Optimization and Control of Decision Making

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)?**
  - **Can think of this in terms of optimization** ***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

Reward Rate Optimization

$$\frac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$$

Reward Rate Optimization

- **Reward rate (RR):** $\dfrac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$

Reward Rate Optimization

- Reward rate (RR): $\dfrac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$

- Re-express RT and ER in terms of DDM parameters:

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

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

Reward Rate Optimization

- Reward rate (RR): $\dfrac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$

- Re-express RT and ER in terms of DDM parameters:

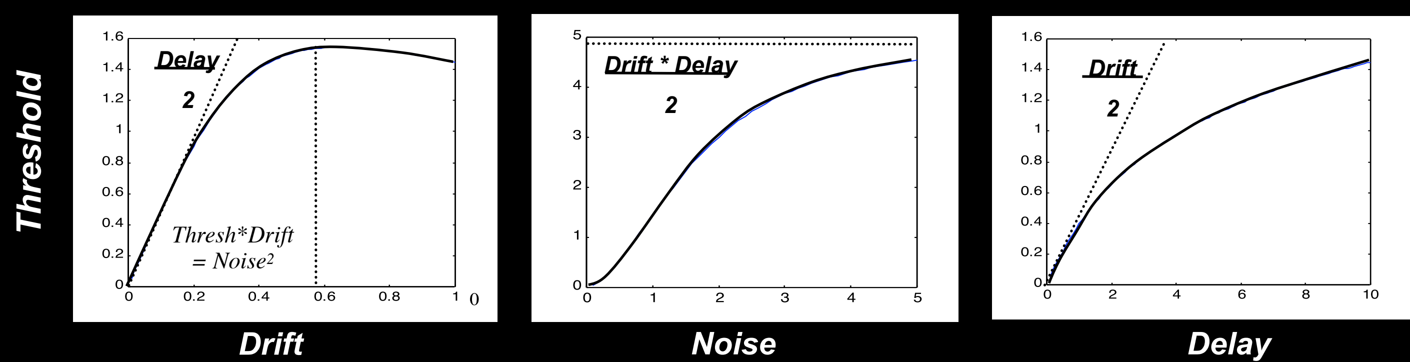
$$\text{Error Rate (ER)} = \frac{1}{1 + e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}}$$

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

$$\text{RR} = \frac{\text{Drift}}{\text{Threshold}} - \text{Delay} - \left(\text{Delay} - \frac{\text{Drift}}{\text{Threshold}}\right) e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}$$

Reward Rate Optimization

- Reward rate (RR): $\dfrac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$
- Re-express RT and ER in terms of DDM parameters:
- **Solve for threshold that maximizes RR:**



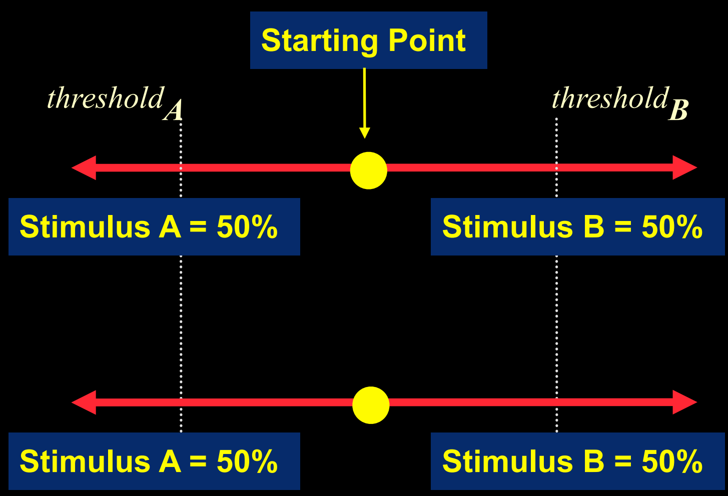
(Bogacz et al., 2006)

Reward Rate Optimization

- Reward rate (RR): $\dfrac{\text{1-Error Rate}}{\text{Reaction Time + Delay}}$
- Re-express RT and ER in terms of DDM parameters:
- Solve for threshold that maximizes RR:
- Predict changes in speed-accuracy tradeoff (threshold) that *optimize RR* as a function of task parameters: delay, drift, and noise

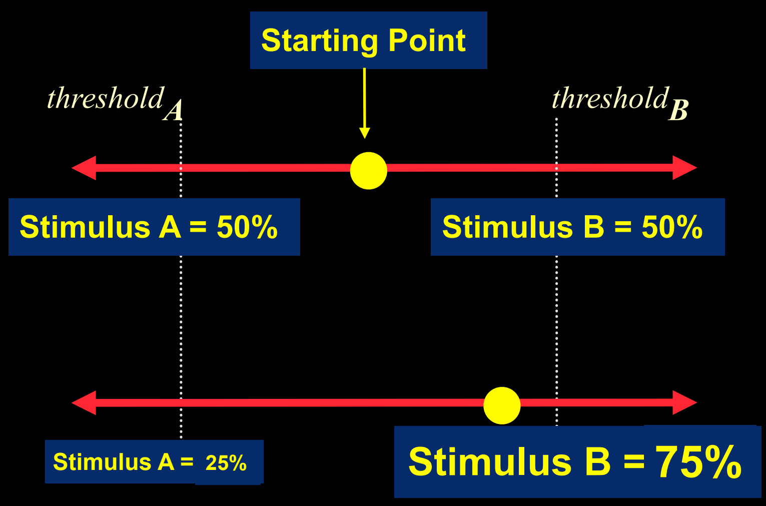
Predictions: Effects of Stimulus Frequency

- If one stimulus is more frequent than the other:
  - it is optimal to move the starting point (“prior”), *not* the threshold *(assuming constant drift [SNR]):*

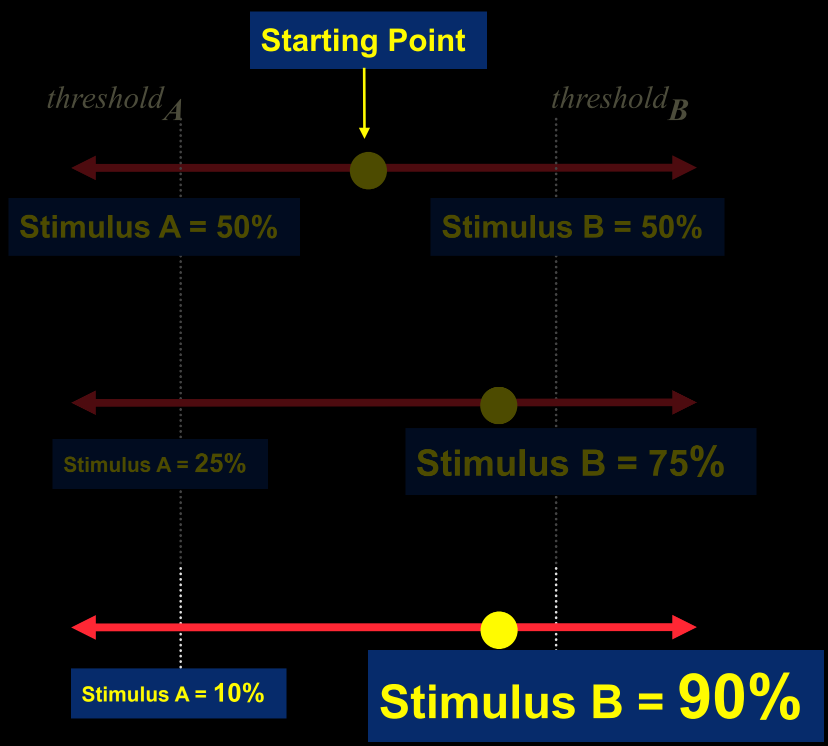


Predictions: Effects of Stimulus Frequency

- **If one stimulus is more frequent than the other:**
  - **it is optimal to move the starting point ("prior"), *not* the threshold** *(assuming constant drift [SNR]):*

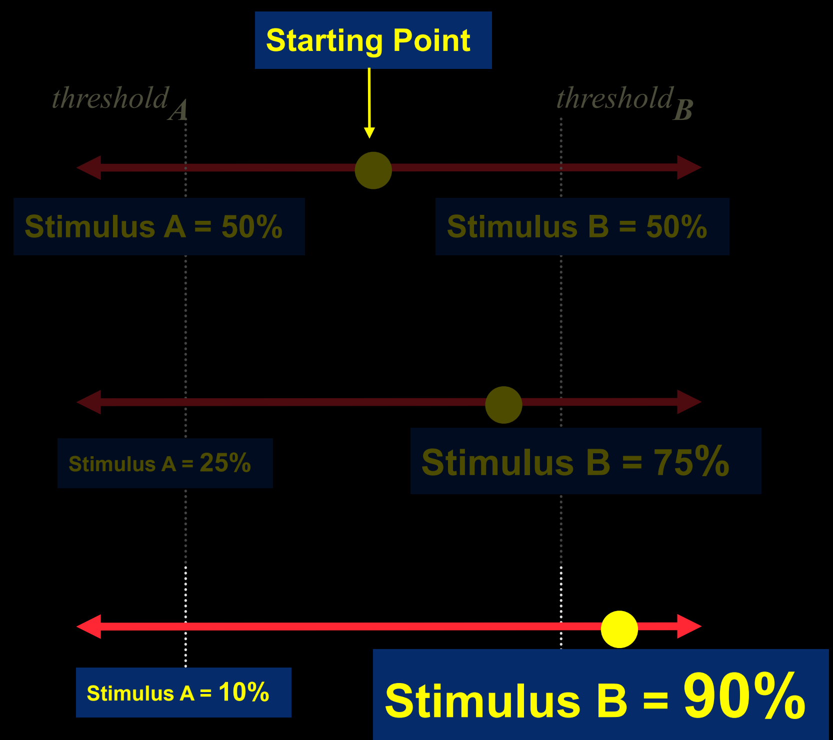


Predictions: Effects of Stimulus Frequency



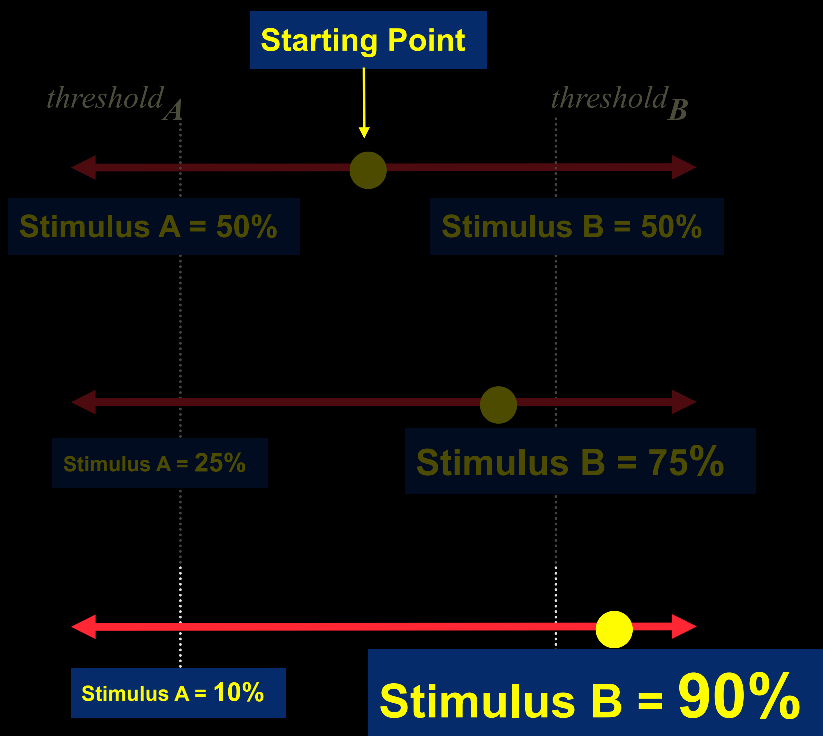
Predictions: Effects of Stimulus Frequency

- For sufficiently extreme frequencies:
  - the optimal starting point exceeds the optimal threshold



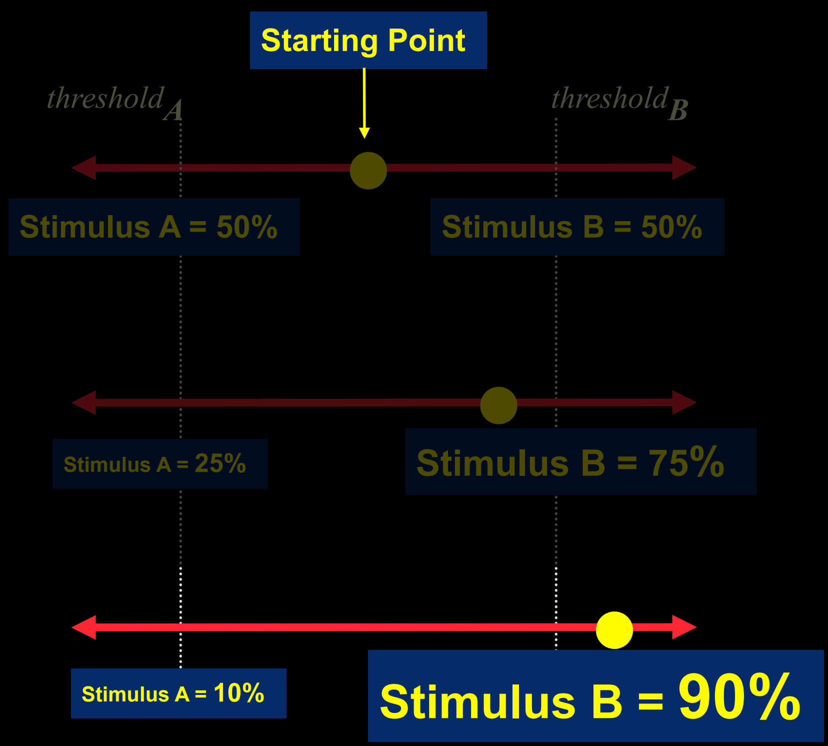
Predictions: Effects of Stimulus Frequency

- For sufficiently extreme frequencies:
  - the optimal starting point exceeds the optimal threshold
  - the model predicts a switch from integration (decision making)



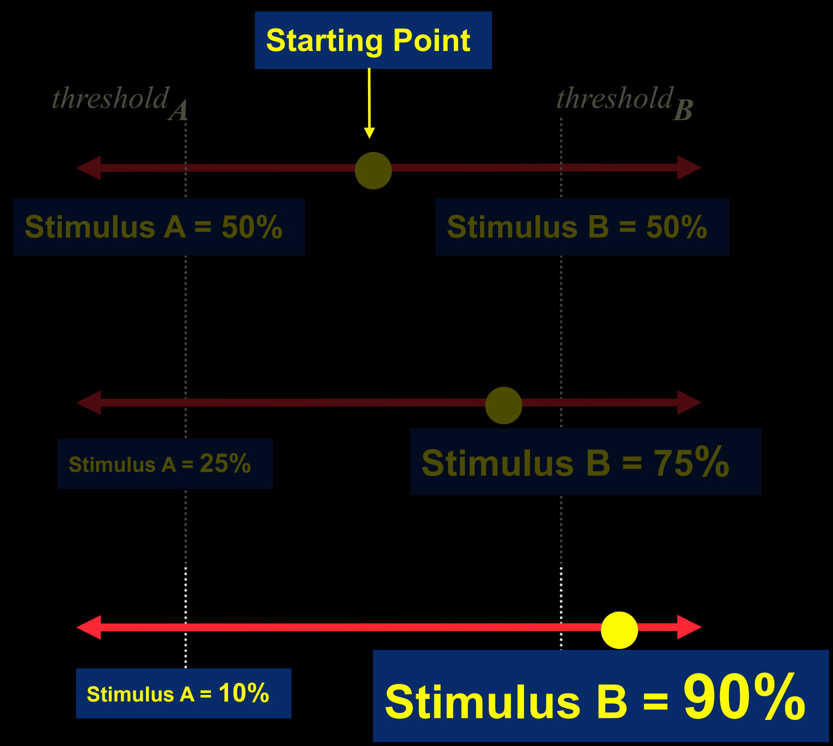
Predictions: Effects of Stimulus Frequency

- For sufficiently extreme frequencies:
  - the optimal starting point exceeds the optimal threshold
  - the model predicts a switch from integration (decision making) to simple signal detection (stereotyped responding)



Predictions: Effects of Stimulus Frequency

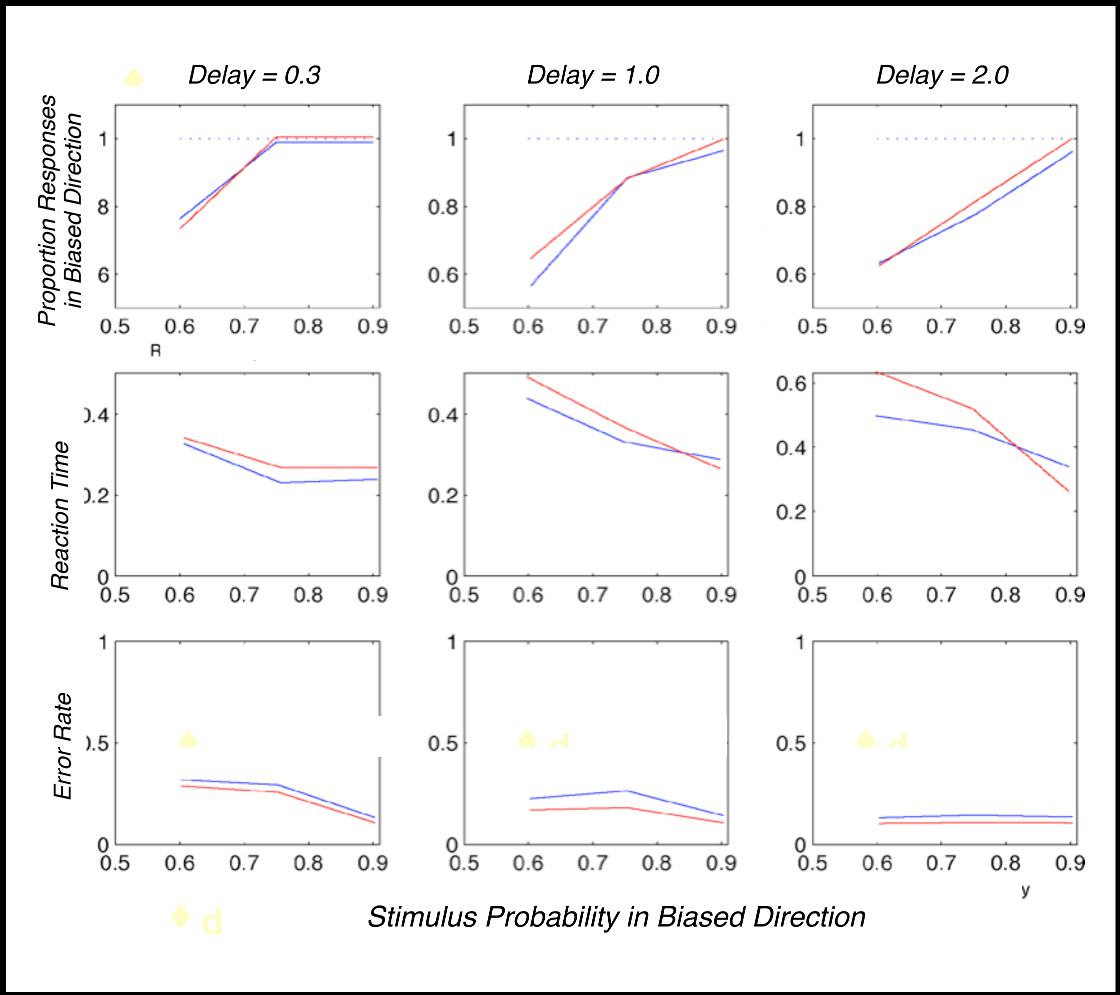
- **For sufficiently extreme frequencies:**
  - the optimal starting point exceeds the optimal threshold
  - the model predicts a switch from integration (decision making) to simple signal detection (stereotyped responding)
  - **the stimulus frequency at which this occurs varies according to delay and drift…**



Data: Effects of Stimulus Frequency

Empirical Data

(Simen et al., 2009)

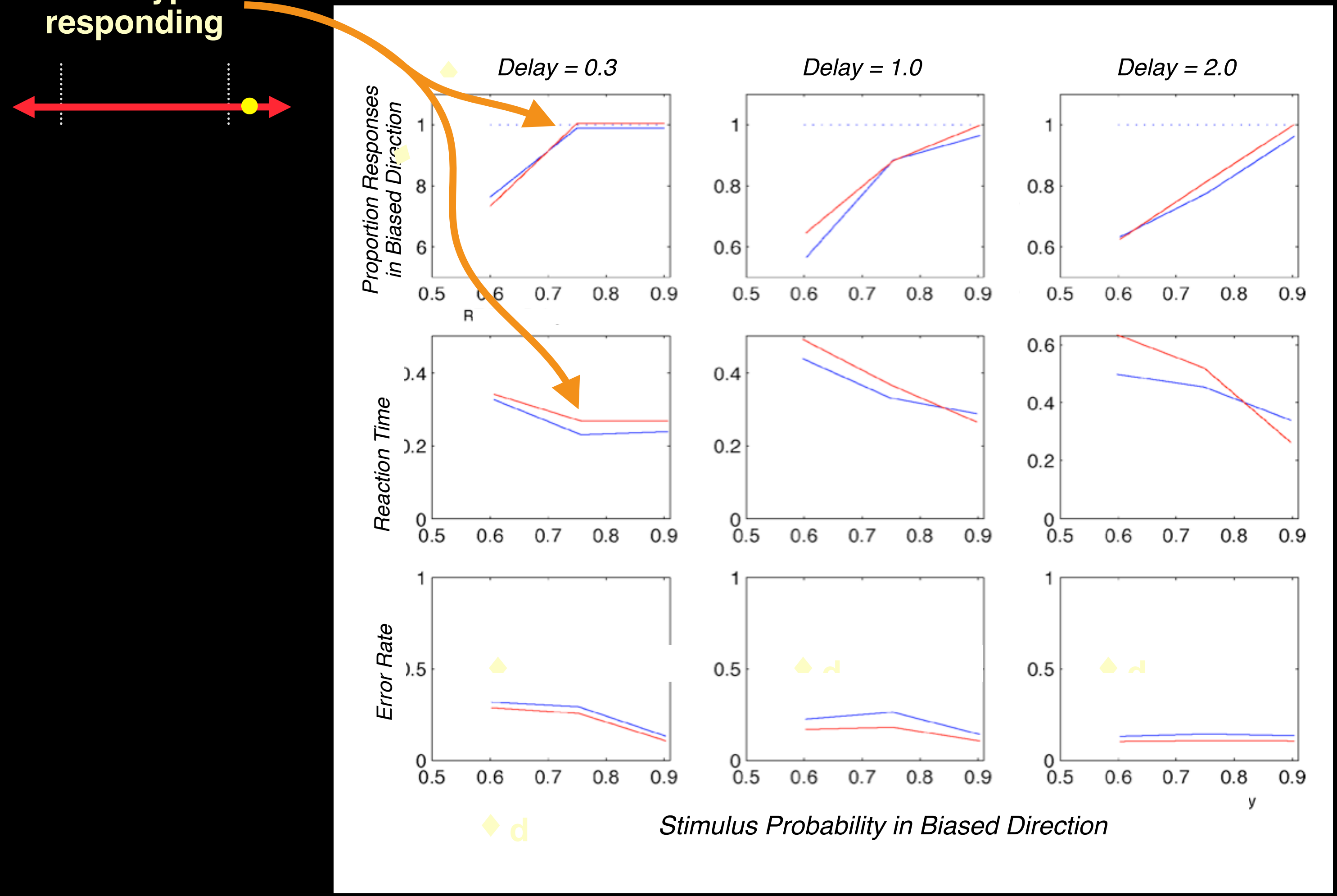


Data: Effects of Stimulus Frequency

Empirical Data

(Simen et al., 2009)

stereotyped responding



Optimal Performance Curve

Optimal Performance Curve

- The fact that there is a ***single optimal threshold*** for a given set of task parameters, means that the DDM equations:

Optimal Performance Curve

- **The fact that there is a *single optimal threshold* for a given set of task parameters, means that the DDM equations:**
  - Error Rate (ER) = $\dfrac{1}{1 + e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}}$
  - Decision Time (DT) = $\dfrac{Threshold}{Drift} \, Tanh\left(\dfrac{Drift \cdot Threshold}{Noise^2}\right)$
  - Reward Rate (RR) = $\dfrac{Drift}{Threshold} - Delay - \left(Delay - \dfrac{Drift}{Threshold}\right) e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}$

Optimal Performance Curve

- The fact that there is a *single optimal threshold* for a given set of task parameters, means that the DDM equations:

  - Error Rate (ER) = $\dfrac{1}{1 + e^{\frac{2 \cdot \text{Drift} \cdot \text{Threshold}}{\text{Noise}^2}}}$
  - Decision Time (DT) = $\dfrac{\text{Threshold}}{\text{Drift}} \, \text{Tanh}\left(\dfrac{\text{Drift} \cdot \text{Threshold}}{\text{Noise}^2}\right)$
  - Reward Rate (RR) = $\dfrac{\text{Drift}}{\text{Threshold}} - \text{Delay} - \left(\text{Delay} - \dfrac{\text{Drift}}{\text{Threshold}}\right) e^{\frac{2 \cdot \text{Drift} \cdot \text{Threshold}}{\text{Noise}^2}}$

  can be solved for DT as a function of ER (under the assumption of optimality):

$$\frac{\text{DT}}{D_{total}} = \left(\frac{1}{\text{ER} \ln \frac{1-\text{ER}}{\text{ER}}} + \frac{1}{1 - 2\text{ER}} \right)^{-1}$$

Optimal Performance Curve

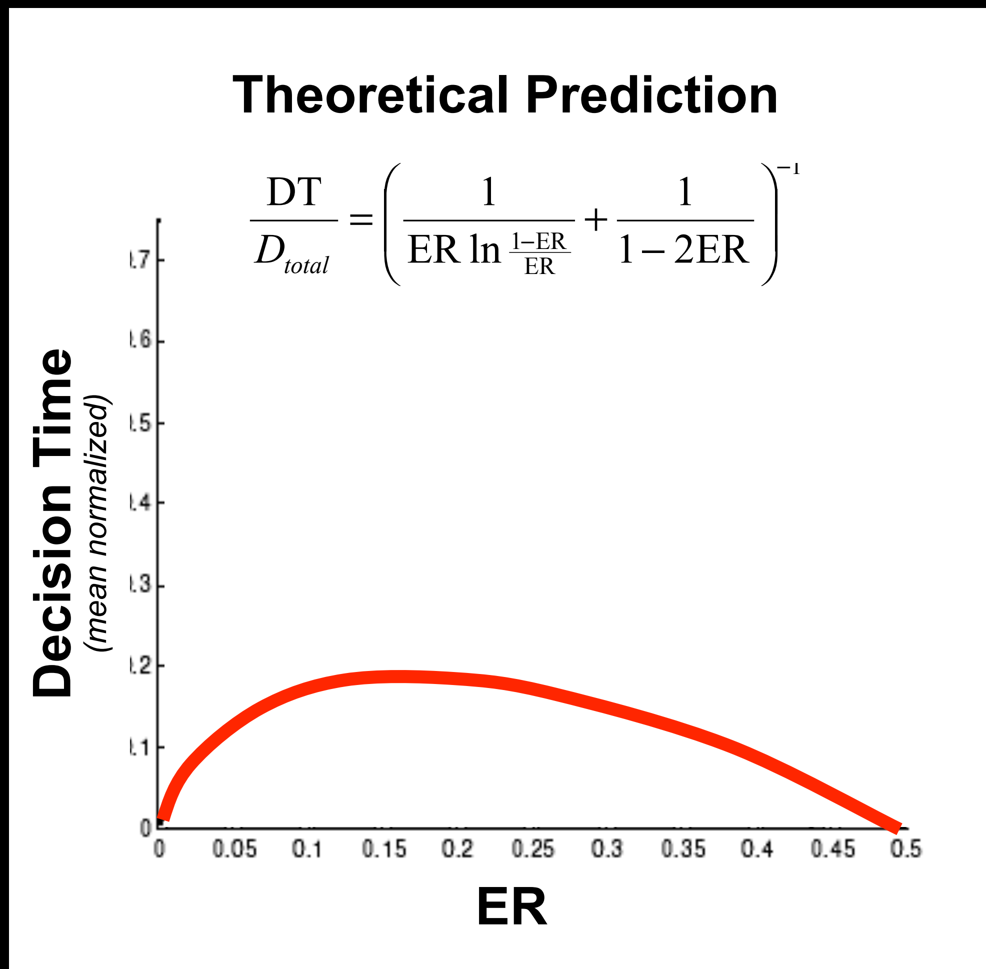
- The fact that there is a *single optimal threshold* for a given set of task parameters, means that the DDM equations:
  - Error Rate (ER) = $\dfrac{1}{1 + e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}}$
  - Decision Time (DT) = $\dfrac{Threshold}{Drift} \operatorname{Tanh}\left(\dfrac{Drift \cdot Threshold}{Noise^2}\right)$
  - Reward Rate (RR) = $\dfrac{Drift}{Threshold} - Delay - \left(Delay - \dfrac{Drift}{Threshold}\right) e^{\frac{2 \cdot Drift \cdot Threshold}{Noise^2}}$

  can be solved for DT as a function of ER (under the assumption of optimality):

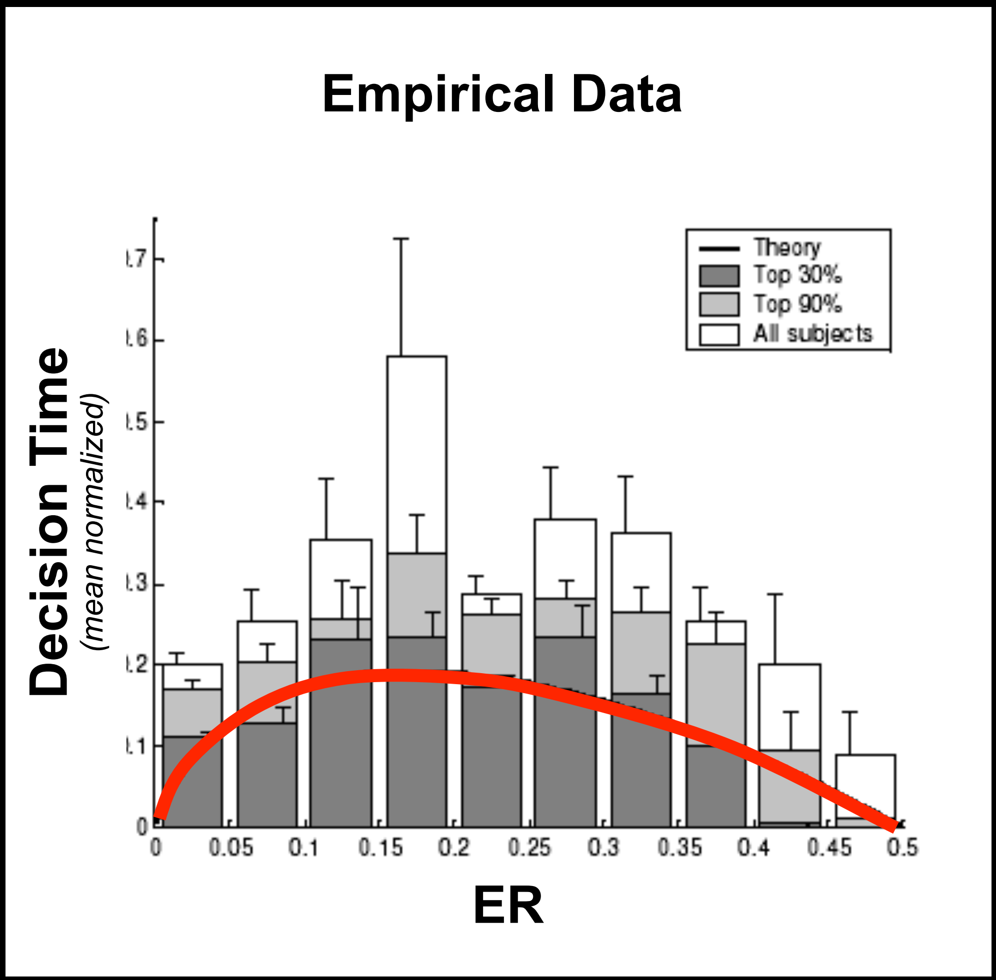
$$\frac{DT}{D_{total}} = \left(\frac{1}{ER \ln \frac{1-ER}{ER}} + \frac{1}{1-2ER} \right)^{-1}$$

- In other words, there is a *single optimal speed-accuracy curve* that should quantitatively define performance…

Optimal Performance Curve

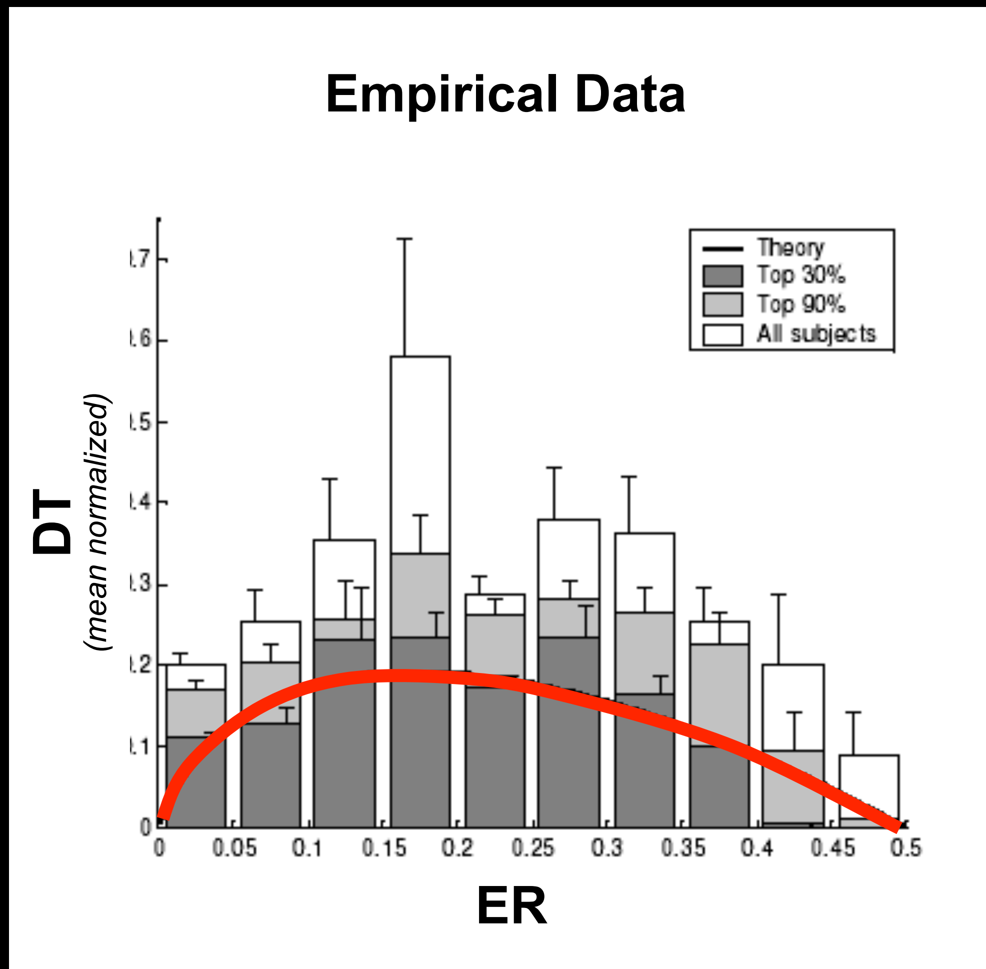


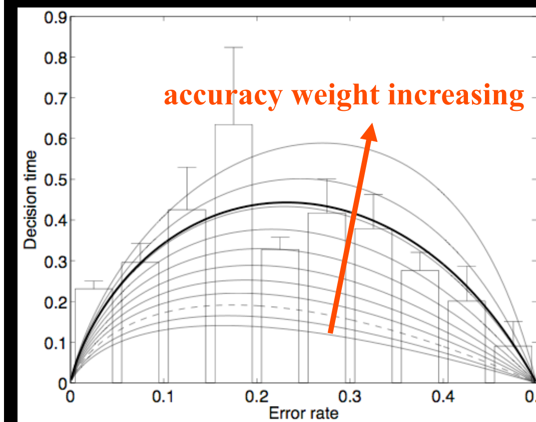
Optimal Performance Curve



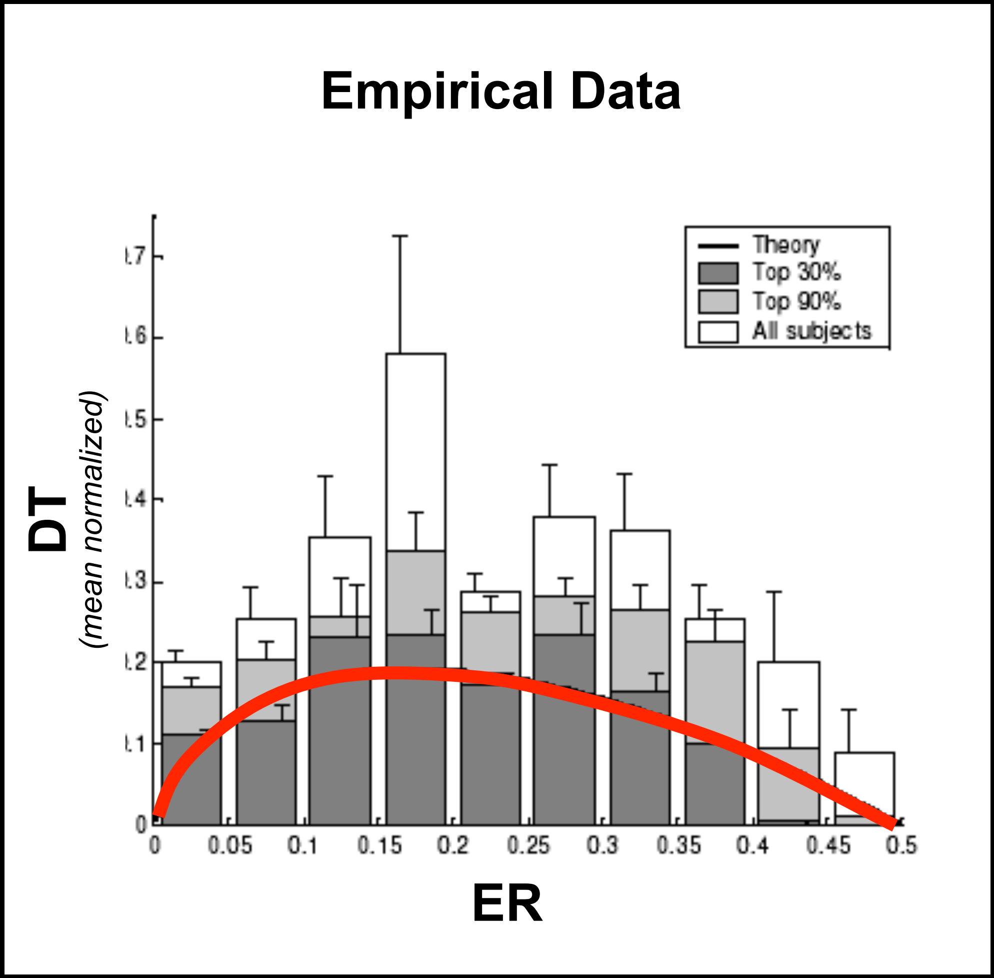
(Bogacz et al., 2006)

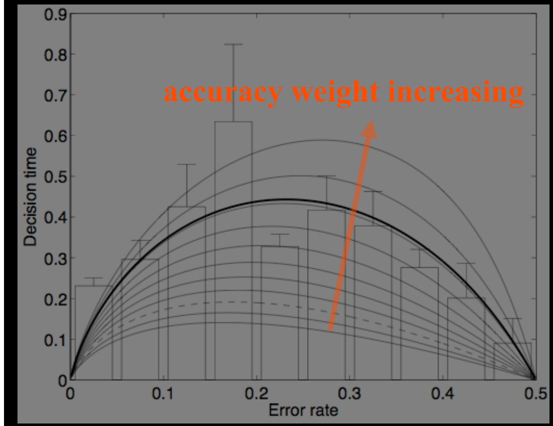
Optimal Performance Curve

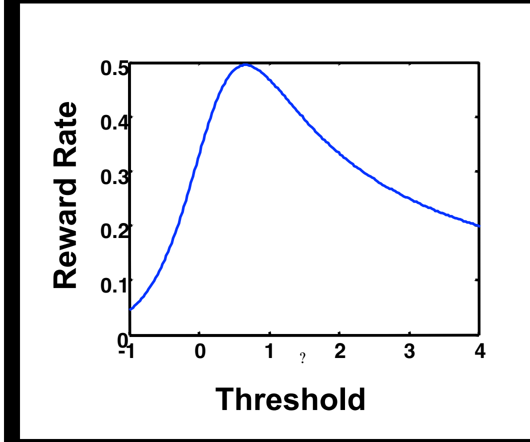




Optimal Performance Curve







Summary

- **Drift Diffusion Model (DDM) can be used to:**

Summary

- **Drift Diffusion Model (DDM) can be used to:**
  - **explain human RT distributions and accuracy in simple decision tasks**

Summary

- **Drift Diffusion Model (DDM) can be used to:**
  - **explain dynamics of neural firing in simple decision tasks**

Summary

- **Drift Diffusion Model (DDM) can be used to:**
  - **formally analyze neural network models of simple decision tasks**

Summary

- **Drift Diffusion Model (DDM) can be used to:**
  - **describe parameters of optimal performance (maximizing reward rate)**

Summary

- **Drift Diffusion Model (DDM) can be used to:**
  - **predict influence of task parameters on speed-accuracy tradeoff that approximate those observed empirically**

Summary

- Drift Diffusion Model (DDM) can be used to:
  - define, in formal terms, the mechanisms underlying decision making in simple two alternative forced choice tasks

Summary

- Drift Diffusion Model (DDM) can be used to:
  - define, in formal terms, variables that are subject to regulation by control mechanisms to optimize outcomes

Summary

- **Drift Diffusion Model (DDM) can be used to:**

  - **be approximated by neurally-plausible mechanism:**
    leaky competitive accumulator (LCA)

Summary

- **Drift Diffusion Model (DDM) can be used to:**

  - **makes contact with abstract, mathematical modeling of psychological function:** *(e.g. normative models based on Bayesian inference)*