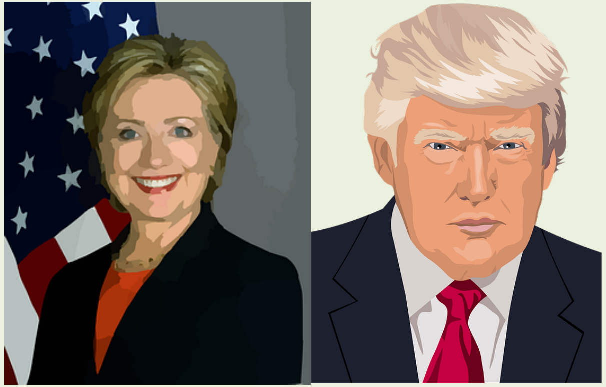
# Models of Action Selection: The Drift Diffusion Model

Mads Lund Pedersen, University of Oslo

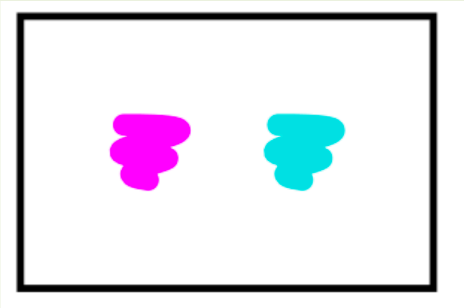
# Binary decision making

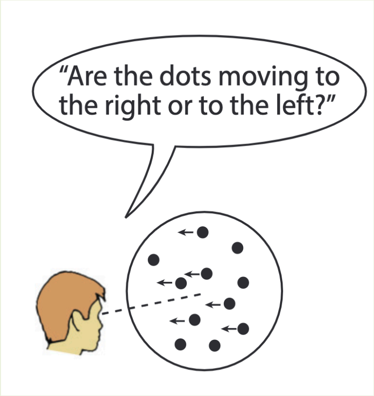

Edible?



- short time-scale
- varying difficulty (between and within decision makers)
- varying outcomes within (and between)

# Decision making tasks



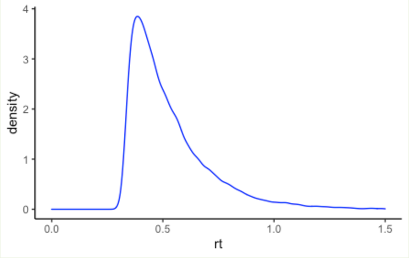



Face or House?

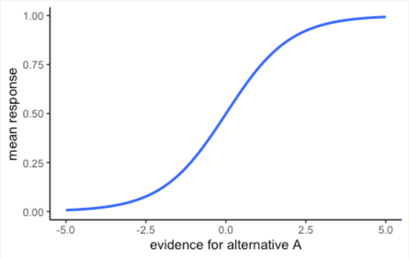
Basten et al., 2010; Mulder et al., 2012; Pedersen et al., 2015

# Behavioral properties of binary decisions

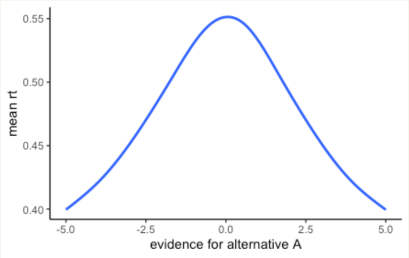
Response times distributions (RT) are right-skewed




Easier stimuli leads to higher accuracy

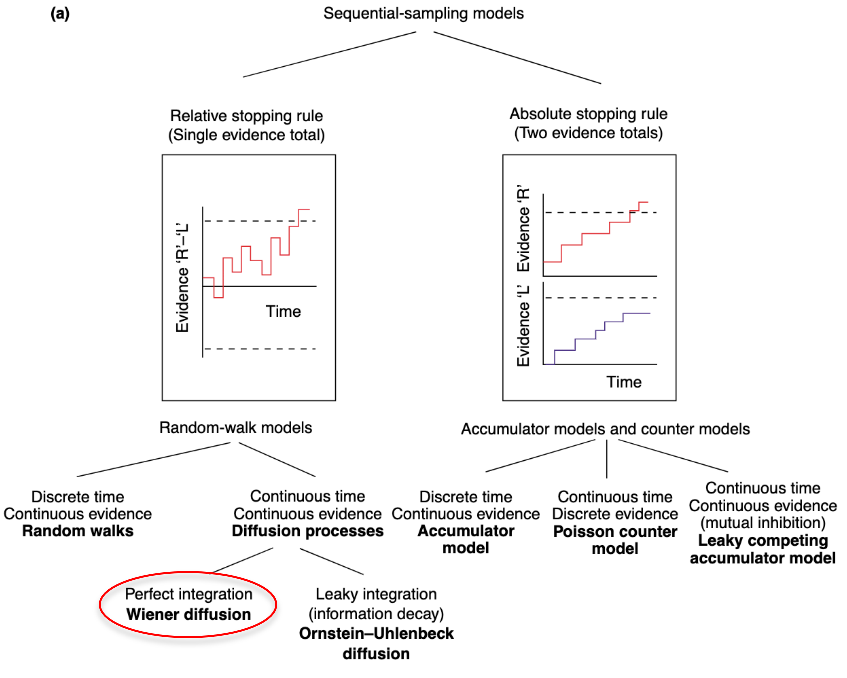



and faster RT



# Sequential sampling models

- Multiple alternative models
    - Accumulation of evidence to bound
    - Single vs. multiple accumulators
- DDM most commonly used




Smith & Ratcliff, 2004

# The drift diffusion model (DDM)

Introduced by Roger Ratcliff (1978)

Builds on work by Wald, Festinger, Turing ++

Assumes evidence is accumulated to one of two bounds

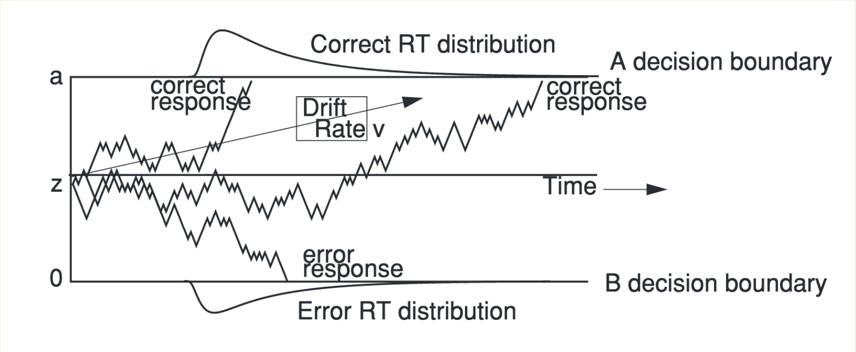
Four main parameters:

Drift rate (v)

Starting point (z)

Decision threshold/boundary separation (a)

Non-decision time (ndt/Ter/t) not shown




Ratcliff & McKoon, 2008

# DDM - Decision variable (DV)

Sampling evidence in favor of option A over option B

Accumulate difference in evidence (A minus B)

Noisy process (internal noise/noise in stimulus)

Momentary evidence (MV) = normal(drift rate, noise)

$DV_{i+1} = DV_i + MV_i$

i = iteration

In reality, DDM assumes continuous evidence:

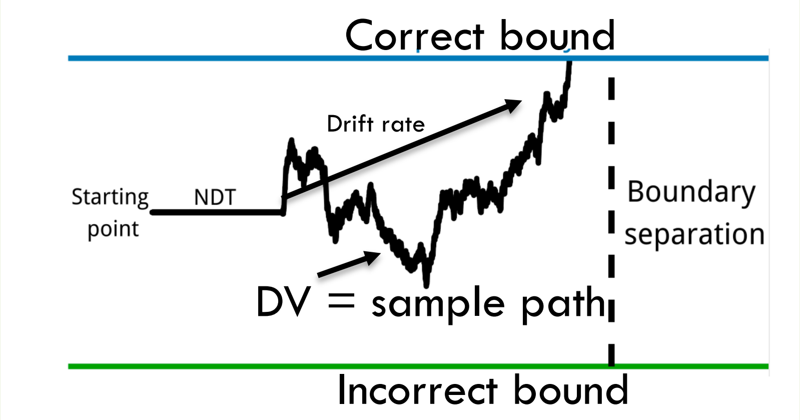
$$\frac{dDV(t)}{dt} = \mu + \sigma W(t)$$

$\mu$ = drift coefficient

$\sigma$ = diffusion coefficient

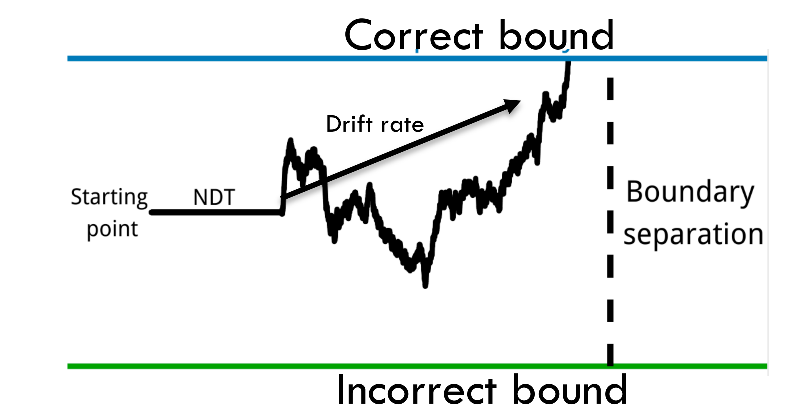
$W$ = Gaussian white noise

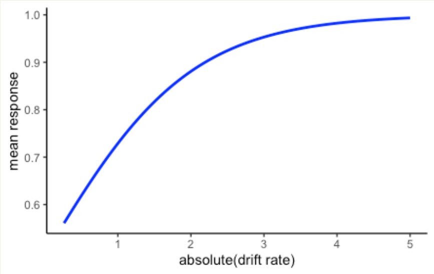
t = timestep



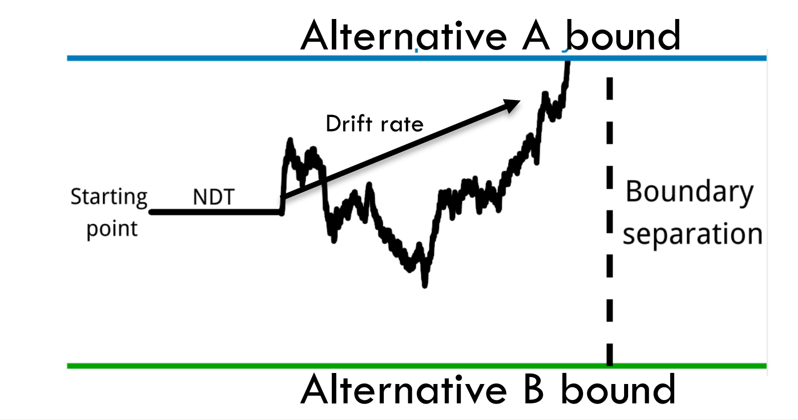
# stimulus coding vs accuracy coding

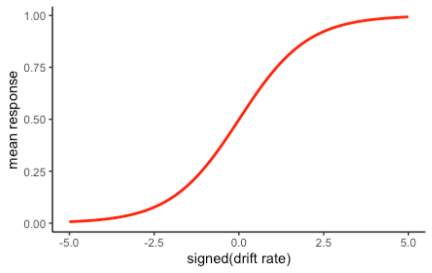
## Accuracy-coding




## Stimulus-coding



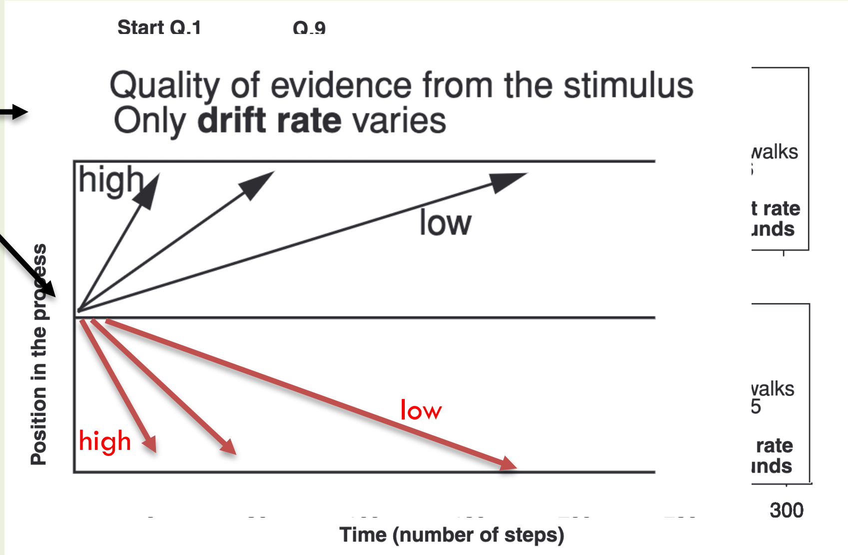
# Drift rate (v)

Rate of evidence accumulation

High absolute drift: fast and consistent choice

Low absolute drift: slow and stochastic choice

Strong influence on tail of distribution (Q.9)




Upper bound correct

Lower bound correct

Ratcliff & McKoon, 2008

# Decision threshold – speed-accuracy trade-off

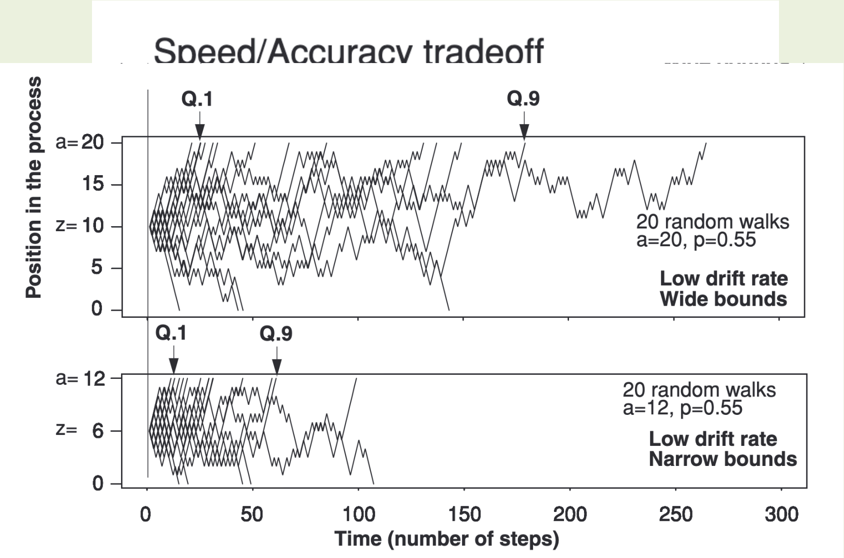
When to stop accumulating evidence?

Wide bounds:

- long decision time
- consistent choice (less impact of noise)

Narrow bounds:

- fast decision time
- stochastic choice (more impact of noise)

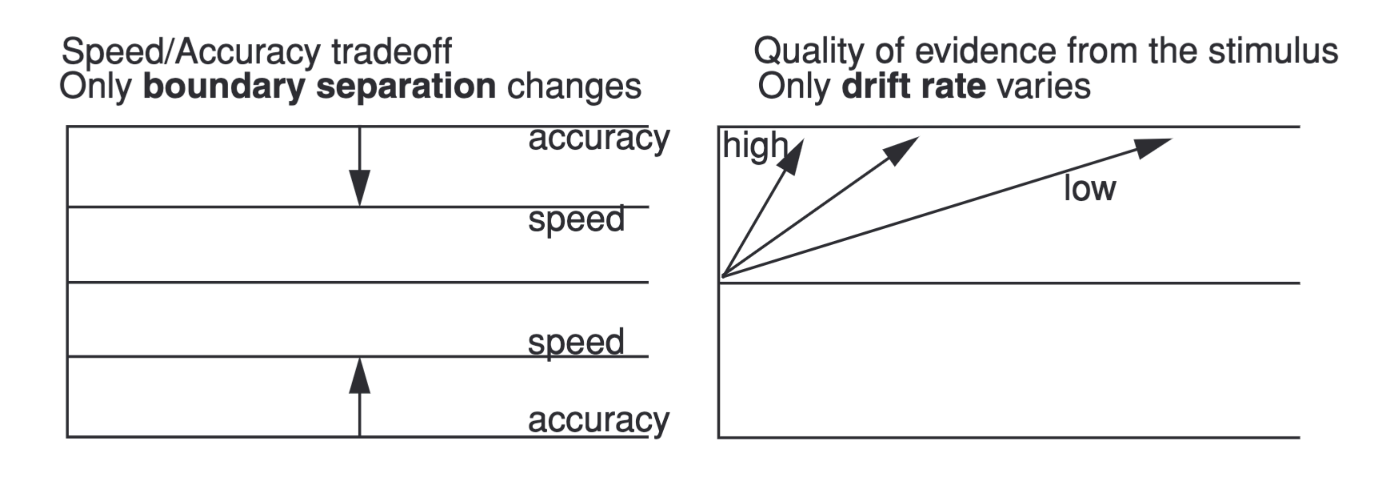



Ratcliff & McKoon, 2008

# Effect of drift and bound on choice and RT

Strong drift: fast and consistent

High bounds: slow and consistent



# Starting point bias

Start point of decision variable

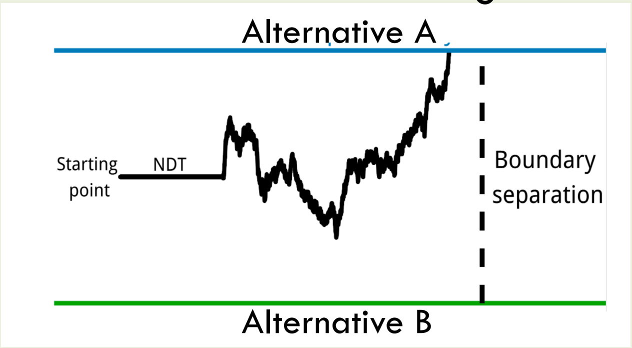
Set a-priori (before stimulus presentation)

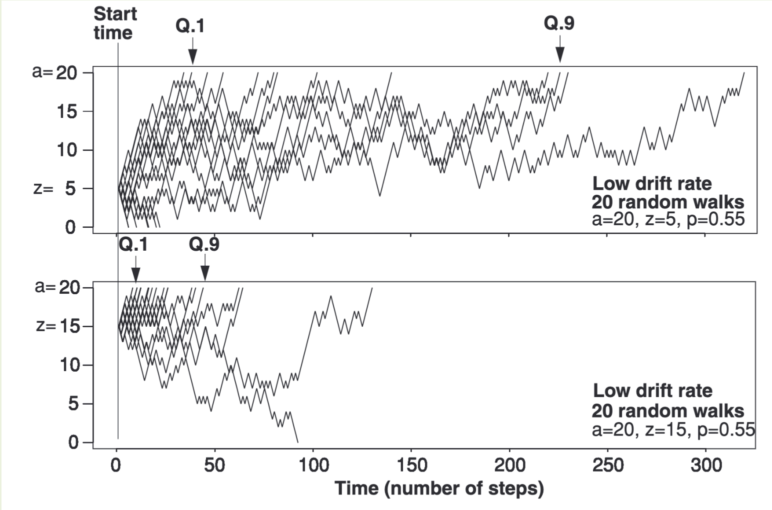
Influenced by:

- preference
- reward for option
- likelihood of correct

Fast RTs for biased option

Stimulus-coding



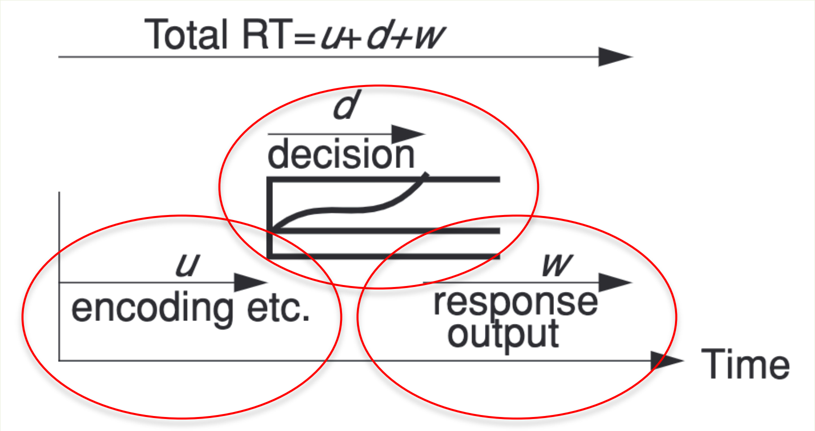

Ratcliff & McKoon, 2008

# Non-decision time

Decision time = accumulation process

Non-decision time = stimulus encoding + motor response

Response time = Decision time + Non-decision time



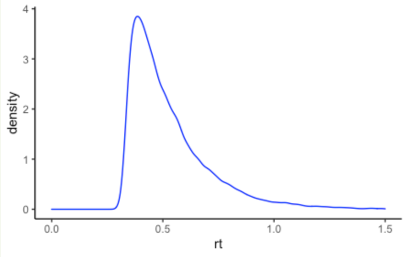

Ratcliff & McKoon, 2008

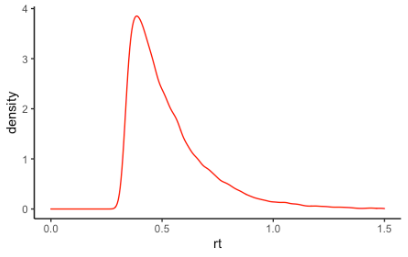
# Revisiting behavioral patterns

Data

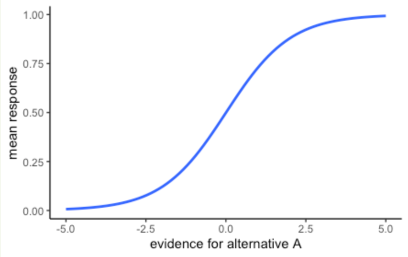
Model

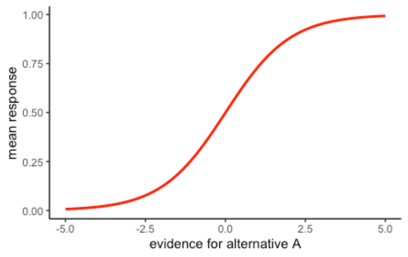
Response times (RT) are right-skewed



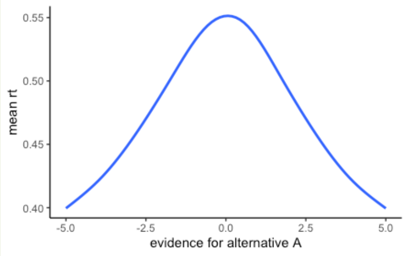



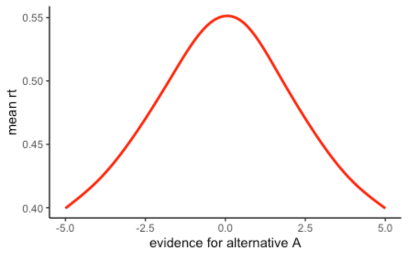

Easier stimuli leads to higher accuracy




and faster RT



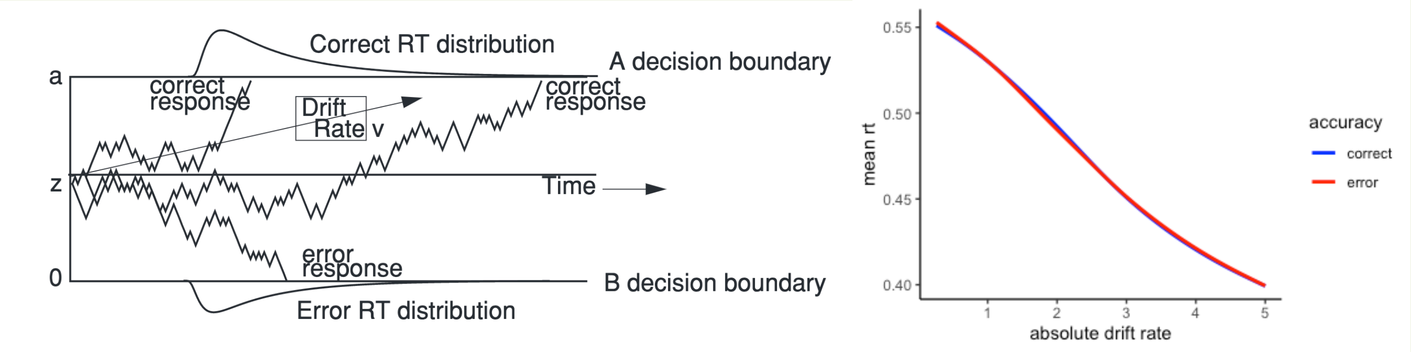
# Error vs correct RT (can) differ

Error RTs sometimes slower than correct RTs, and sometimes faster.

Cannot be accounted for by 'standard' DDM (assuming centered starting point)

observed RT

predicted RT



# Between-trial variability in drift rate

Extensions to DDM include variability in drift rate, starting point and non-decision time (Ratcliff, 1991; Rouder & Ratcliff, 1999)

Variability in drift rate ($S_v$):

Assumes trial-by trial drift comes from normal distribution (drift rate, $S_v$)

**How it accounts for slow errors:**

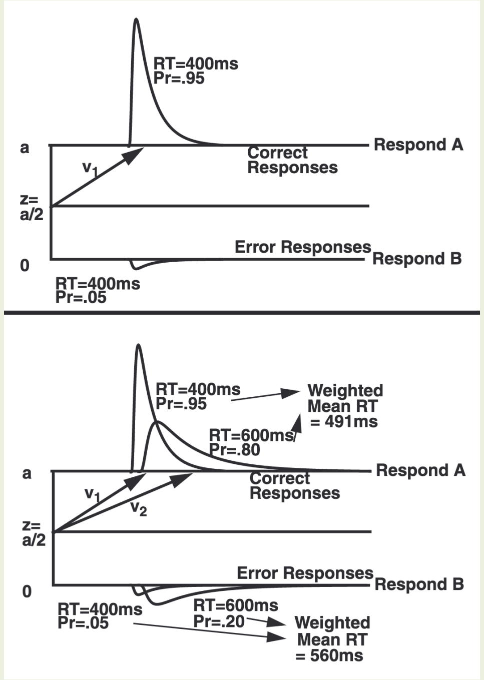
Absolute drift rates closer to 0 ▻ slow + more errors

Absolute drift rates away from 0 ▻ fast + fewer errors

Result: errors slower than correct responses

without Sv

with Sv




Ratcliff & McKoon, 2008

# Between-trial variability

Variability in starting point ($S_z$):

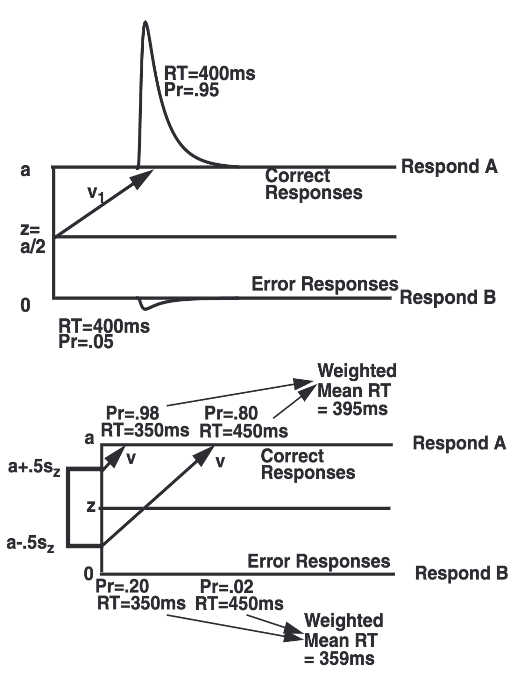
Assumes trial-by trial starting point comes from uniform distribution ($z-S_z$, $z+S_z$)

Fast error responses arise when starting point closer to incorrect bound

Result: errors faster than correct responses

without Sz

with Sz




Ratcliff & McKoon, 2008

# Model estimation

Wiener first passage time (wfpt):

$$\text{Wiener}(y|\alpha, \tau, \beta, \delta) = \frac{\alpha^3}{(y-\tau)^{3/2}} \exp\left(-\delta\alpha\beta - \frac{\delta^2(y-\tau)}{2}\right) \sum_{k=-\infty}^{\infty} (2k+\beta)\phi\left(\frac{2k\alpha+\beta}{\sqrt{y-\tau}}\right)$$

Simplified:

probability of choice and reaction time given parameter values:

wfpt(choice+rt|parameters)

goal: find parameters that best explain observed data (choice+rt)

# Model fitting

Analytical solution to wiener process – likelihood function

Chi-square quantile CDF

Likelihood function (PDF):

- Maximum likelihood
- Bayesian parameter estimation
  - MCMC: approximate posterior distribution
  - Variational Bayes

# Model fitting toolboxes

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

HBayesDM: https://github.com/CCS-Lab/hBayesDM

BRMS/STAN: http://paul-buerkner.github.io/brms/reference/Wiener.html

Rwiener: https://journal.r-project.org/archive/2014/RJ-2014-005/RJ-2014-005.pdf

Jags-wiener: https://sourceforge.net/projects/jags-wiener/

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

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

NB! List not exhaustive

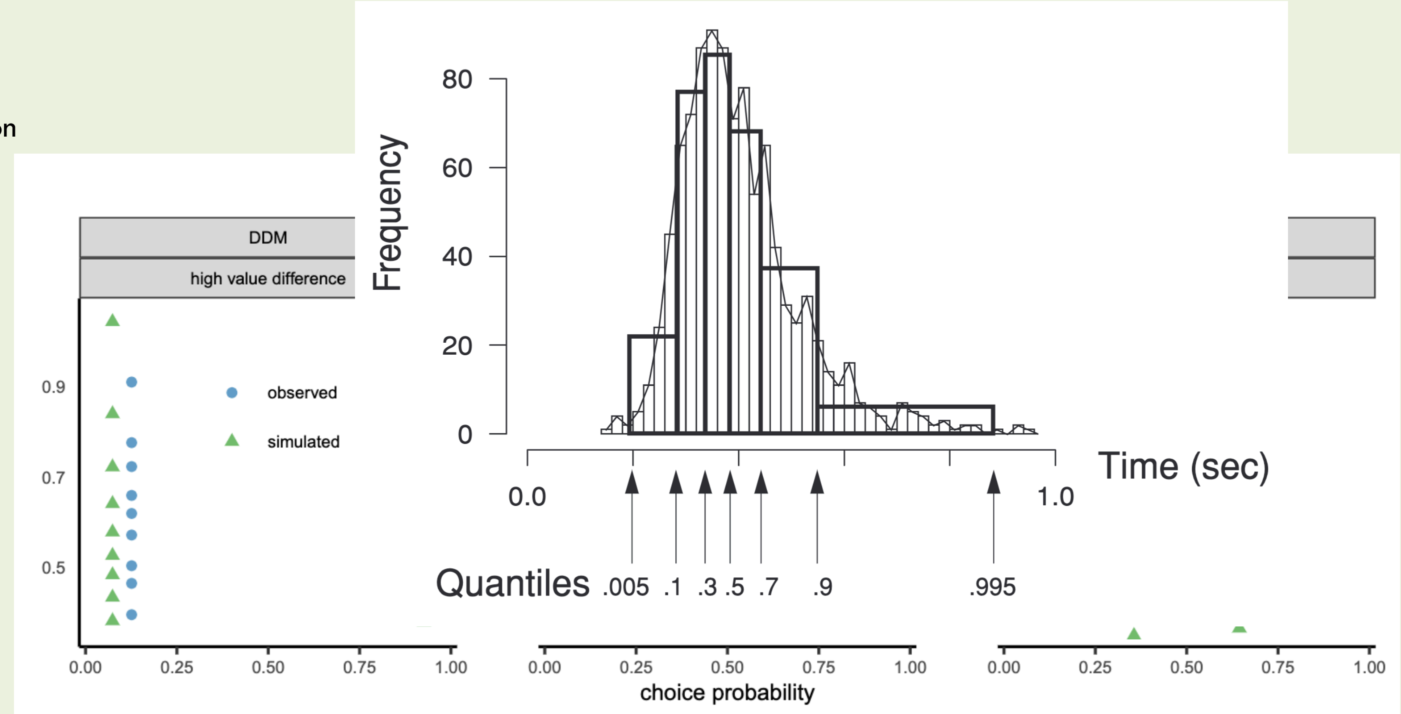
# Model validation

## Does the model recreate observed patterns?

### Account for:

Proportion choice

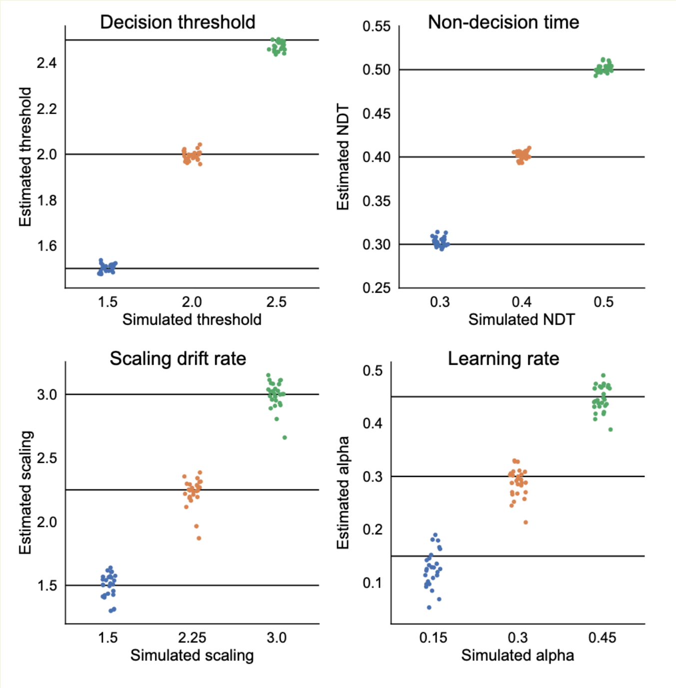
Correct and error RT distribution




Ratcliff & McKoon, 2008; Pedersen & Frank, 2020

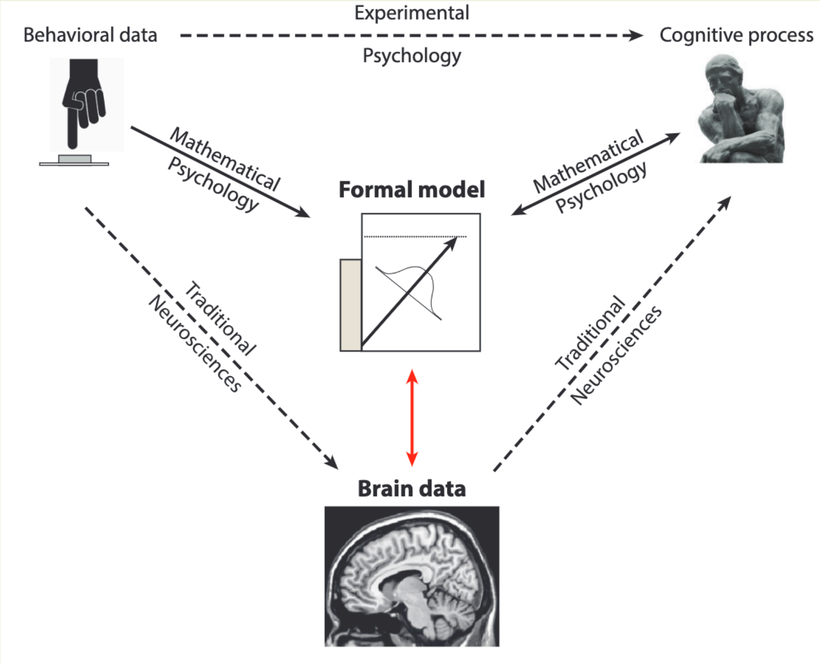
# Model recovery

Can we recover parameter values used to generate data?




Pedersen & Frank, 2020

# Linking brain and behavior




Forstmann et al., 2011

# DDM as neural activity – non-human primates

Random dot motion task

Momentary evidence in area middle temporal area (MT)

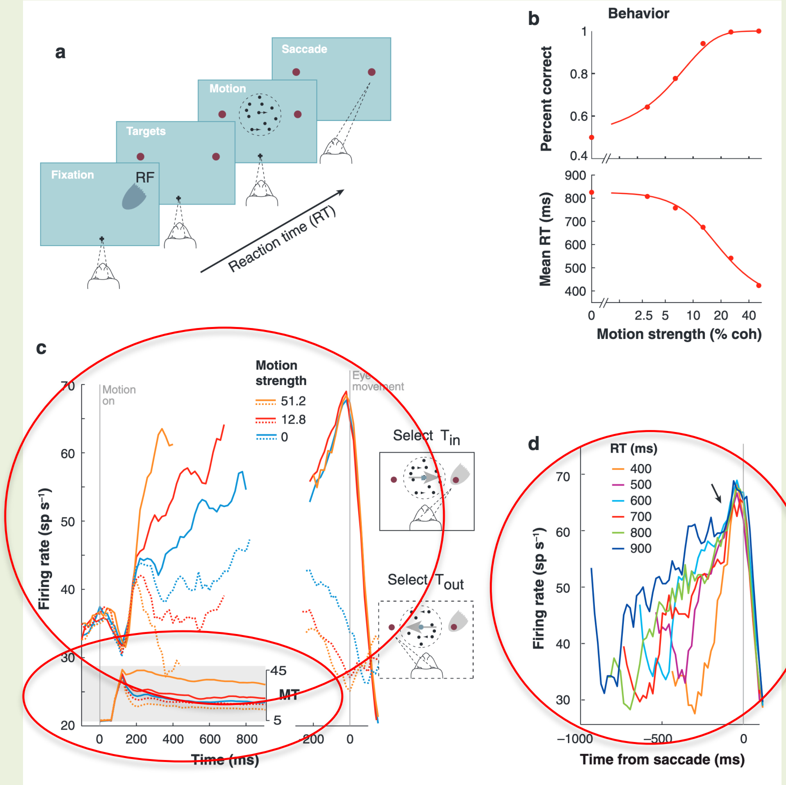
Integration of evidence in lateral intraparietal cortex (LIP)

Rises earlier for easier decisions

Reaches threshold in firing rate

Best explained as a race-process (separate accumulators)

Leaky competing accumulator (Usher & McClelland, 2001)




Gold & Shadlen, 2007

# DDM as neural activity - humans

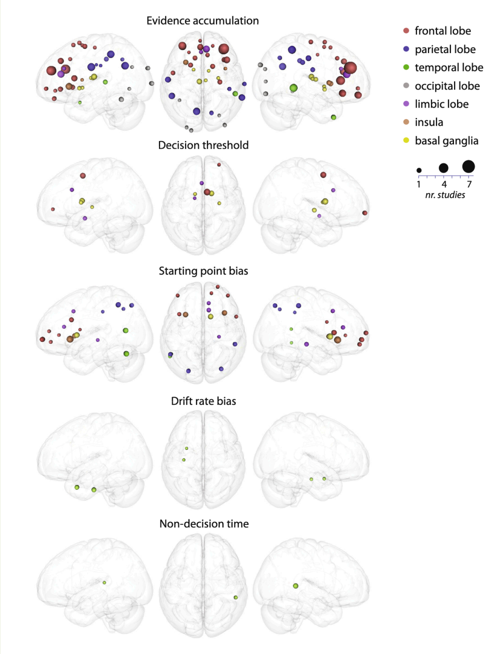
Identify regions associated with decision process

Option 1:

extract parameters ▻ inform BOLD ▻ GLM(BOLD ~ parameters)

Option 2 (HDDM):

extract ROIs ▻ inform parameters ▻ GLM(parameter ~ ROI)




Mulder et al., 2014

# Applied to mental disorders

## Benefits:

Decompose behavior into latent processes

Increase sensitivity to detect group differences

Link to neural activity

Improve on reliability paradox for cognitive tasks (Hedge et al., 2017), see Haines et al. (PsyArXiv) using generative models (less complex than DDM) to increase reliability.

## Example:

Meta-analyses of perceptual DM in ADHD (Mowinckel et al., 2015, Karalunas et al., 2014) find reduced drift rate (attention?) but not narrower bound (impulsivity?)

# Extensions – collapsing bounds

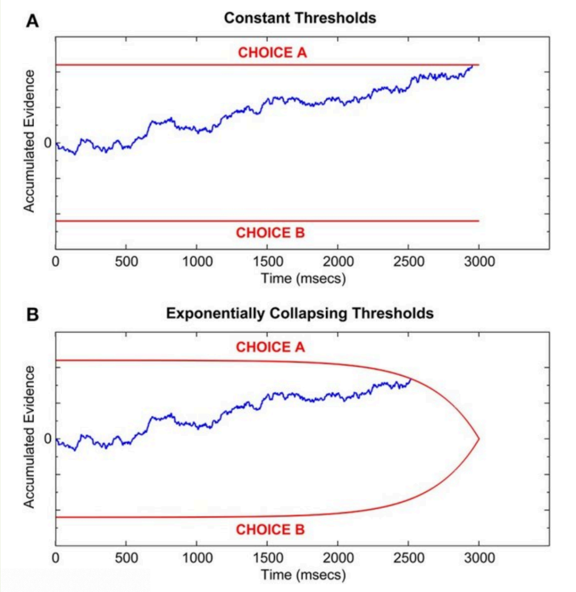
Alternative to explaining slow errors (late in accumulation more likely to cross error bound)

Produces shorter tails of RT distribution

No analytical solution (but extension coming to HDDM)

Comparison standard vs collapsing: Hawkins et al., 2015

collapsing model best for highly trained subjects ▻ NHP

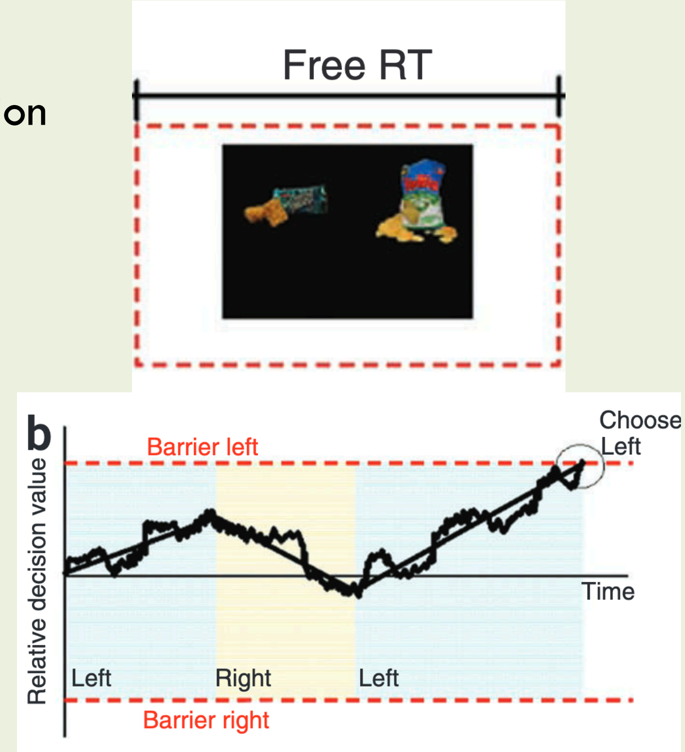



Karsilar et al., 2014

# Extensions – preference-based DM

Choice between e.g. food items

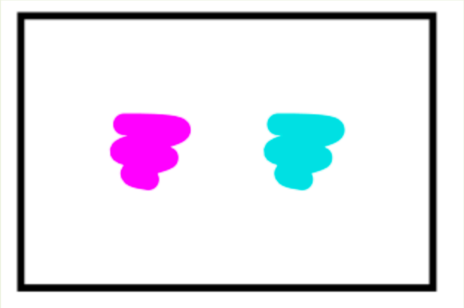
Alternative model: accumulation dependent on fixation




Krajbich et al., 2010

# Extensions - value-based DM

Choice between learned value items (e.g. Basten et al., 2010)



Explain choice in RL with DDM (RLDDM; e.g. Frank et al., 2015; Pedersen et al., 2017)

Extension in HDDM (Pedersen & Frank, 2020)

Drift $\text{rate}_i = (Q_a - Q_b) * v$

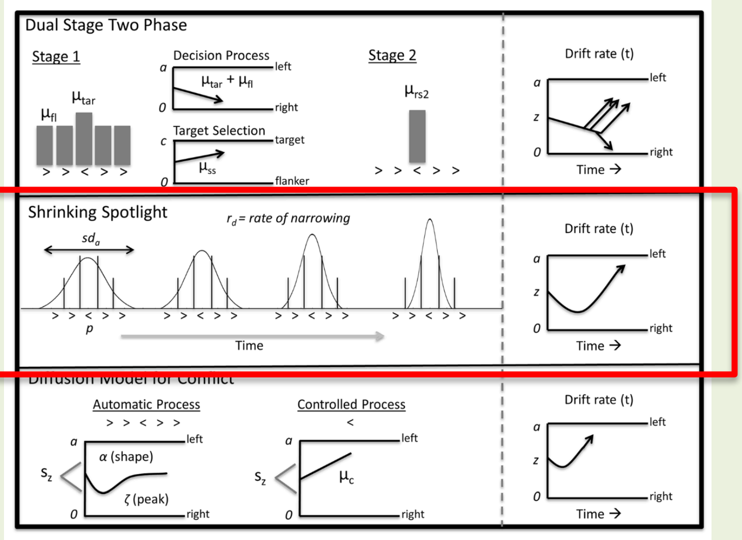
signed RT ~ wfpt($a$, $ndt$, $z$, $v_i$)

# Extensions - Conflict

**Flanker task (Eriksen & Eriksen, 1974**

Report direction of center stimulus:

>><>>

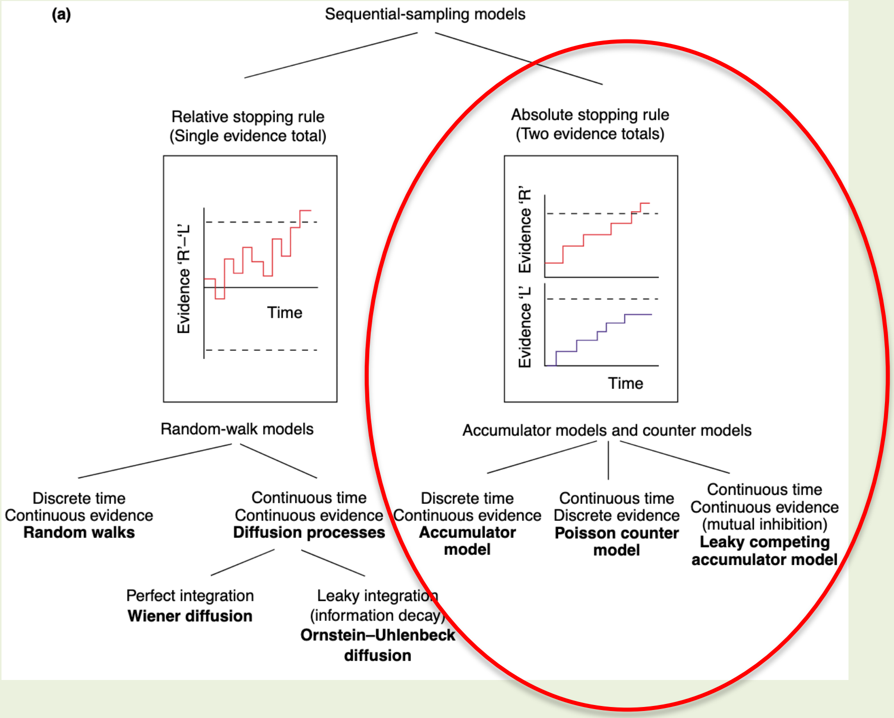



White et al., 2017

# Multialternative decision making

Linear ballistic accumulator
(Brown & Heathcote, 2008)

Leaky competing accumulator
(Usher & McClelland, 2001)




Smith & Ratcliff, 2004

# Summary

**Modeling binary decision making**

behavioral phenomena to be accounted for

**Drift diffusion model**

capturing behavioral phenomena with parameters

variability account for error vs. correct RT

**Model fitting**

Estimation – validation – recovery

Toolboxes

**Neurophysiology**

Single-cell in non-human primates

Imaging in humans

**Applied to mental disorders**

**Extensions**

# Thanks for listening!

**Supervisors and PIs:**

Guido Biele, Norwegian Insitute of Public Health

Michael J. Frank, Brown University

Lars T. Westlye, University of Oslo

# Suggested reading

Introduction:

Ratcliff & McKoon, 2008

Reviews:

Gold & Shadlen, 2007

Forstmann, Ratcliff & Wagenmakers, 2015

Ratcliff, Smith, Brown & McKoon, 2016