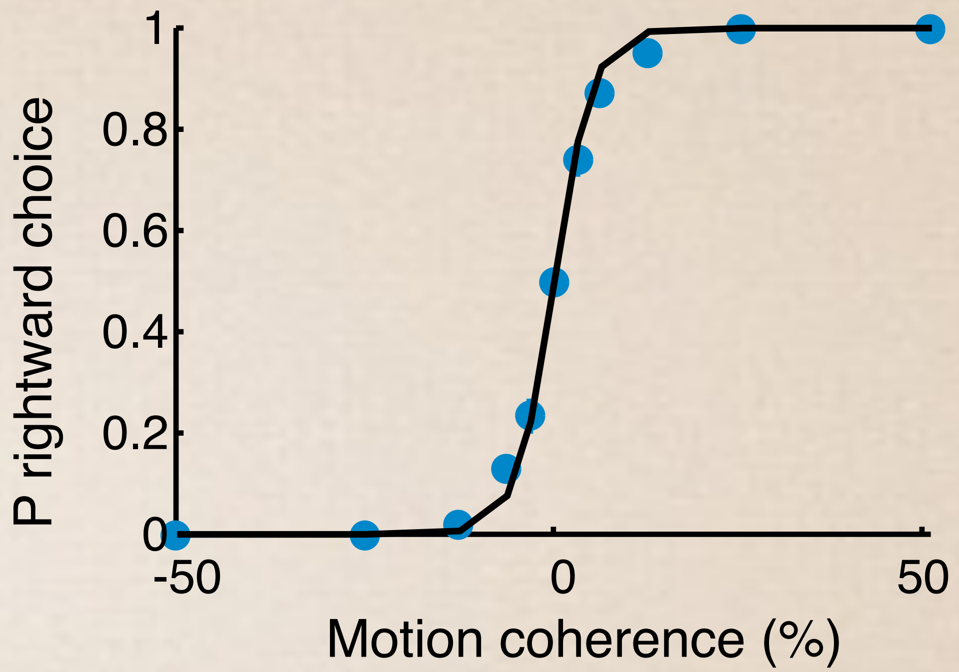


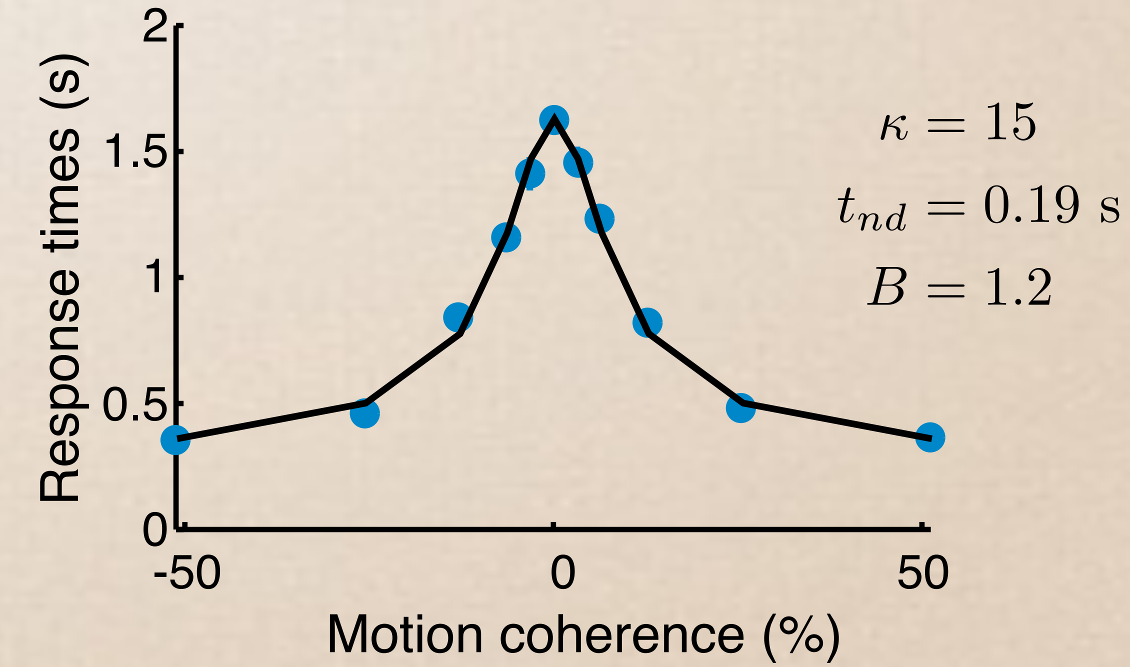
Model fitting

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019

test_data.mat

Motion Coherence (%)	Choice	RT (s)	Correct
-3.2	0	2.32	1
-12.8	0	0.76	1
0.0	0	1.31	0
51.2	1	0.39	1
-51.2	0	0.38	1
-25.6	0	0.55	1
-12.8	0	0.52	1
0.0	0	1.40	0
51.2	1	0.31	1
-3.2	0	2.05	1





fit_means/main.m

Model fitting

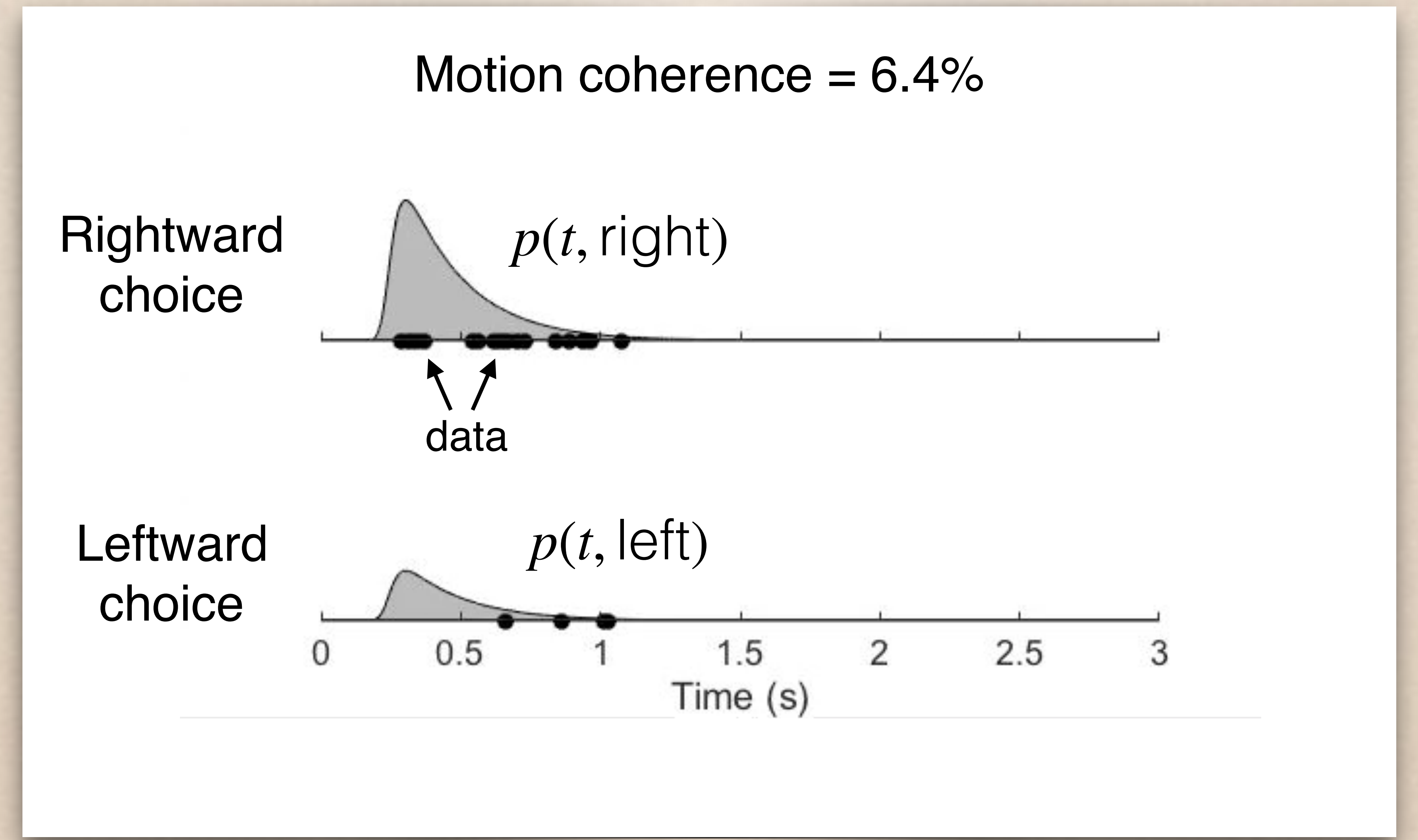
$$\hat{\theta}_{\mathrm{mle}} = \arg\max_{\theta} \left(\sum_{i=1}^{n_{tr}} p(\mathrm{RT}^{(i)}, \mathrm{choice}^{(i)} | c^{(i)}, \theta) \right)$$

motion
coherence

Model fitting

$$\hat{\theta}_{\text{mle}} = \arg\max_{\theta} \left(\sum_{i=1}^{n_{tr}} p(\text{RT}^{(i)}, \text{choice}^{(i)} | c^{(i)}, \theta) \right)$$

motion coherence



Analytic solution to the choice and response time distribution

Probability of terminating at the lower bound at time t:

$$p(t, \text{left}|\mu, B, \sigma, x_0) = p_-(\mu, B, \sigma, x_0) - \frac{\pi\sigma^2}{4B^2} e^{-\frac{(B+x_0)\mu}{\sigma^2}} \times \sum_{k=1}^{\infty} \frac{2k \ \sin(\frac{k\pi(B+x_0)}{2B}) e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2} + \frac{\pi^2 k^2 \sigma^2}{4B^2})t}}{\left(\frac{\mu^2}{\sigma^2} + \frac{\pi^2 k^2 \sigma^2}{4B^2}\right)}$$

, where

$$p_-(\mu, B, \sigma, x_0) = \frac{e^{-(4\mu B/\sigma^2)} - e^{-(2\mu(B+x_0)/\sigma^2)}}{e^{-(4\mu B/\sigma^2)} - 1}$$

Ratcliff 1978, Feller 1968

Fitting the DDM with flat bounds

1. HDDM

   http://ski.clps.brown.edu/hddm_docs/

2. DMAT

   https://ppw.kuleuven.be/okp/software/dmat/

3. fast-dm

   https://www.psychologie.uni-heidelberg.de/ae/meth/fast-dm/

4. EZ

   https://www.ejwagenmakers.com/2007/EZ.pdf

When analytical solutions are not available (e.g., collapsing bounds)

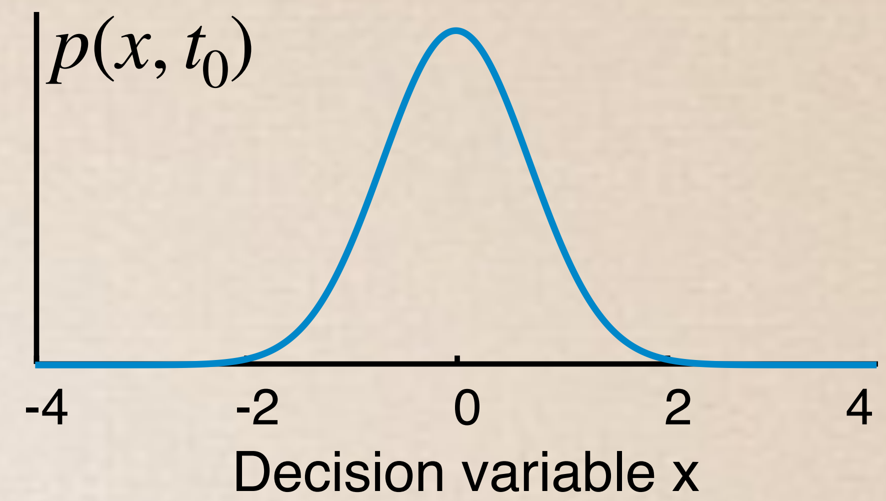
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}$$

When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

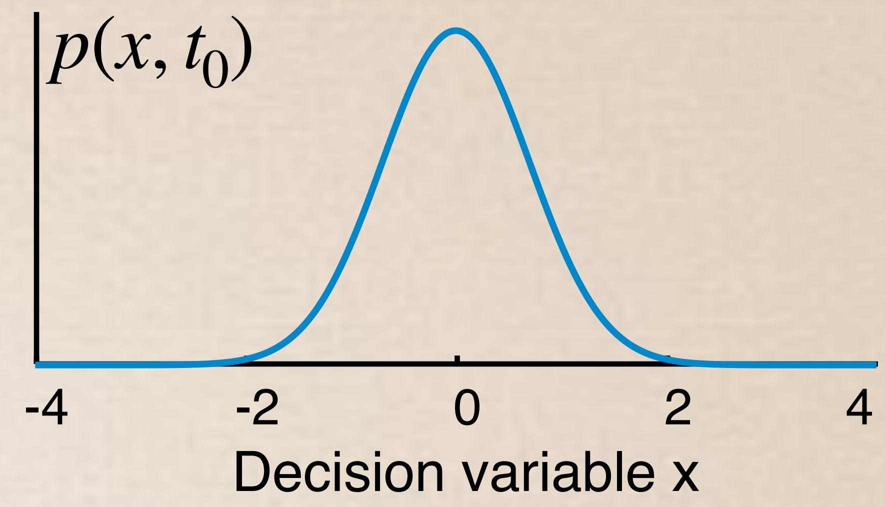
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}$$



When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

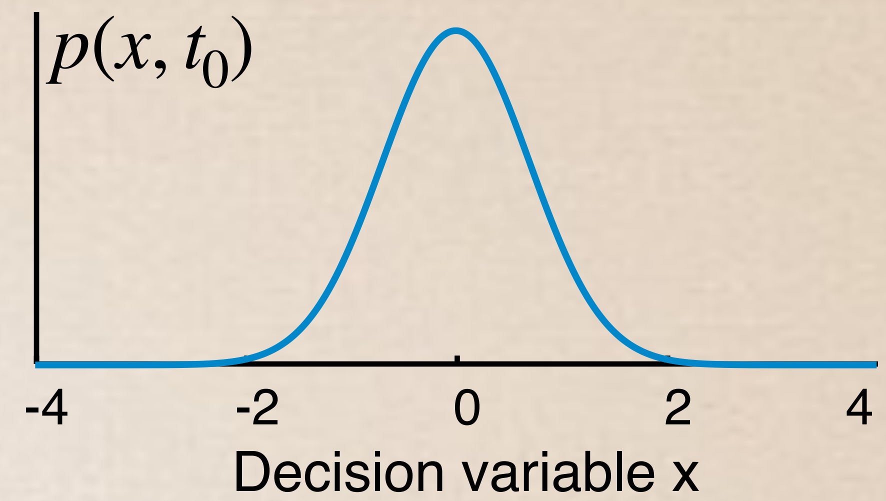
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \cancel{\frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}}$$



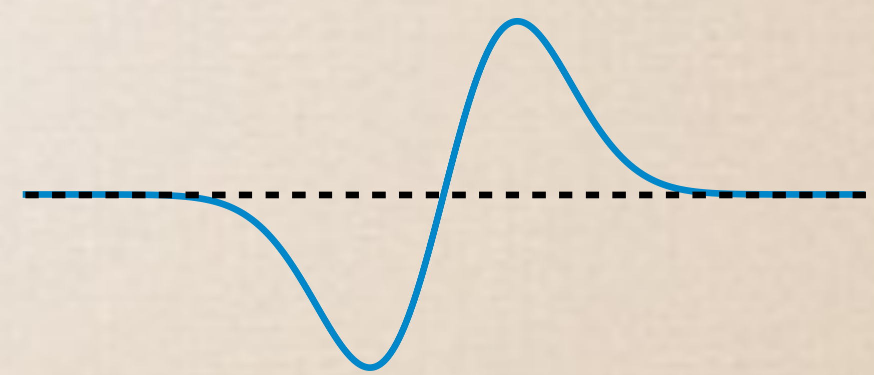
When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \cancel{\frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}}$$



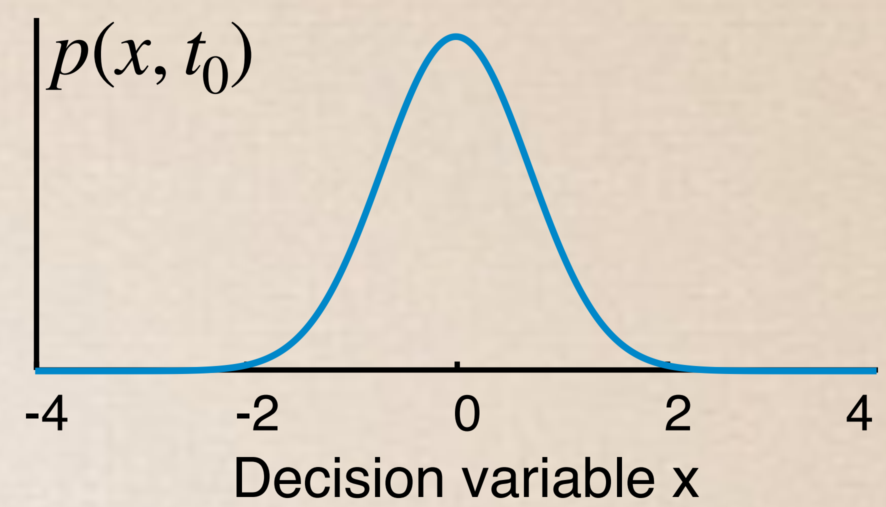
$$-\mu \frac{\partial p(x,t)}{\partial x}$$



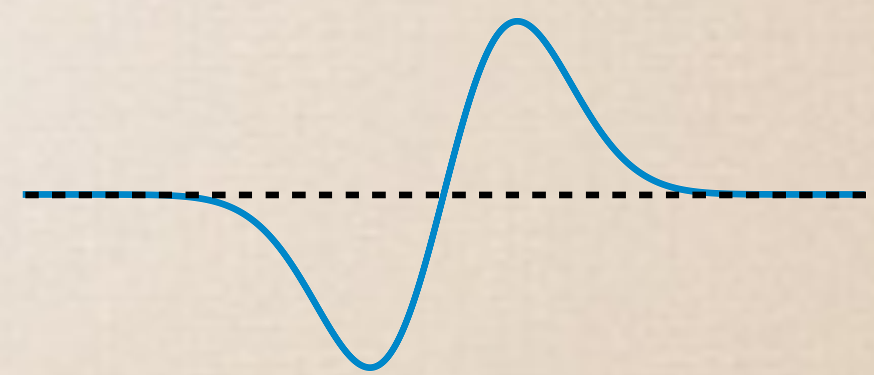
When analytical solutions are not available (e.g., collapsing bounds)

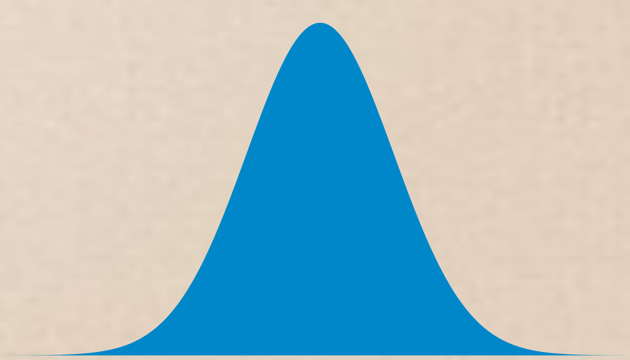
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \cancel{\frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}}$$



$$-\mu \frac{\partial p(x,t)}{\partial x}$$

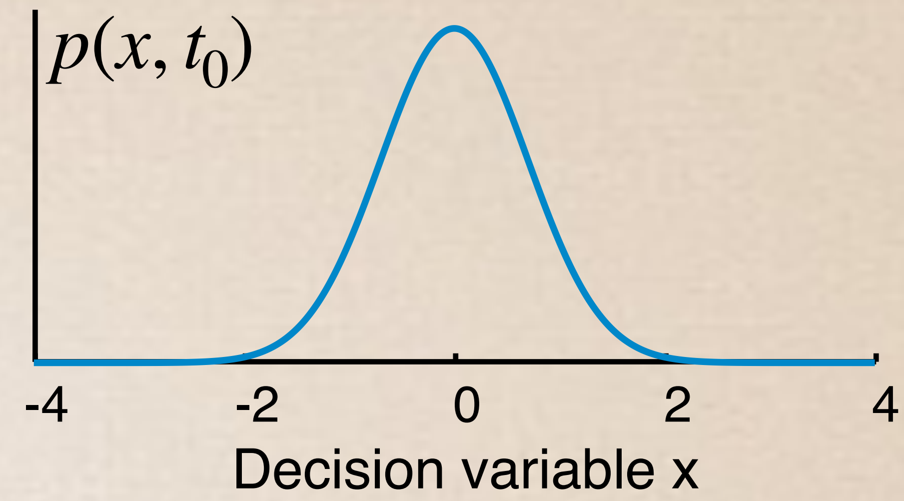




When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

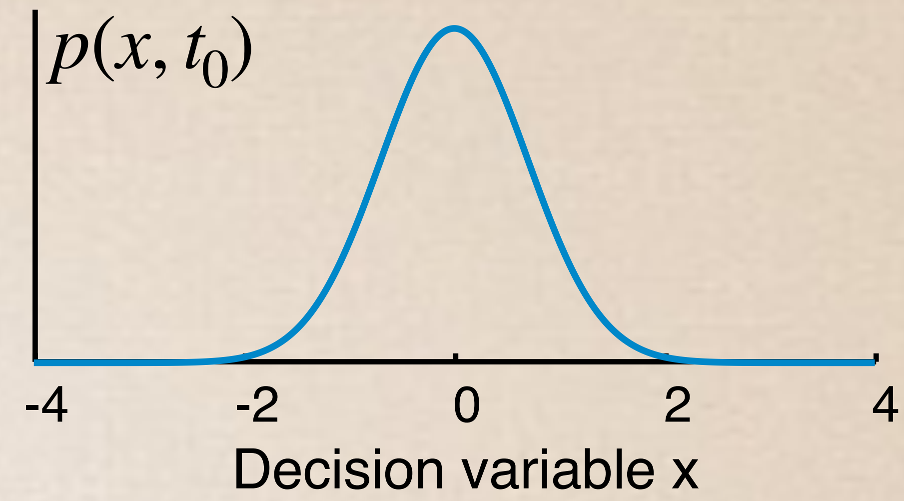
$$\frac{\partial p(x,t)}{\partial t} = \cancel{-\mu \frac{\partial p(x,t)}{\partial x}} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



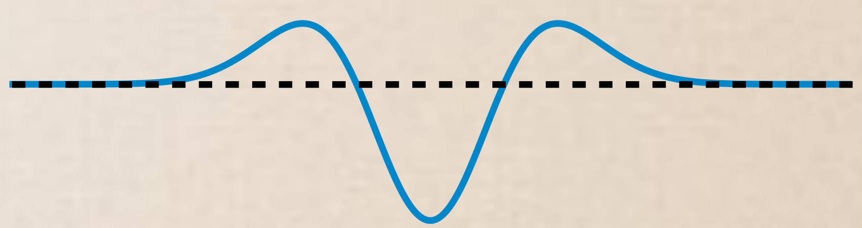
When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = \cancel{-\mu \frac{\partial p(x,t)}{\partial x}} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



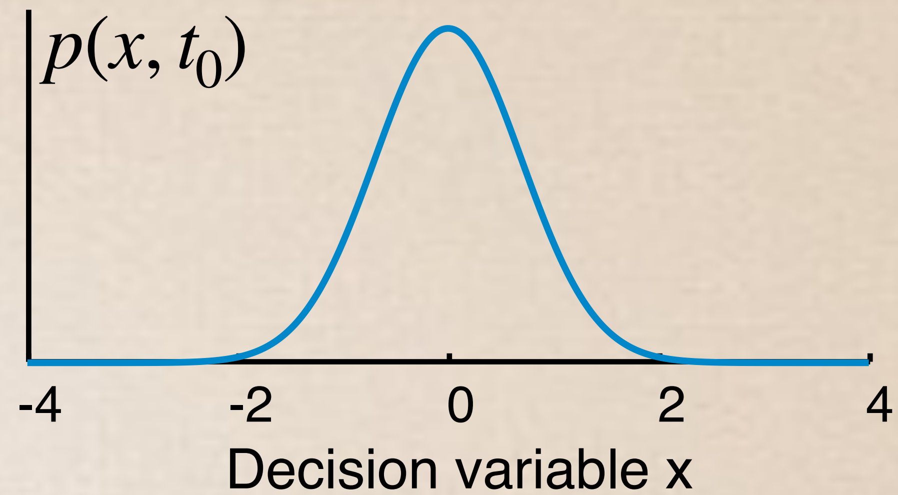
$$\frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



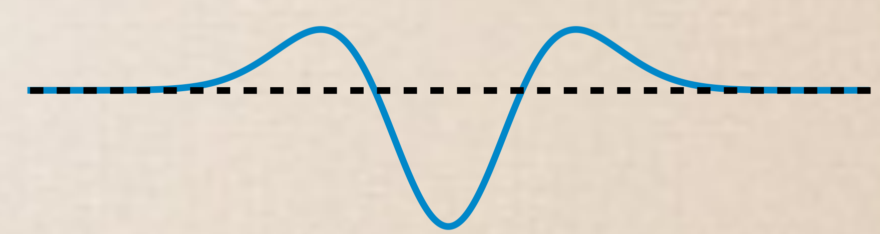
When analytical solutions are not available (e.g., collapsing bounds)

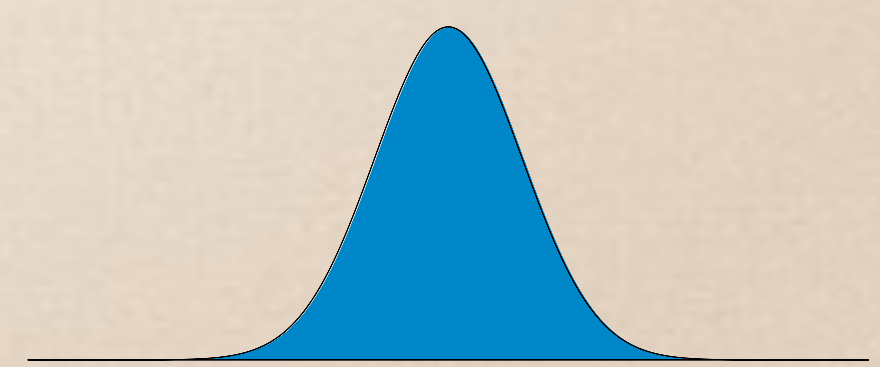
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = \cancel{-\mu \frac{\partial p(x,t)}{\partial x}} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



$$\frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$

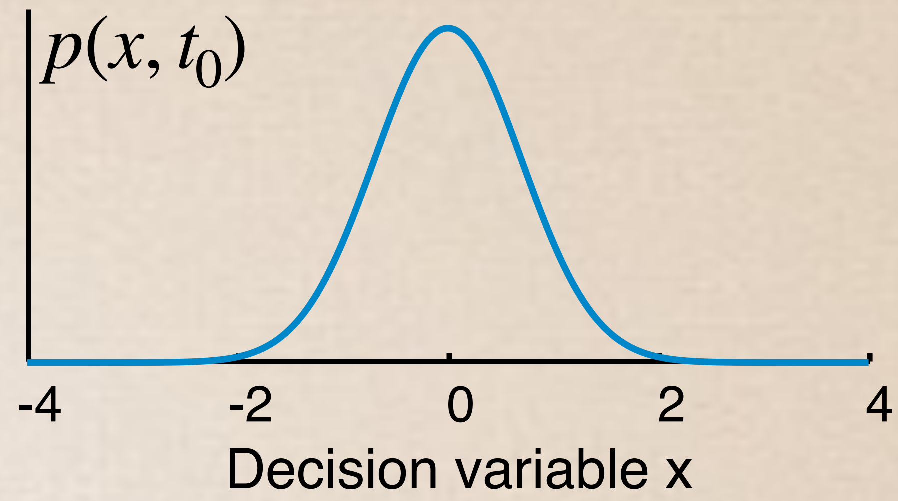




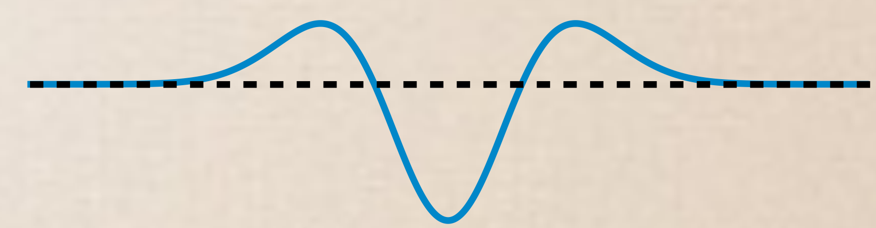
When analytical solutions are not available (e.g., collapsing bounds)

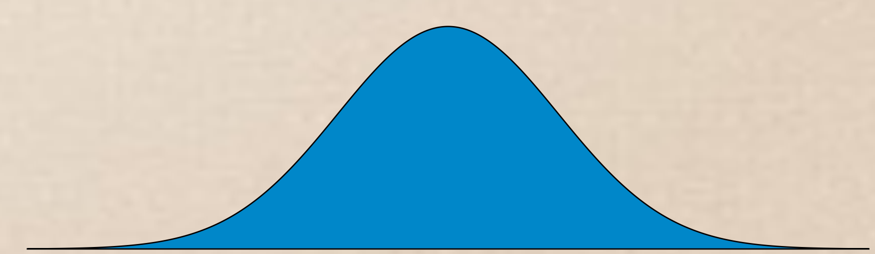
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = \cancel{-\mu \frac{\partial p(x,t)}{\partial x}} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



$$\frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$

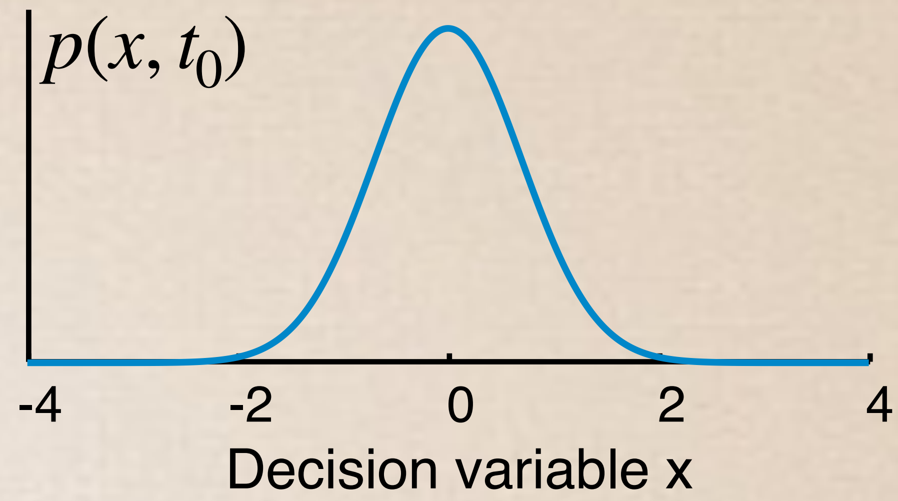




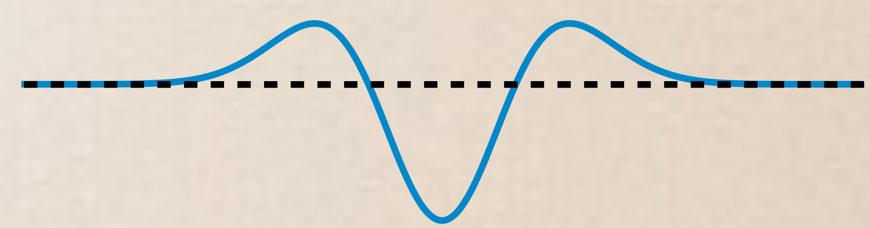
When analytical solutions are not available (e.g., collapsing bounds)

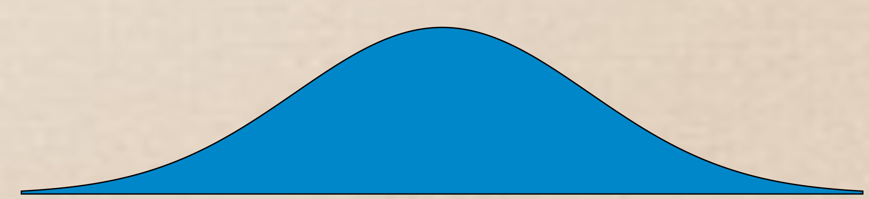
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\cancel{\mu \frac{\partial p(x,t)}{\partial x}} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



$$\frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$

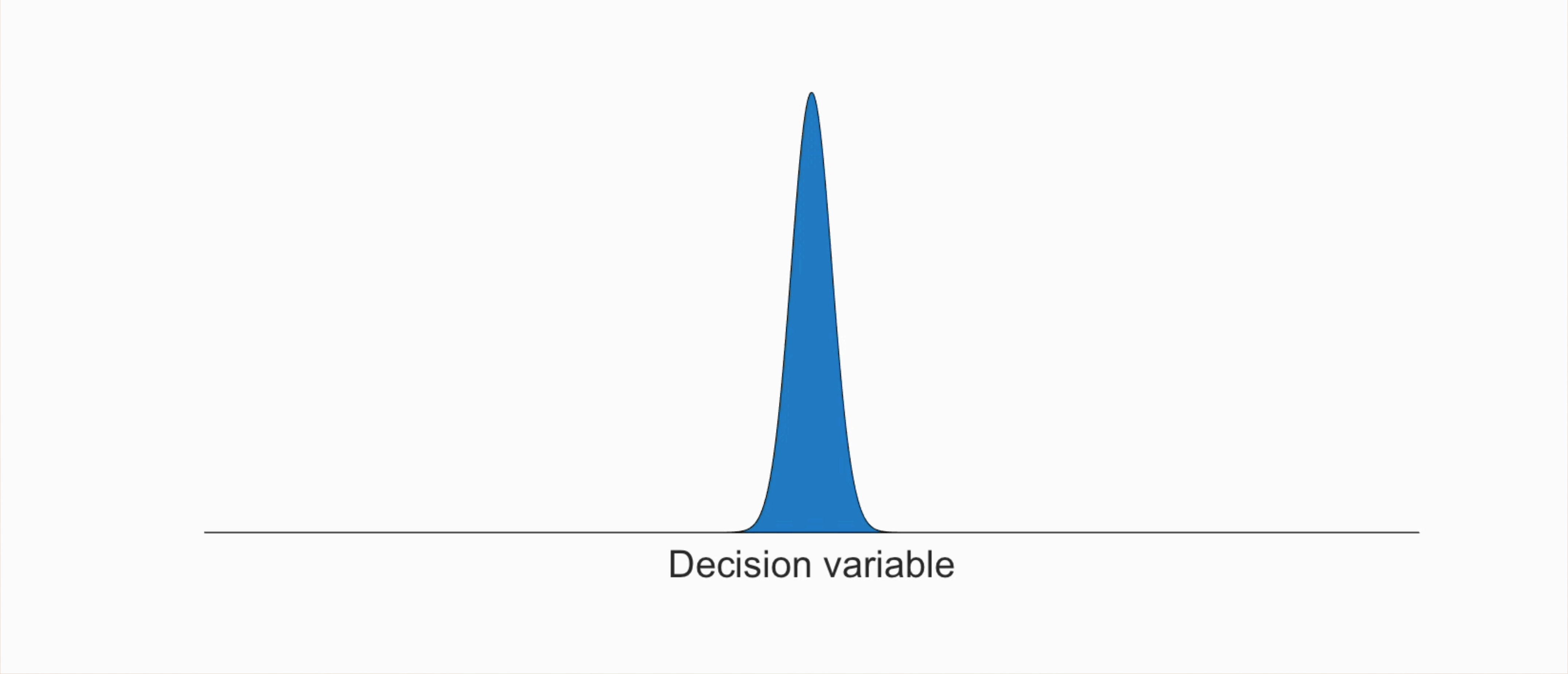




When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

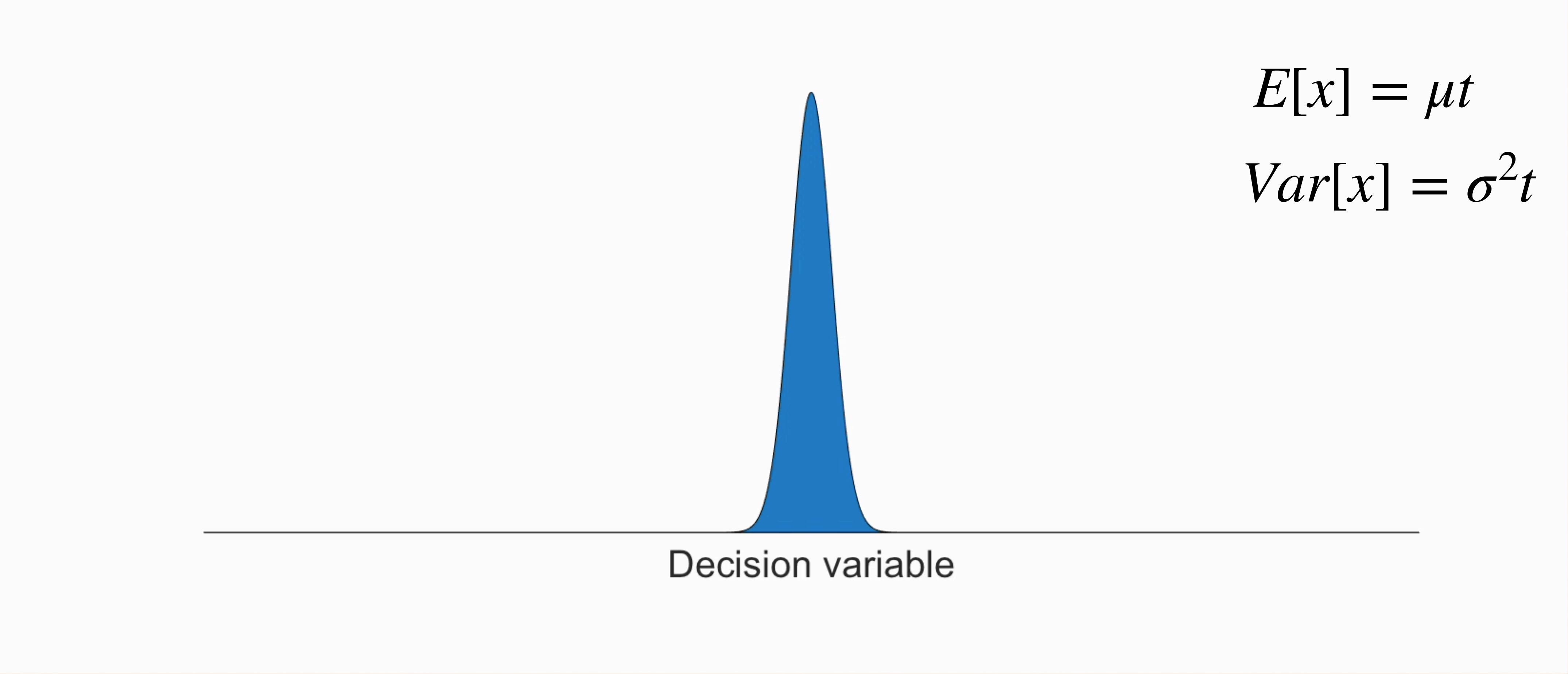
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 p(x,t)}{\partial^2 x}$$



When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

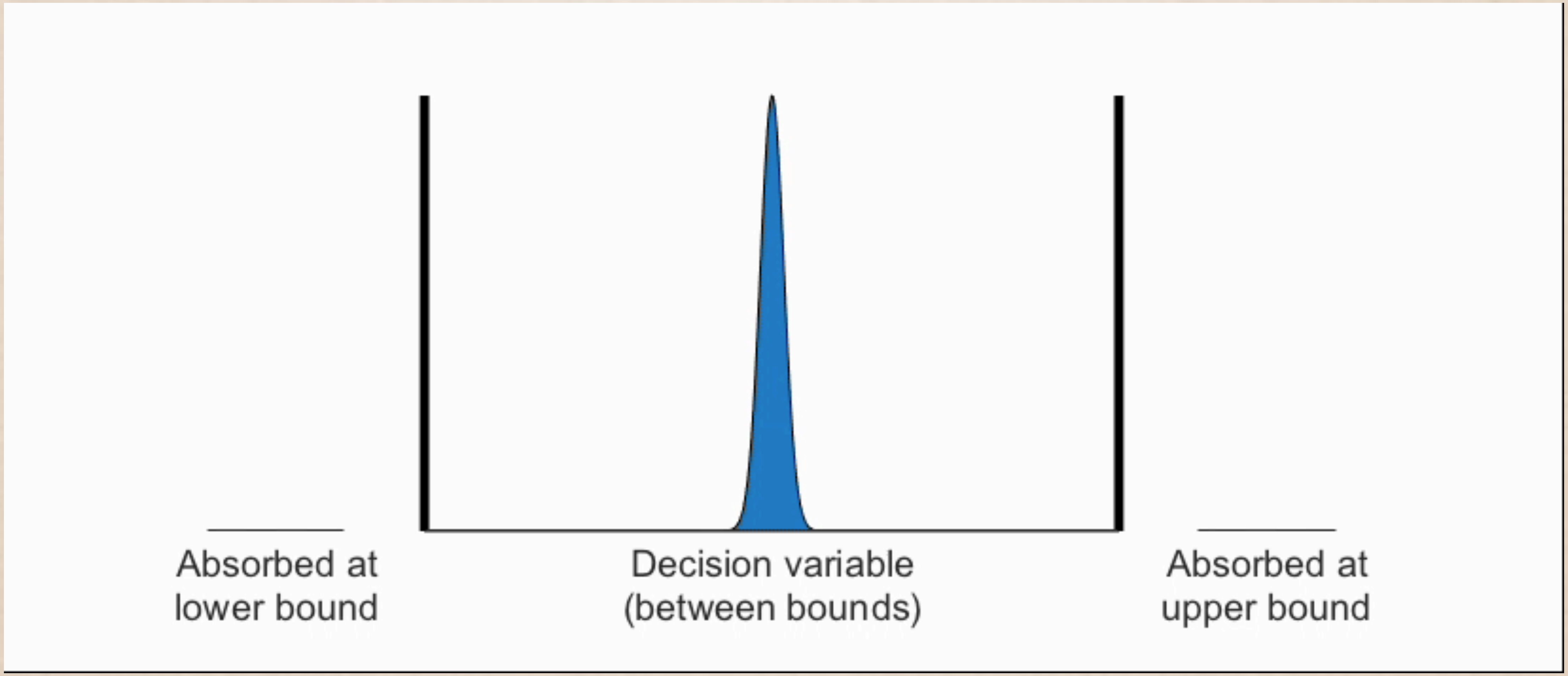
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}$$



When analytical solutions are not available (e.g., collapsing bounds)

Fokker Planck equation:

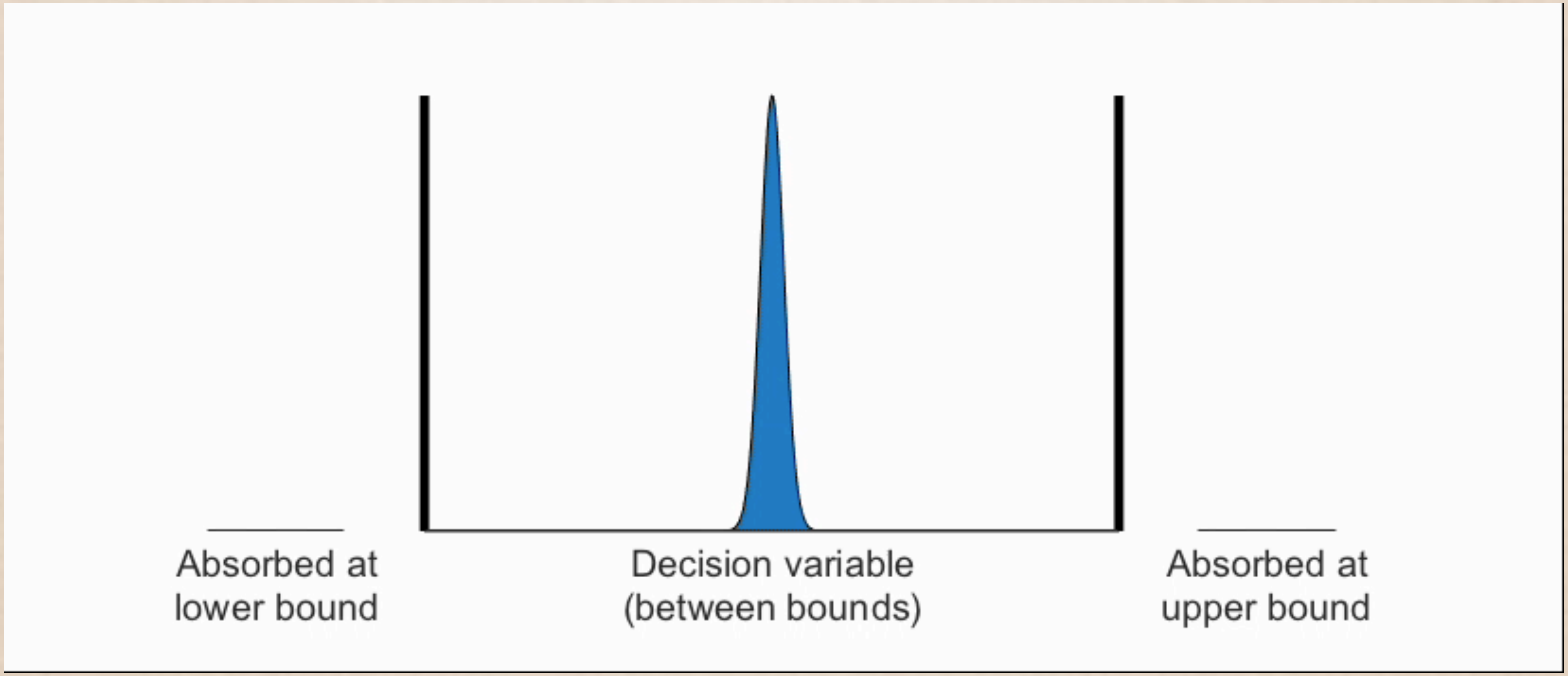
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}$$



When analytical solutions are not available (e.g., collapsing bounds)

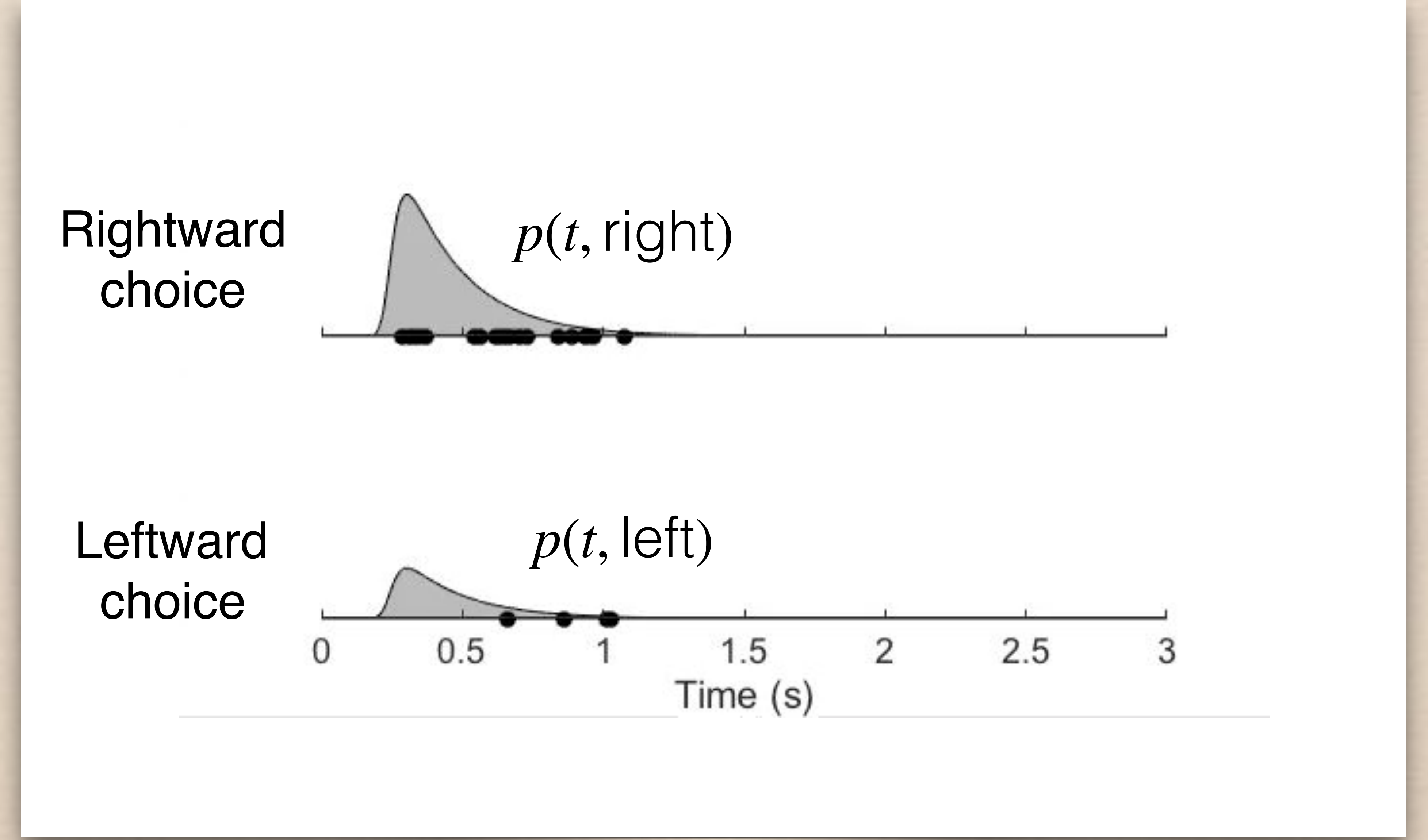
Fokker Planck equation:

$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 p(x,t)}{\partial^2 x}$$



Model fitting

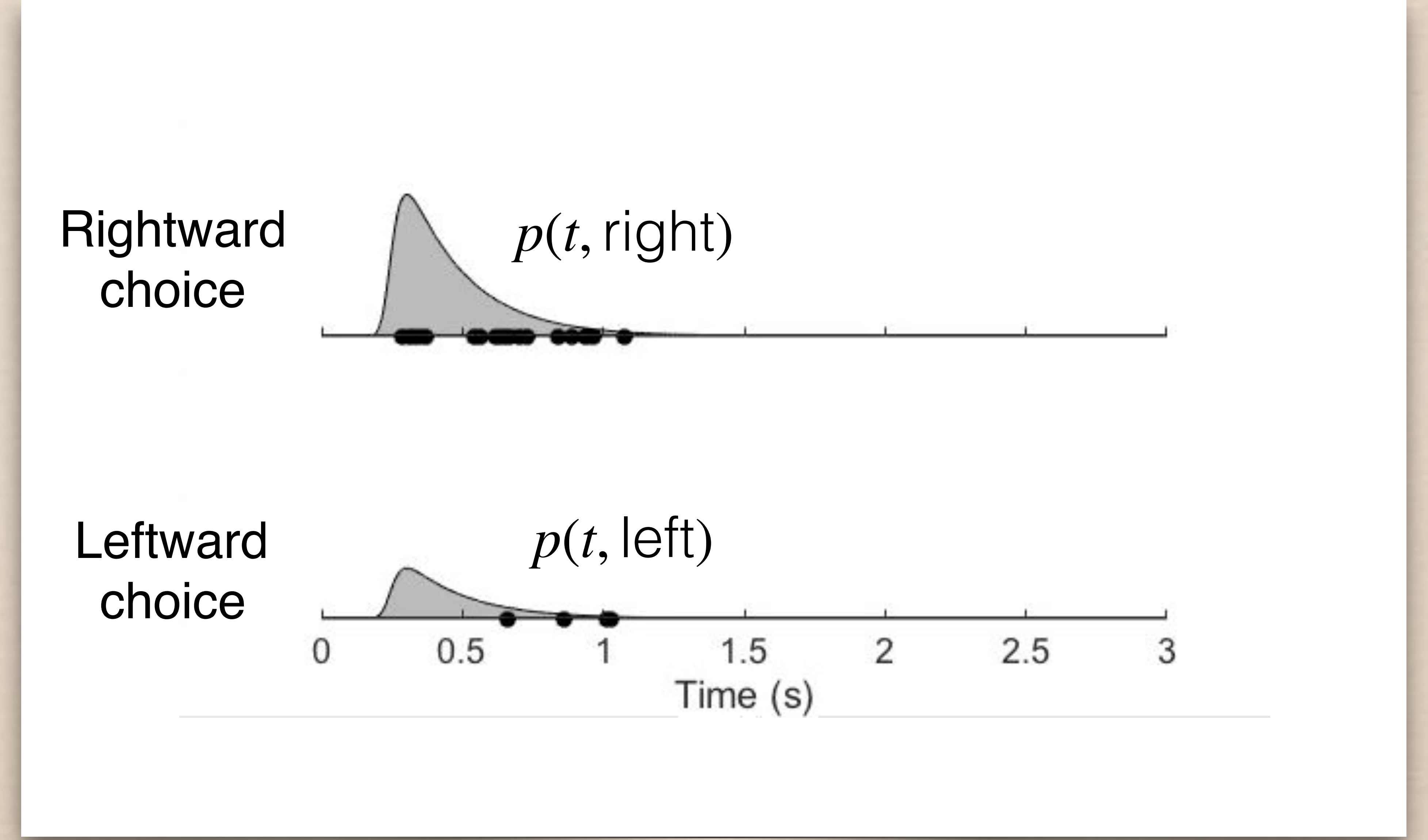
$$\hat{\theta}_{\mathrm{mle}} = \arg\max_{\theta} \left(\sum_{i=1}^{n_{tr}} p(\mathrm{RT}^{(i)}, \mathrm{choice}^{(i)} | c^{(i)}, \theta) \right)$$



Code in Github uses Chang & Cooper's (1970) implicit integration method.

Model fitting

$$\hat{\theta}_{\mathrm{mle}} = \arg\max_{\theta} \left(\sum_{i=1}^{n_{tr}} p(\mathrm{RT}^{(i)}, \mathrm{choice}^{(i)} | c^{(i)}, \theta) \right)$$



Code in Github uses Chang & Cooper’s (1970) implicit integration method.

Parameter optimization

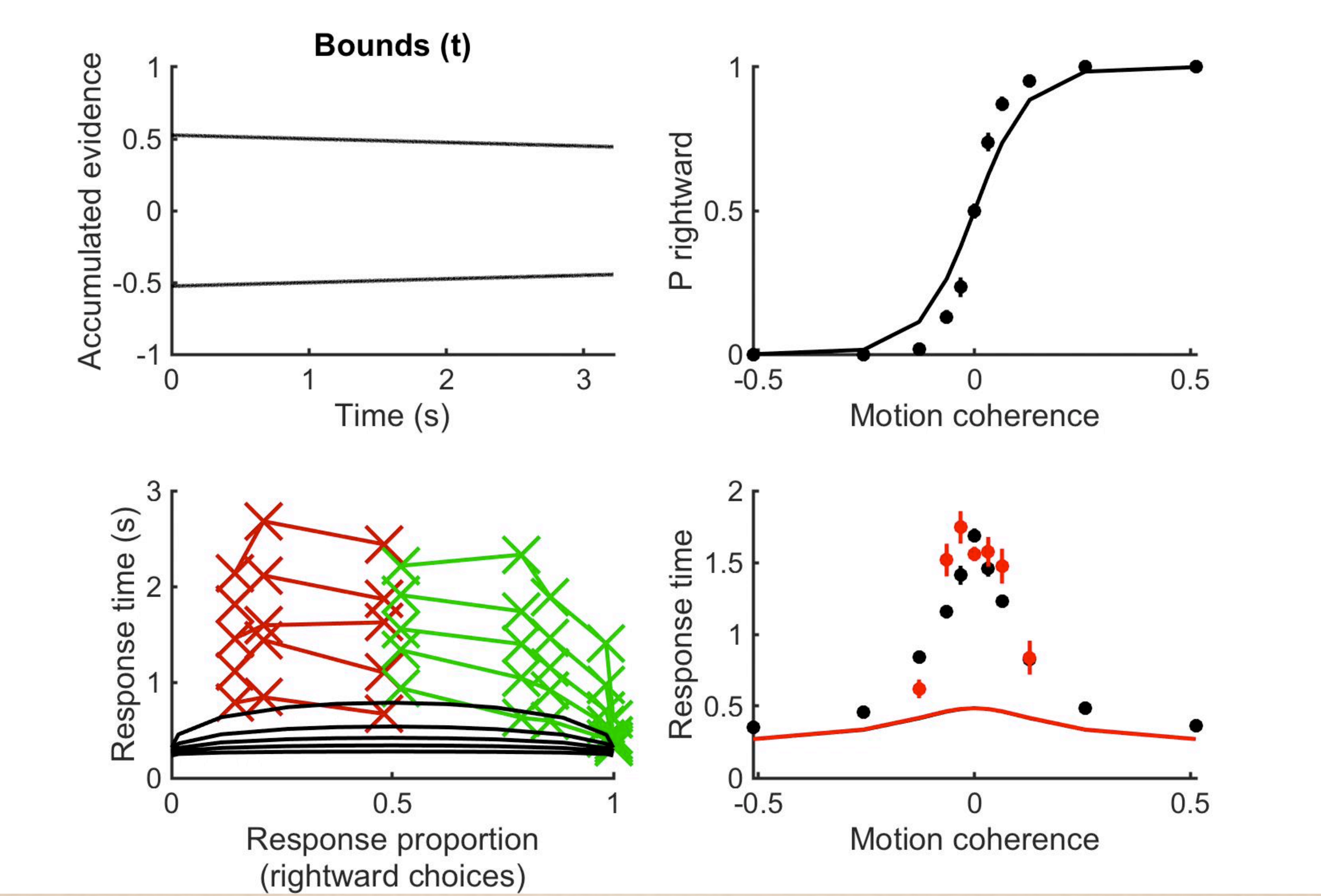
Parameter optimization

- fminsearch (Nelder–Mead simplex method)
- fmincon
- Bayesian Adaptive Direct Search (BADS)
  - Luigi Acerbi (https://github.com/lacerbi/bads)

Parameter optimization

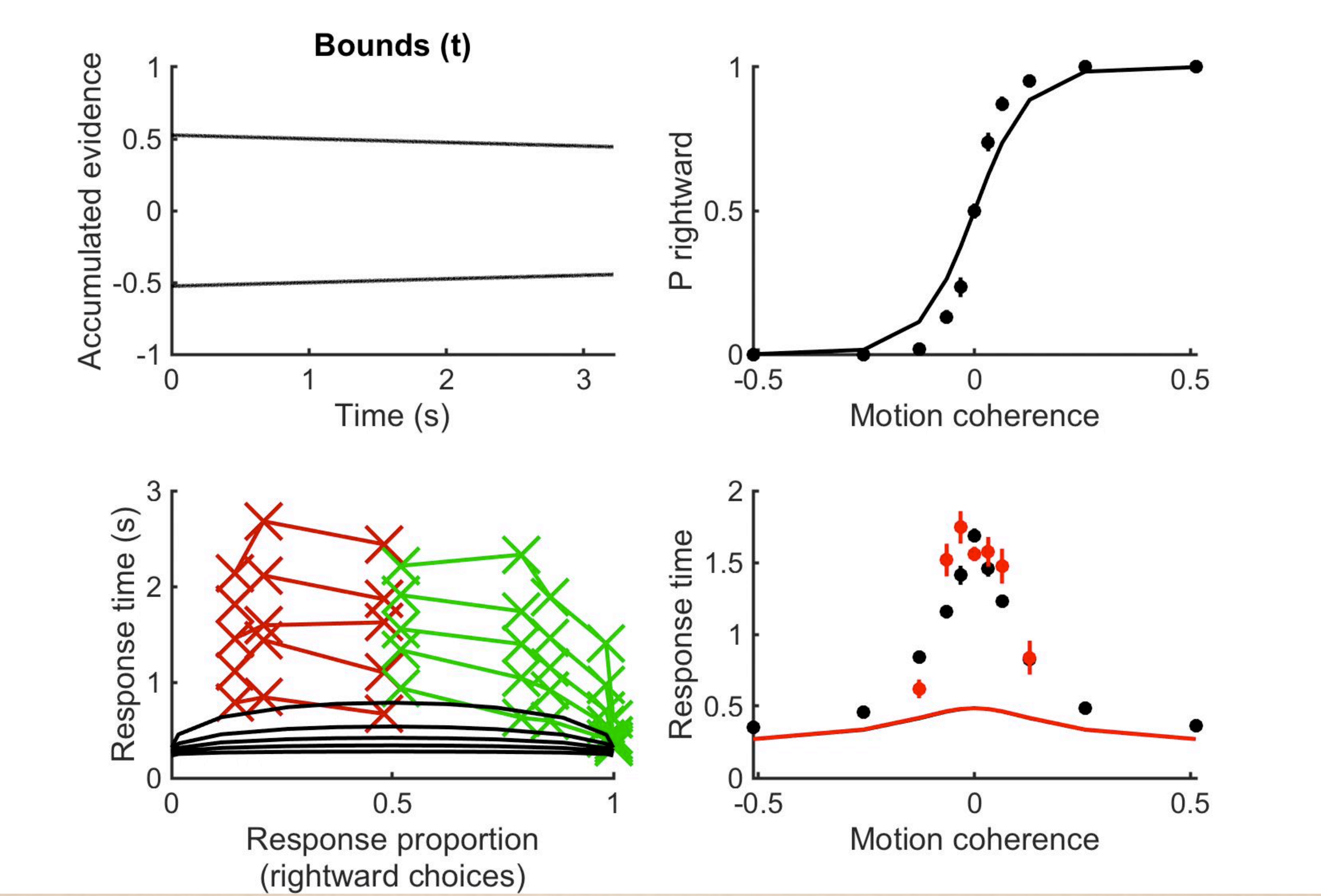
- fminsearch (Nelder–Mead simplex method)
- fmincon
- Bayesian Adaptive Direct Search (BADS)
  - Luigi Acerbi (https://github.com/lacerbi/bads)

https://github.com/arielzylberberg/CompPsychCourse_Zurich2019/fit_single_trials/run_example.m



Six parameters:

signal-to-noise (κ), bound shape (3 params), non-decision time distribution (2 params) https://github.com/arielzylberberg/CompPsychCourse_Zurich2019/fit_single_trials/run_example.m



Six parameters:

signal-to-noise (κ), bound shape (3 params), non-decision time distribution (2 params)

Parameter recovery

Parameter recovery

» Fit model to experimental data

Parameter recovery

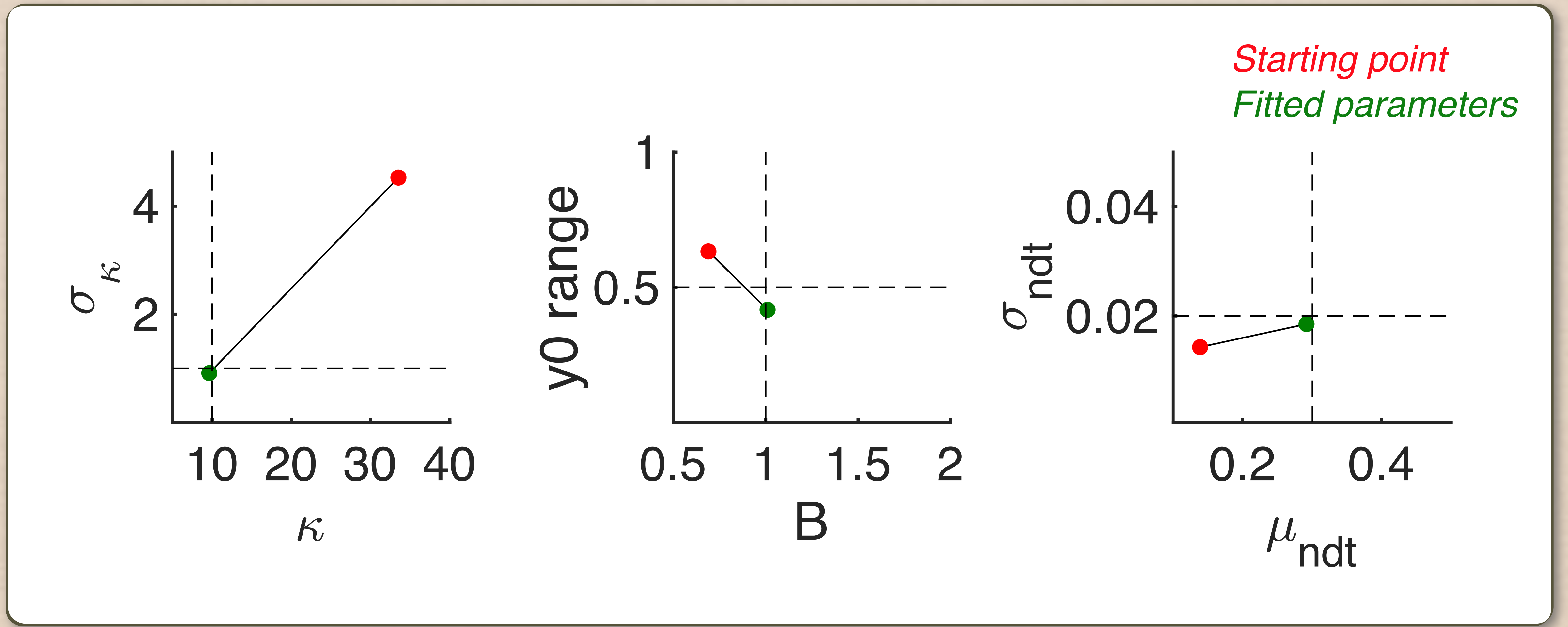
- Fit model to experimental data
- Simulate data with the fitted parameters

Parameter recovery

- Fit model to experimental data
- Simulate data with the fitted parameters
- Fit the simulated data and see if the optimization procedure retrieves the parameters used to generate the data

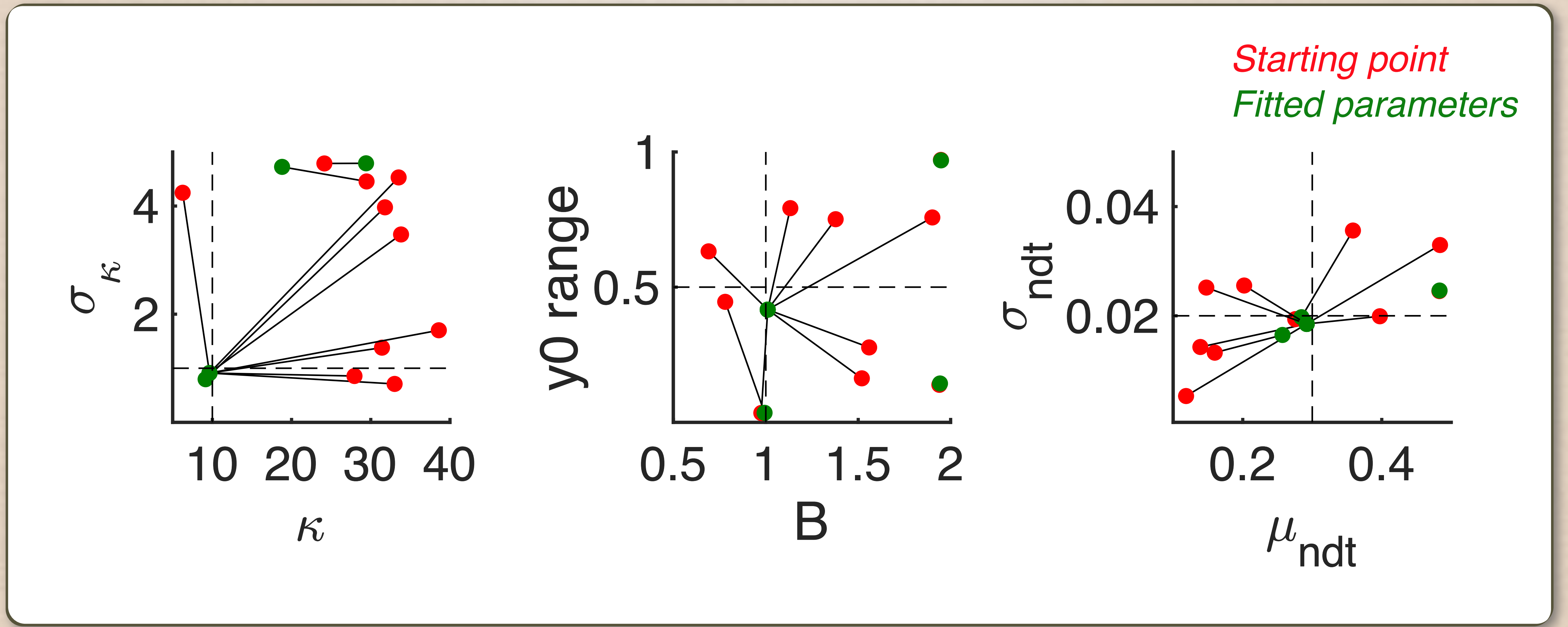
Parameter recovery

- Fit model to experimental data
- Simulate data with the fitted parameters
- Fit the simulated data and see if the optimization procedure retrieves the parameters used to generate the data

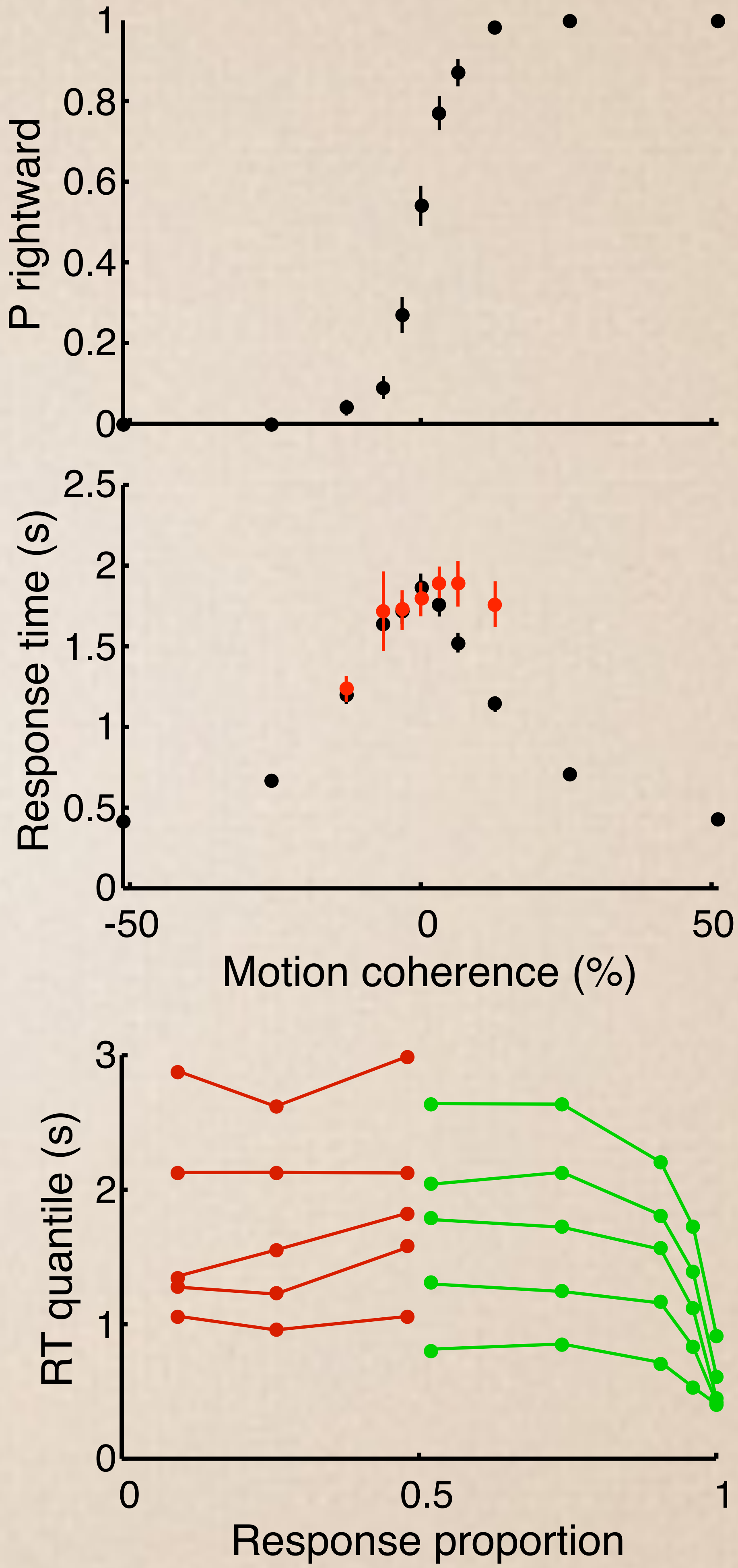


Parameter recovery

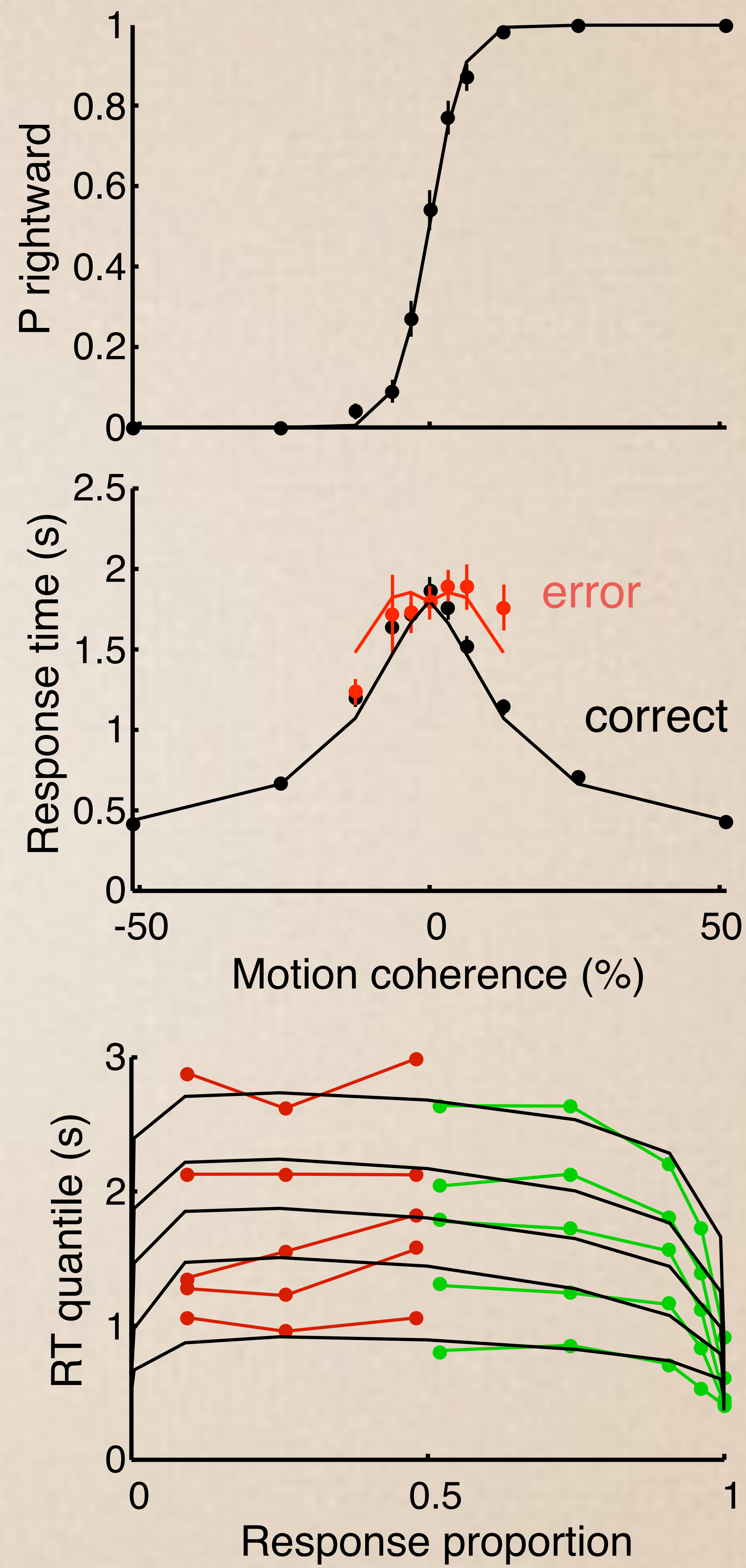
- Fit model to experimental data
- Simulate data with the fitted parameters
- Fit the simulated data and see if the optimization procedure retrieves the parameters used to generate the data



Model mimicry

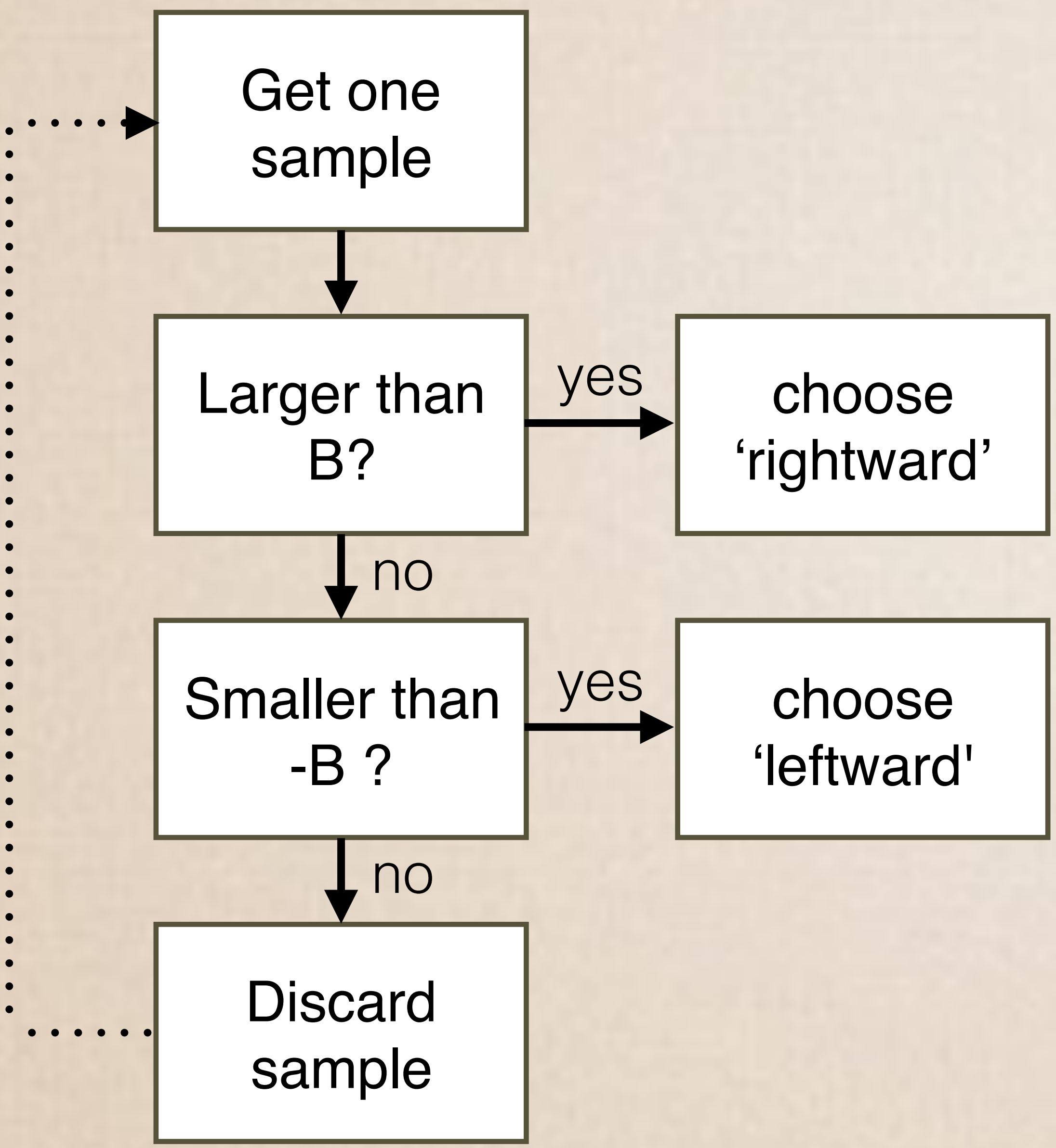


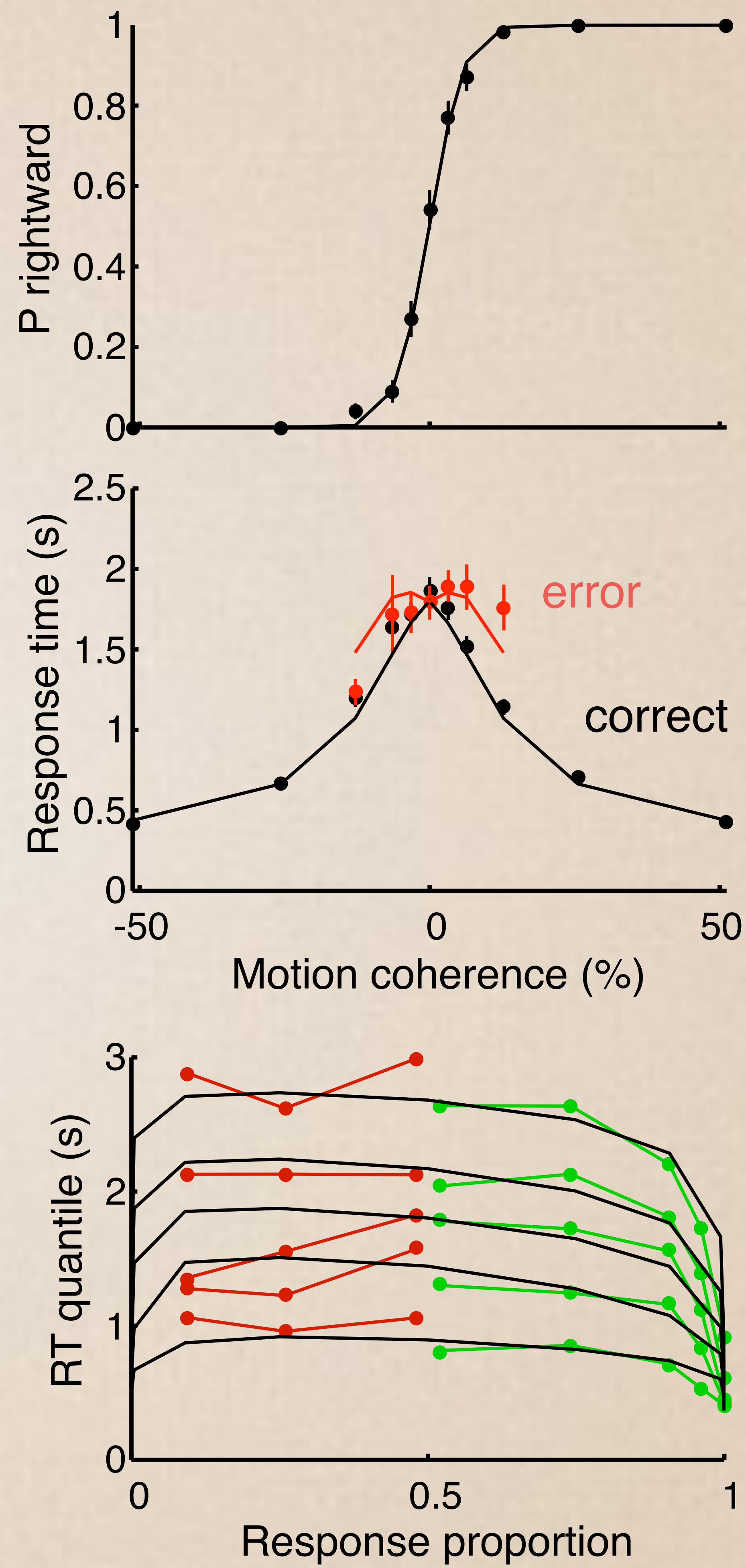
Model mimicry



Model mimicry

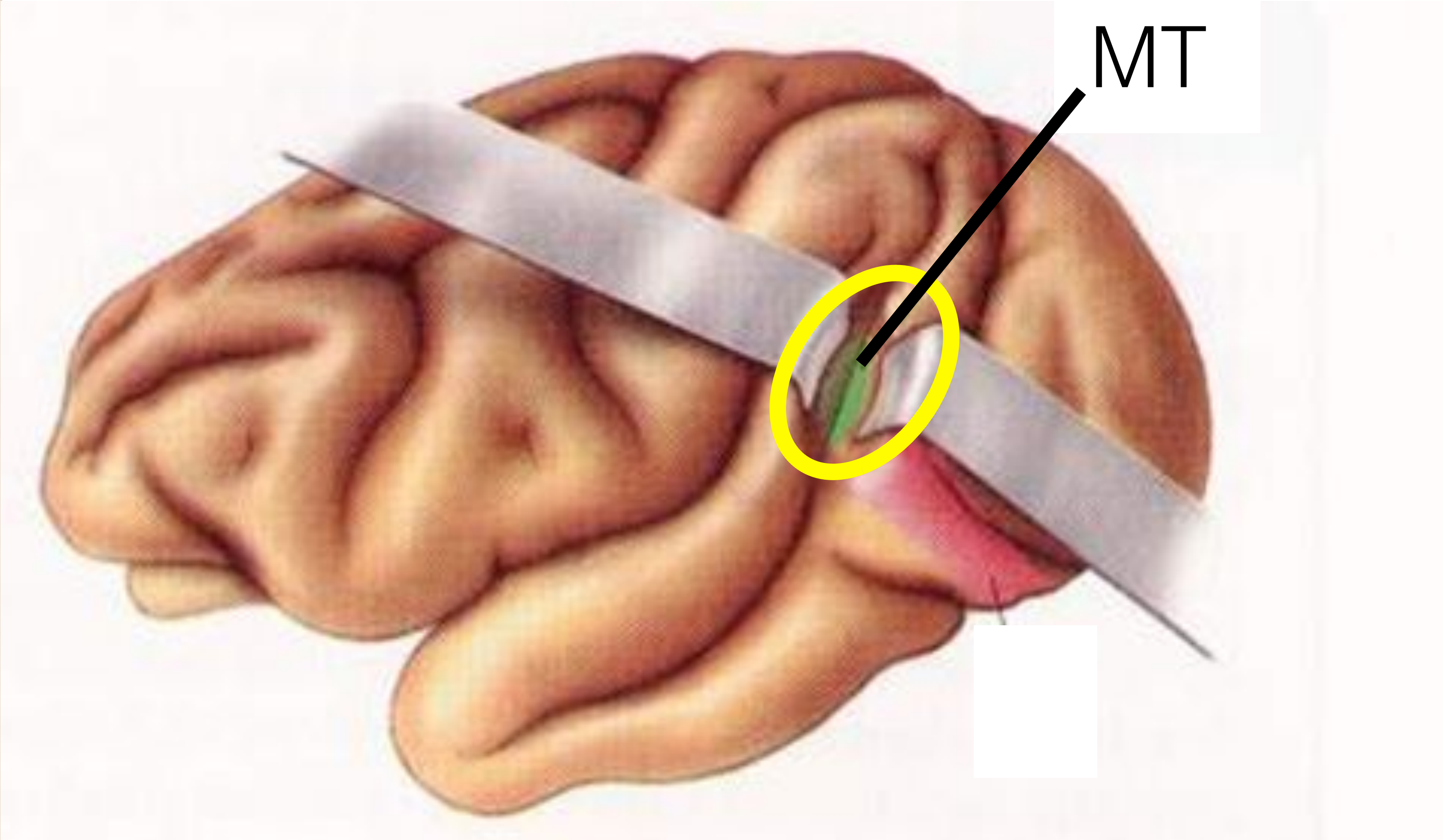
Cartwright and Festinger 1943





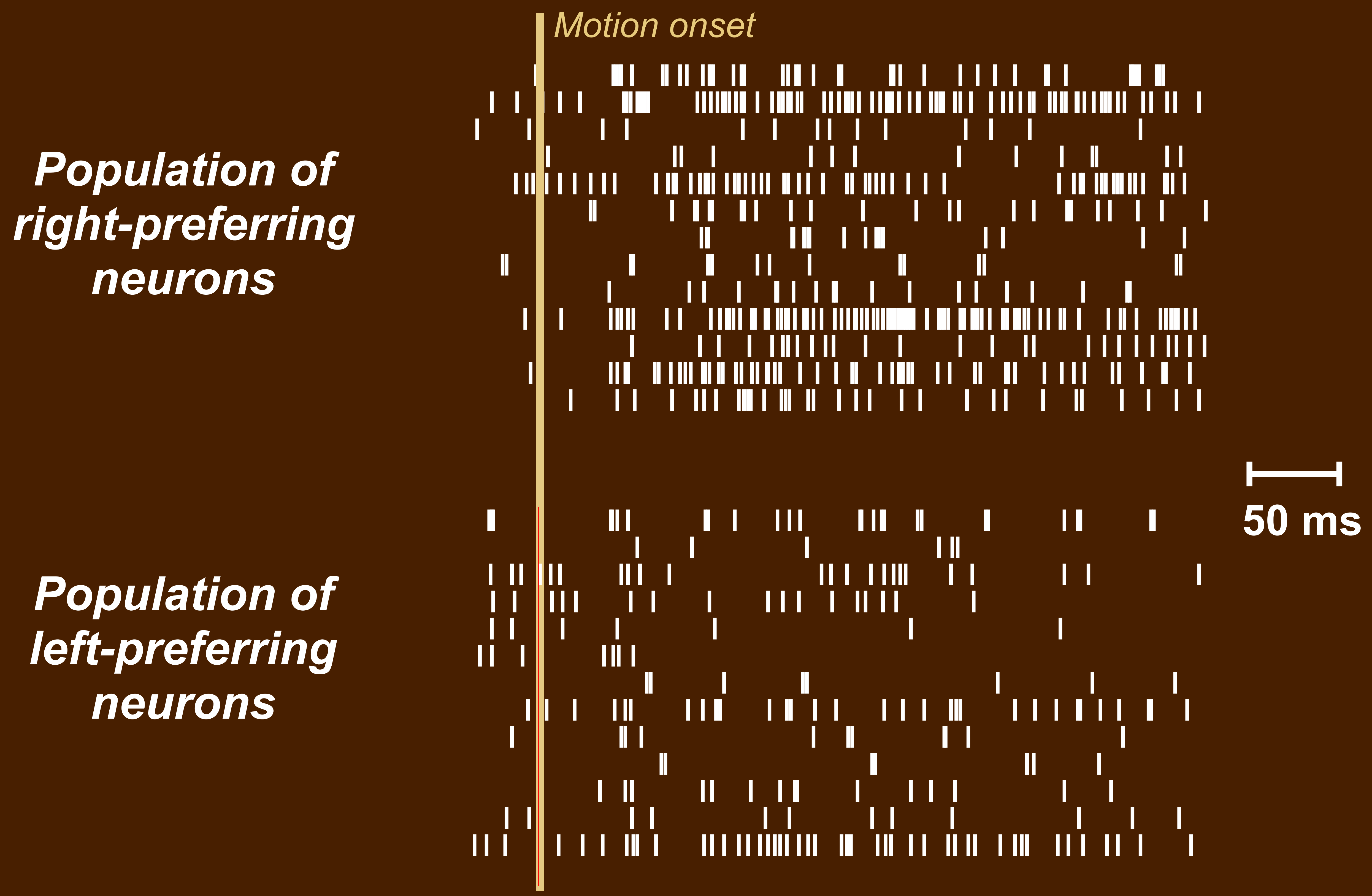
Neurophysiology

Direction-selective neurons in the visual cortex



Albright, Movshon, Newsome, Britten, …

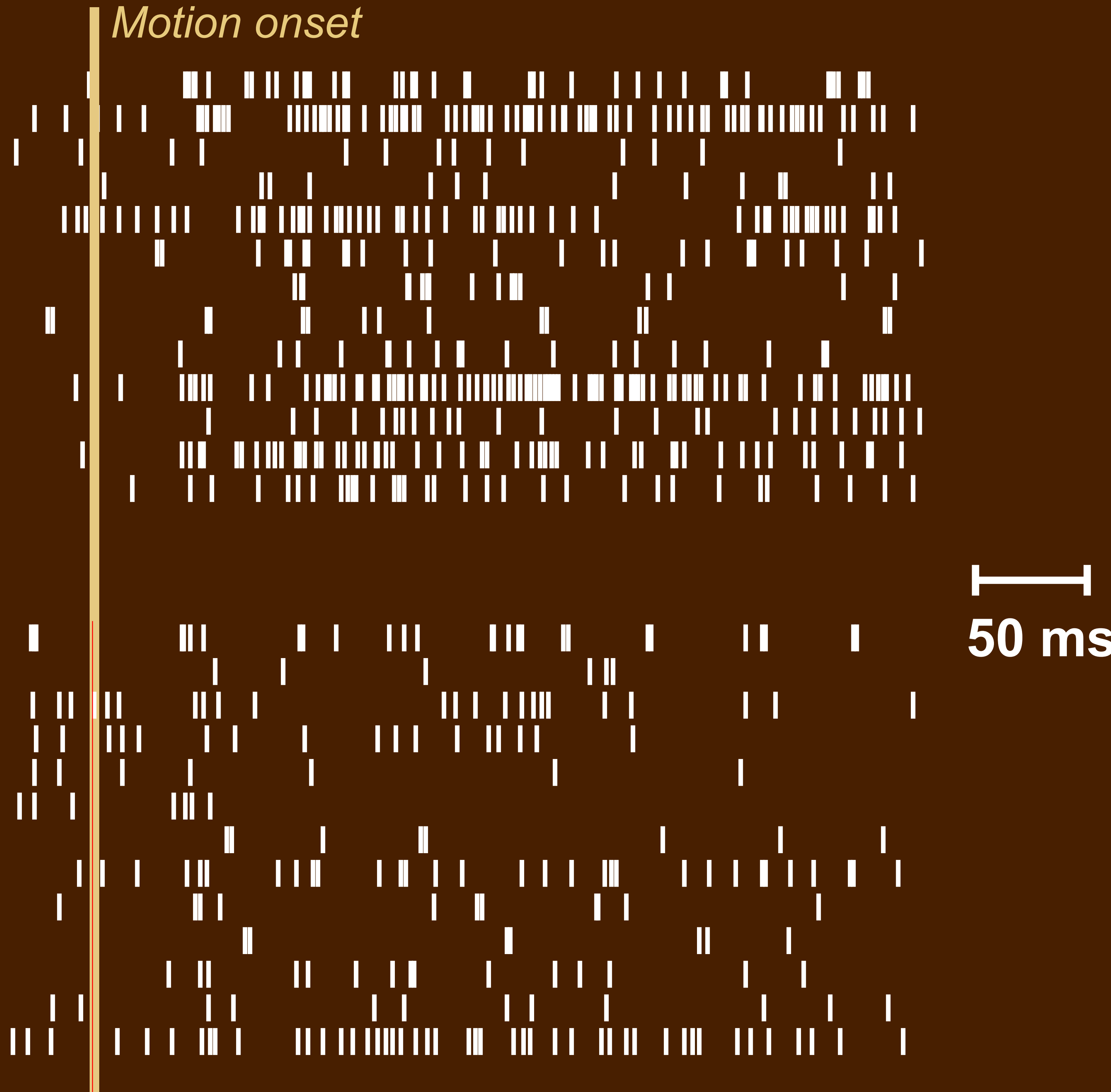
Is the motion rightward or leftward?



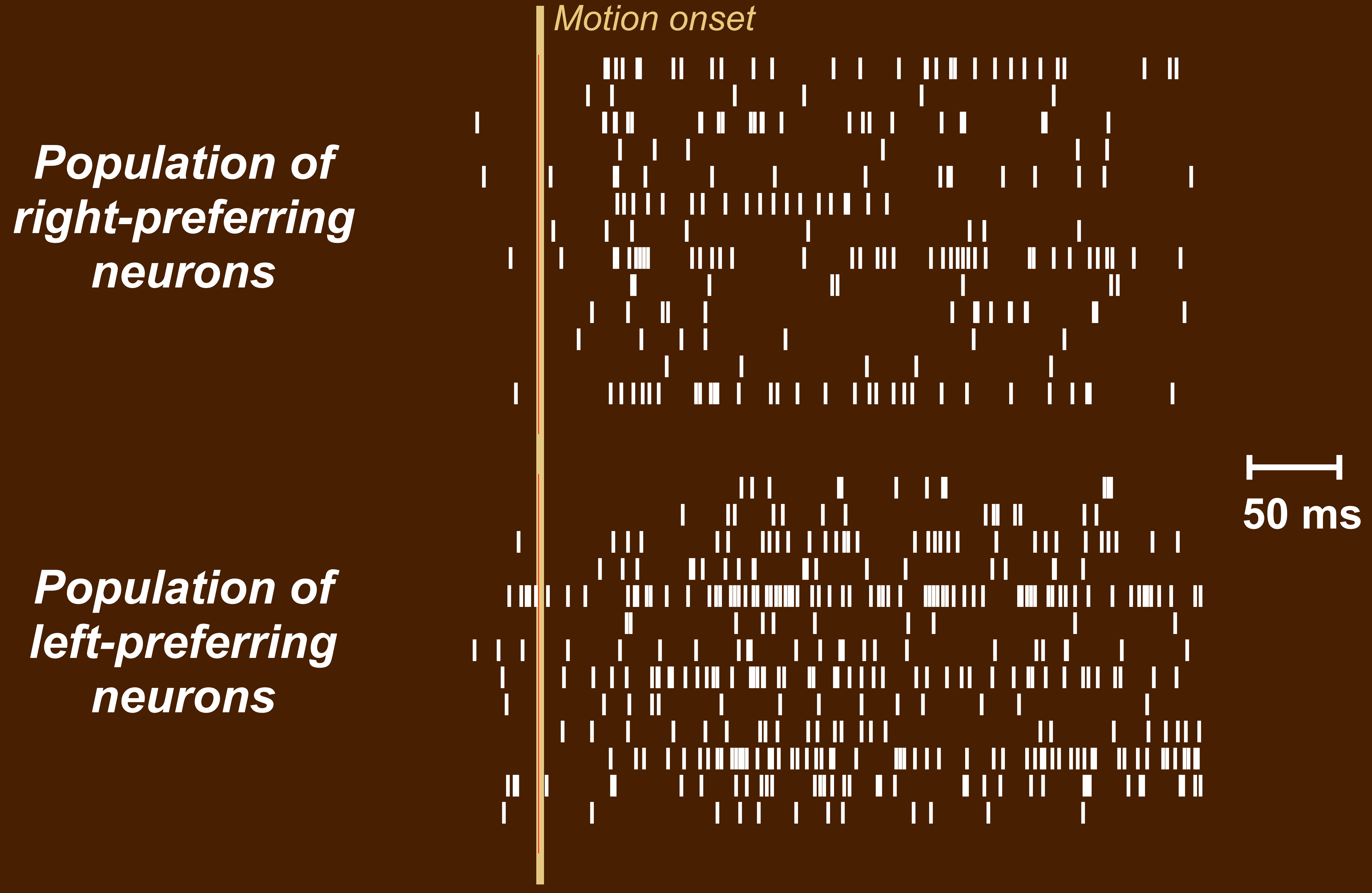
Is the motion rightward or leftward?

Population of right-preferring neurons

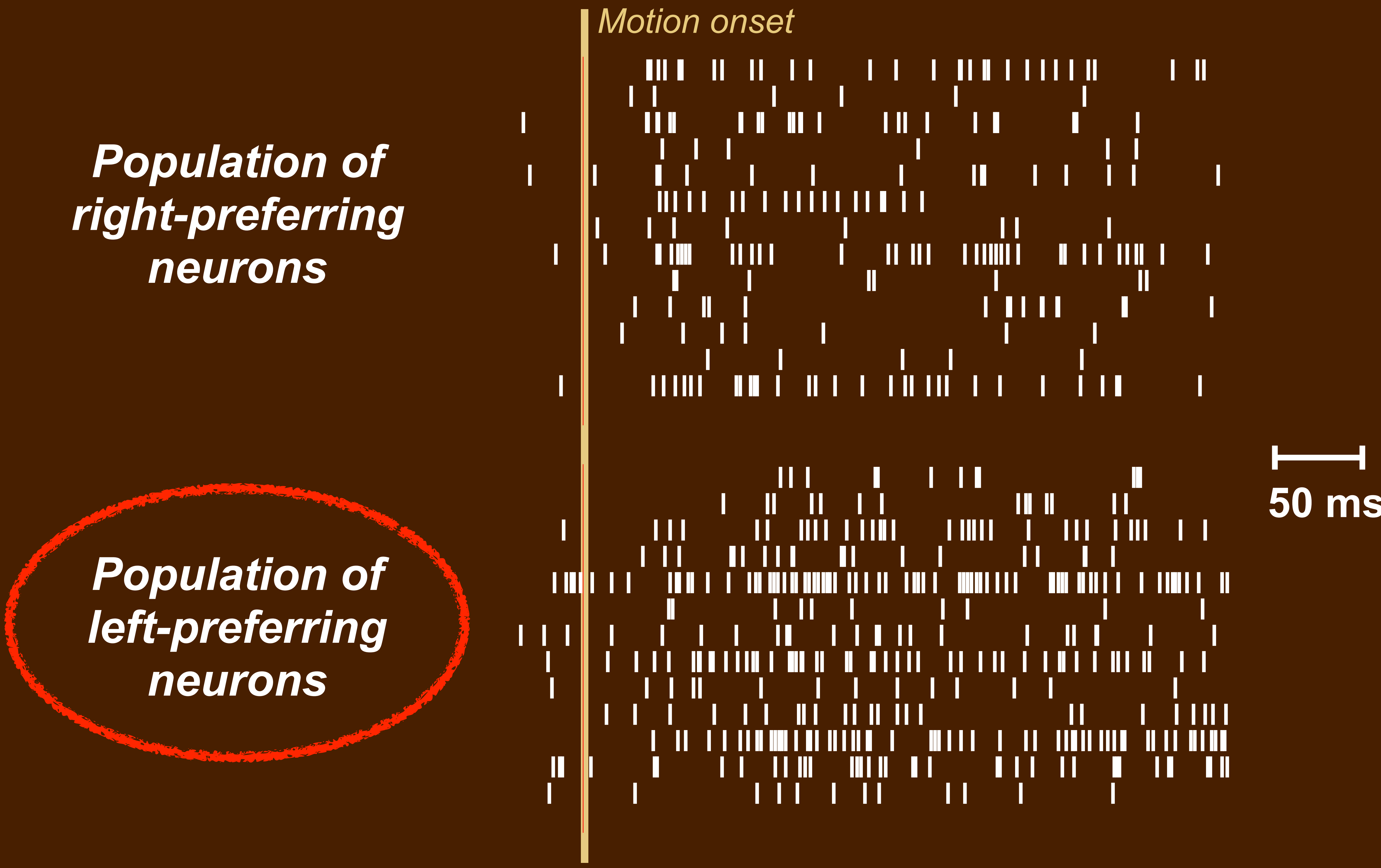
Population of left-preferring neurons



Is the motion rightward or leftward?



Is the motion rightward or leftward?



Decisions based on evidence streams

Population of right-preferring neurons

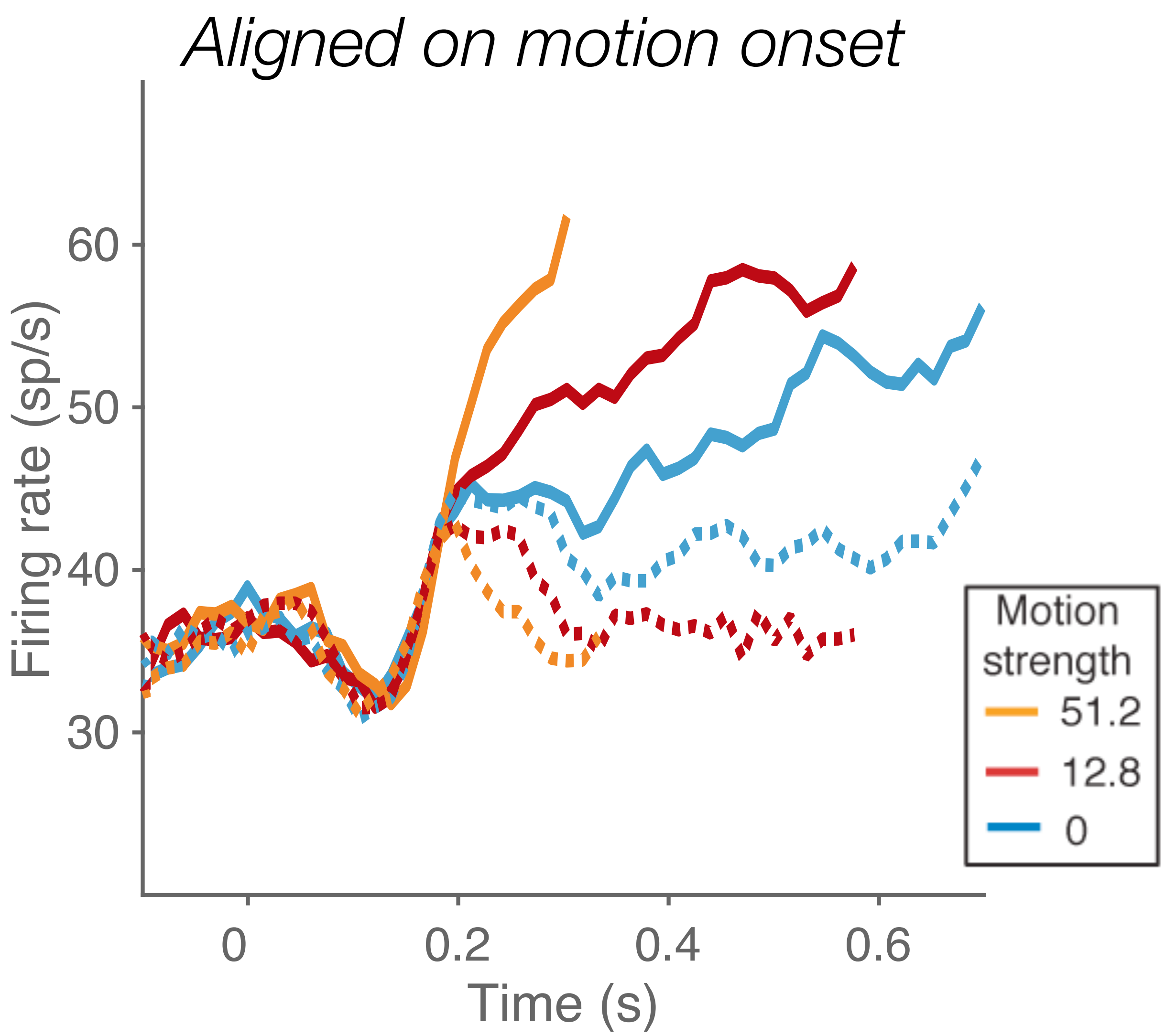
Population of left-preferring neurons

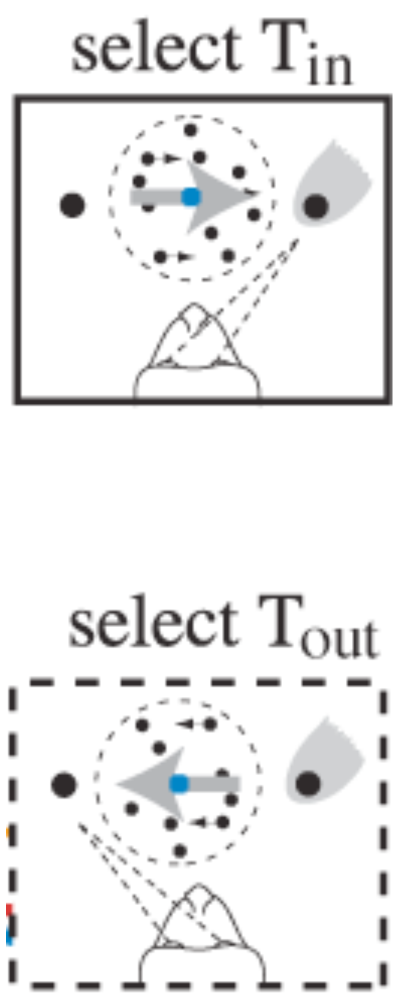
Decisions based on evidence streams

Population of right-preferring neurons

Population of left-preferring neurons

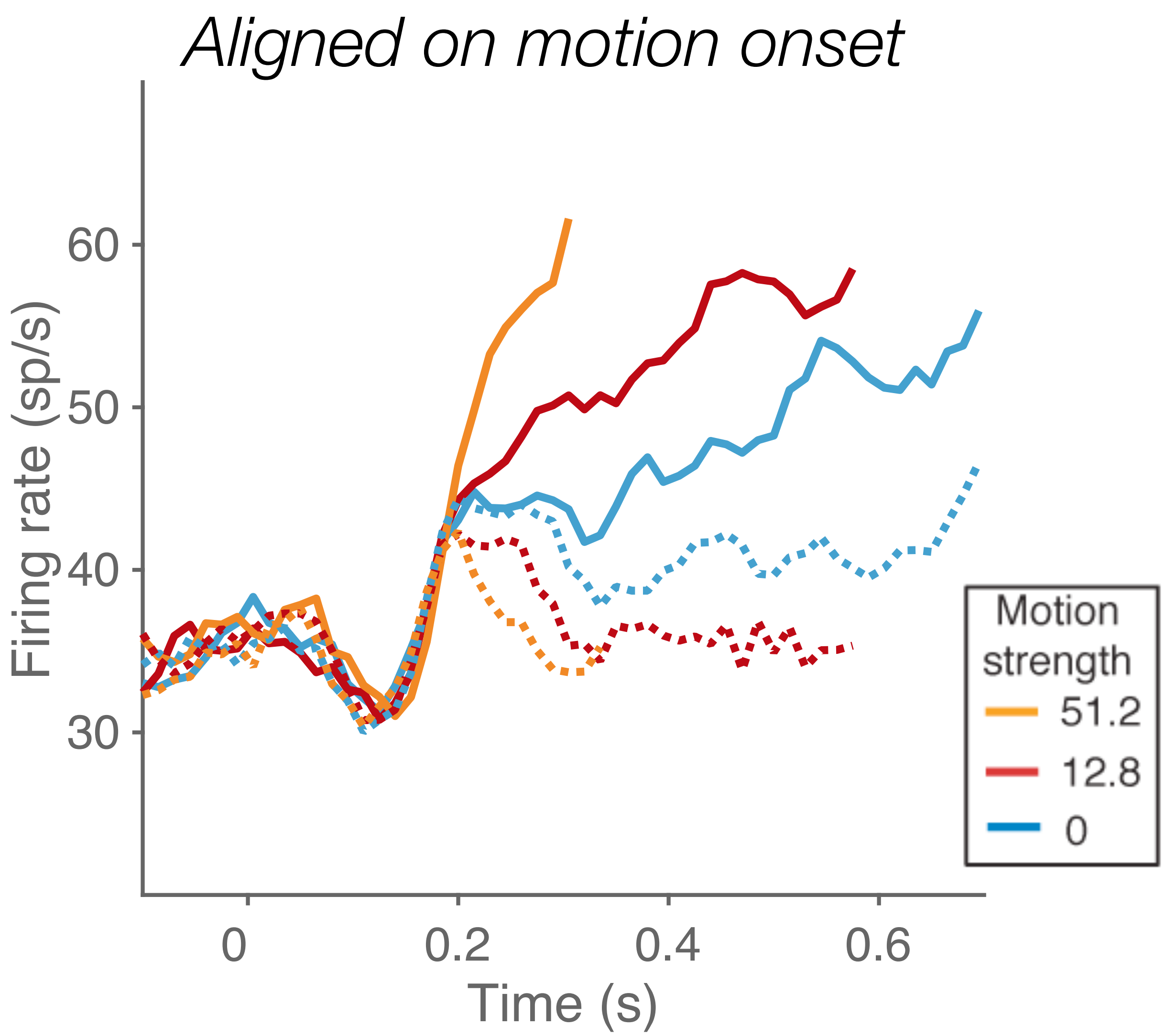
Average LIP activity in the random-dot motion task

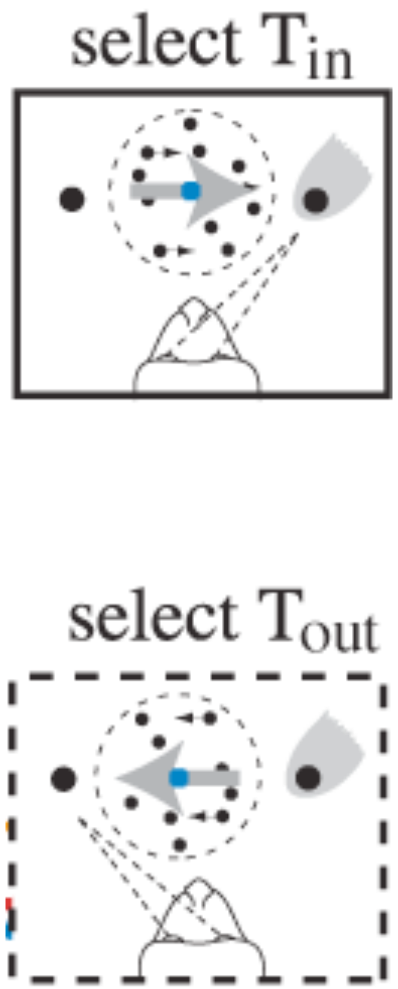




Roitman & Shadlen, 2002 J. Neurosci.

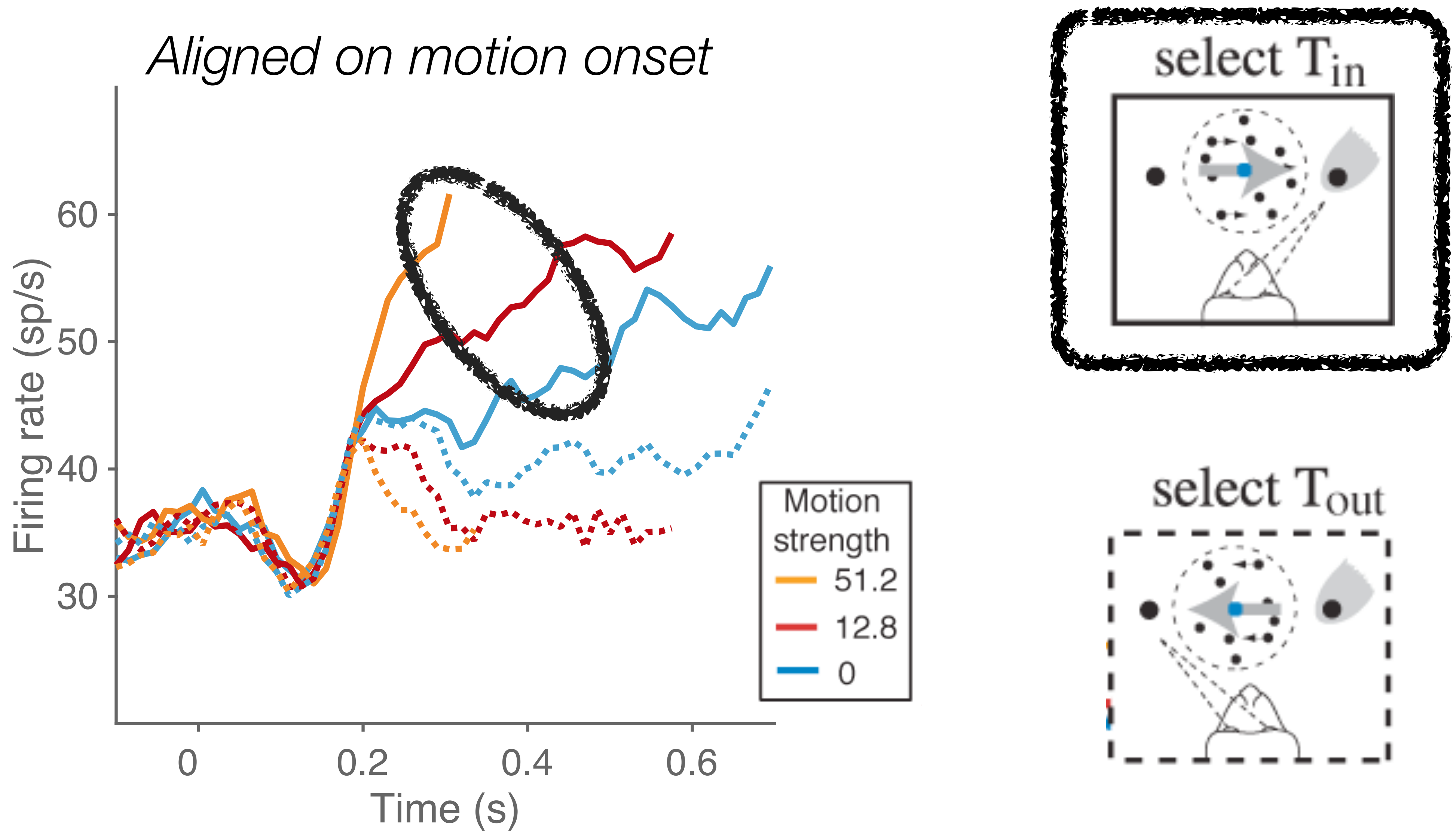
Average LIP activity in the random-dot motion task





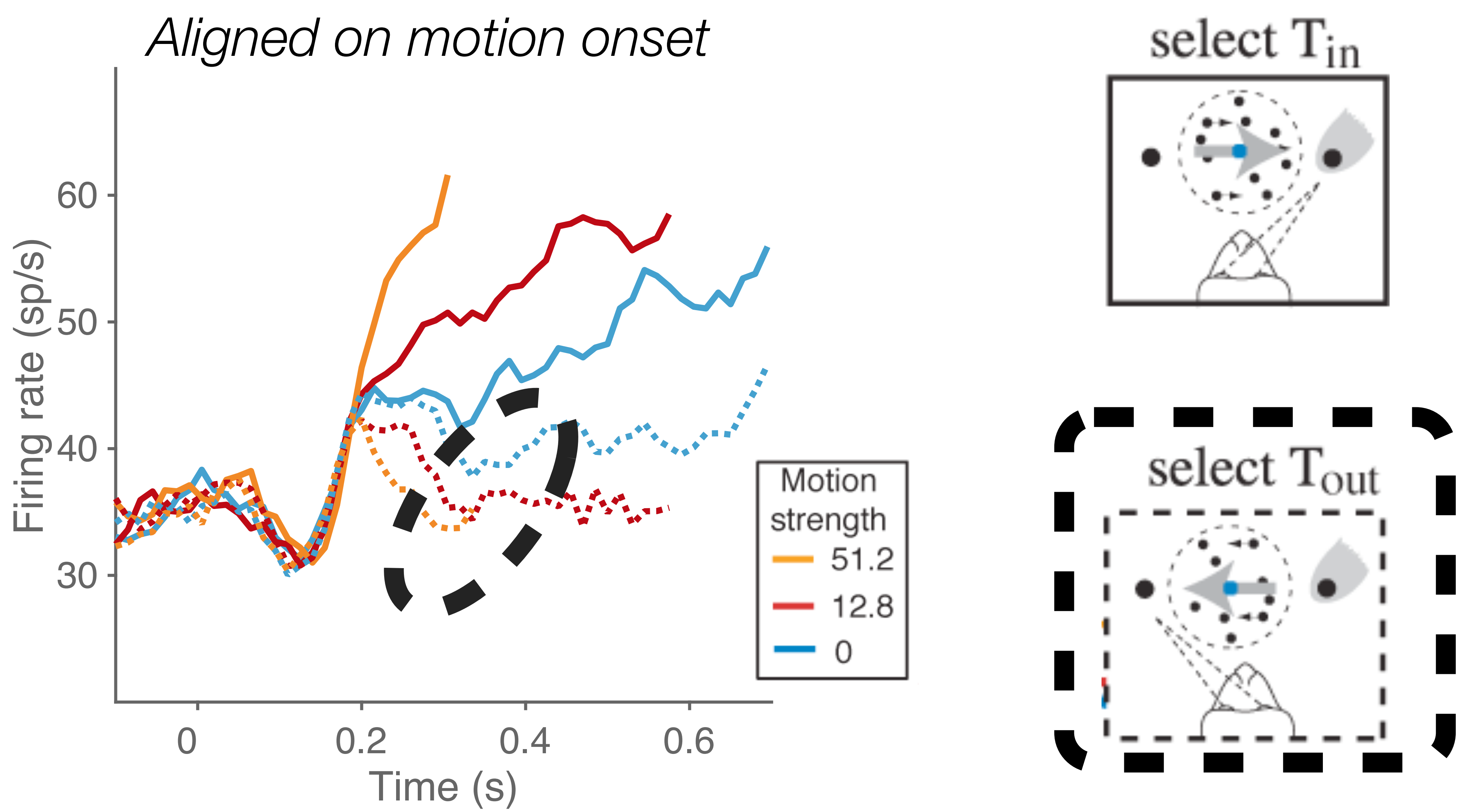
Roitman & Shadlen, 2002 J. Neurosci.

Average LIP activity in the random-dot motion task



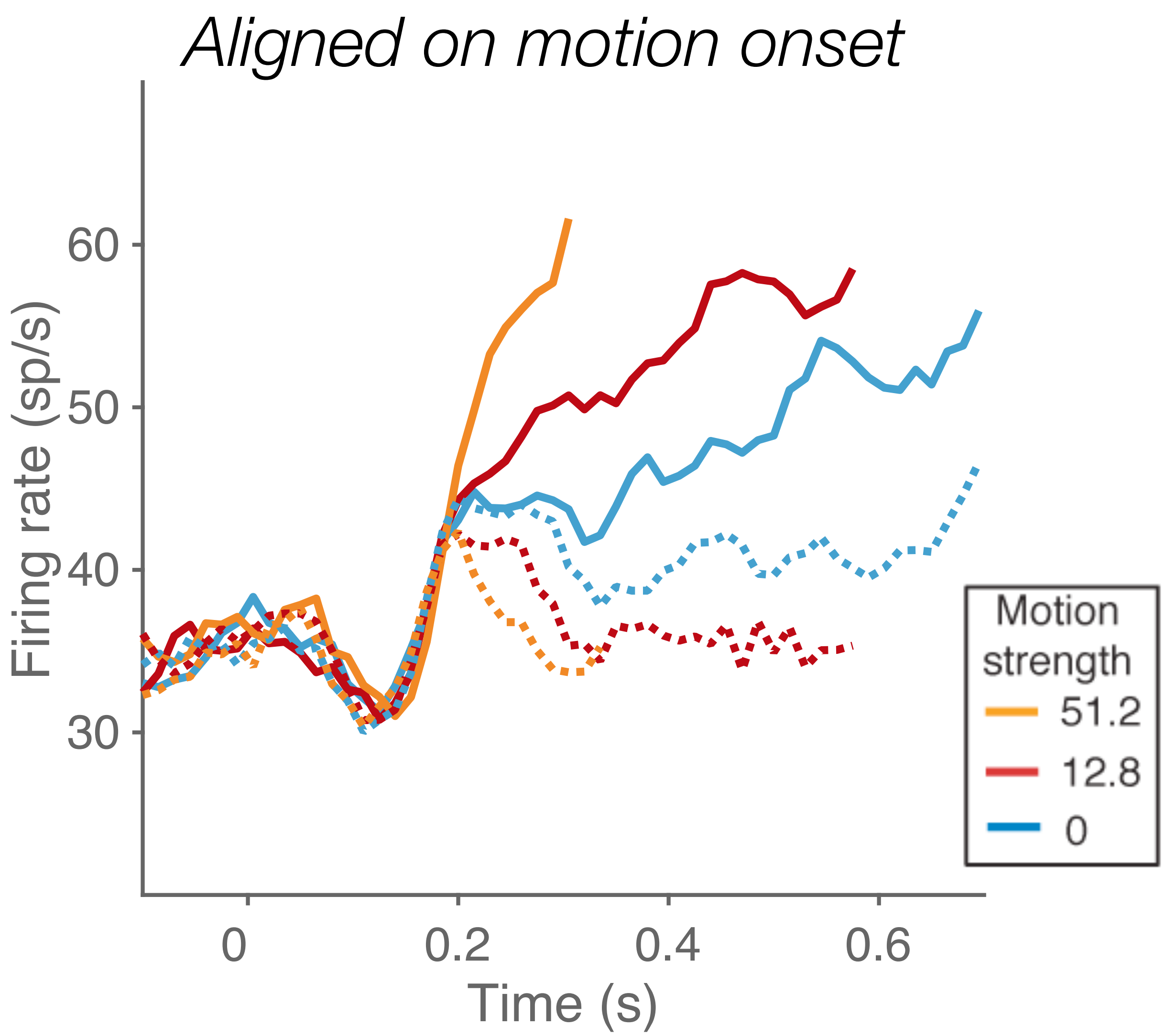
Roitman & Shadlen, 2002 J. Neurosci.

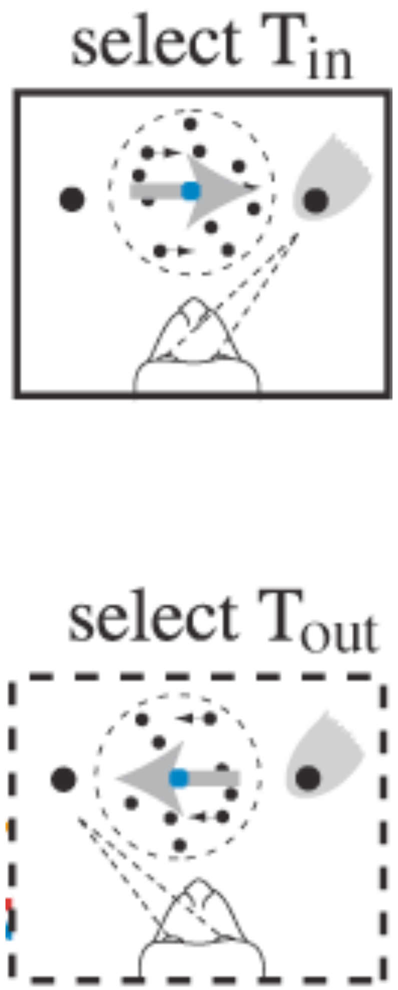
Average LIP activity in the random-dot motion task



Roitman & Shadlen, 2002 J. Neurosci.

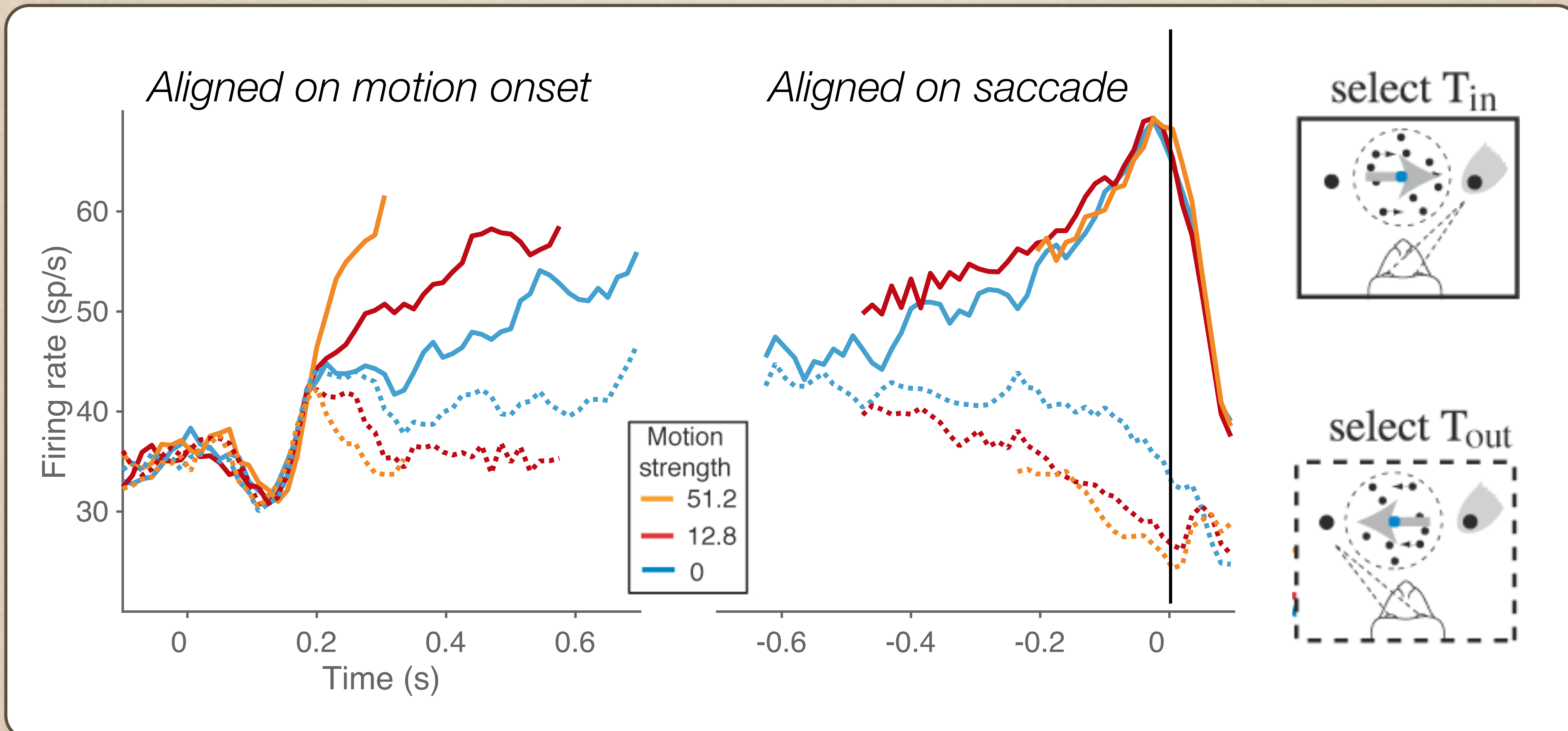
Average LIP activity in the random-dot motion task





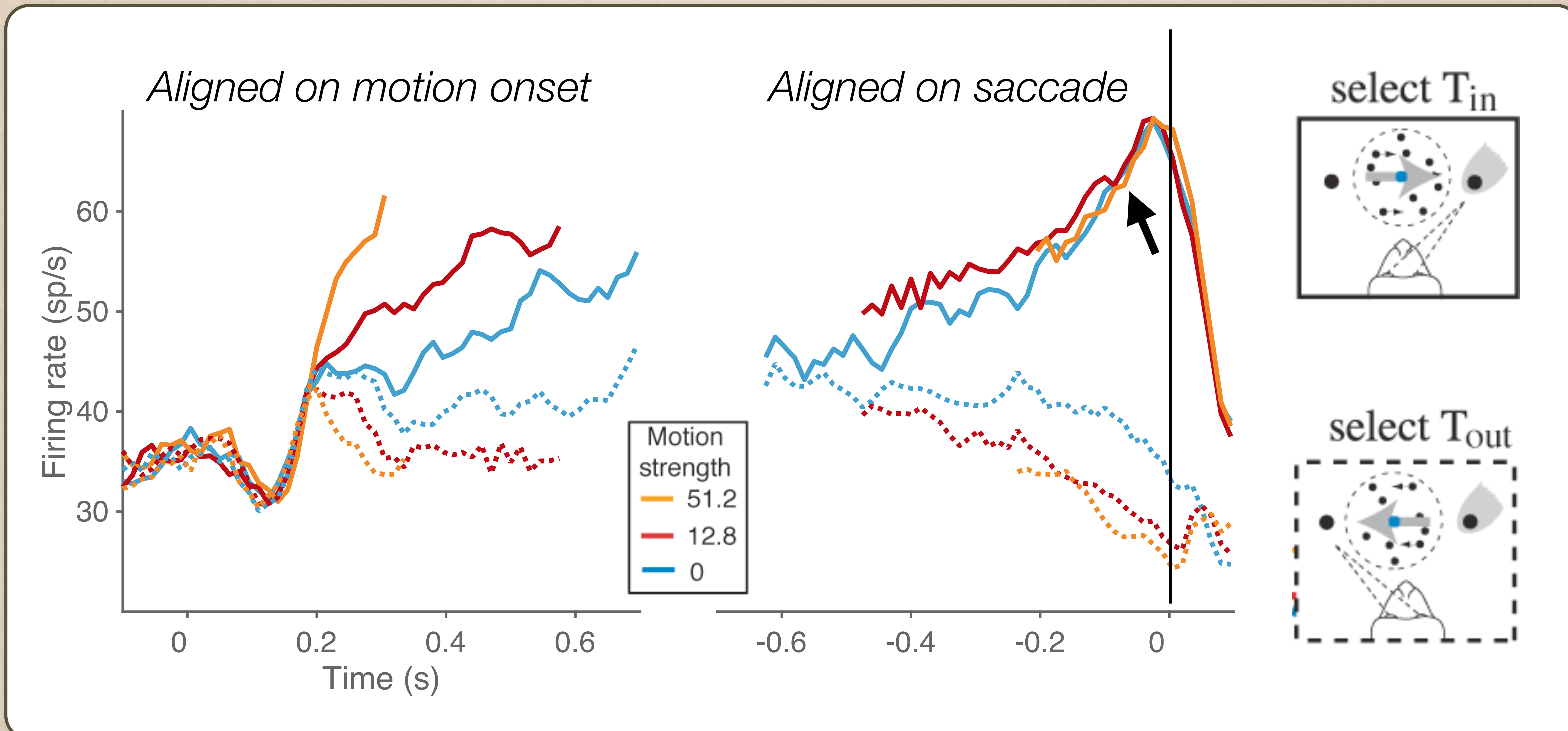
Roitman & Shadlen, 2002 J. Neurosci.

Average LIP activity in the random-dot motion task



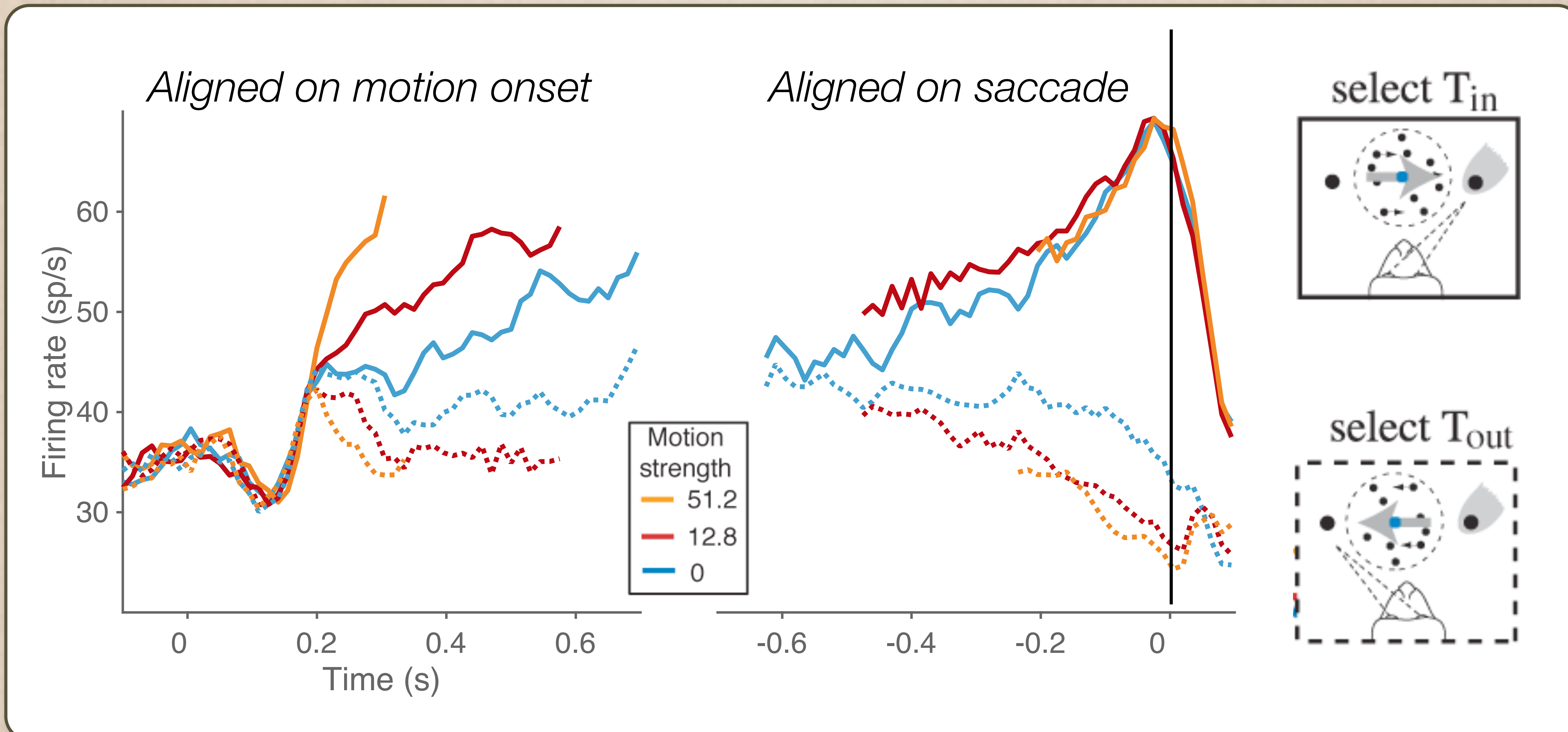
Roitman & Shadlen, 2002 J. Neurosci.

Average LIP activity in the random-dot motion task



Roitman & Shadlen, 2002 J. Neurosci.

Average LIP activity in the random-dot motion task

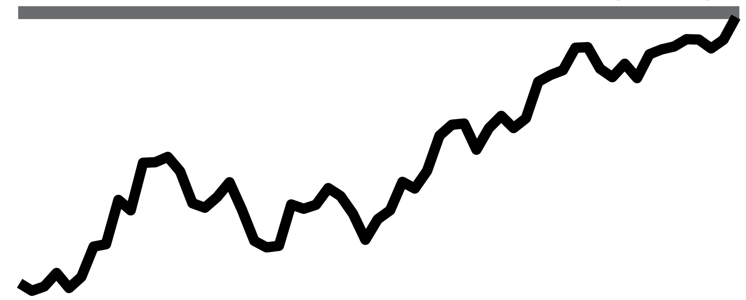


Roitman & Shadlen, 2002 J. Neurosci.

Competing races

$$e_{right} - e_{left} \xrightarrow{\int dt}$$

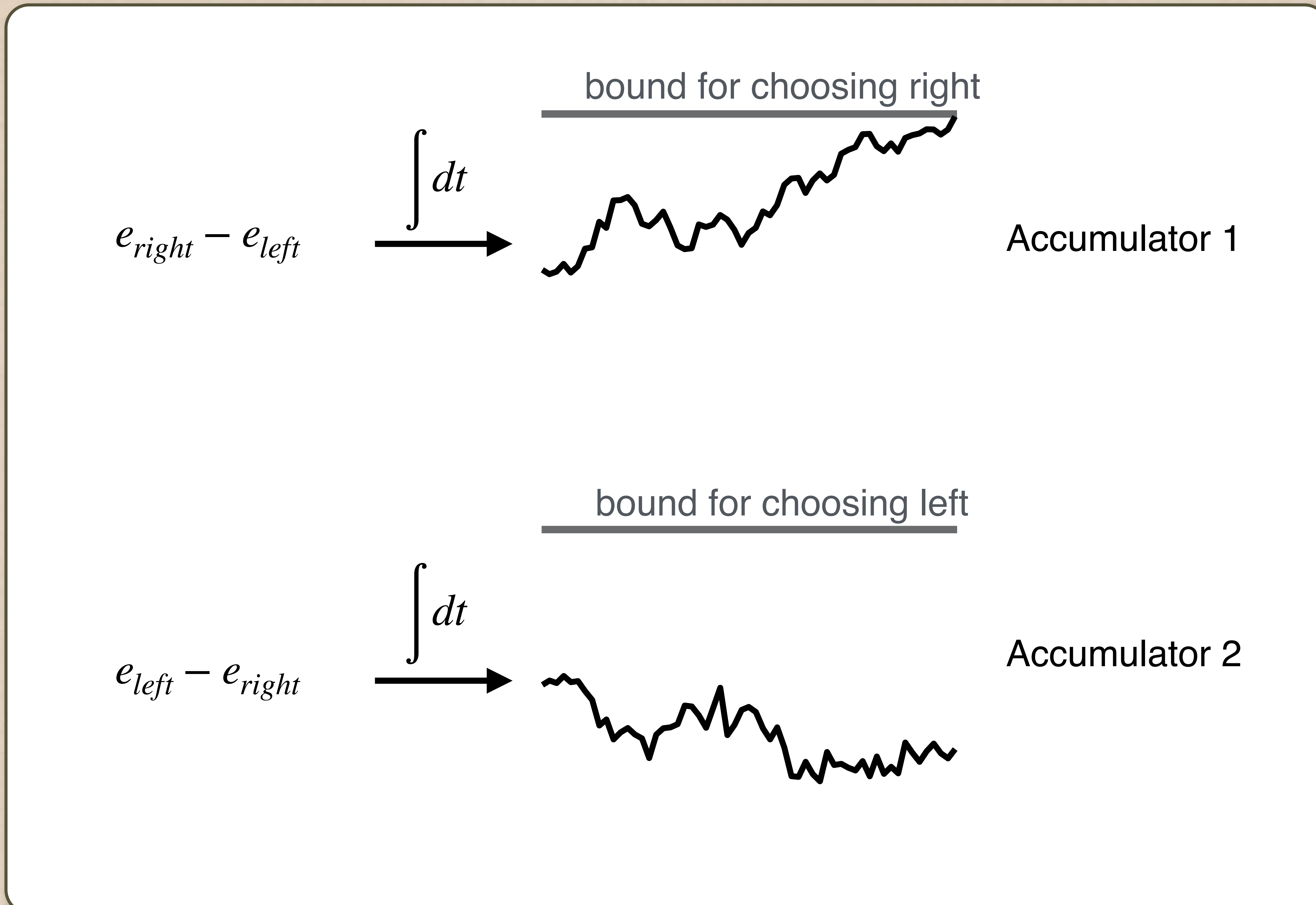
bound for choosing right



Accumulator 1

e.g., Usher and McClelland 2001, Kiani & Shadlen 2014

Competing races



e.g., Usher and McClelland 2001, Kiani & Shadlen 2014

Final summary

Final summary

Background

Final summary

Background

Early sequential sampling model

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Influence of the different parameters

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Influence of the different parameters

The quantile-probability plot

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Influence of the different parameters

The quantile-probability plot

Differences in RT between error and correct responses

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Influence of the different parameters

The quantile-probability plot

Differences in RT between error and correct responses

Model fitting

Final summary

Background

Early sequential sampling model

Sequential probability ratio test

The drift-diffusion model

Influence of the different parameters

The quantile-probability plot

Differences in RT between error and correct responses

Model fitting

Probability distributions over choice and RT

Final summary

Background

- Early sequential sampling model
- Sequential probability ratio test

The drift-diffusion model

- Influence of the different parameters
- The quantile-probability plot
- Differences in RT between error and correct responses

Model fitting

- Probability distributions over choice and RT
- Parameter optimization & recovery

Final summary

Background

- Early sequential sampling model
- Sequential probability ratio test

The drift-diffusion model

- Influence of the different parameters
- The quantile-probability plot
- Differences in RT between error and correct responses

Model fitting

- Probability distributions over choice and RT
- Parameter optimization & recovery
- Model mimicry

Final summary

Background

- Early sequential sampling model
- Sequential probability ratio test

The drift-diffusion model

- Influence of the different parameters
- The quantile-probability plot
- Differences in RT between error and correct responses

Model fitting

- Probability distributions over choice and RT
- Parameter optimization & recovery
- Model mimicry

Neurophysiology

Final summary

Background

- Early sequential sampling model
- Sequential probability ratio test

The drift-diffusion model

- Influence of the different parameters
- The quantile-probability plot
- Differences in RT between error and correct responses

Model fitting

- Probability distributions over choice and RT
- Parameter optimization & recovery
- Model mimicry

Neurophysiology

- Neural responses from areas MT and LIP

Thank you!

Suggested reviews:

Gold and Shadlen, Neuron 2002,
"Banburismus and the Brain: Decoding Review the Relationship between Sensory Stimuli, Decisions, and Reward “

Ratcliff, Smith, Brown and McKoon, TiCS 2016,
“Diffusion Decision Model: Current Issues and History”

extra slides

Attempt to disambiguate between these alternatives using model comparison

Hawkings et al. (JNeurosci. 2015) “Revisiting the Evidence for Collapsing Boundaries and Urgency Signals in Perceptual Decision-Making”.