Integration and Perceptual Decision Making

Decision Making
(writ “small”)

Decision Making
(writ "small")

- Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:
  - perception and psychophysics *(e.g., Sperling task)*
  - attention *(e.g., Stroop task, Posner paradigm)*
  - short term memory *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - long term memory *(e.g., priming, familarity effects, etc.)*

Decision Making
(writ "small")

- Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:
  - perception and psychophysics *(e.g., Sperling task)*
  - attention *(e.g., Stroop task, Posner paradigm)*
  - short term memory *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - long term memory *(e.g., priming, familarity effects, etc.)*
- These often *(though not always)* involve a decision about how to respond between *two alternatives:*
  - old or new?

Decision Making
(writ "small")

- **Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:**
  - **perception and psychophysics** *(e.g., Sperling task)*
  - **attention** *(e.g., Stroop task, Posner paradigm)*
  - **short term memory** *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - **long term memory** *(e.g., priming, familarity effects, etc.)*
- **These often** *(though not always)* **involve a decision about how to respond between *two alternatives:***
  - **old or new?**
  - **yes or no?**

Decision Making
(writ "small")

- **Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:**
  - **perception and psychophysics** *(e.g., Sperling task)*
  - **attention** *(e.g., Stroop task, Posner paradigm)*
  - **short term memory** *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - **long term memory** *(e.g., priming, familarity effects, etc.)*
- **These often** *(though not always)* **involve a decision about how to respond between *two alternatives:***
  - **old or new?**
  - **yes or no?**
  - **red or green?**

Decision Making
(writ "small")

- Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:
  - perception and psychophysics *(e.g., Sperling task)*
  - attention *(e.g., Stroop task, Posner paradigm)*
  - short term memory *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - long term memory *(e.g., priming, familarity effects, etc.)*
- These often *(though not always)* involve a decision about how to respond between *two alternatives:*
  - old or new?
  - yes or no?
  - red or green?
  - ⬅ or ➡?

Decision Making
(writ "small")

- **Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:**
  - **perception and psychophysics** *(e.g., Sperling task)*
  - **attention** *(e.g., Stroop task, Posner paradigm)*
  - **short term memory** *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - **long term memory** *(e.g., priming, familarity effects, etc.)*
- **These often** *(though not always)* **involve a decision about how to respond between *two alternatives:***
  - **old or new?**
  - **yes or no?**
  - **red or green?**
  - **⬅ or ➡?**
- **What are the *dynamics* that underlie these decisions…**

Decision Making
(writ "small")

- Empirical data in cognitive psychology rely on measures of *speed and accuracy* of response:
  - perception and psychophysics *(e.g., Sperling task)*
  - attention *(e.g., Stroop task, Posner paradigm)*
  - short term memory *(e.g., Sternberg paradigm, Brown-Petersen paradigm)*
  - long term memory *(e.g., priming, familarity effects, etc.)*
- These often *(though not always)* involve a decision about how to respond between *two alternatives:*
  - old or new?
  - yes or no?
  - red or green?
  - ⬅ or ➡?
- What are the *dynamics* that underlie these decisions… and how are the *controlled*?

Outline

- **Simple behavioral task:**
  - **Two alternative forced choice (2AFC) decision task**

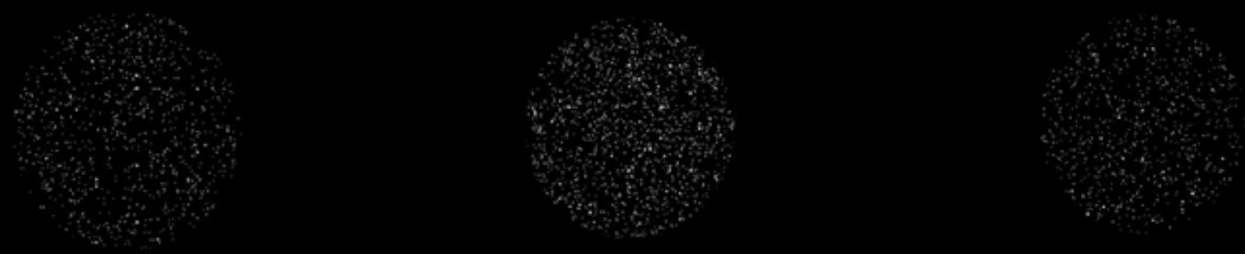
- **Current "state of play:"**
  - **Behavioral findings**
  - *(Some informative)* **neurobiological findings**
  - **PDP models**

- **Formal analysis:**
  - **Drift diffusion model (DDM)**

  Sets the stage for:
  - **Control as optimization**

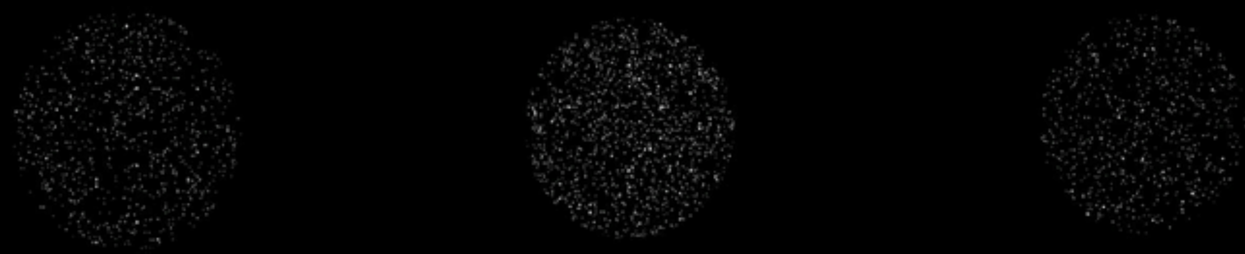
Two Alternative Forced Choice Task

Are the dots moving left or right?



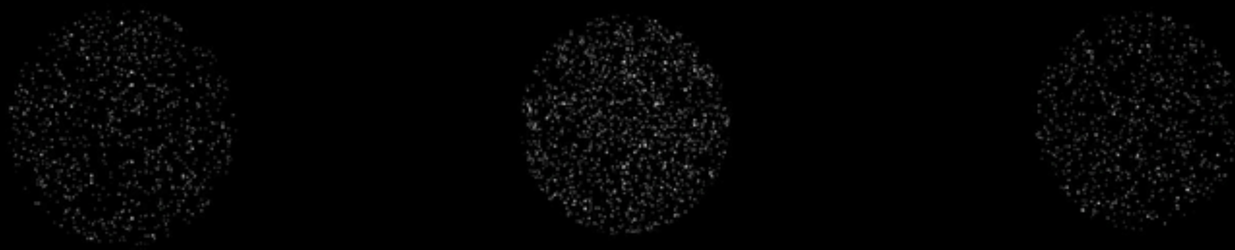
Two Alternative Forced Choice Task

Are the dots moving left or right?



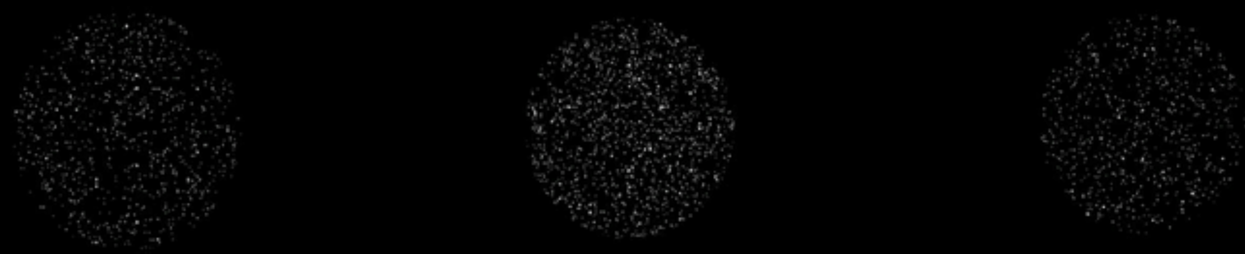
Two Alternative Forced Choice Task

Are the dots moving left or right?



Two Alternative Forced Choice Task

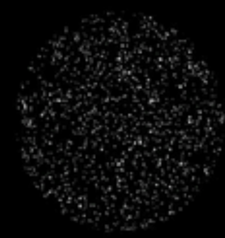
Are the dots moving left or right?



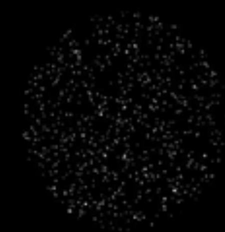
Two Alternative Forced Choice Task

Are the dots moving left or right?

High coherence



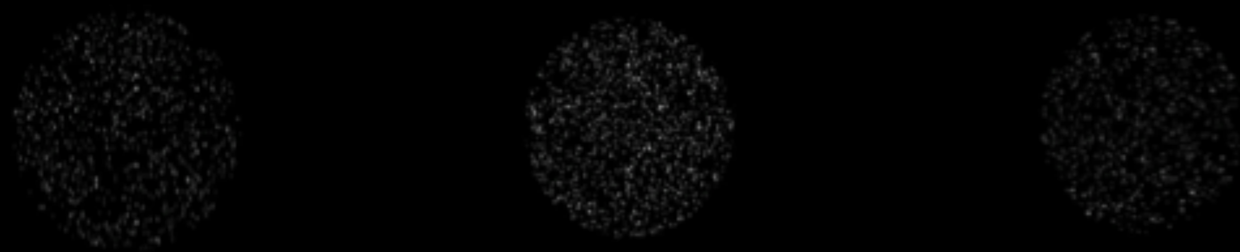
High coherence



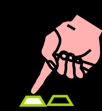
Low coherence

Two Alternative Forced Choice Task

Are the dots moving left or right?

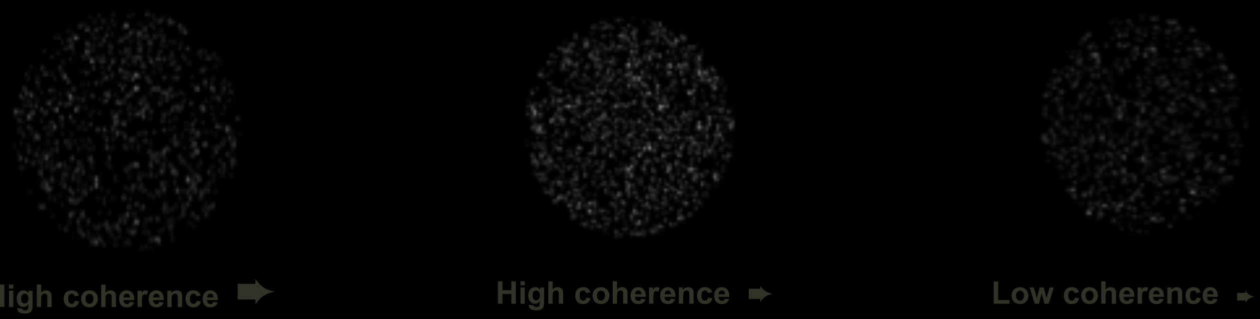


High coherence ➡ High coherence ➡ Low coherence ➡

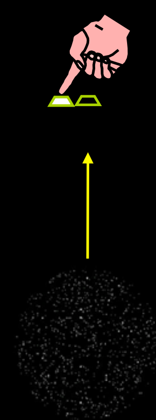


Two Alternative Forced Choice Task

Are the dots moving left or right?

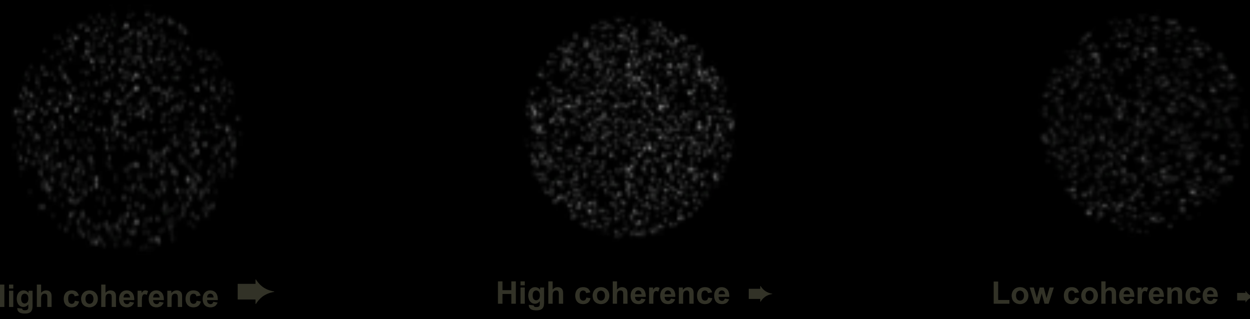


See dots moving left ( ),
press the left button



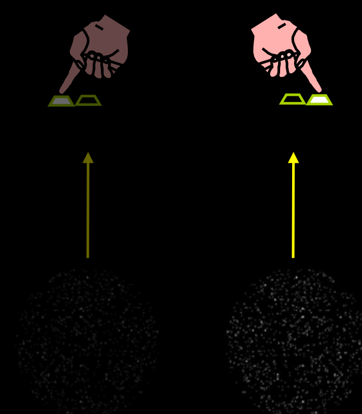
Two Alternative Forced Choice Task

Are the dots moving left or right?



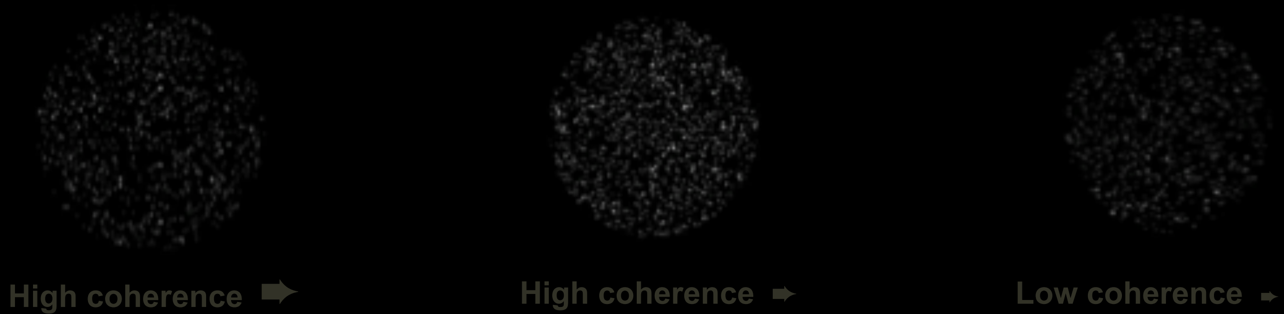
See dots moving left ( ),
press the left button

See dots moving right ( ),
press the right button



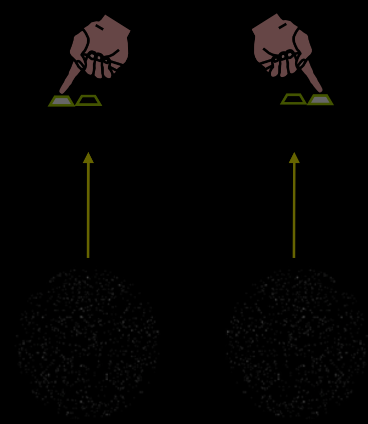
Two Alternative Forced Choice Task

Are the dots moving left or right?



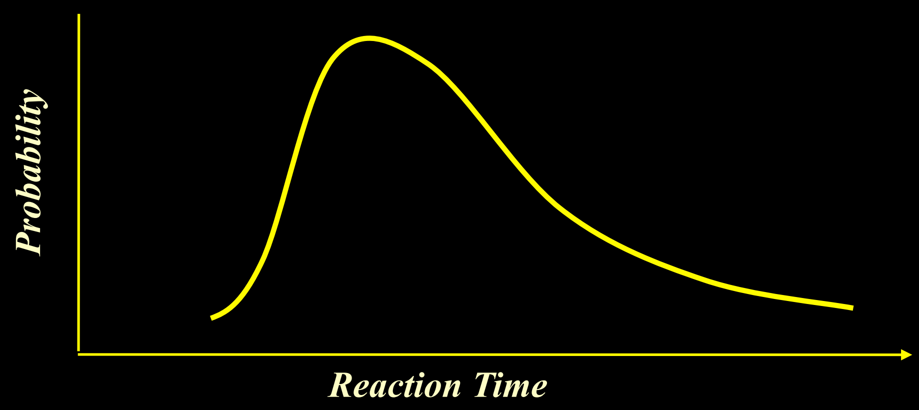
See dots moving left ( ),
press the left button

See dots moving right ( ),
press the right button



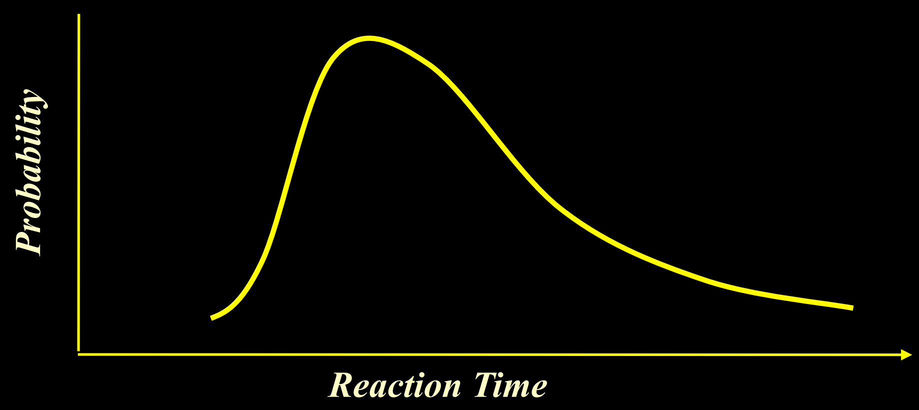
Measure reaction time (RT) and accuracy

Behavioral Findings



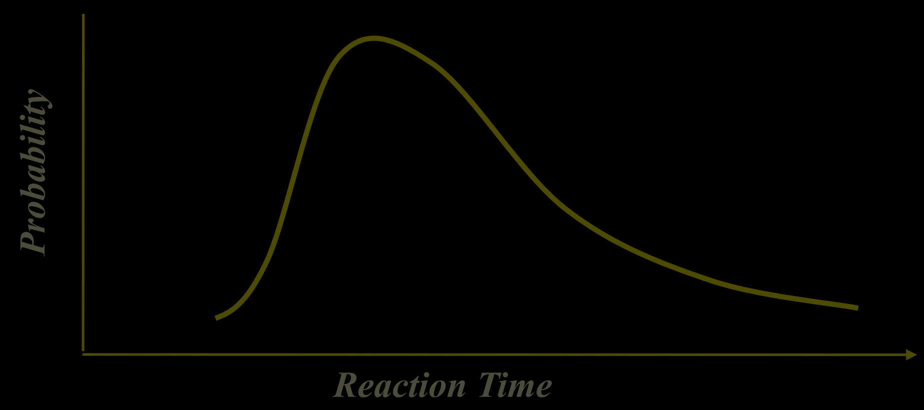
Behavioral Findings

- Characteristically skewed RT distribution:



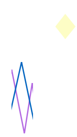
Behavioral Findings

- Characteristically skewed RT distribution:

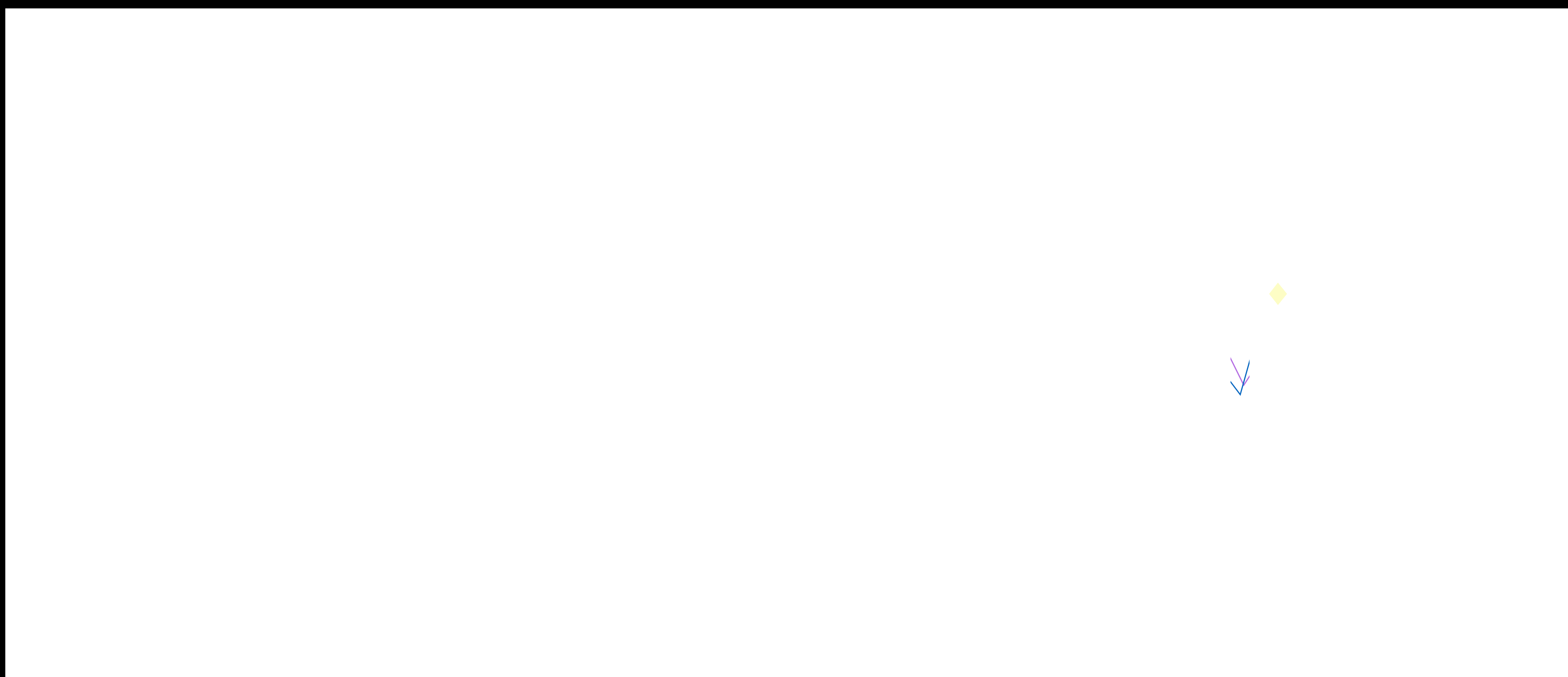


- Speed/accuracy tradeoff:
  - – faster response → less accurate
  - – more accurate → slower response

Neural Findings



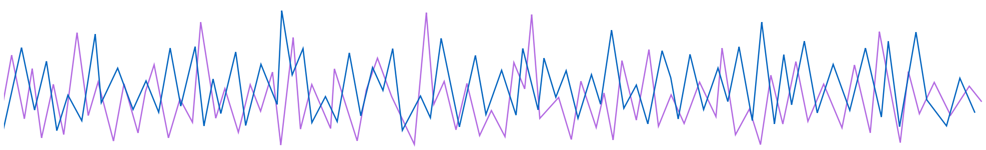
Neural Findings



Neural Findings

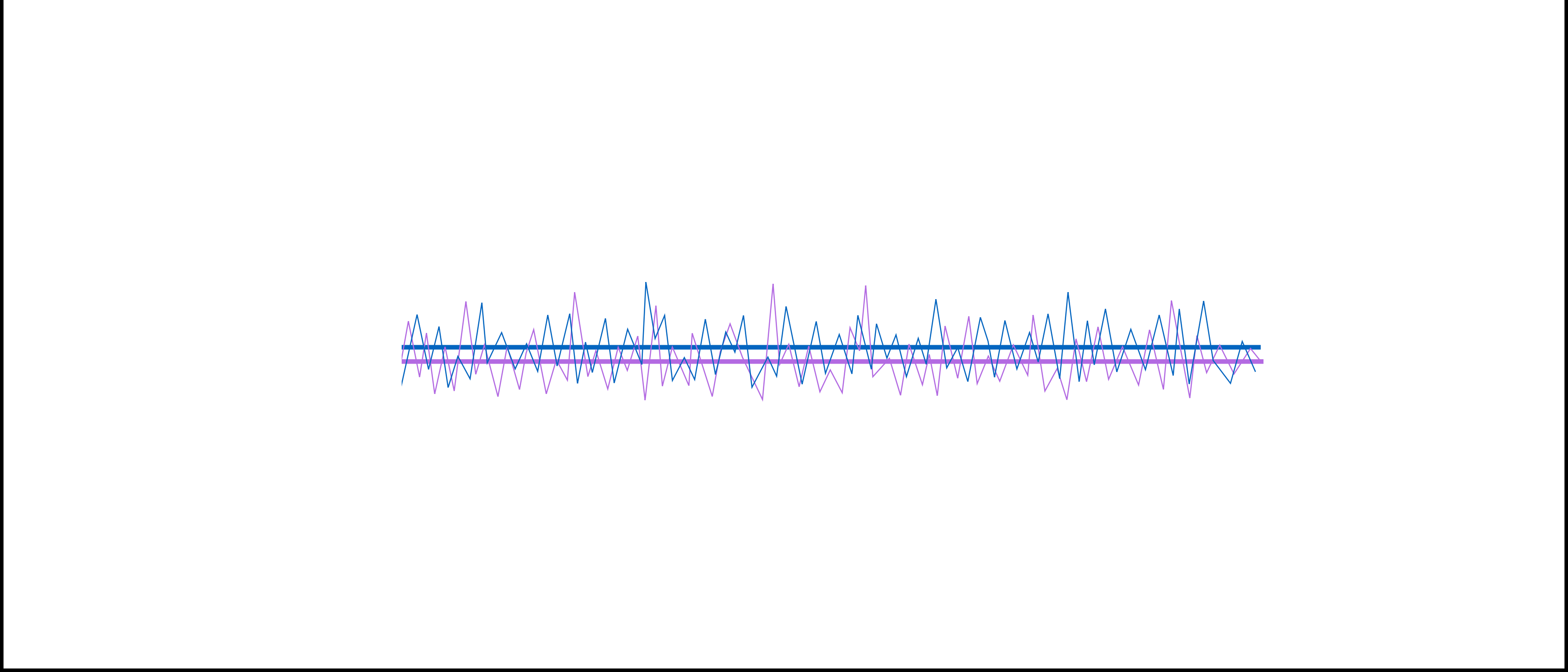
Which was higher, the blue or the purple?

Neural Findings



Which was higher, the blue or the purple?

Neural Findings



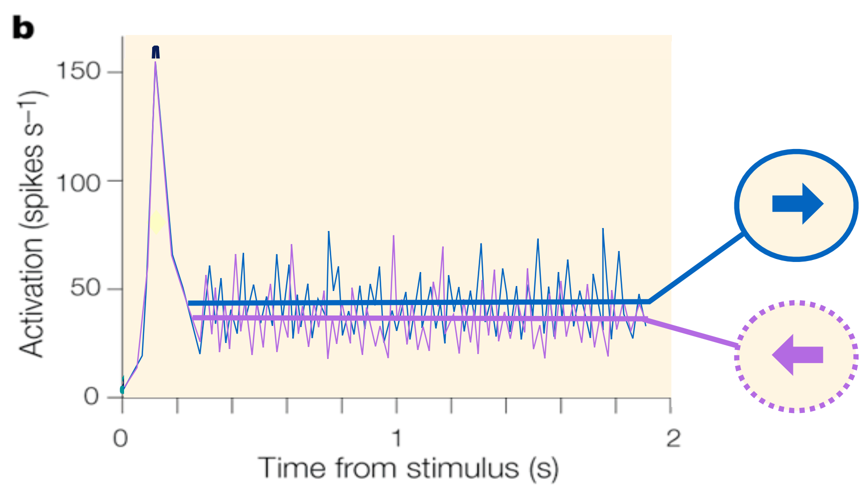
Which was higher, the blue or the purple?

Neural Findings

Motion Detection:

Area MT
(temporal cortex)

motion sensitive
visual cortex



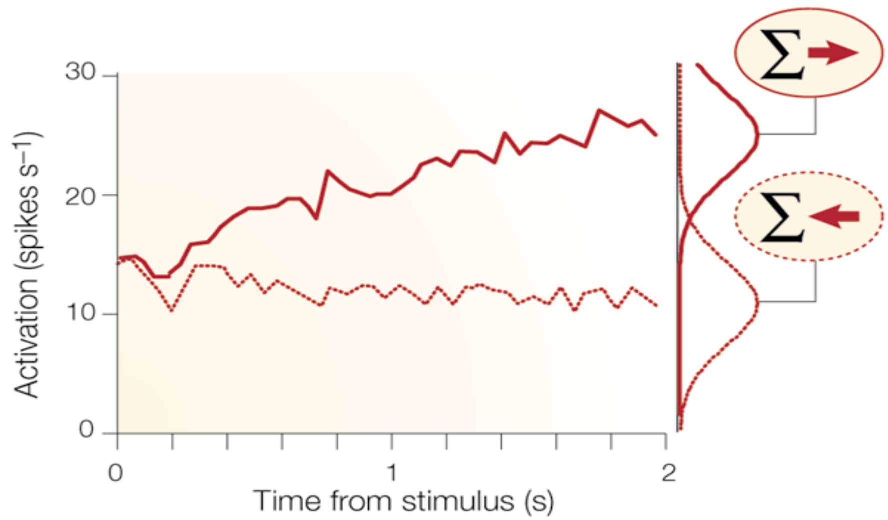
Shadlen & Newsome, 1998

Neural Findings

Integration:

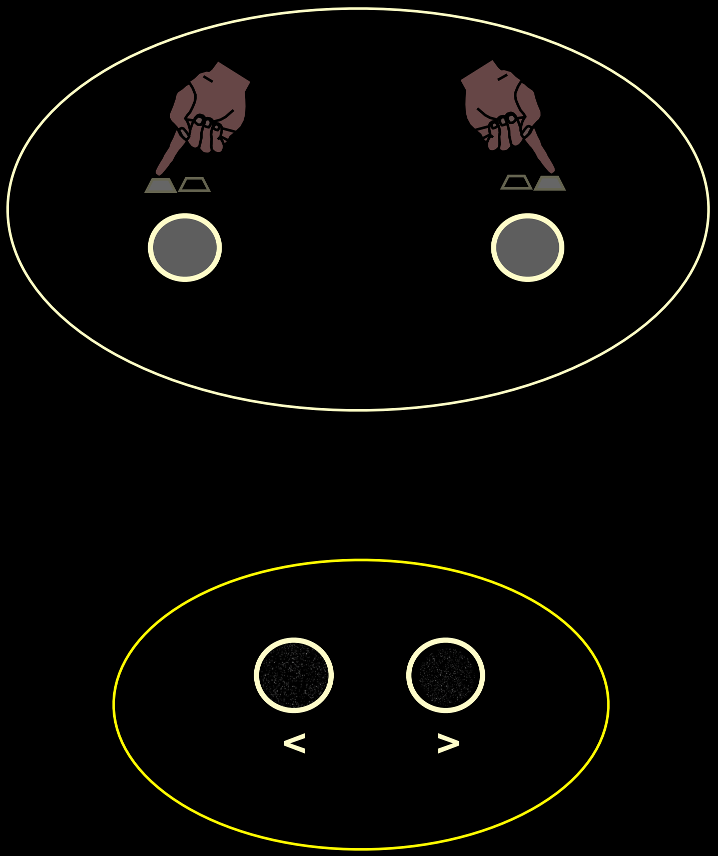
Areas LIP
(parietal cortex)
and SEF
(supplementary eye fields):

control of
eye movements



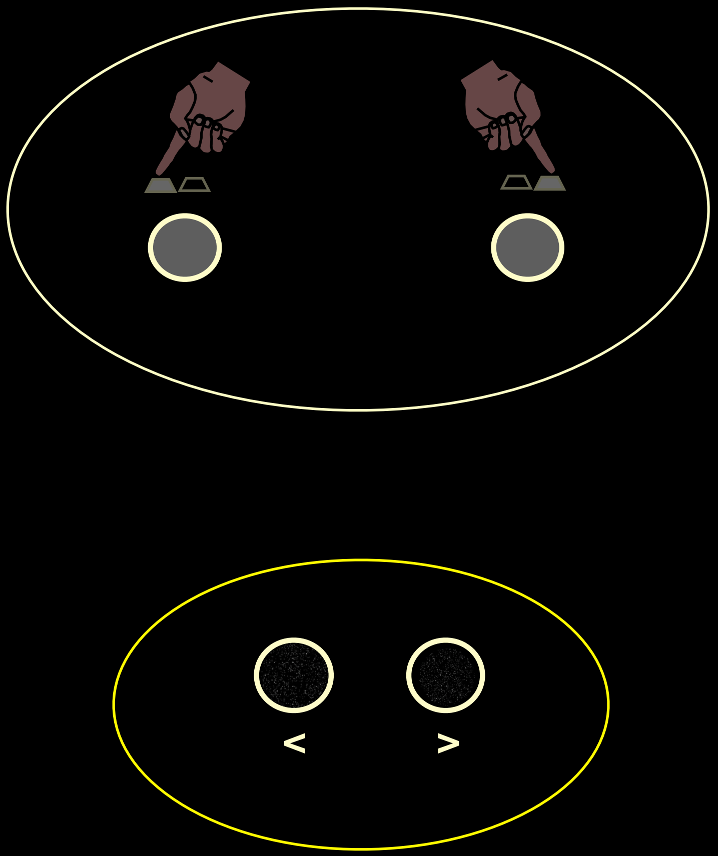
Shadlen & Newsome, 1998

Simple PDP Model of Two Alternative Decision Task



Usher & McClelland, 2001

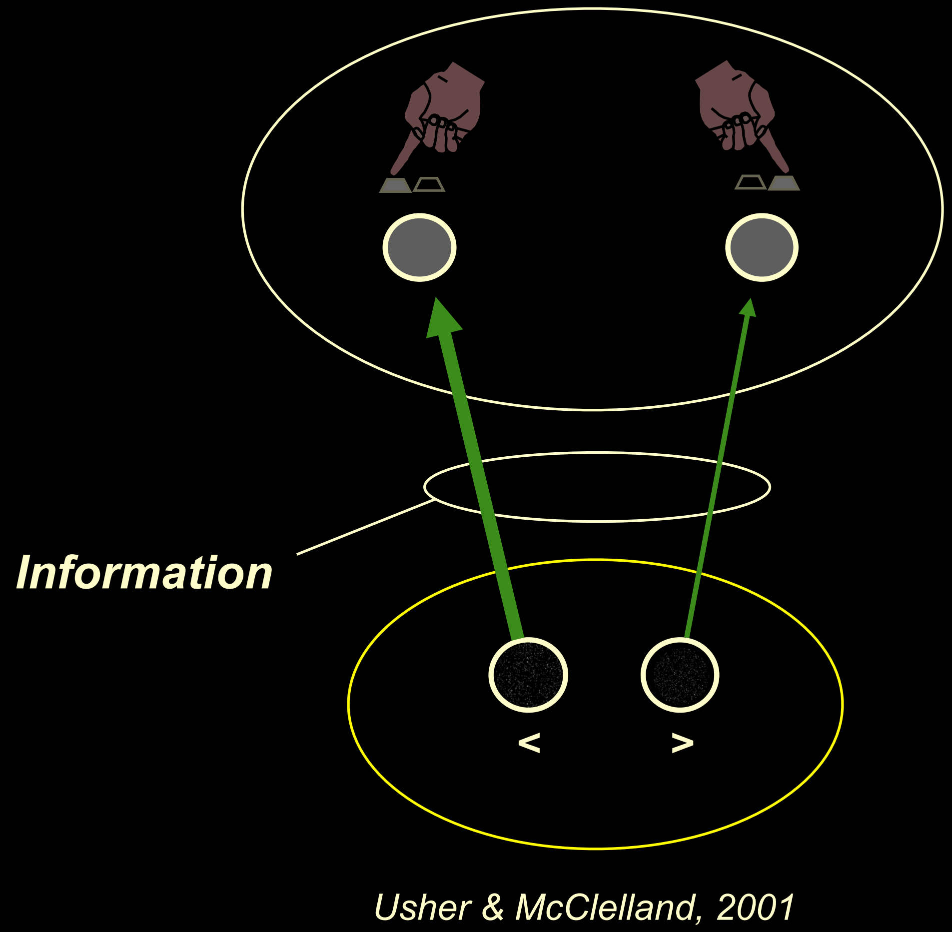
Simple PDP Model of Two Alternative Decision Task



Usher & McClelland, 2001

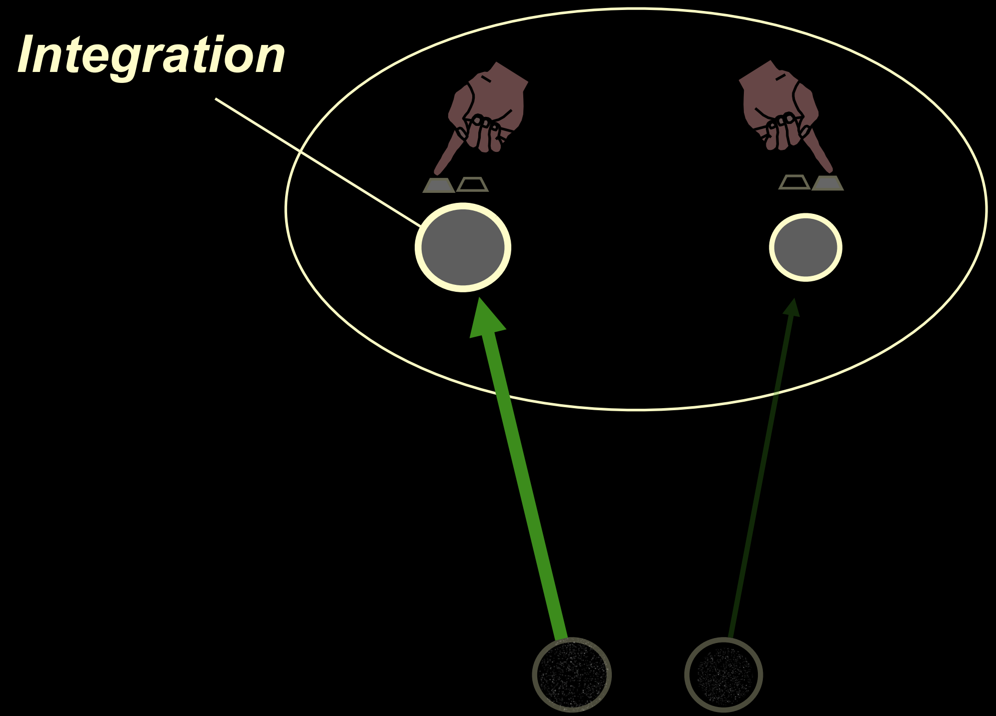
Simple PDP Model of Two Alternative Decision Task

- **Processing:**
  - **Flow of activity from stimulus inputs to a pair *decision units***



Simple PDP Model of Two Alternative Decision Task

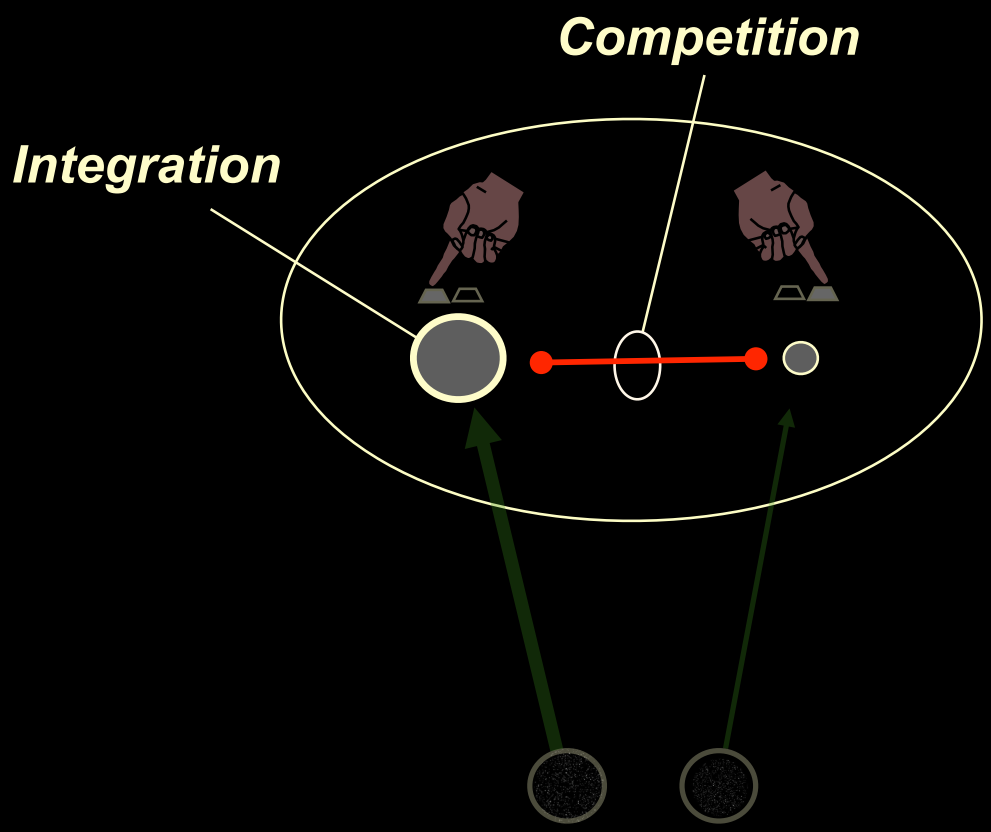
- **Processing:**
  - **Flow of activity from stimulus inputs to a pair *decision units***
  - **Each decision unit accumulates / *integrates* input**



Usher & McClelland, 2001

Simple PDP Model of Two Alternative Decision Task

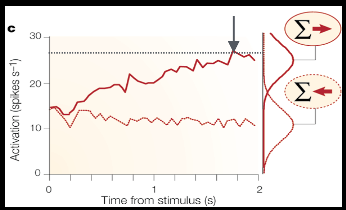
- **Processing:**
  - Flow of activity from stimulus inputs to a pair *decision units*
  - Each decision unit accumulates / *integrates* input
  - Decision units *compete*

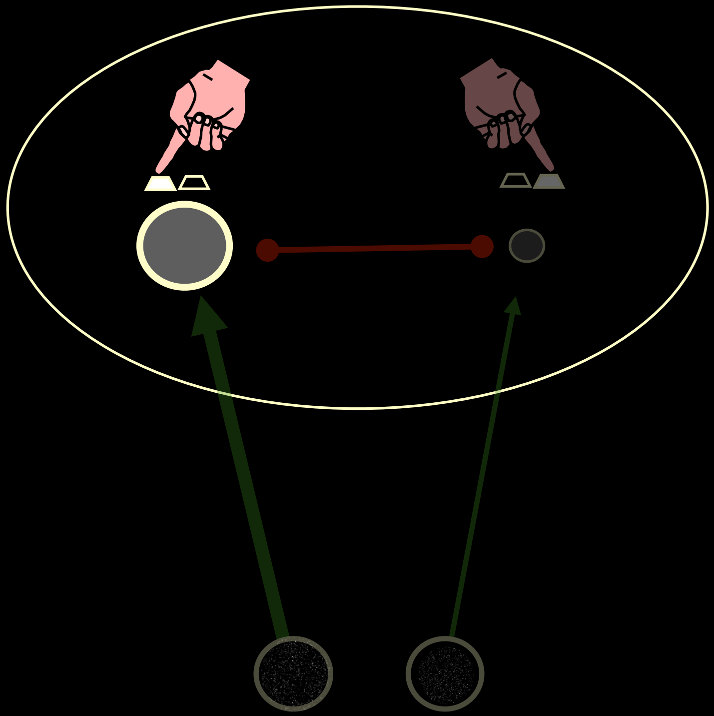


Usher & McClelland, 2001

Simple PDP Model of Two Alternative Decision Task

- Processing:
  - Flow of activity from stimulus inputs to a pair *decision units*
  - Each decision unit accumulates / *integrates* input
  - Decision units *compete*
- Decision:
  - Occurs when activity of one decision unit exceeds a specified *threshold*

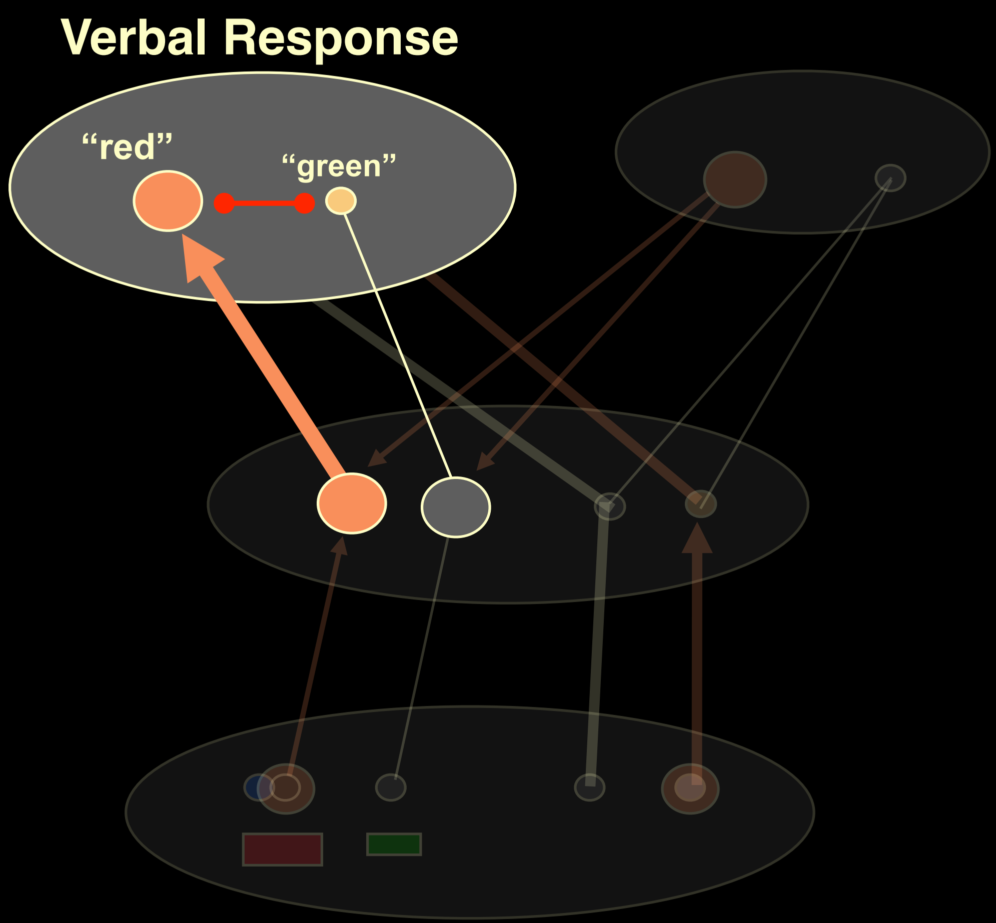




Usher & McClelland, 2001

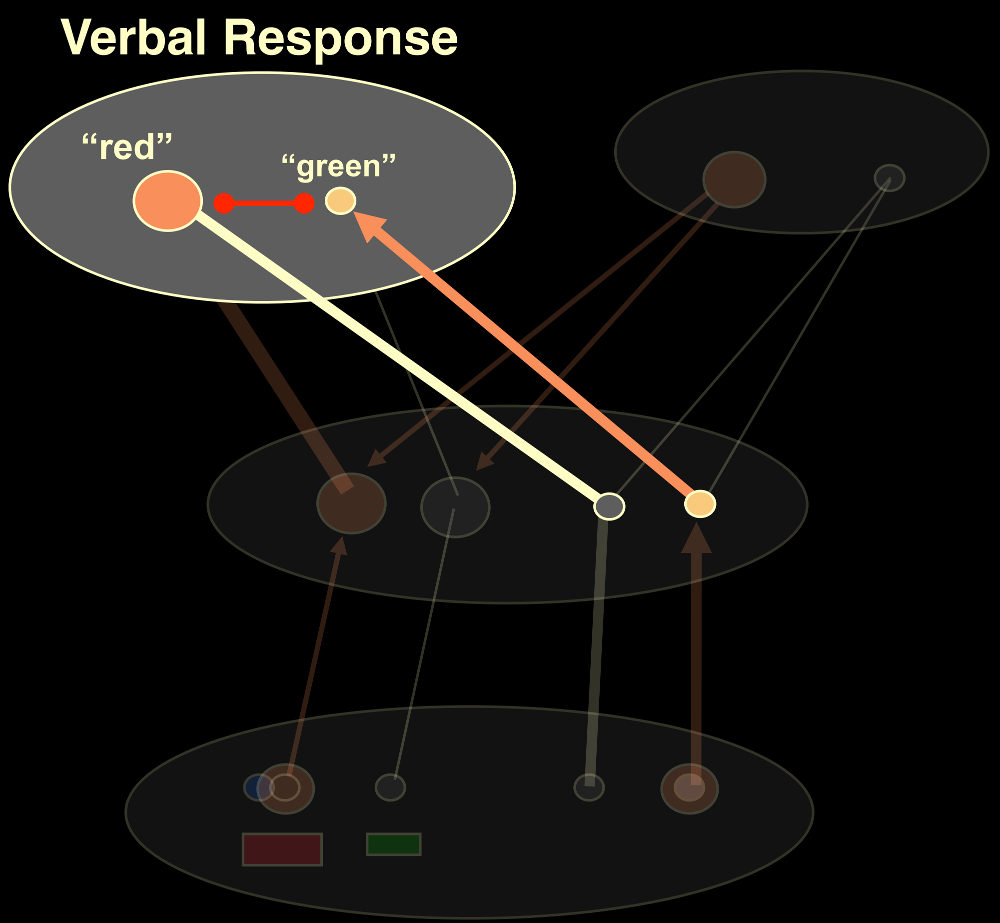
Fundamental Mechanism of Decision Making

- **Processing:**
  - Flow of activity from stimulus inputs to a pair *decision units*
  - Each decision unit accumulates / *integrates* input
  - Decision units *compete*
- **Decision:**
  - Occurs when activity of one decision unit exceeds a specified *threshold*



Fundamental Mechanism of Decision Making

- **Processing:**
  - Flow of activity from stimulus inputs to a pair ***decision units***
  - Each decision unit accumulates / ***integrates*** input
  - Decision units ***compete***
- **Decision:**
  - Occurs when activity of one decision unit exceeds a specified ***threshold***



Some Problems with PDP Models

Some Problems with PDP Models

- **Large parameter space:** *difficult to parameterize*
  - **how to set stimulus strength, connection weights, thresholds, etc.**

Some Problems with PDP Models

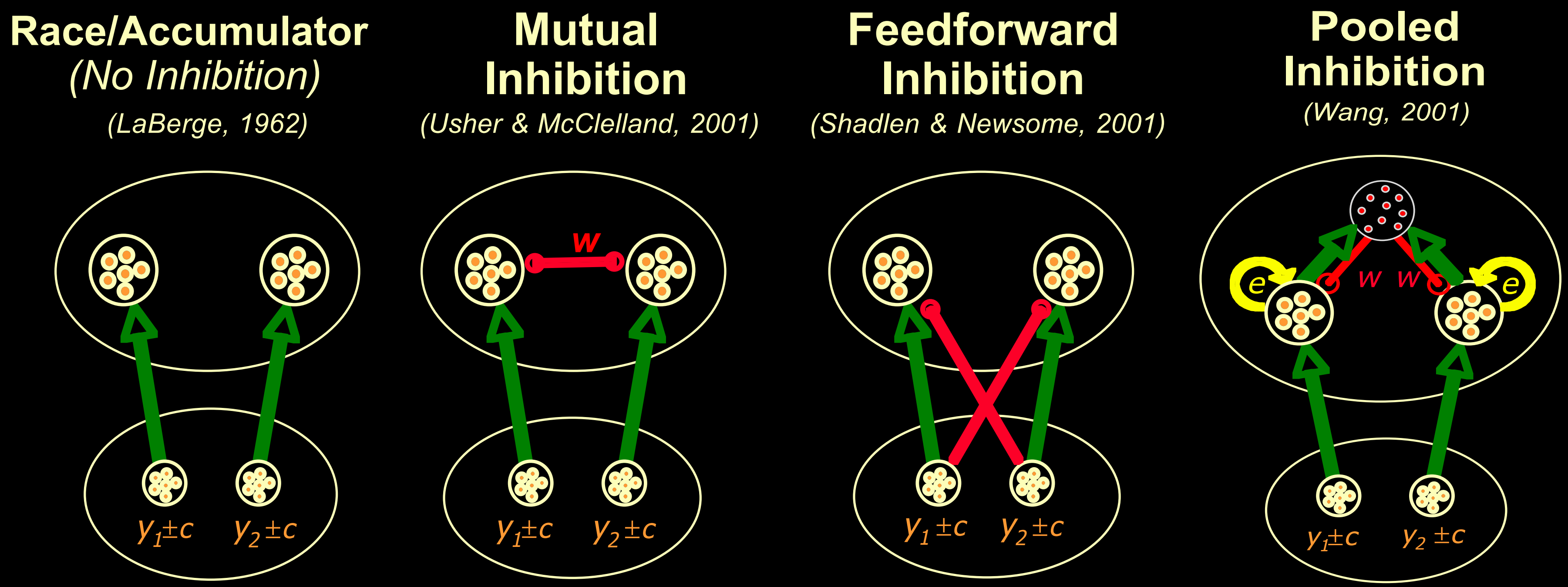
- **Large parameter space:** *difficult to parameterize*
  - **how to set stimulus strength, connection weights, thresholds, etc.**
- **Complex dynamics:** *hard to characterize and compare*

Some Problems with PDP Models

- **Large parameter space:** *difficult to parameterize*
  - **how to set stimulus strength, connection weights, thresholds, etc.**
- **Complex dynamics:** *hard to characterize and compare*
- **Theoretical degeneracy:** *proliferation of models…*

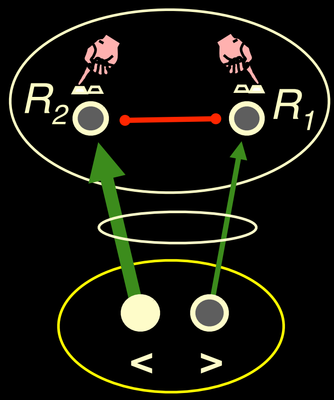
Some Problems with PDP Models

- **Large parameter space:** *difficult to parameterize*
  - **how to set stimulus strength, connection weights, thresholds, etc.**
- **Complex dynamics:** *hard to characterize and compare*
- **Theoretical degeneracy:** *proliferation of models…*



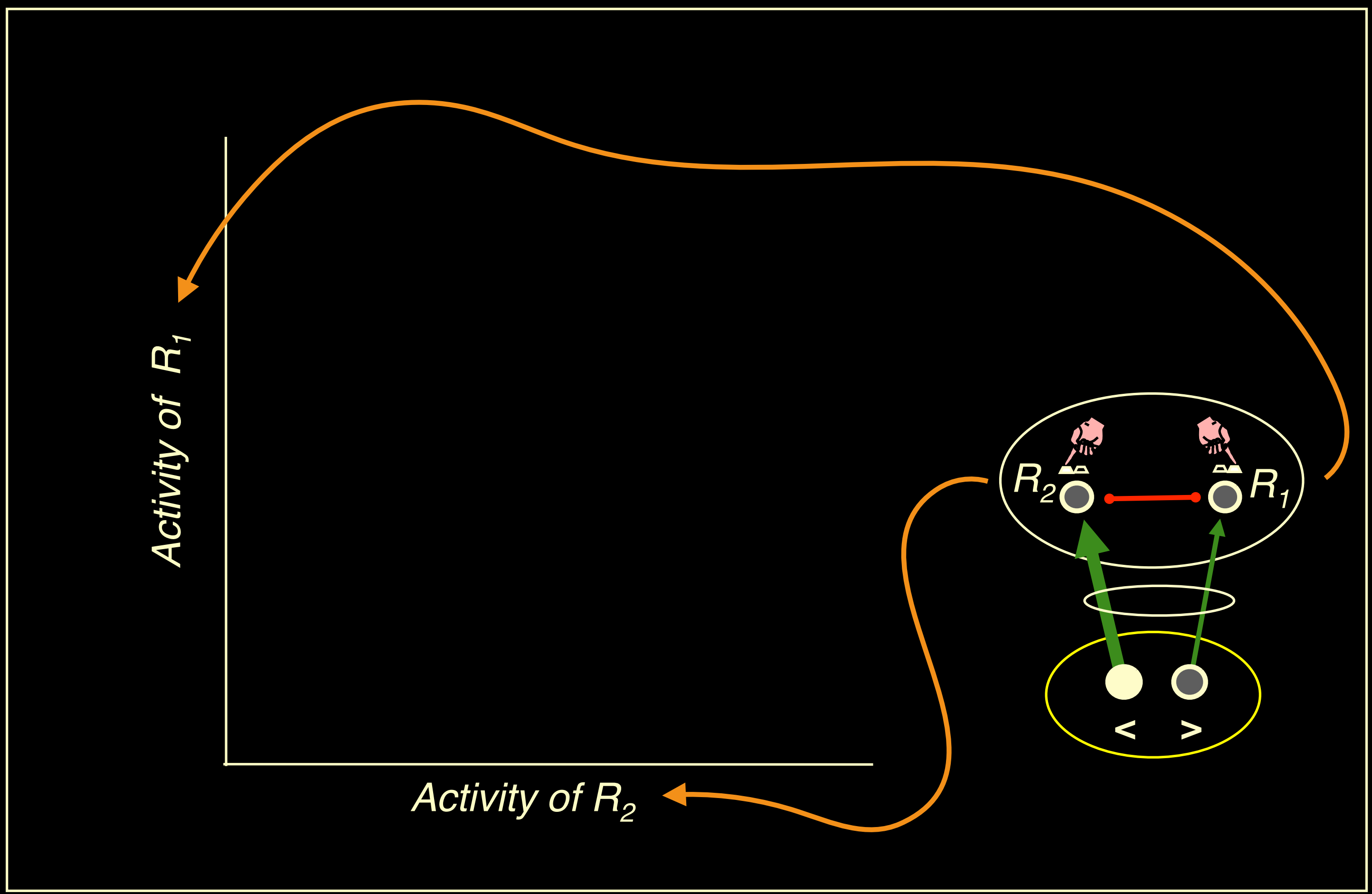
Simplification & Analysis

Step 1: Construct geometric representation of model's behavior



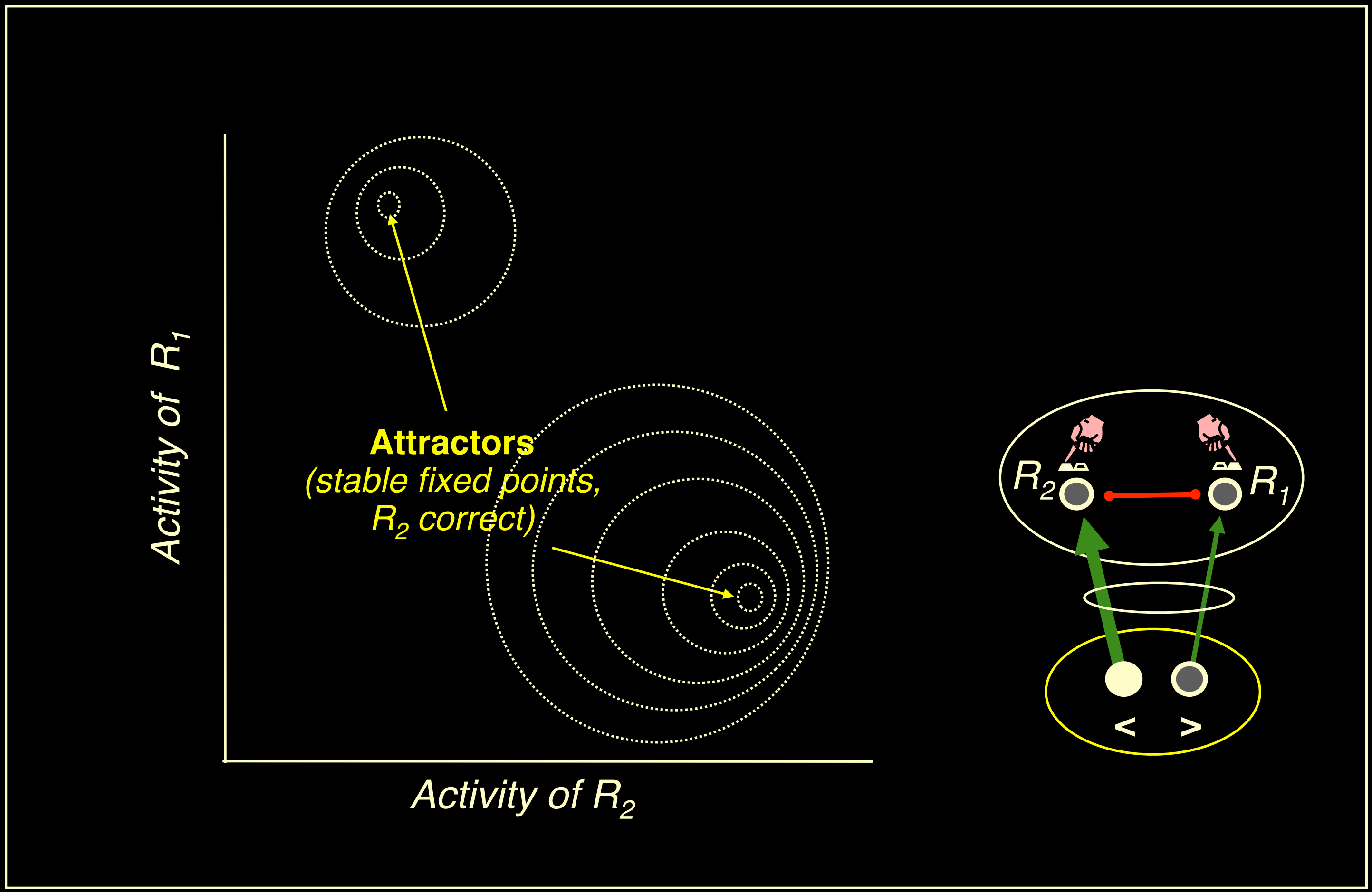
Simplification & Analysis

Step 1: Construct geometric representation of model's behavior



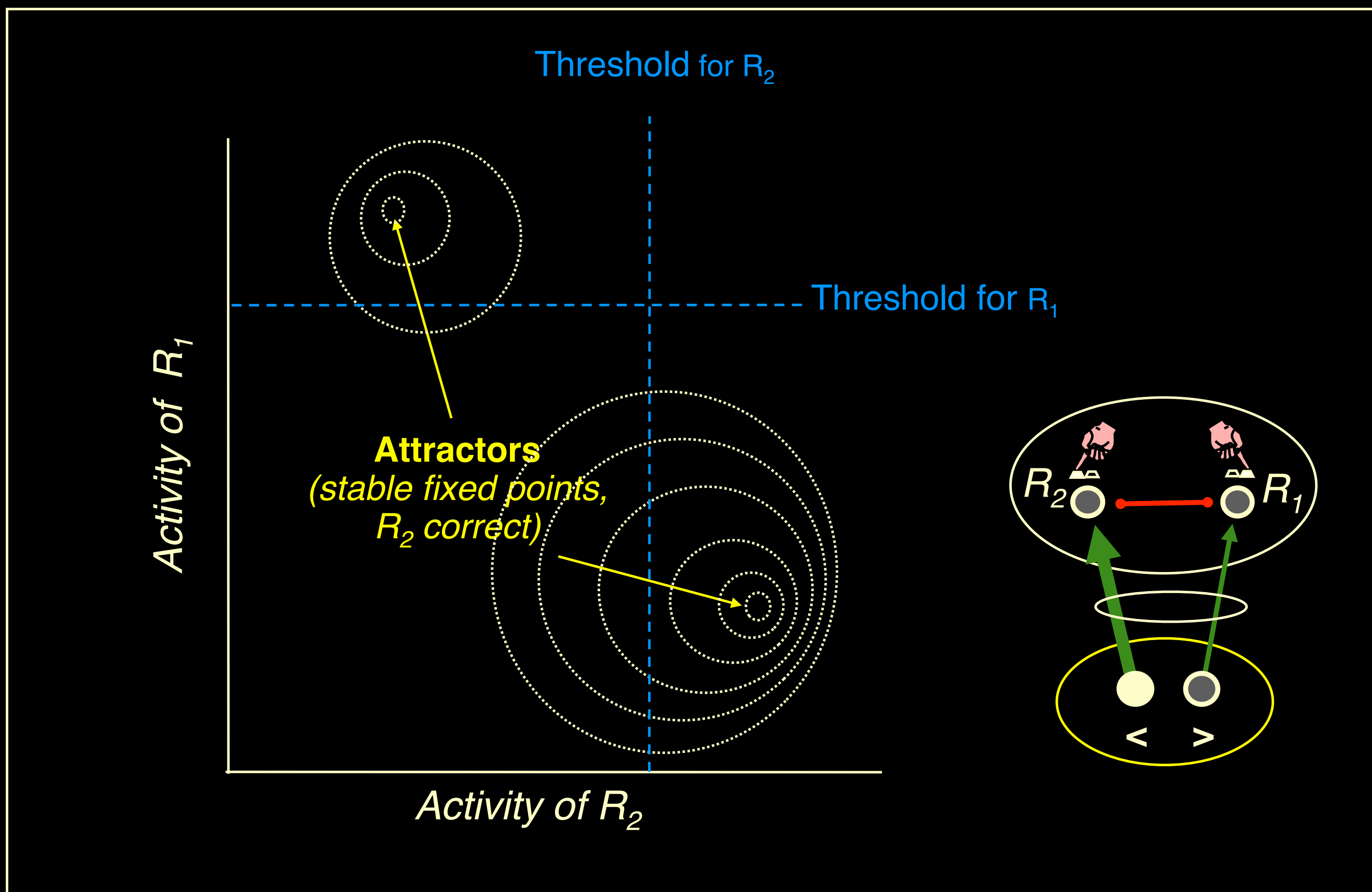
Simplification & Analysis

Step 1: Construct geometric representation of model's behavior



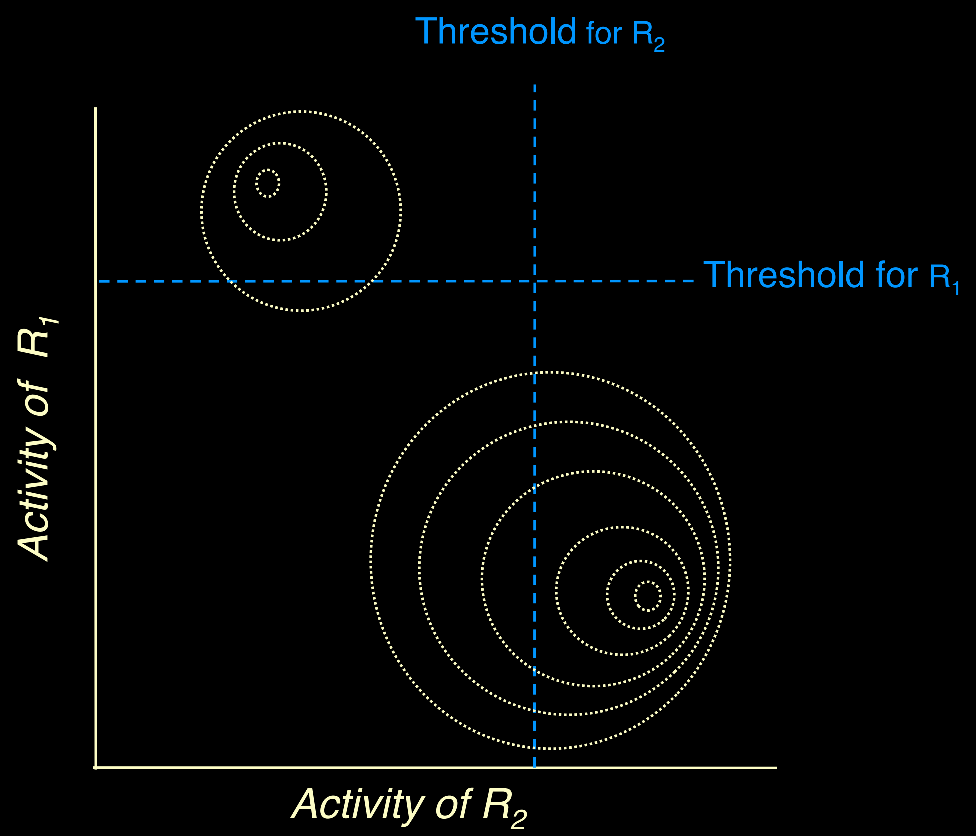
Simplification & Analysis

Step 1: Construct geometric representation of model's behavior



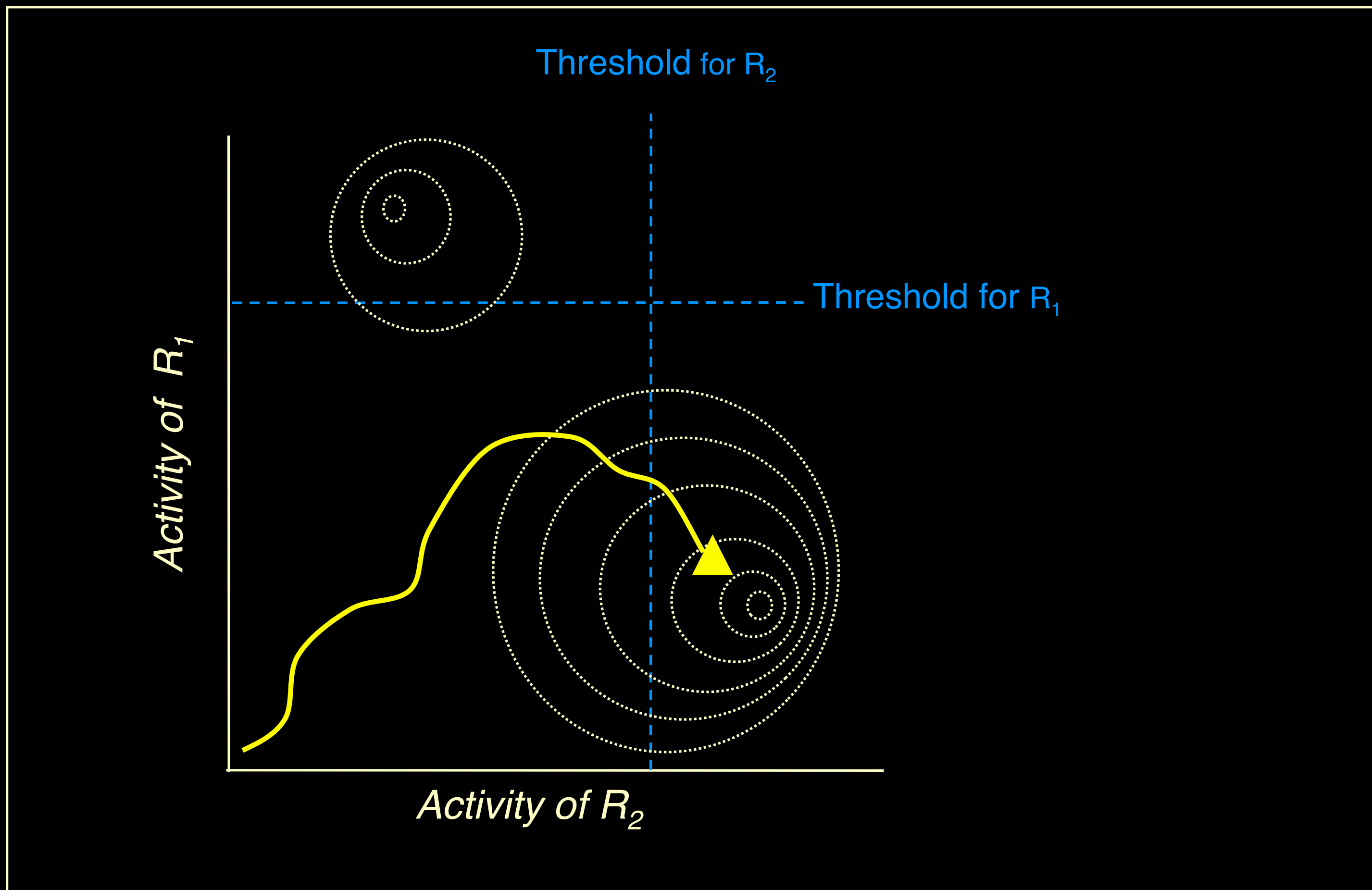
Simplification & Analysis

Step 2: Examine Dynamics



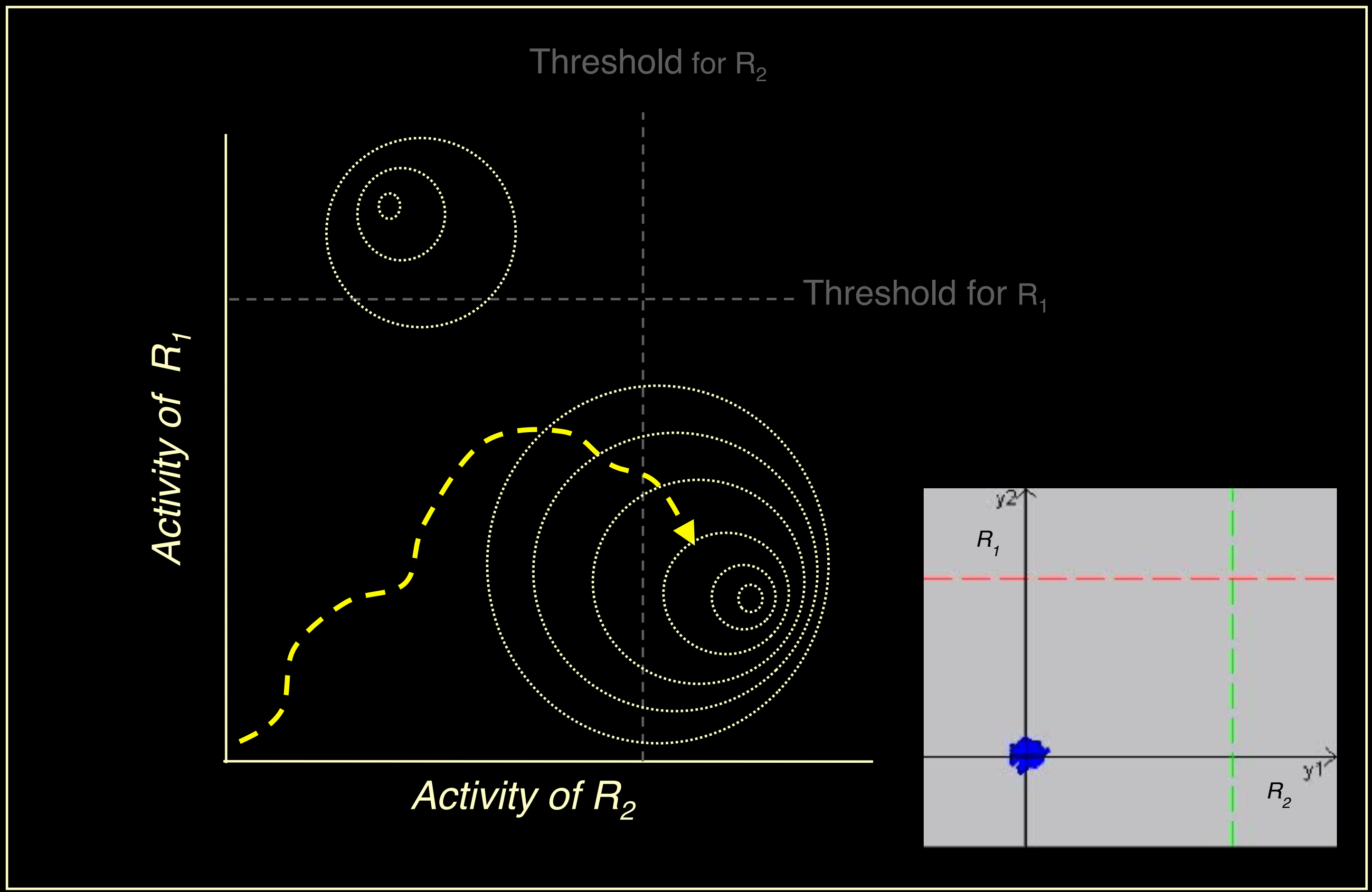
Simplification & Analysis

Step 2: Examine Dynamics



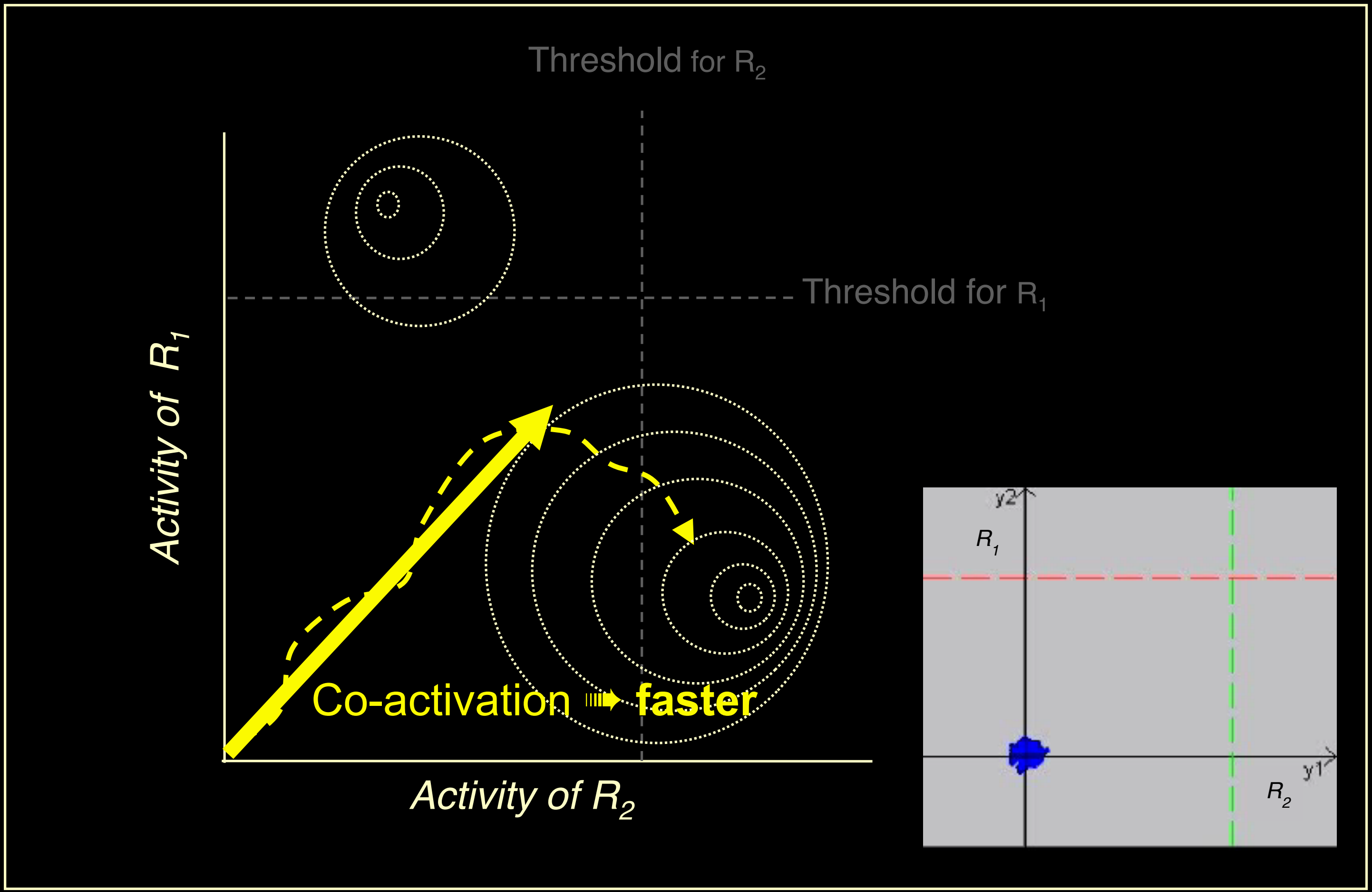
Simplification & Analysis

Step 3: Note that there are two components of the trajectory…



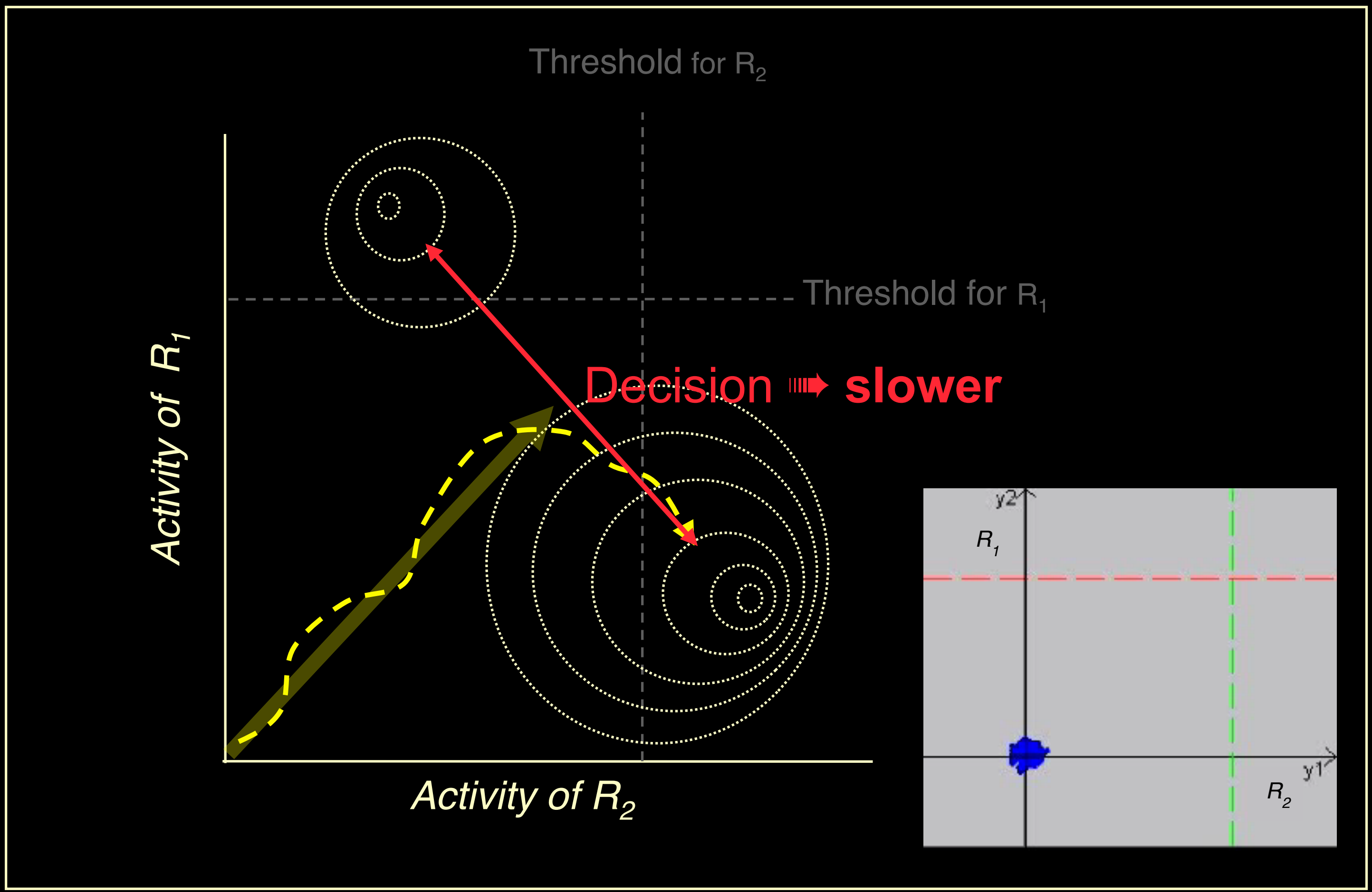
Simplification & Analysis

Step 3: Note that there are two components of the trajectory…



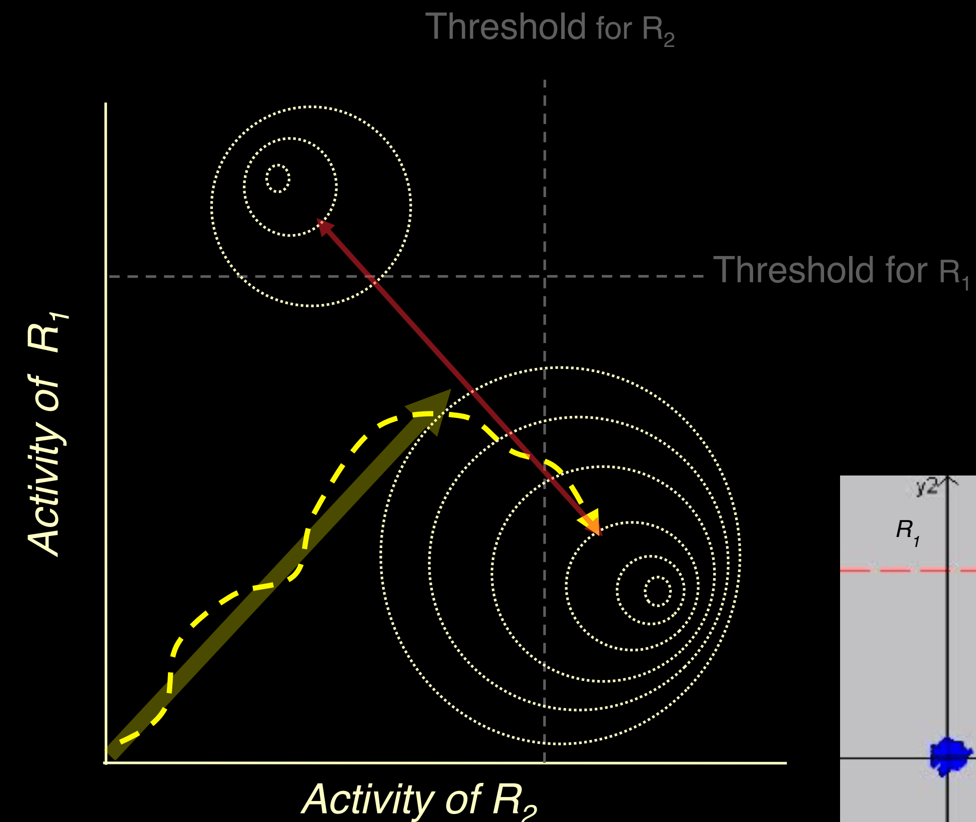
Simplification & Analysis

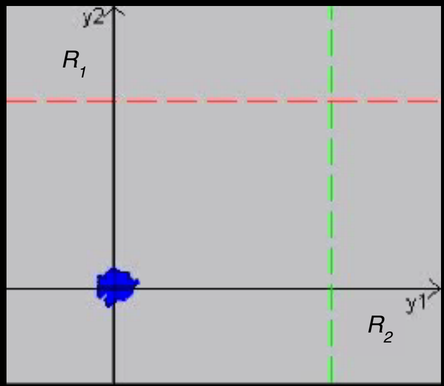
Step 3: Note that there are two components of the trajectory…



Simplification & Analysis

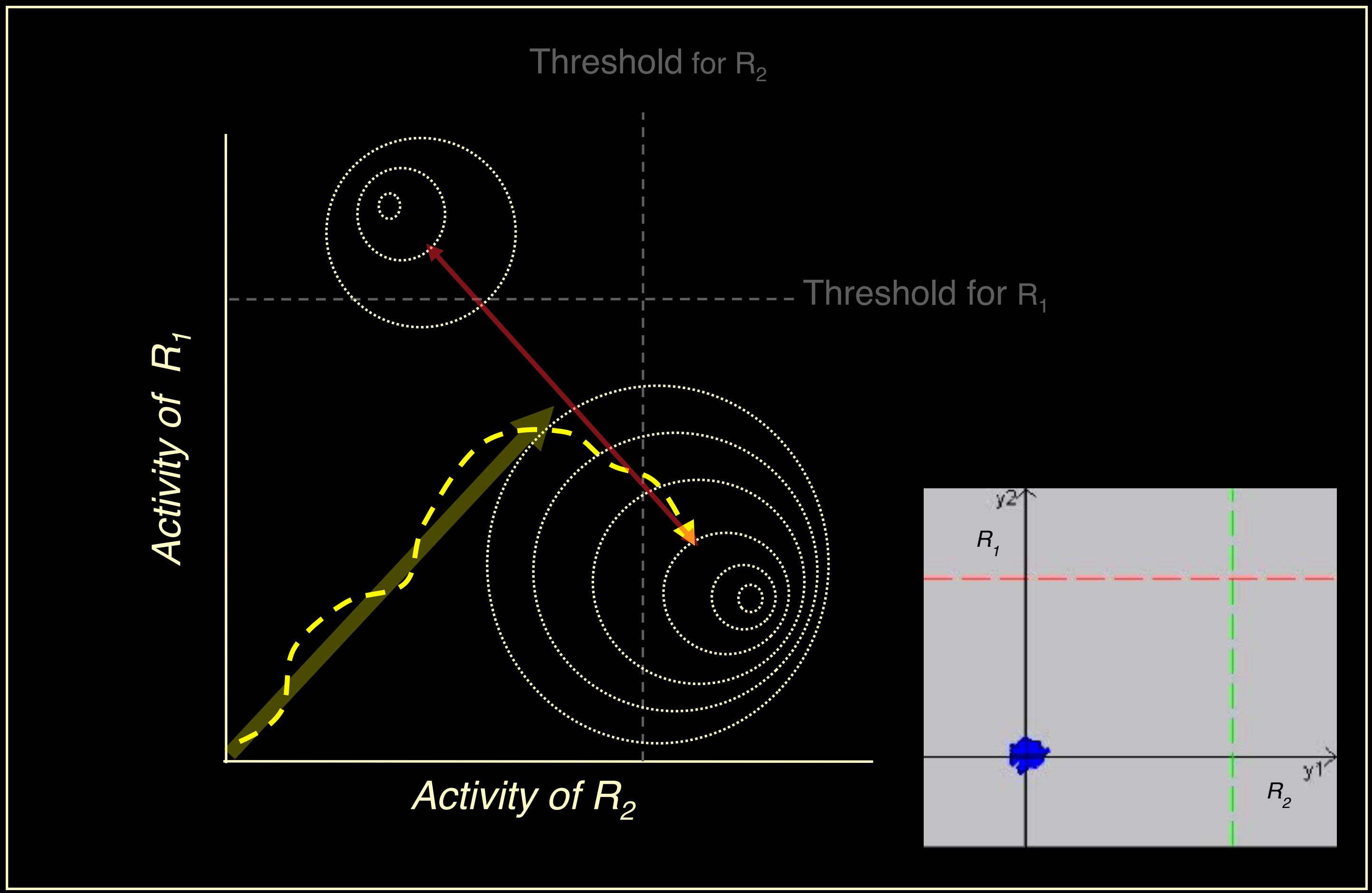
and that they have different dynamics





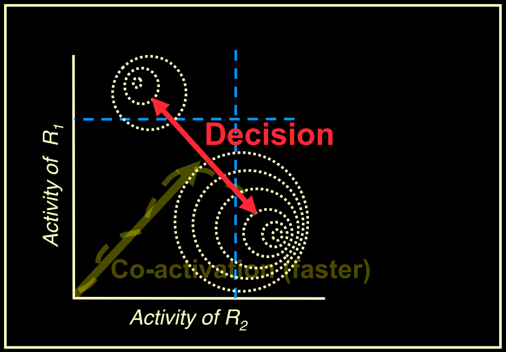
Simplification & Analysis

and that they have different dynamics



Simplification & Analysis

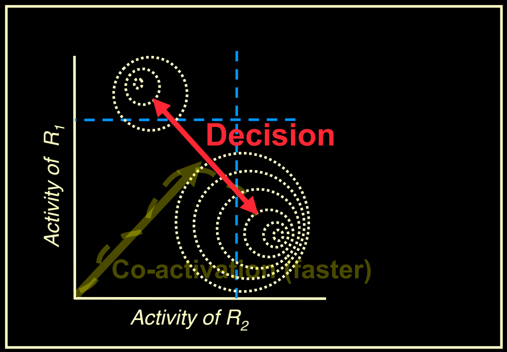
Step 4: Dimension Reduction



Simplification & Analysis

Step 4: Dimension Reduction

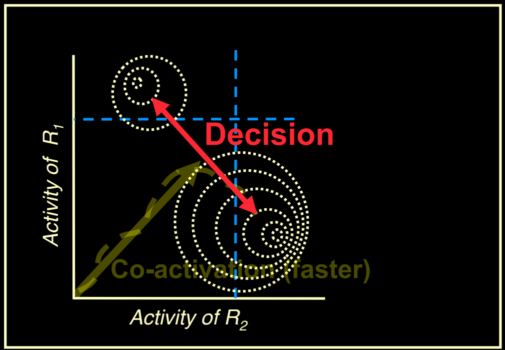
- **Focus only on one dimension** *(Brown & Holmes, 2001)*



Simplification & Analysis

Step 4: Dimension Reduction

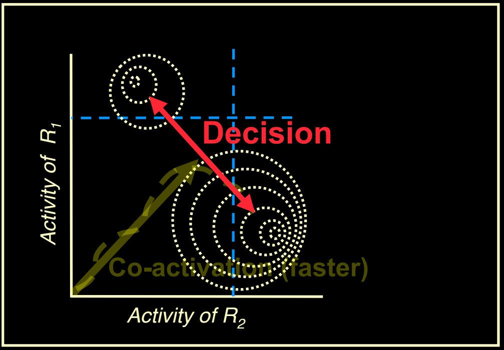
- **Focus only on one dimension** *(Brown & Holmes, 2001)*
  - **assume that most of the "action" is along the decision line**



Simplification & Analysis

Step 4: Dimension Reduction

- **Focus only on one dimension** *(Brown & Holmes, 2001)*
  - **assume that most of the "action" is along the decision line**
  - **therefore, decision process can be approximated by a *one-dimensional process* (*difference in activity*)**



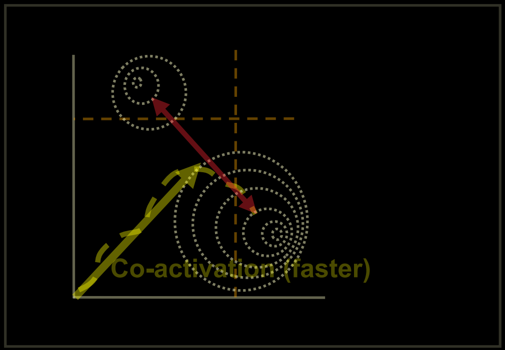
Simplification & Analysis

Step 5: Linearization

- **Focus only on one dimension** *(Brown & Holmes, 2001)*

assume that most of the "action" is along the decision line

therefore, decision process can be approximated by a *one-dimensional process* (*difference in activity*



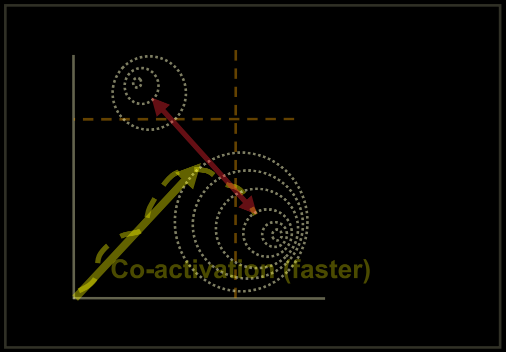
Simplification & Analysis

Step 5: Linearization

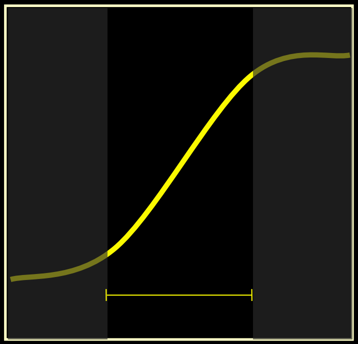
- **Focus only on one dimension** *(Brown & Holmes, 2001)*

  assume that most of the "action" is along the decision line

  therefore, decision process can be approximated by a *one-dimensional process* (*difference in activity*



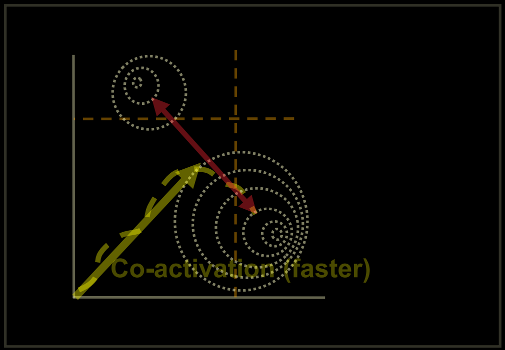
- **Focus on linear range of function**



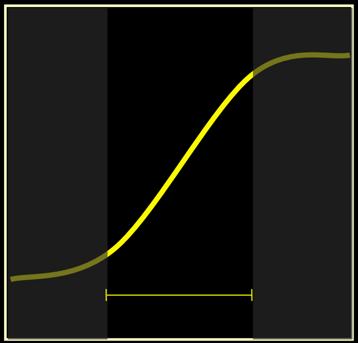
Simplification & Analysis

Step 5: Linearization

- **Focus only on one dimension** *(Brown & Holmes, 2001)*
  assume that most of the "action" is along the decision line
  therefore, decision process can be approximated by a *one-dimensional process* (*difference in activity*



- **Focus on linear range of function**
  - **assume that units in the "focus of attention" are on the linear part of their activation function (i.e., most sensitive part of their dynamic range)**
    Cohen et al (Psychological Review, 1990)



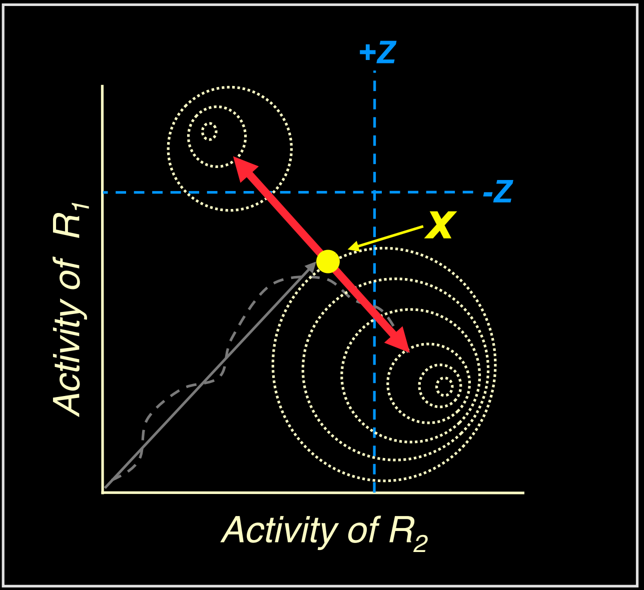
Simplification

- **Drift Diffusion Model (DDM):**

$$\dot{x} = A + c$$

A = drift rate
c = noise
P(x,t) = N(At, c√t)

Process ends when x exceeds ± z



Simplification

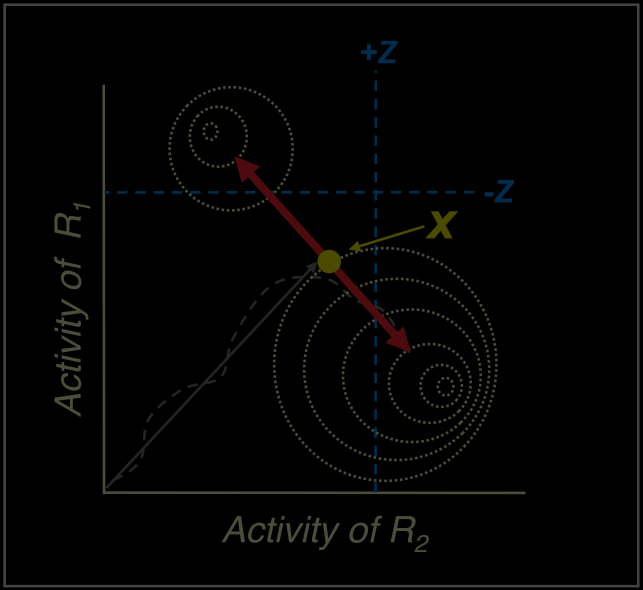
- **Drift Diffusion Model (DDM):**

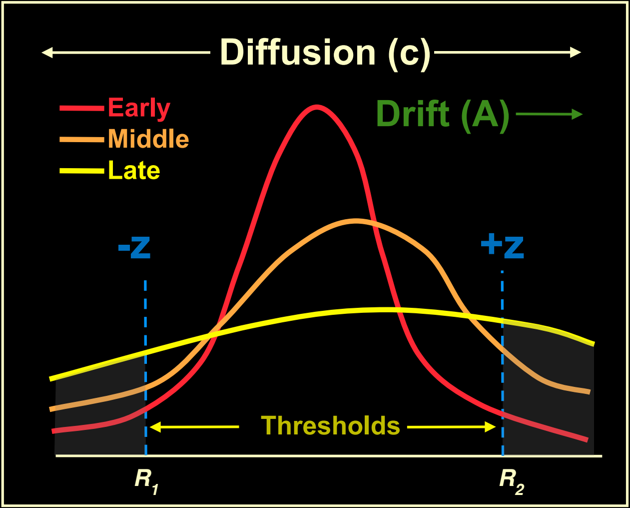
$$\dot{x} = A + c$$

A = drift rate
c = noise

$$P(x,t) = N(At, c\sqrt{t})$$

Process ends when x exceeds ± z





  - This process is described by the **Fokker-Planck equation** for the evolution of a Gaussian probability distribution toward a pair of boundaries

Simplification

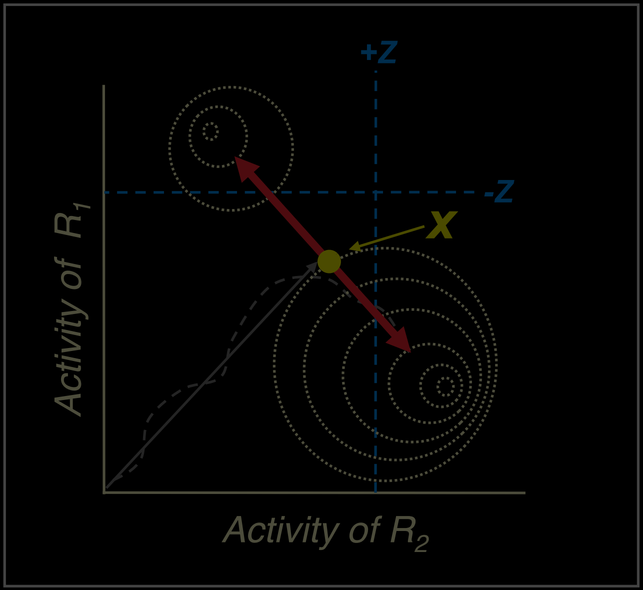
- Drift Diffusion Model (DDM):

$$\dot{x} = A + c$$

A = drift rate
c = noise

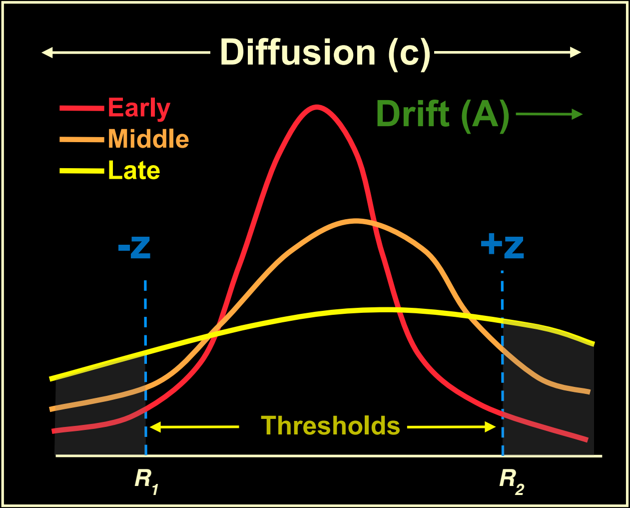
$$P(x,t) = N(At, c\sqrt{t})$$

Process ends when x exceeds ± z



- This process is described by the Fokker-Planck equation for the evolution of a Gaussian probability distribution toward a pair of boundaries

  = mathematical description of diffusion of an ideal gas



Simplification

- **Drift Diffusion Model (DDM):**

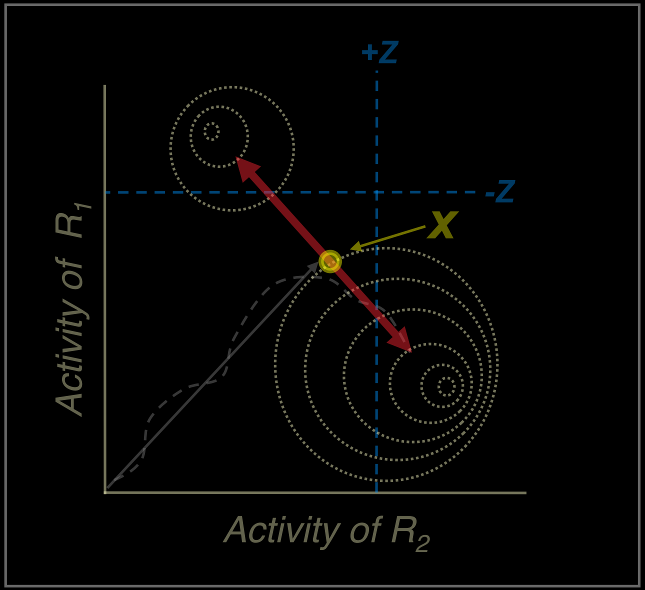
$$x = A + c$$

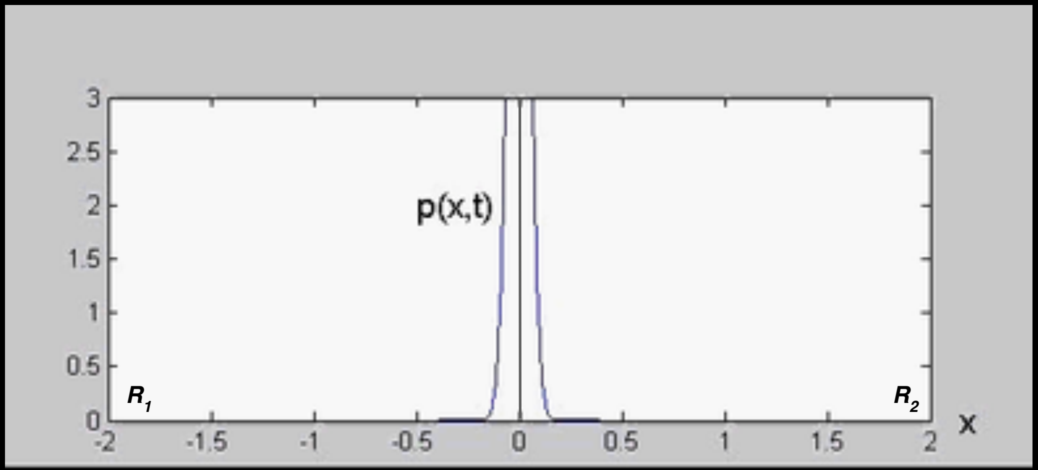
A = drift rate

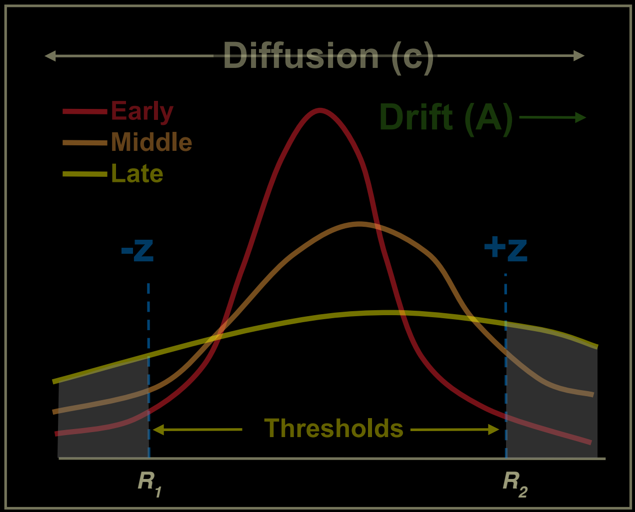
c = noise

$P(x,t) = N(At, c\sqrt{t})$

Process ends when x exceeds ± z







Simplification

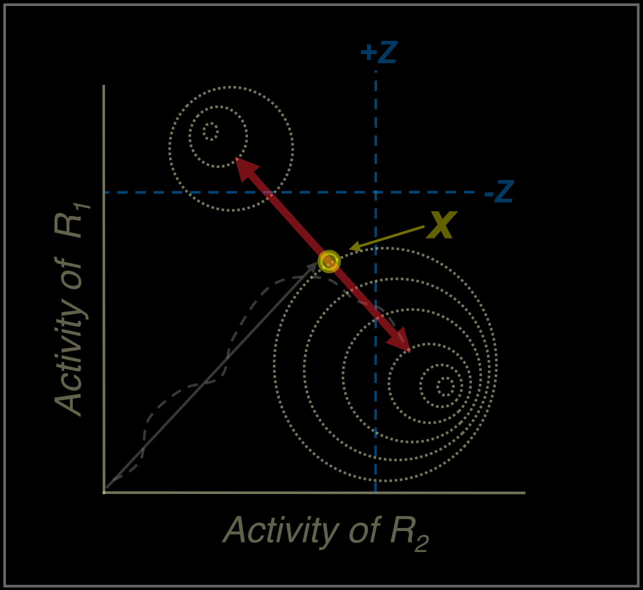
- **Drift Diffusion Model (DDM):**

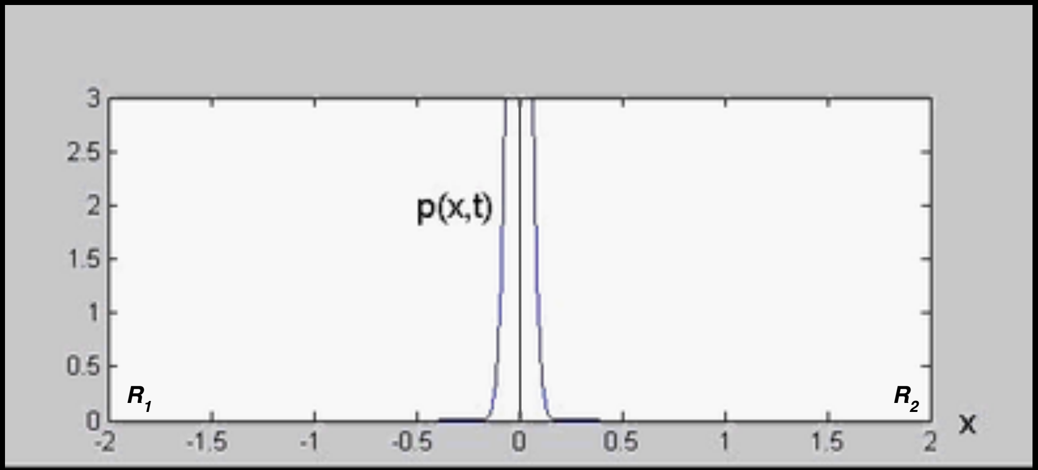
$$x = A + c$$

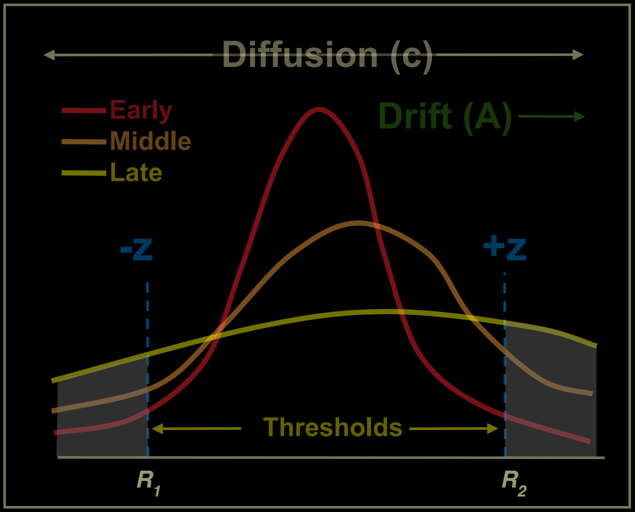
A = drift rate
c = noise

$$P(x,t) = N(At, c\sqrt{t})$$

Process ends when x exceeds ± z







Simplification

- **Drift Diffusion Model (DDM):**

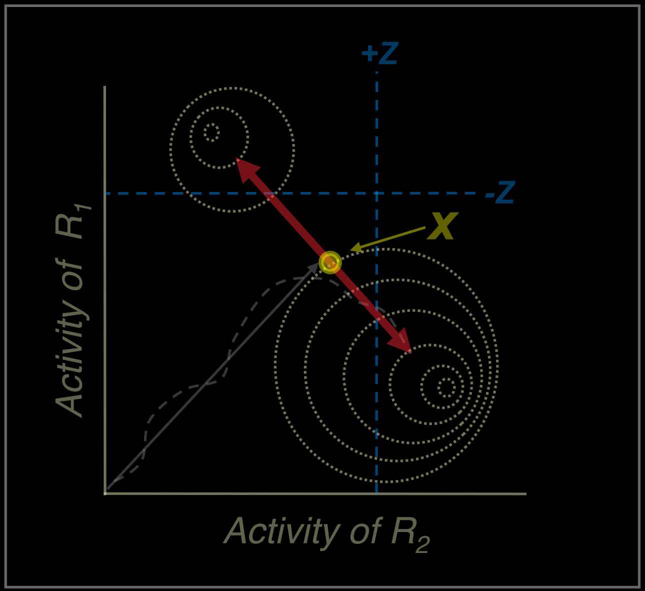
$$x = A + c$$

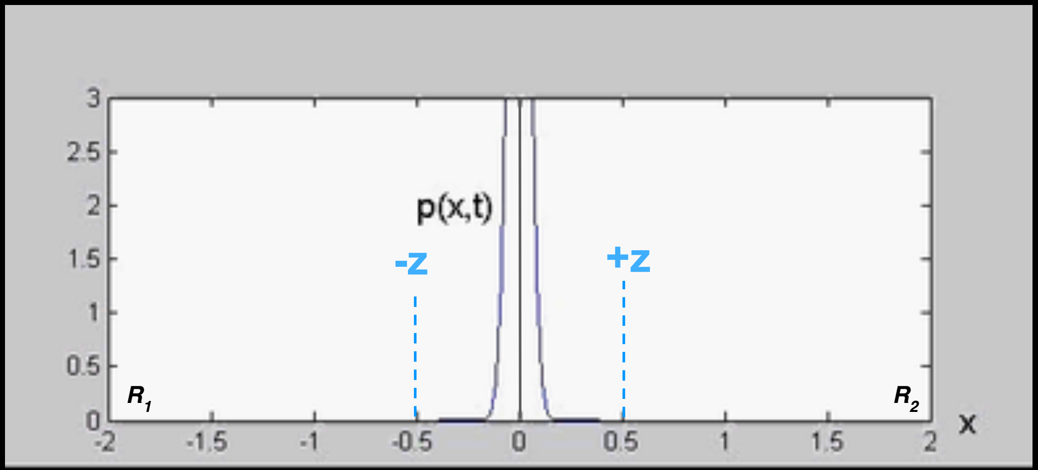
A = drift rate

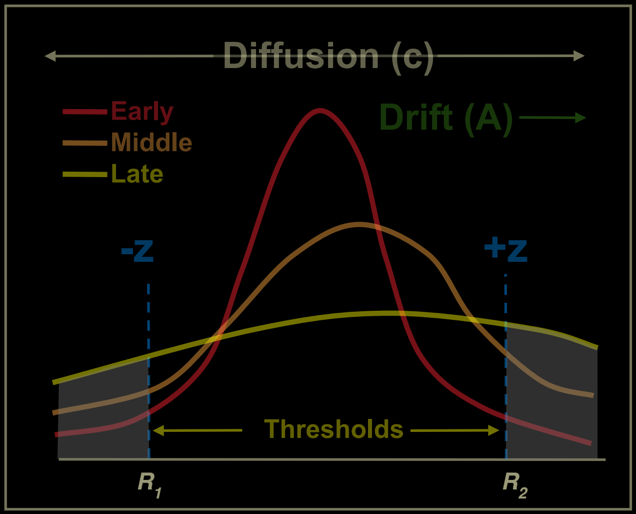
c = noise

$$P(x,t) = N(At, c\sqrt{t})$$

Process ends when x exceeds $\pm$ z







Simplification

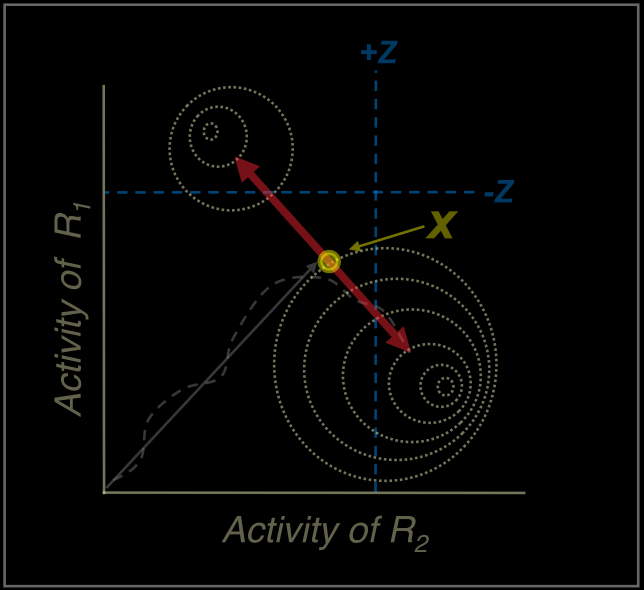
- **Drift Diffusion Model (DDM):**

$$x = A + c$$

A = drift rate
c = noise

$$P(x,t) = N(At, c\sqrt{t})$$

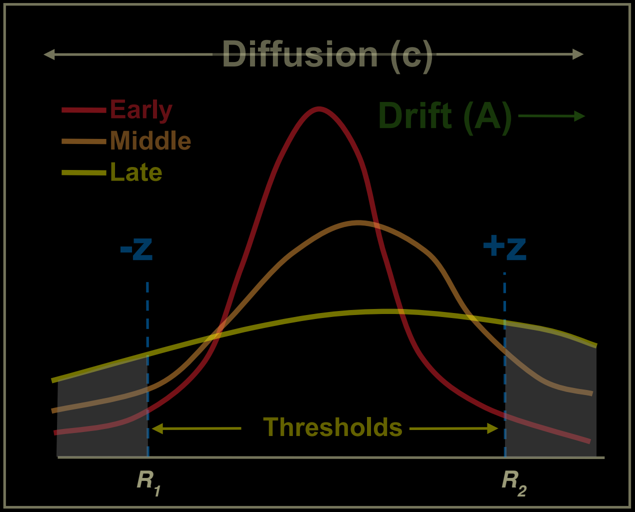
Process ends when x exceeds ± z



- **Can analytically solve for Error Rate and Decision Time:**

$$\text{Error Rate (ER)} = \frac{1}{1 + e^{\frac{2 \cdot \text{Drift} \cdot \text{Threshold}}{\text{Noise}^2}}}$$

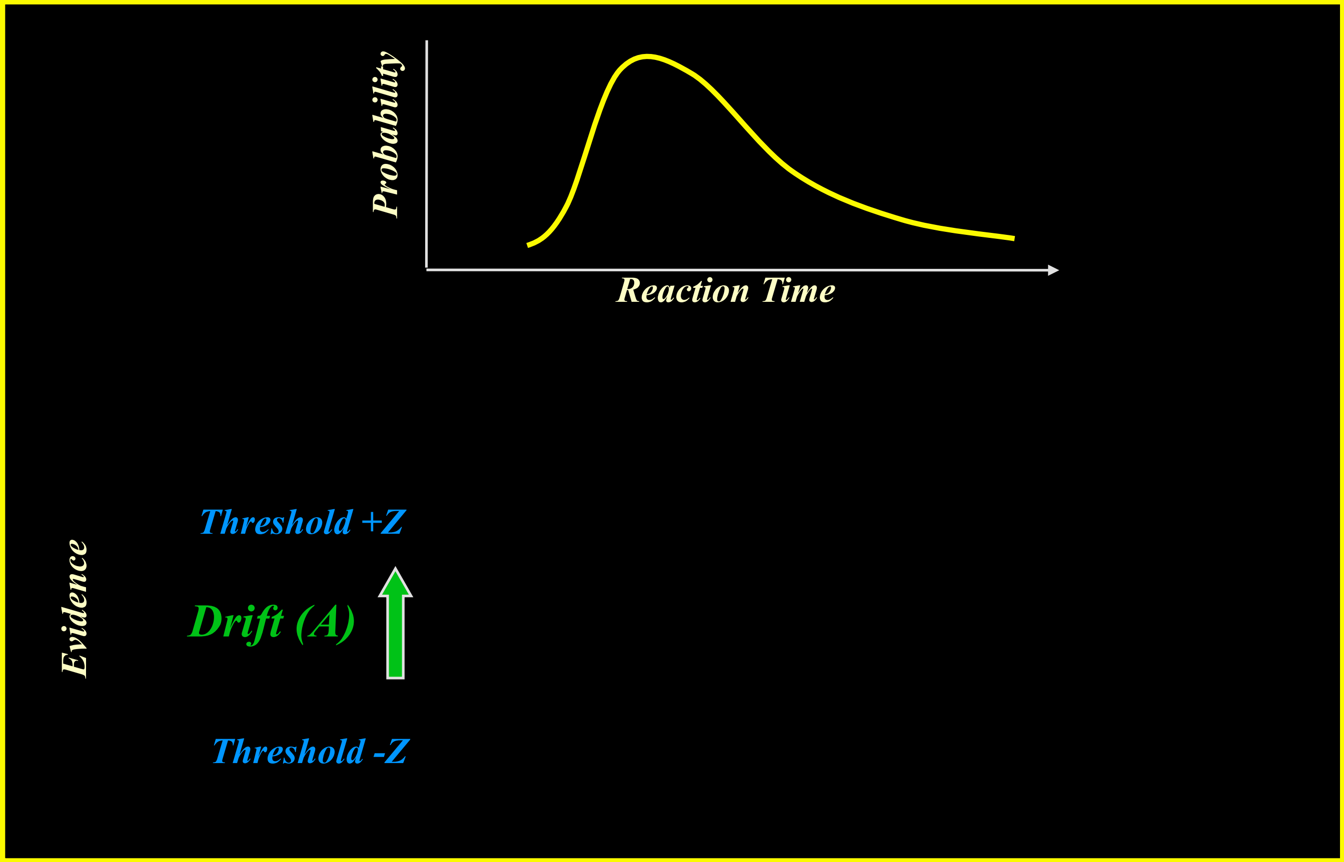
$$\text{Decision Time (DT)} = \frac{\text{Threshold}}{\text{Drift}} \tanh\left(\frac{\text{Drift} \cdot \text{Threshold}}{\text{Noise}^2}\right)$$



Drift Diffusion Model

- **Accurately describes reaction time distributions and error rates in simple decision making tasks**

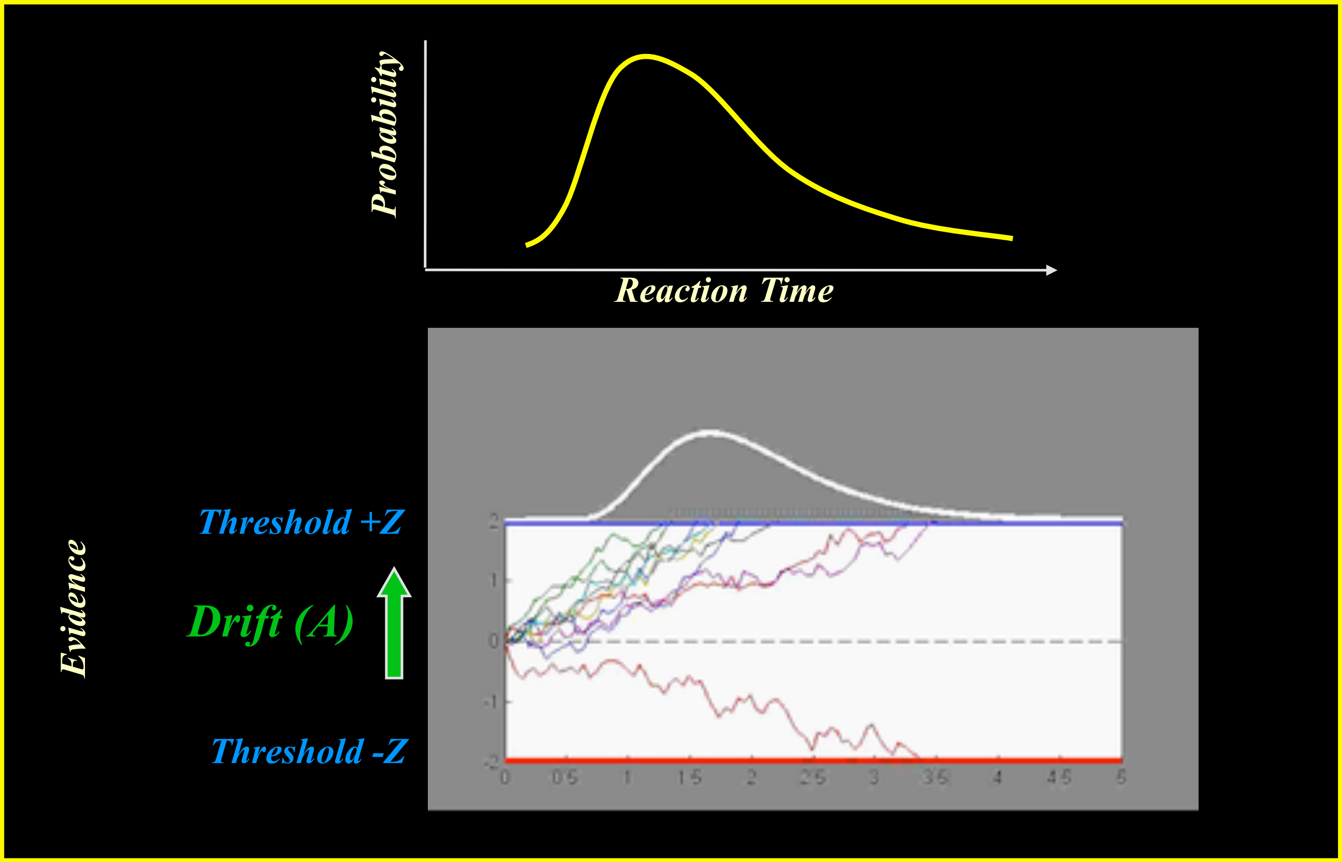
(Ratcliff, 1978, 1999)



Drift Diffusion Model

- **Accurately describes reaction time distributions and error rates in simple decision making tasks**

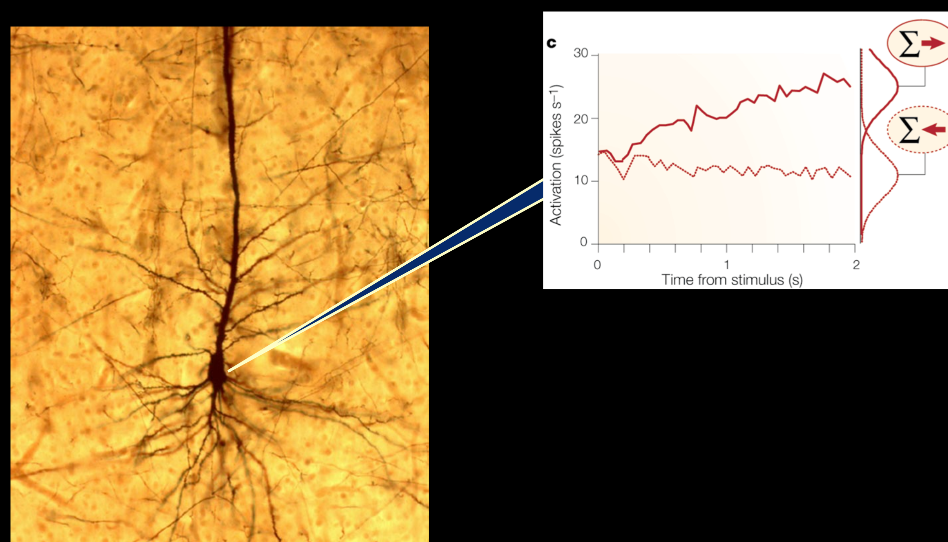
  (Ratcliff, 1978, 1999)



Drift Diffusion Model

- **Accurately describes reaction time distributions and error rates in simple decision making tasks**
  (Ratcliff, 1978, 1999)

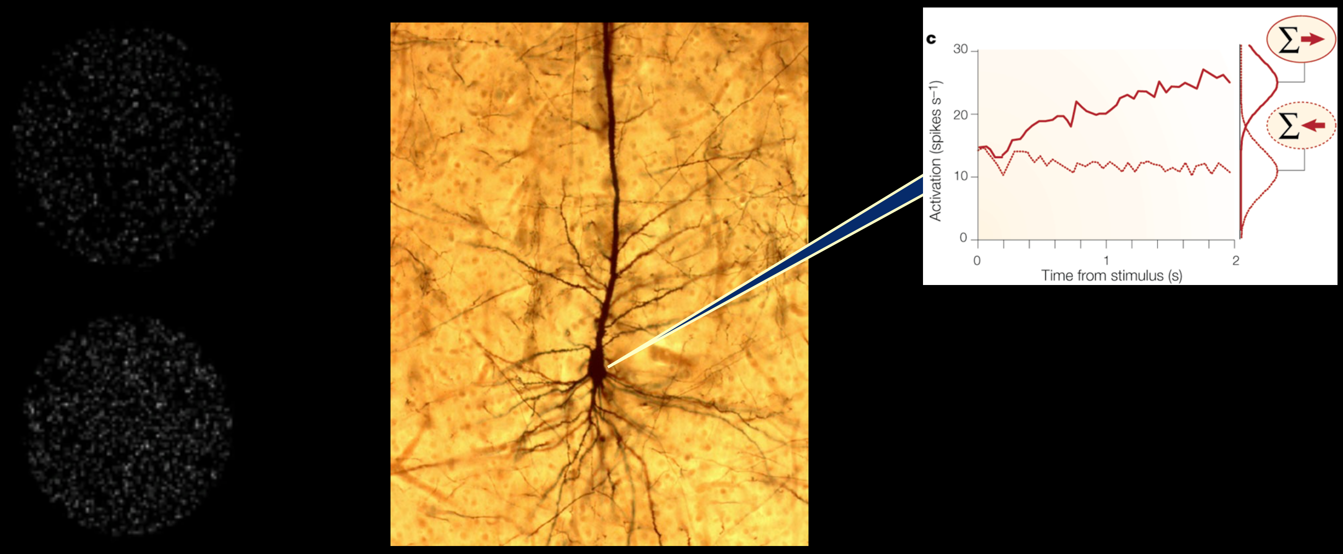
- **Accurately describes dynamics of firing among stimulus- and response-selective neurons in simple decision making tasks**
  (Schall, 1994; Shadlen & Newsome, 1998; Gold & Schadlen, 2002)



Drift Diffusion Model

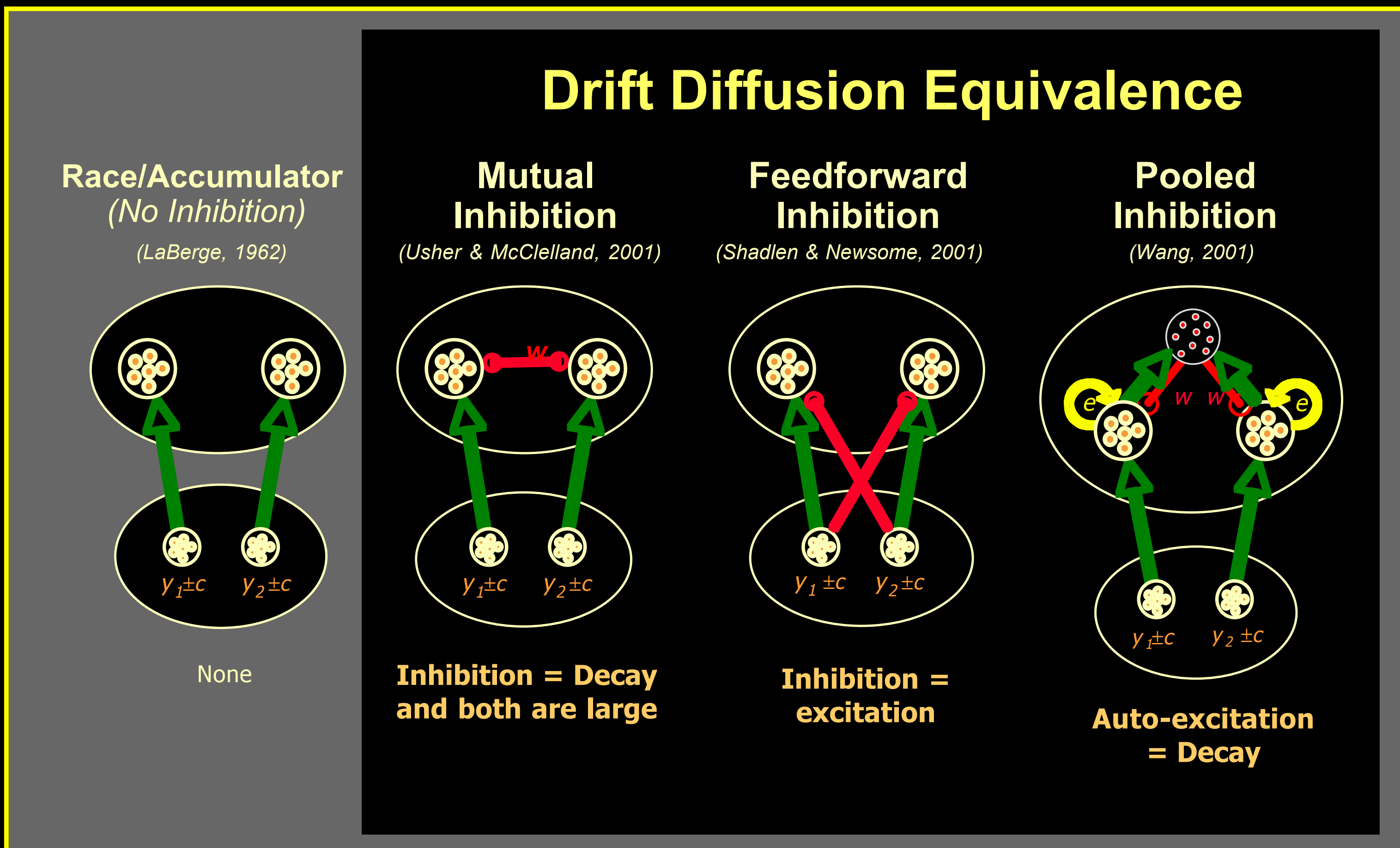
- Accurately describes reaction time distributions and error rates in simple decision making tasks
  (Ratcliff, 1978, 1999)

- Accurately describes dynamics of firing among stimulus- and response-selective neurons in simple decision making tasks
  (Schall, 1994; Shadlen & Newsome, 1998; Gold & Schadlen, 2002)



Theoretical Traction

- **Formal reduction of neural network models:**
  (Bogacz et al., 2006)



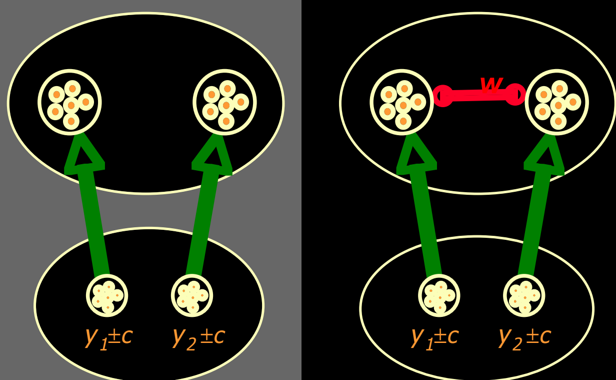
Theoretical Traction

- **Formal reduction of neural network models:**
  (Bogacz et al., 2006)

Drift Diffusion Equivalence

Race/Accumulator
(No Inhibition)
(LaBerge, 1962)

Mutual Inhibition
(Usher & McClelland, 2001)



None

Inhibition = Decay and both are large

Theoretical Traction

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)
  - ***Fastest*** **to reach a decision for given threshold and error rate and** ***most accurate*** **for a given decision time**
    (Wald, 1948; Turing [Good, 1979]; Rouder, 1996)

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)
  - ***Fastest*** **to reach a decision for given threshold and error rate and** ***most accurate*** **for a given decision time**
    (Wald, 1948; Turing [Good, 1979]; Rouder, 1996)
  - **Guarantees arbitrarily low error rate as threshold is increased**
    (Bogacz et al., 2006)

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)
  - ***Fastest*** **to reach a decision for given threshold and error rate and** ***most accurate*** **for a given decision time**
    (Wald, 1948; Turing [Good, 1979]; Rouder, 1996)
  - **Guarantees arbitrarily low error rate as threshold is increased**
    (Bogacz et al., 2006)

- **However, presents optimization problem of its own:**

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)
  - ***Fastest*** **to reach a decision for given threshold and error rate and** ***most accurate*** **for a given decision time**
    (Wald, 1948; Turing [Good, 1979]; Rouder, 1996)
  - **Guarantees arbitrarily low error rate as threshold is increased**
    (Bogacz et al., 2006)

- **However, presents optimization problem of its own:**
  - **How to set parameters (e.g., threshold and starting point)?**

Theoretical Traction

- **Formal reduction of neural network models**
  (Bogacz et al., 2006)

- **Optimal decision making process**
  - **Continuous** (NLDS) **analog of the sequential probability ratio test** (SPRT)
    (used by Turing to crack German Enigma code in WWII)
  - ***Fastest*** **to reach a decision for given threshold and error rate and** ***most accurate*** **for a given decision time**
    (Wald, 1948; Turing [Good, 1979]; Rouder, 1996)
  - **Guarantees arbitrarily low error rate as threshold is increased**
    (Bogacz et al., 2006)

- **However, presents optimization problem of its own:**
  - **How to set parameters (e.g., threshold and starting point)?**
  - **Here is where** ***control*** **comes in…**

DDM and Control

DDM and Control

- DDM specifies psychologically relevant ***control parameters***

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)
  - Drift rate: signal strength / *attention*

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)
  - Drift rate: signal strength / *attention*
  - Threshold: *speed-accuracy trade-off*

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)
  - Drift rate: signal strength / *attention*
  - Threshold: *speed-accuracy trade-off*
- **Empirical question:**
  - Do people in fact adjust these parameters to optimize performance?

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)
  - Drift rate: signal strength / *attention*
  - Threshold: *speed-accuracy trade-off*
- **Empirical question:**
  - Do people in fact adjust these parameters to optimize performance?
  - Analyze DDM to determine optimal parameters under various experimental conditions
    → generate testable predictions
    (for example, what is the optimal threshold, and do people use this?)

DDM and Control

- DDM specifies psychologically relevant *control parameters*
  - Starting point: *expectations* (priors)
  - Drift rate: signal strength / *attention*
  - Threshold: *speed-accuracy trade-off*
- Empirical question:
  - Do people in fact adjust these parameters to optimize performance?
  - Analyze DDM to determine optimal parameters under various experimental conditions
    → generate testable predictions
    (for example, what is the optimal threshold, and do people use this?)
  - First, however, must define *"objective function"*
    → the function that control seeks to optimize