

Introduction to the diffusion decision model

Blair R K Shevlin

NYCPW

— NEW YORK COMPUTATIONAL PSYCHIATRY WORKSHOP —

who am I?

Blair Shevlin, PhD

- BA in Psychology [Goucher College]
 - Mentor: Jennifer McCabe
 - Focus: pedagogical uses of mnemonic techniques
- MA in Experimental Psychology [Towson University]
 - Mentor: Kerri Goodwin
 - Focus: beliefs and motivations surrounding distracted driving
- PhD in Decision Psychology [The Ohio State University]
 - Mentors: Ian Krajbich, Roger Ratcliff
 - Focus: economic decisions; sequential sampling models
- Postdoc in Computational Psychiatry [Mount Sinai]
 - Mentors: Laura Berner, Xiaosi Gu
 - Focus: compulsive-use disorders; human voltammetry

Acknowledgments

- Kianté Fernandez (UCLA)
- Ian Krajbich (UCLA)
- Laura Fontanesi (Basel)
- Robert (Bob) Wilson (Arizona, Georgia Tech)

Goals

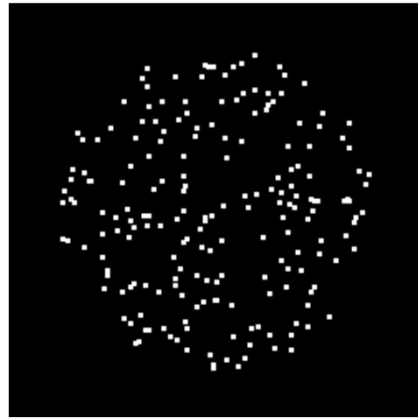
- **Describe** the background behind the development of the diffusion decision model
- **Explain** the intuitions behind the diffusion decision model
- **Describe** the application of the diffusion decision model to value-based choice
- **Demonstrate** the practical implementation of the diffusion decision model

3-2-1 Exercise

For three minutes, list the following about modeling the decision-making process:

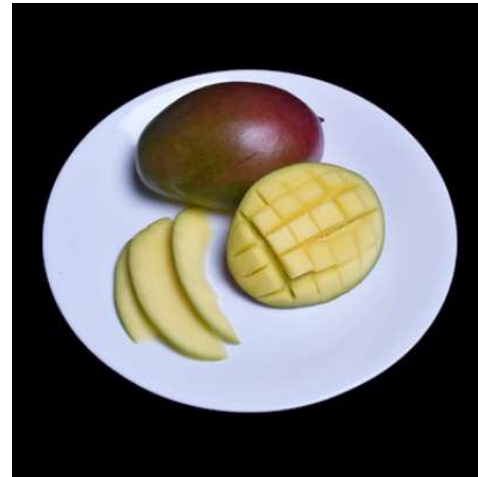
- 3 things you know
- 2 things you would like to know
- 1 question you have

how to understand the decision-making process?



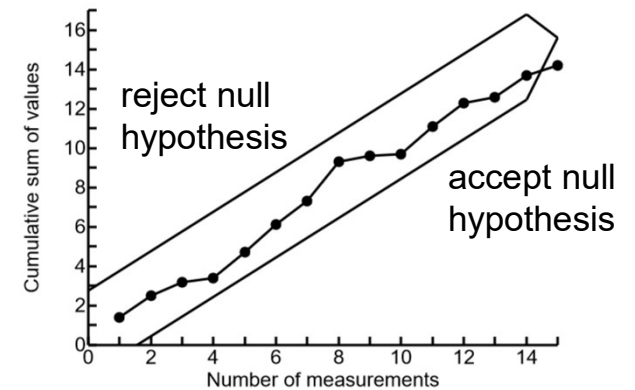
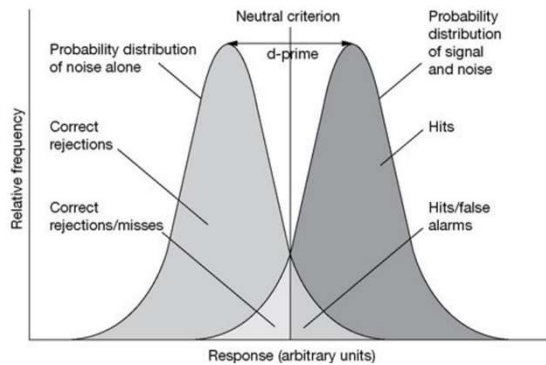
which way are the dots moving?

how to understand the decision-making process?



which food should I choose?

history



Signal detection theory

h1: motion left; **h2:** motion right

$$DV = P(e | h_1) / P(e | h_2)$$

Decision rule: choose after 1* observation based on 1 criterion value β

- $\beta = P(h_2) / P(h_1)$
- Choose h1 when $DV \geq \beta$
- Choose h2 when $DV < \beta$

Sequential probability ratio test

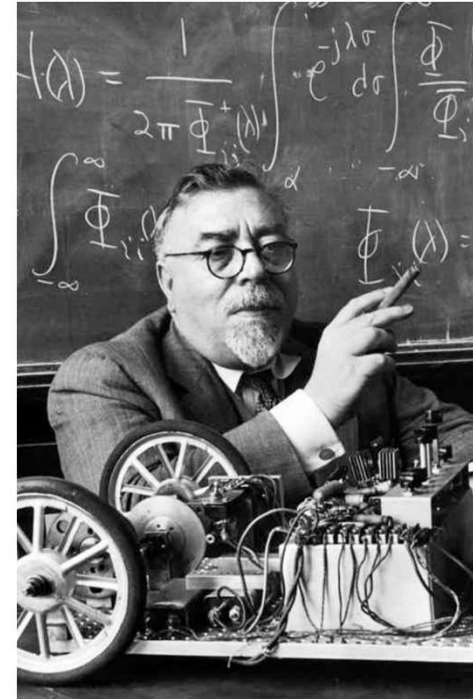
$$DV = \log LR_{12} \equiv \log \frac{P(e_1, e_2, \dots, e_n | h_1)}{P(e_1, e_2, \dots, e_n | h_2)}$$

Decision rule: continue sampling until DV hits either of 2 criterion values $\beta, -\beta$

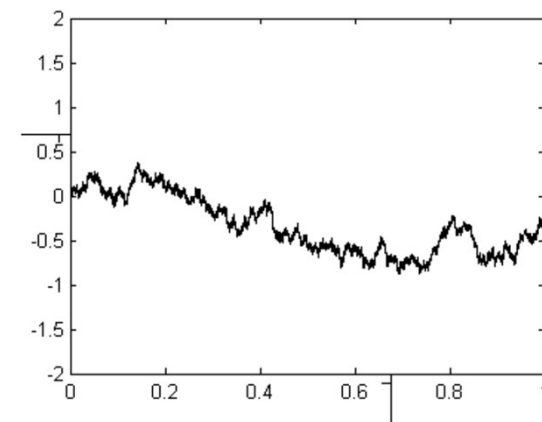
- $\beta = |-\beta|$
- Choose h1 when $DV \geq \beta$
- Choose h2 when $DV \leq -\beta$

wiener diffusion process (1-D Brownian motion)

- mathematics, physics, evolutionary biology, economics, finance
- **properties**
 - starts at $x = 0$
 - at each step, x changes by a **Gaussian** increment $N(0, \sigma)$
 - each increment is **independent**
 - process is **continuous in time**



Norbert Wiener (1894 - 1964)



the diffusion decision model (DDM)

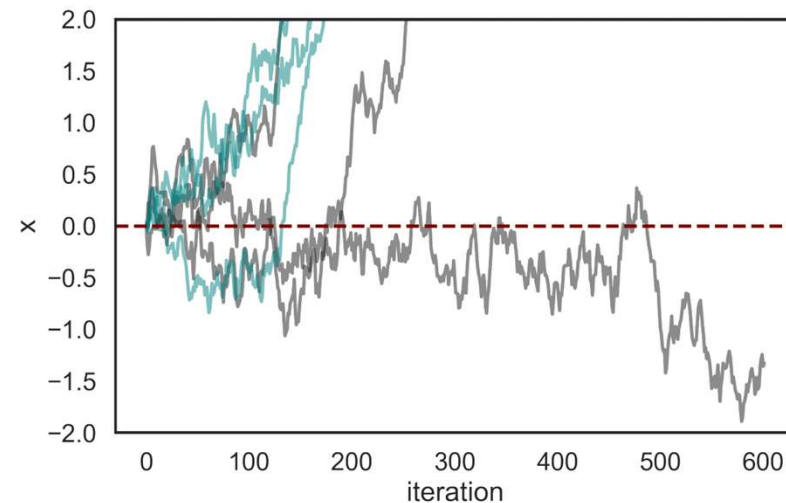
- process terminates at fixed **thresholds**
- process **drifts** towards positive or negative values
- process can be **biased** *a priori*
- constant **non-decision time** is added to total time to threshold
- parameters can **vary across trials**

Psychological Review

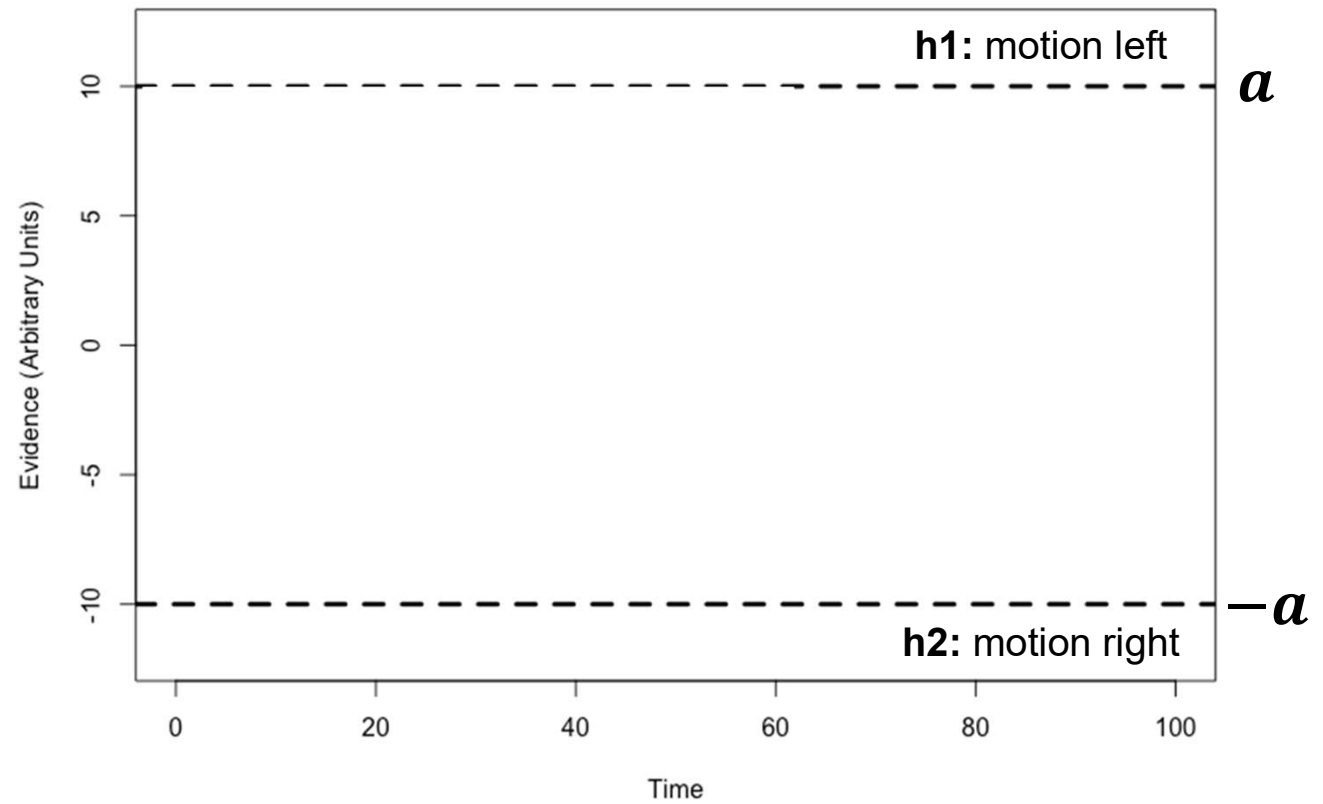
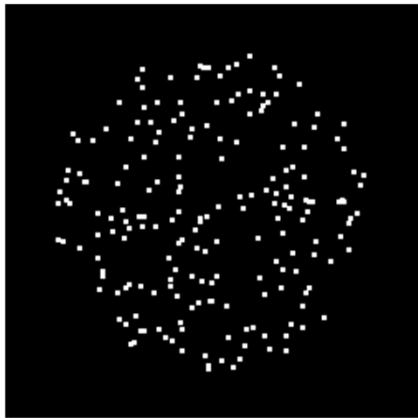
VOLUME 85 NUMBER 2 MARCH 1978

A Theory of Memory Retrieval

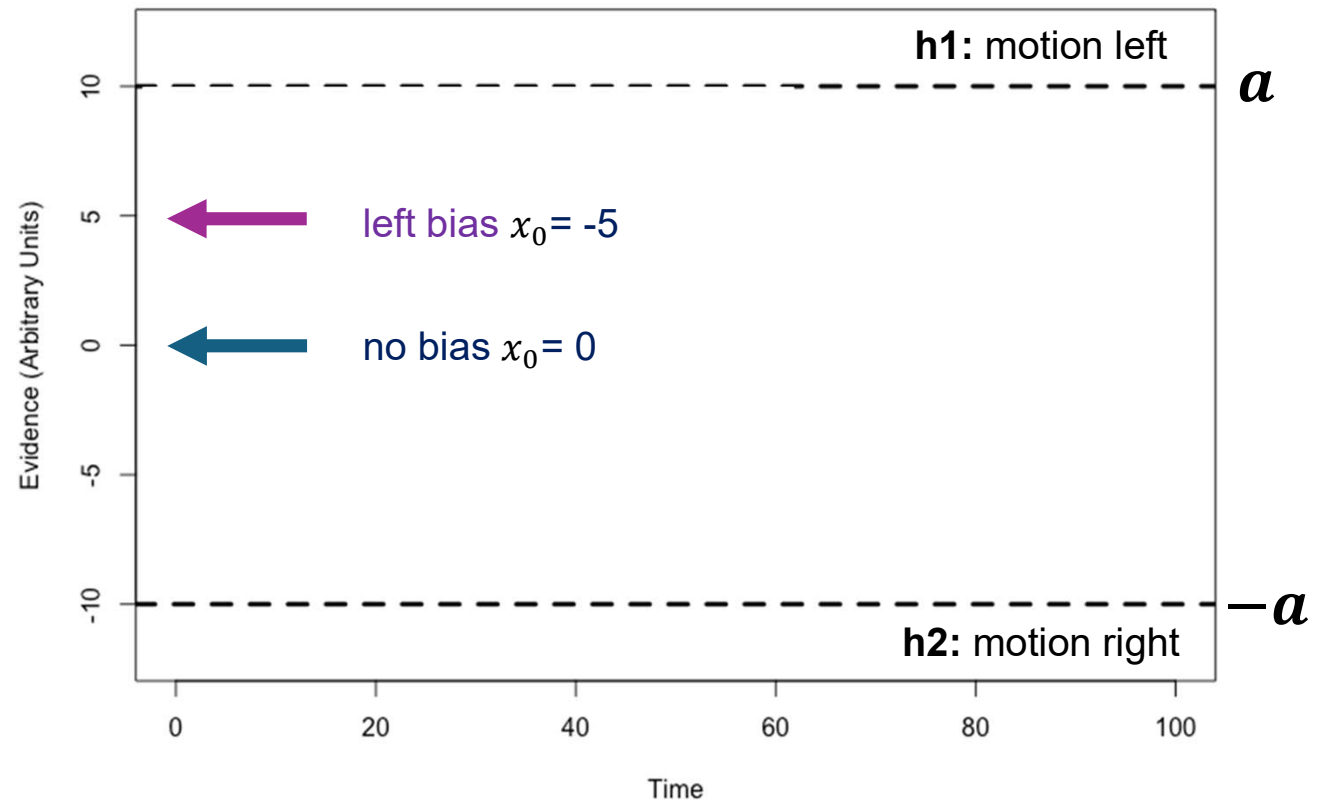
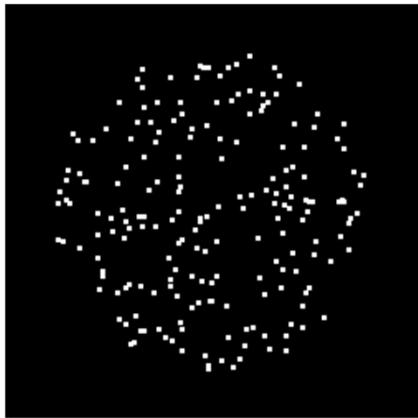
Roger Ratcliff
University of Toronto, Ontario, Canada



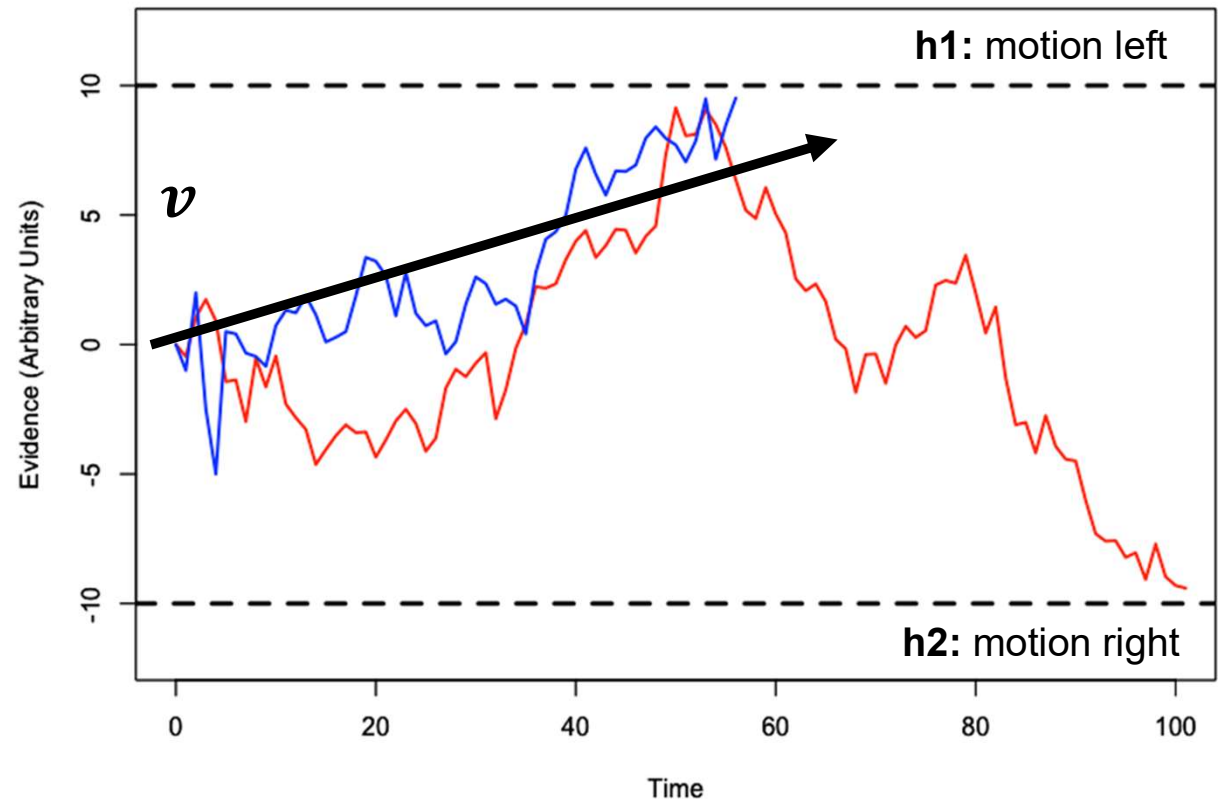
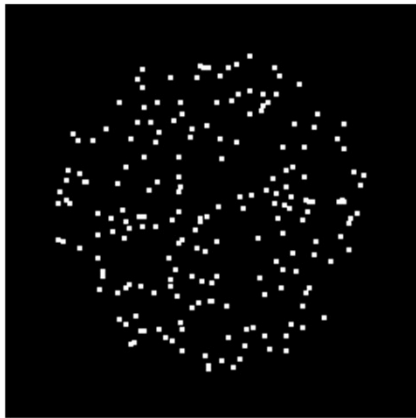
decision thresholds



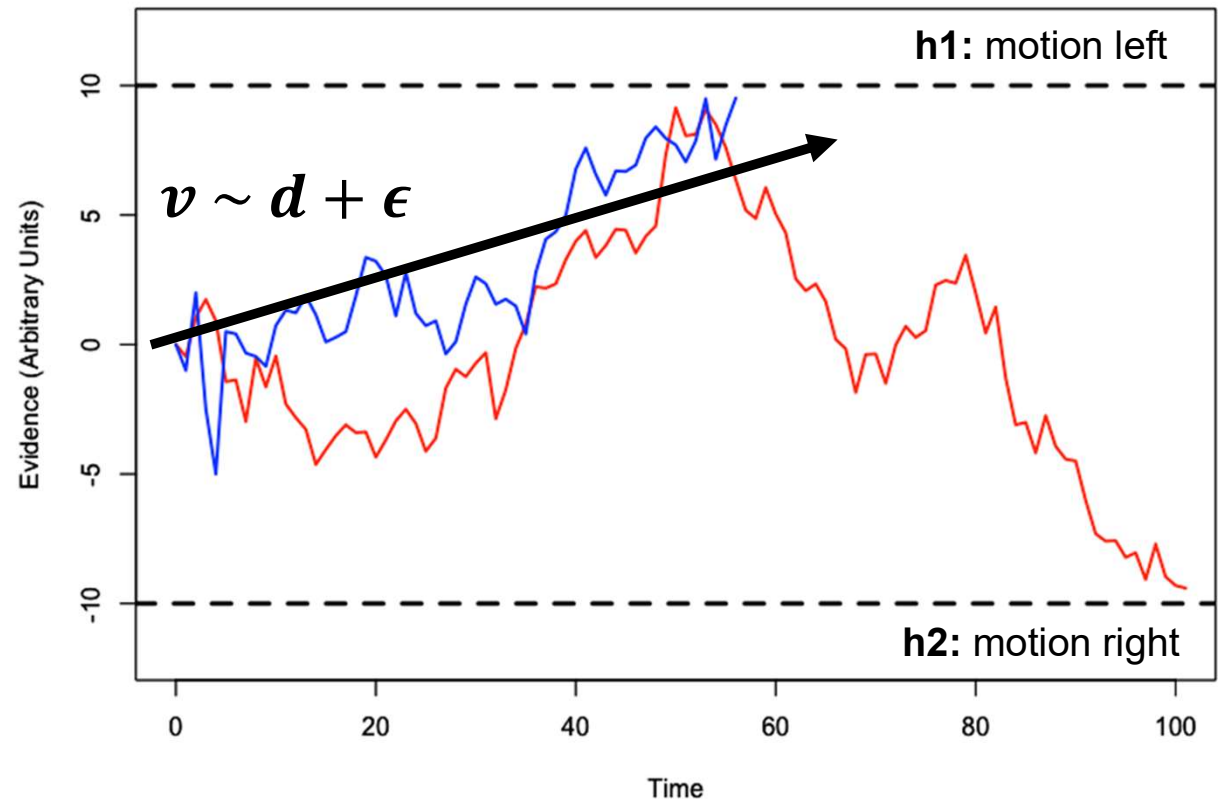
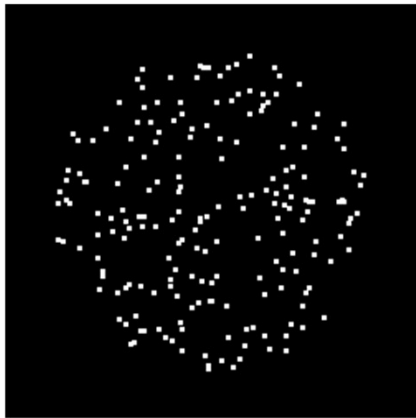
starting point



drift rate

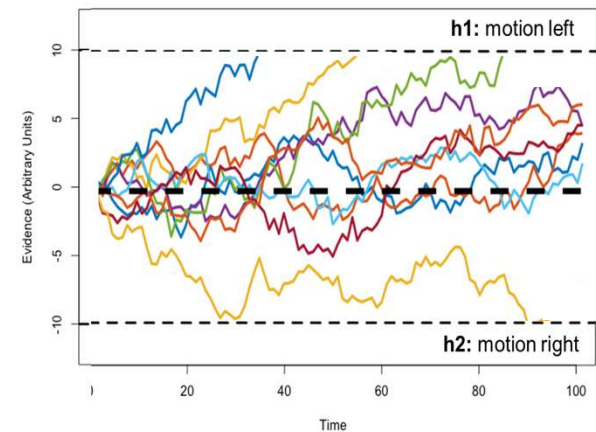


drift rate



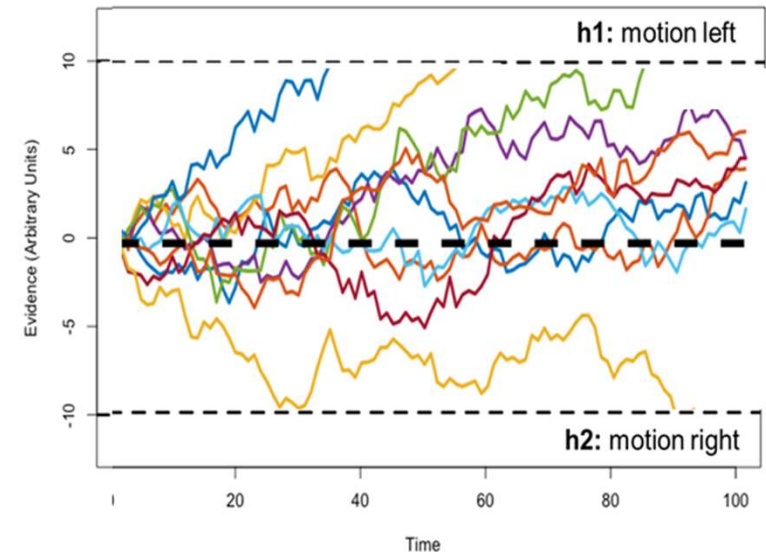
evidence contains both signal and noise

- **signal**: on average, evidence points in the correct direction
- **noise**: randomness in the stimulus and the brain
- Over time, the accumulated evidence **drifts** (signal) and **diffuses** (noise) in the correct direction
- Because noise is random, every trial has a different trajectory

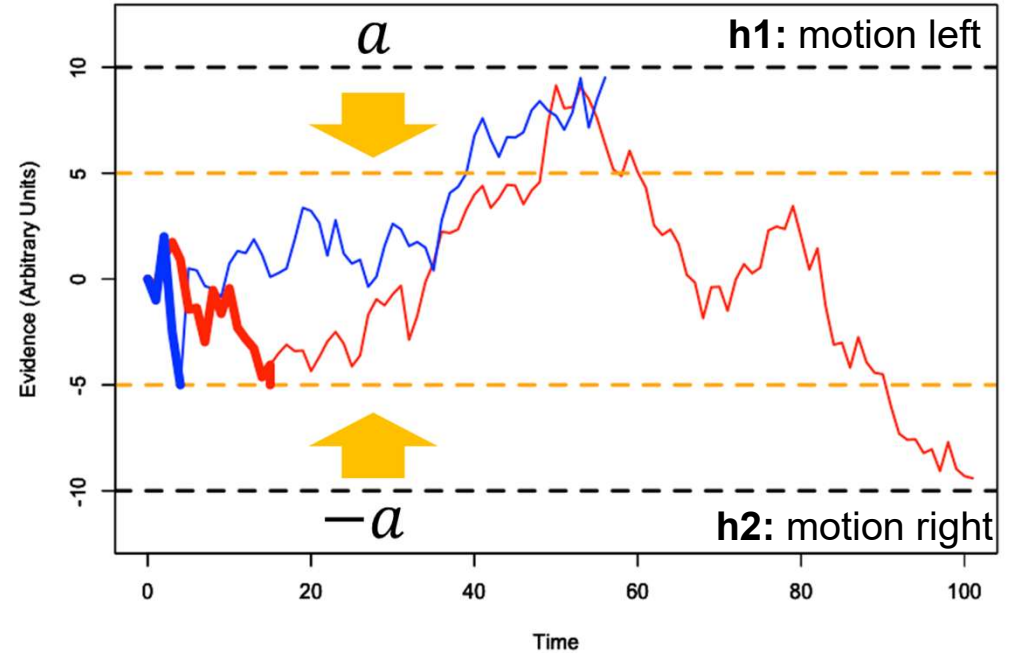
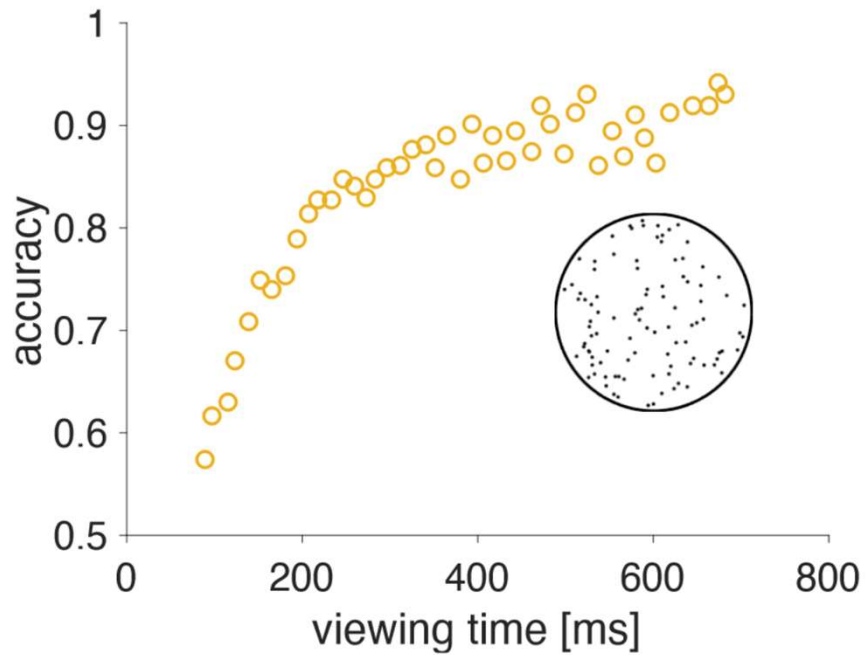


evidence contains both signal and noise

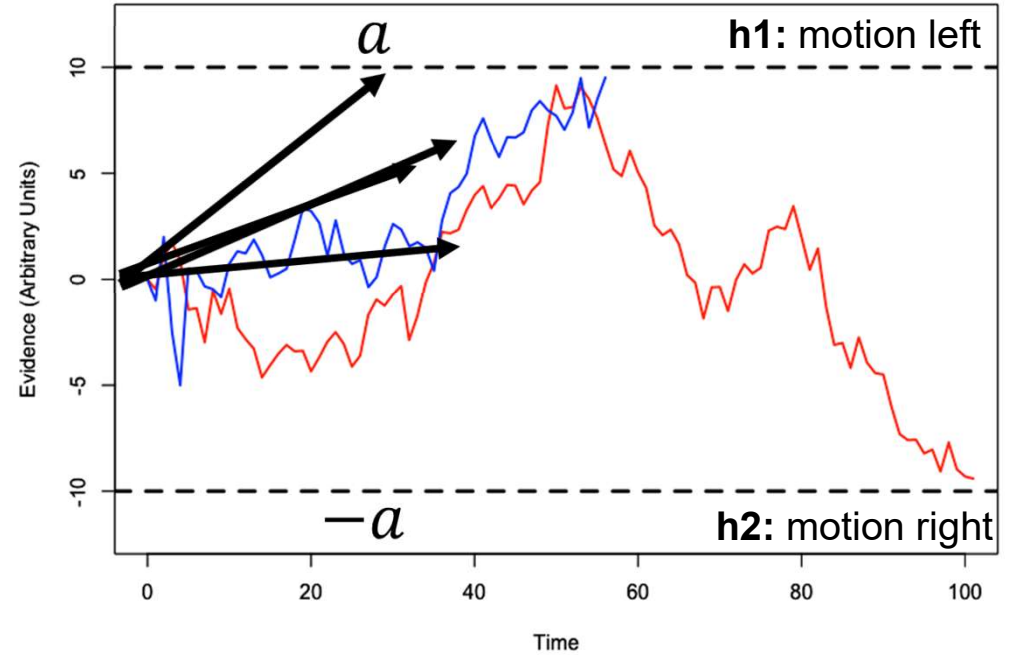
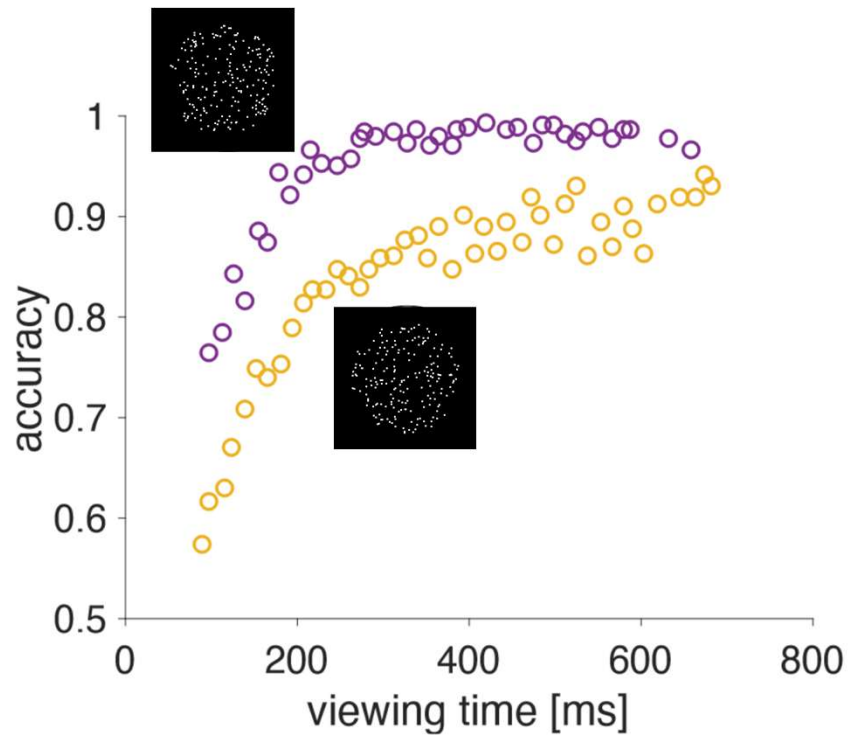
- **signal**: on average, evidence points in the correct direction
- **noise**: randomness in the stimulus and the brain
- Over time, the accumulated evidence **drifts** (signal) and **diffuses** (noise) in the correct direction
- Because noise is random, every trial has a different trajectory



decreasing viewing time decreases accuracy

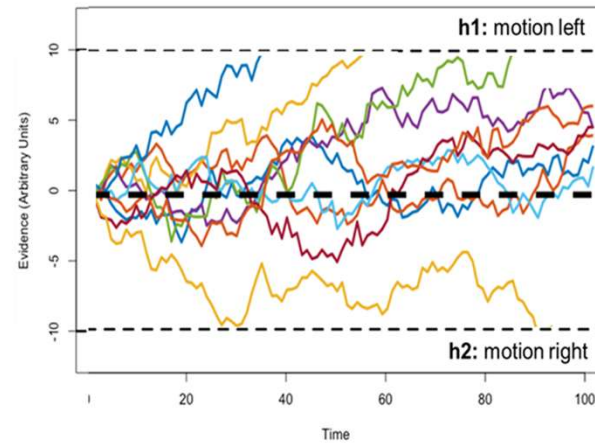


increasing drift rate increases accuracy



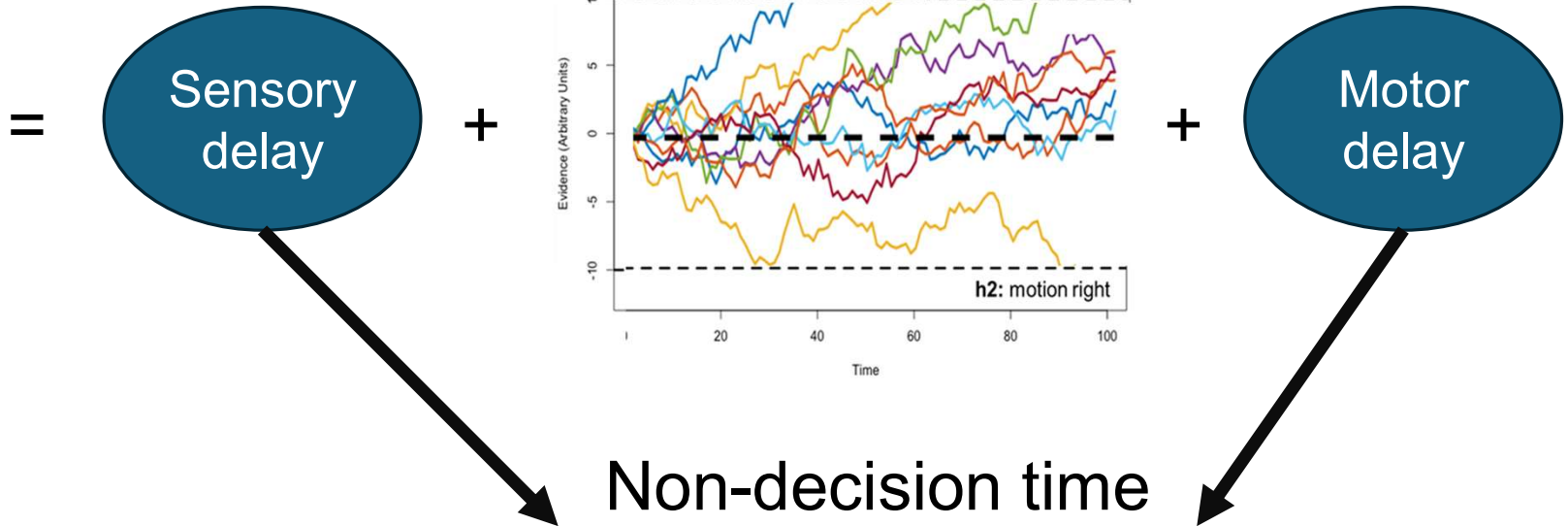
non-decision time

Decision time (DT)



Non-decision time (T_0)

Response time



stochastic equation

- The stochastic differential equation for evidence accumulation:

Change in evidence
at time t

drift

diffusion

$$dE(t) = vdt + cdW(t)$$

signal

noise

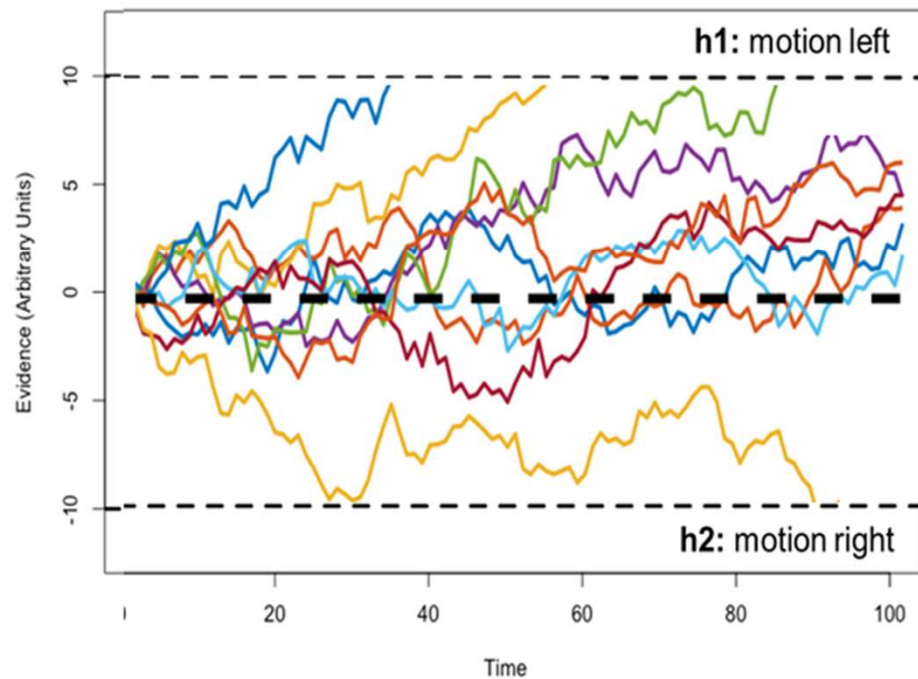
- Evidence is integrated over time, starting at an initial bias x_0
- Accumulation is terminated when evidence crosses the threshold at a or $-a$
- Response time is $T_0 + DT$

parameters

- non-decision time t_0
- starting point bias x_0
- drift rate (signal-to-noise ratio) v
- threshold a
- noise c
 - In practice, often set to 1 to .1
- variability parameters sv, sz, st
 - Across-trial fluctuations stimuli and physiological states

why variability parameters?

Noise causes variability in choices and decision times



what have we learned so far?

- what cognitive processes do the DDM parameters map onto?
 - **threshold, drift rate, starting point, non-decision time**

- in what sorts of tasks would the DDM be a useful tool?

analytic expressions

$$P(\text{Left}) = \frac{1}{1 + \exp(2av)} - \frac{1 - \exp(-2x_0v)}{\exp(2av) - \exp(-2av)}$$

$$RT = t_0 + \frac{a}{v} \text{tahn}(av) + \frac{a}{v} \times \frac{2(1 - \exp(-2x_0v))}{\exp(2av) - \exp(-2av)} - \frac{x_0}{v}$$

- Useful for visualizing results
- Calculate the reward rate (reward per unit of time)
- Equates to softmax equation for value-based choice

application to value-based decisions

- Value-based decisions: decisions where **each option** have **different values**
- Two-alternative case:
 - Choose between option 1 with value R_1 and option 2 with value R_2
 - Connect to DDM by setting the drift rate proportional to difference in value

$$v = d(R_1 - R_2) + \varepsilon = d\Delta R + \varepsilon$$

application to value-based decisions

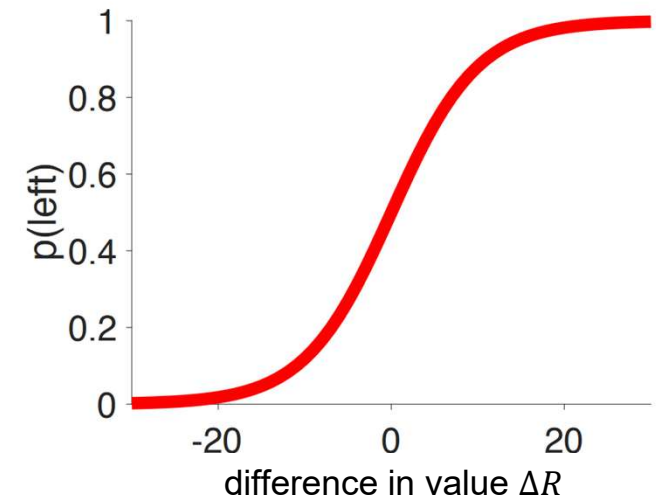
- Choice probabilities

$$p(\text{left}) = \frac{1}{1 + \exp(2ad\Delta R)} - \frac{1 - \exp(-2x_0d\Delta R)}{\exp(2ad\Delta R) - \exp(-2ad\Delta R)}$$

- Special case with unbiased starting point ($x_0 = 0$)

$$p(\text{left}) = \frac{1}{1 + \exp(2ad\Delta R)}$$

- This is the **softmax probability function!**



softmax :: DDM connection

- Compare the two:

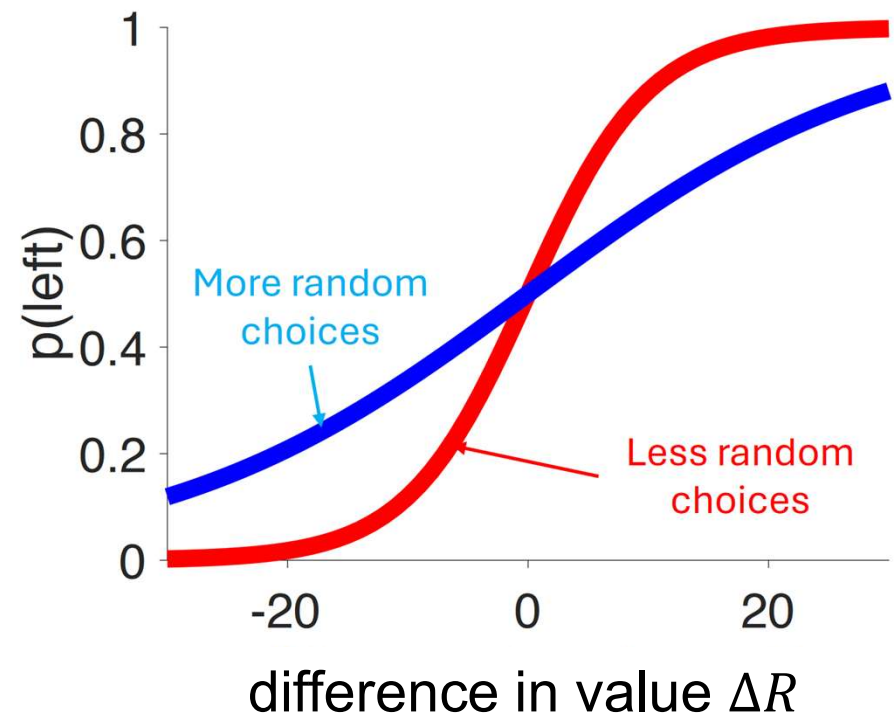
$$\text{DDM: } p(\text{left}) = \frac{1}{1 + \exp(2ad\Delta R)} \quad \text{softmax: } p(\text{left}) = \frac{1}{1 + \exp(\beta\Delta R)}$$

- Softmax's inverse temperature parameter (β) is controlled by two DDM parameters: threshold (a) and signal-to-noise ratio (d)

$$\beta = 2ad$$

softmax :: DDM connection

- In the DDM, stochasticity in choice can be generated by:
 - Reduced signal-to-noise ratio (d)
 - Lower threshold (a)
- Different mechanisms cannot be distinguished by choices alone



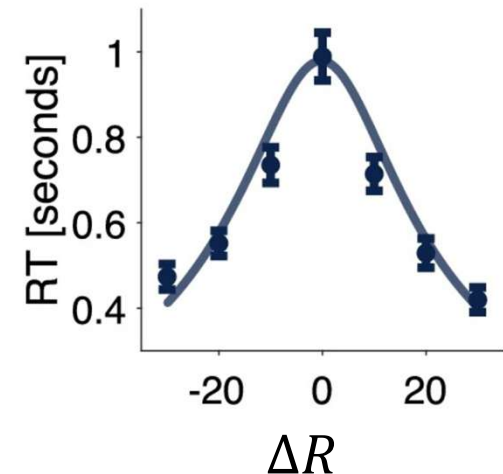
Response times in the value-based DDM

- Response time formula

$$RT = t_0 + \frac{a}{\Delta R} \tanh(a\Delta R) + \frac{a}{\Delta R} \times \frac{2(1 - \exp(-2x_0\Delta R))}{\exp(2a\Delta R) - \exp(-2a\Delta R)} - x_0\Delta R$$

- Special case with unbiased starting bias $x_0 = 0$

$$RT = t_0 + \frac{a}{\Delta R} \tanh(a\Delta R)$$



Response times in the value-based DDM

- Changes to drift rate and threshold have **opposite** effects on response times

Value-based DDM Summary

- Drift rate is proportional to difference between options
- Approximates softmax choice probabilities when initial bias is set to 0
- Stochasticity in choice is influenced by two mechanisms
 - Drift rate
 - Threshold
- These mechanisms **can only** be distinguished using response times

reflection

- what did you already know?
- what have you learned so far?
- what is still confusing?

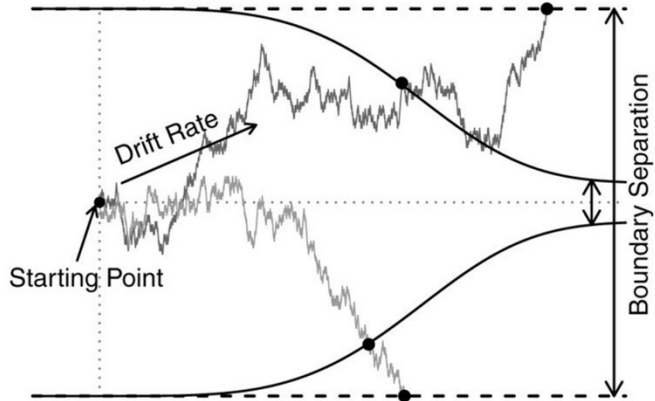
when can we use the DDM?

- only **two options**
- task involves **relative evidence**
- there is **perfect inhibition**
- there is no **leakage** of information
- process is **continuous** in time
- process is **single-stage**

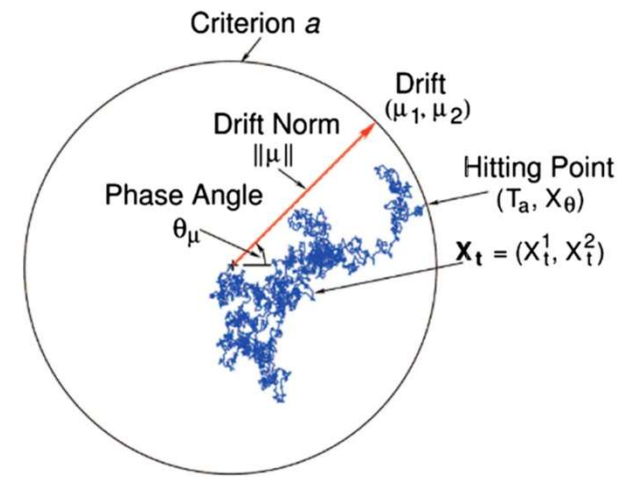
otherwise: Leaky Competing Accumulator Model, Gaze-Weighted Accumulator Model, Linear Ballistic Accumulator Model, Piecewise Diffusion Model, Racing Diffusion Model, Circular Diffusion Model, etc.

Variations of the DDM

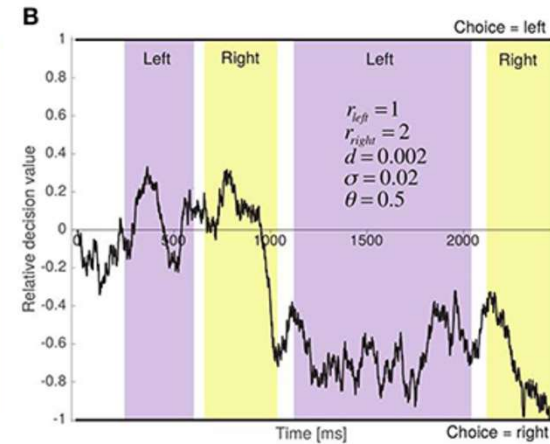
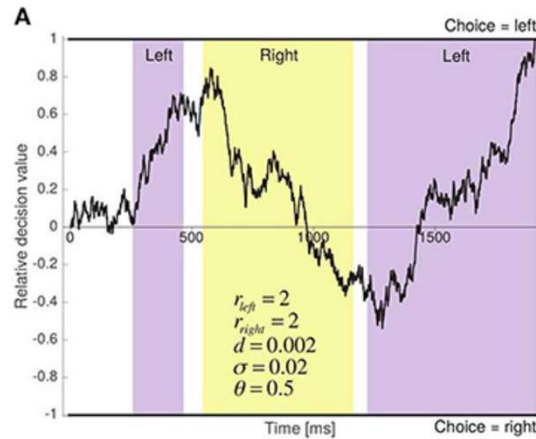
Collapsing bounds



Circular diffusion model



Attentional drift-diffusion model



Variations of the DDM

- Reinforcement learning

- RLDDM (Fontanesi et al., 2019)

$$v = m(Q_{correct} - Q_{incorrect}) \qquad v = \frac{2v_{max}}{1 + \exp(m(Q_{correct} - Q_{incorrect}))} - v_{max}$$

- RLLBA (McDougle & Collins, 2021)

$$v = mQ_i$$

- RLARDM (Miletic et al., 2021)

$$v_1 = V_0 + w_d(Q_1 - Q_2) + w_s(Q_1 + Q_2)$$

$$v_2 = V_0 + w_d(Q_2 - Q_1) + w_s(Q_1 + Q_2)$$

essential bibliography

- **History of the DDM and theory on how the brain might implement it.** Gold, J. I. & Shadlen, M. N. (2002). Banburismus and the Brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36(2), 299-308
- **Understanding the role of the different parameters and how to fit the DDM to behavioral data:** Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, 20(4), 873-922.
- **Understanding the relationship between the DDM and neuropsychological data:** Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in cognitive sciences*, 20(4), 260-281.

software for parameter estimation

- Python
 - HDDM/HSSM
 - rISSM
 - GLAMbox
 - PyDDM
- R
 - **RJAGS**
 - RStan/BRMS
- Julia
 - SequentialSamplingModels.jl

Hierarchical Bayesian Estimation

Models individual-level parameters within a group structure

Key components

- Group-level parameters (hyperparameters): capture overall trends
- Individual-level parameters: capture individual differences

Advantages

- Combines individual- and group-level information to produce accurate estimates
- Useful when data are sparse or noisy at individual level

JAGS (Just Another Gibbs Sampler)

JAGS is a software tool for Bayesian inference using Markov Chain Monte Carlo (MCMC) sampling

Why JAGS?

- Flexible framework for Bayesian modeling
- Can handle complex models and multiple distributions
- Easy to use in R via *runjags* package

JAGS (Just Another Gibbs Sampler)

How it works:

- Define the model in separate model file
 - Likelihood
 - Priors
 - Hierarchical structure
- Run MCMC sampling
 - Initialize distributions
 - Burn-in samples
 - Sample from the posterior

Workflow overview

Step 1: Define the model

write model in JAGS syntax (specify likelihood, priors, hierarchical structure)

Workflow overview

Step 1: Define the model

Step 2: Load data in R

prepare dataset and ensure it is in a format that can easily be processed by JAGS (*list* format in R)

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

use *rjags* to compile and run MCMC sampling

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

Step 4: Assess convergence

use diagnostics such as trace plots, Gelman-Rubin statistic (R-hat), and effective sample size

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

Step 4: Assess convergence

Step 5: Summarize posterior distributions

extract and summarize posterior estimates (e.g., means, credible intervals)

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

Step 4: Assess convergence

Step 5: Summarize posterior distributions

Step 6: Interpret results

interpret the posterior distribution to make inferences

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

Step 4: Assess convergence

Step 5: Summarize posterior distributions

Step 6: Interpret results

Step 7: Model comparisons

generate alternative model specifications and compare using model fit metrics (e.g., DIC, WAIC)

Workflow overview

Step 1: Define the model

Step 2: Load data in R

Step 3: Compile and run

Step 4: Assess convergence

Step 5: Summarize posterior distributions

Step 6: Interpret results

Step 7: Model comparisons

Step 8: Posterior predictive checks

simulate data and compare to empirical data

Workflow overview

Step 1: Define the model ***

Step 2: Load data in R ***

Step 3: Compile and run ***

Step 4: Assess convergence ***

Step 5: Summarize posterior distributions ***

Step 6: Interpret results ***

Step 7: Model comparisons

Step 8: Posterior predictive checks

DDM estimation with JAGS

Downloads

- JAGS: <https://sourceforge.net/projects/mcmc-jags/files/JAGS/4.x/>
- JAGS-WIENER: <https://github.com/yeagle/jags-wiener>
- R-TOOLS: <https://cran.r-project.org/bin/windows/Rtools/>

additional resources