Measures and models of covert visual attention in neurotypical function and ADHD

Andra Mihali

April 9, 2018

Divided attention

…the fewer objects we consider at once, the clearer and more distinct will be our knowledge of them. Hence the more vividly we will or desire that a certain object should be clearly and distinctly known, the more we concentrate consciousness through some special faculty upon it.

Selective attention

William Hamilton, 1859

Covert visual attention

Part 1

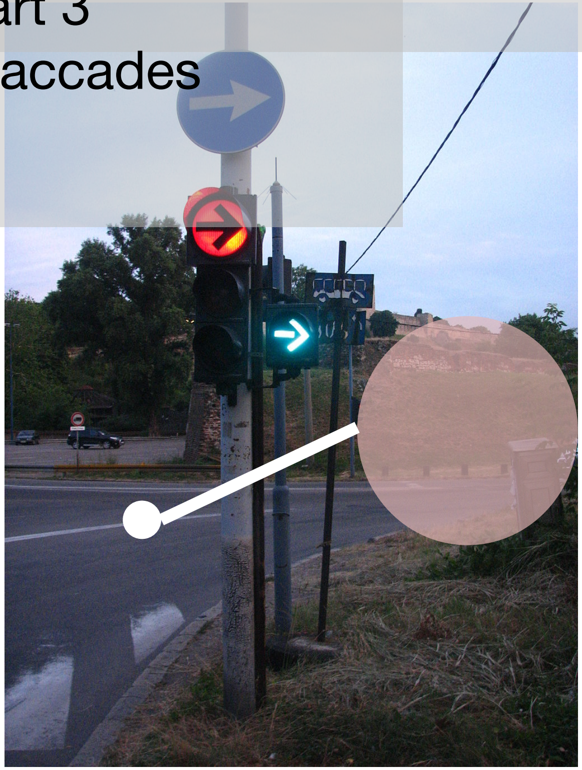
1. Divided attention

Part 2

2. Selective attention

Part 3

Microsaccades



Outline

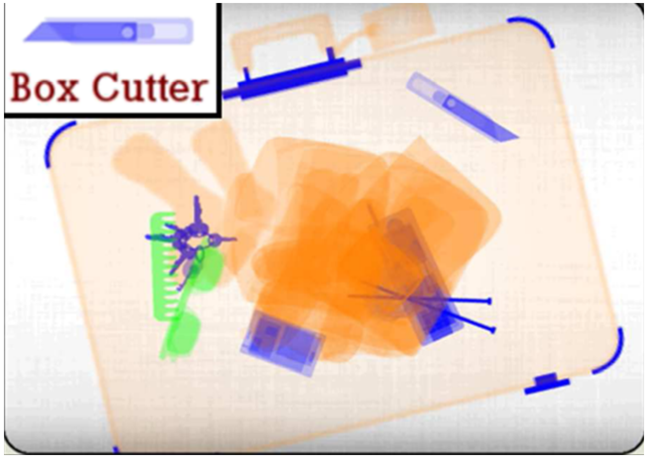
1. Divided attention: visual search with heterogenous distractors
   In preparation

2. Selective attenti eature cueing and task-switching in neu DHD
   with Allison G Yo Adler, Michael H Halassa – bioRxiv preprint, Computational Psychiatry

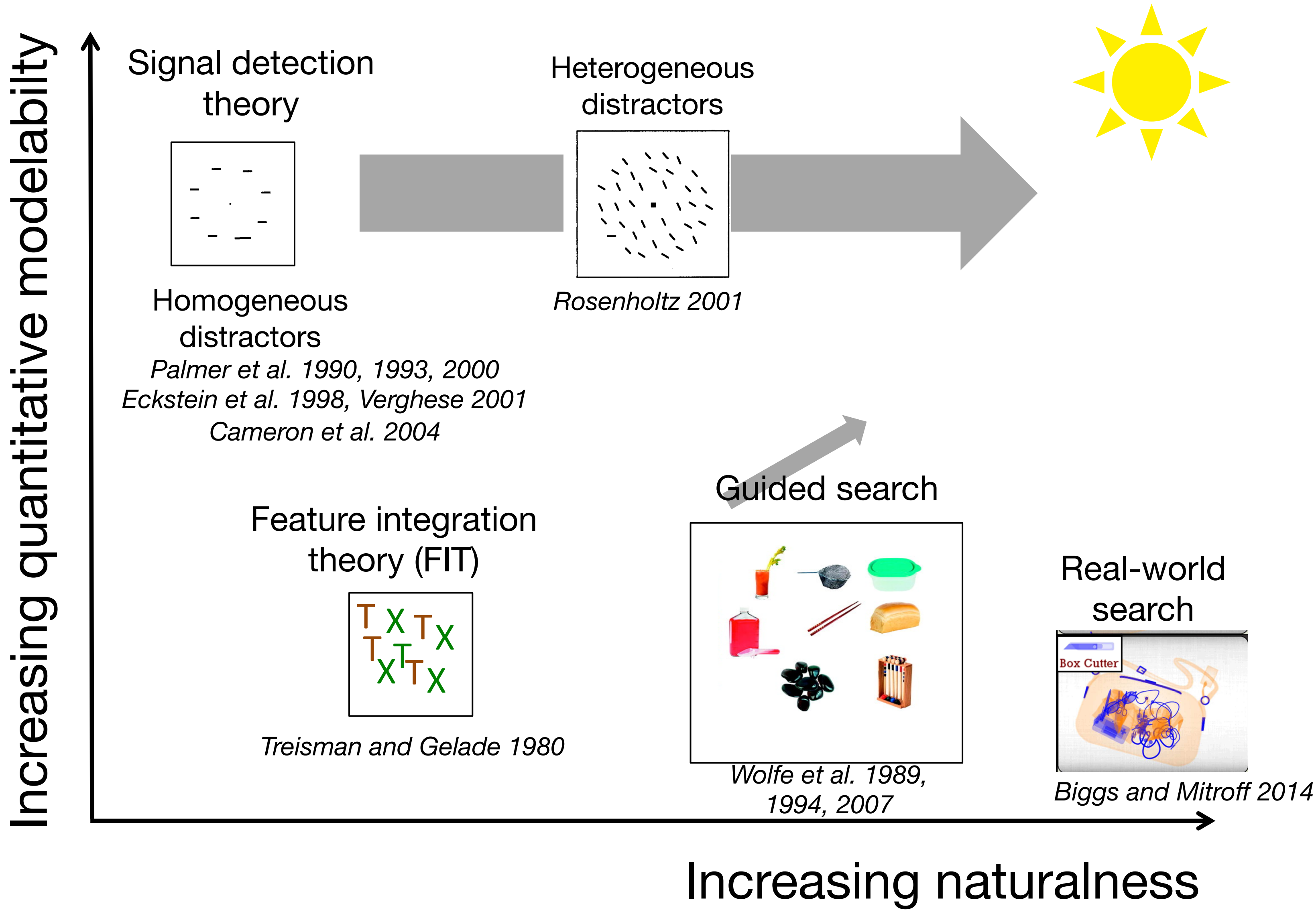
3. Bayesian micros n, with an application to ADHD
   *with Bas van Opheusden**
   Journal of Vision, 17 (1):13

Part 1: visual search with heterogeneous distractors

Visual search

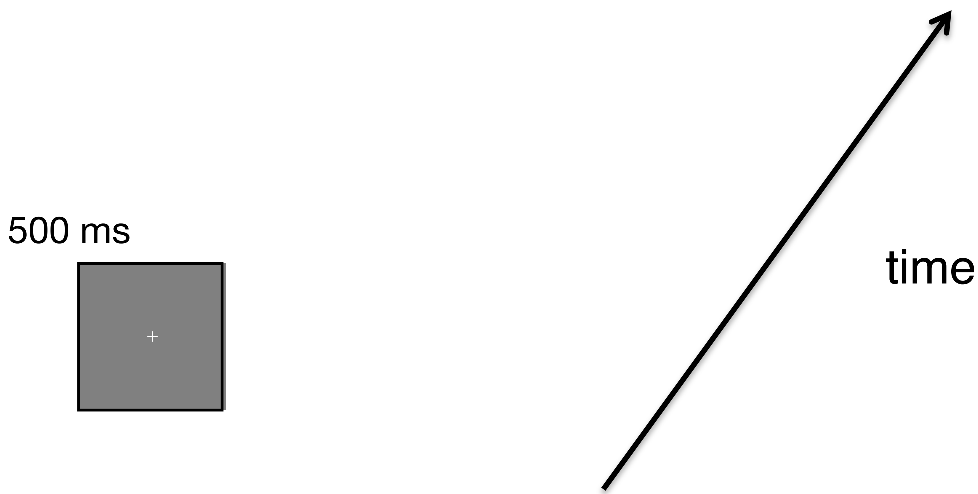


Biggs and Mitroff 2014

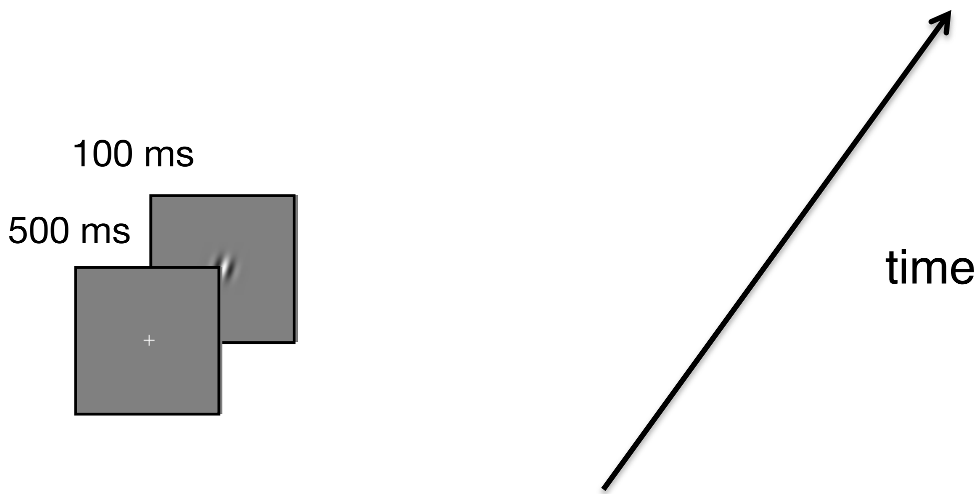


Increasing quantitative modelabilty
Signal detection theory
Heterogeneous distractors
Homogeneous distractors
Palmer et al. 1990, 1993, 2000
Eckstein et al. 1998, Verghese 2001
Cameron et al. 2004
Rosenholtz 2001
Feature integration theory (FIT)
T X T X T T X T X
Treisman and Gelade 1980
Guided search
Wolfe et al. 1989, 1994, 2007
Real-world search
Box Cutter
Biggs and Mitroff 2014
Increasing naturalness

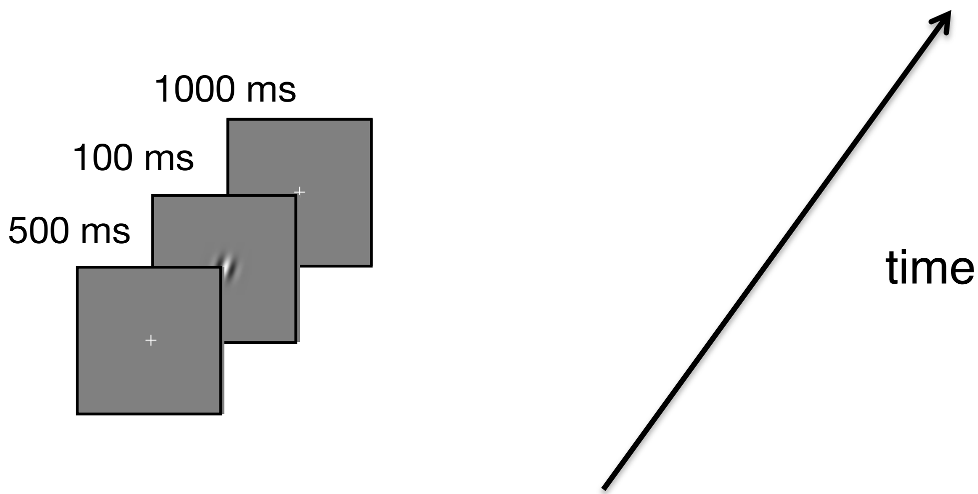
n-AFC localization with heterogenous distractors



n-AFC localization with heterogenous distractors

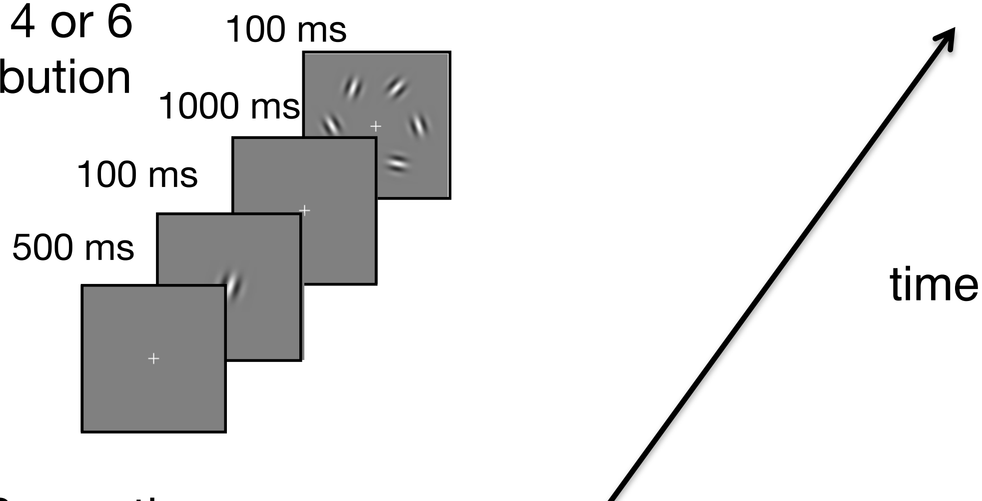


n-AFC localization with heterogenous distractors



n-AFC localization with heterogenous distractors

Set size: 2, 3, 4 or 6
Uniform distribution

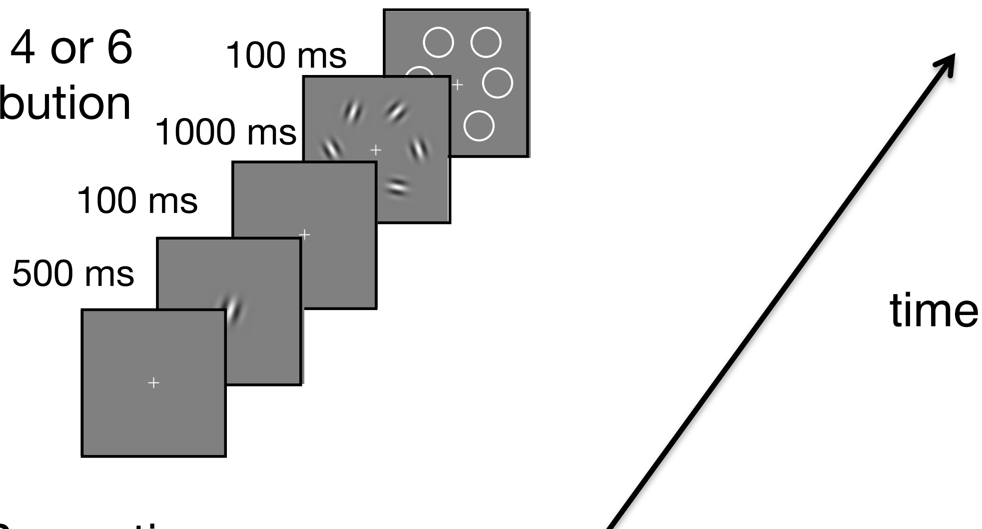


Perception

n-AFC localization with heterogenous distractors

Where was the target?
Mouse click

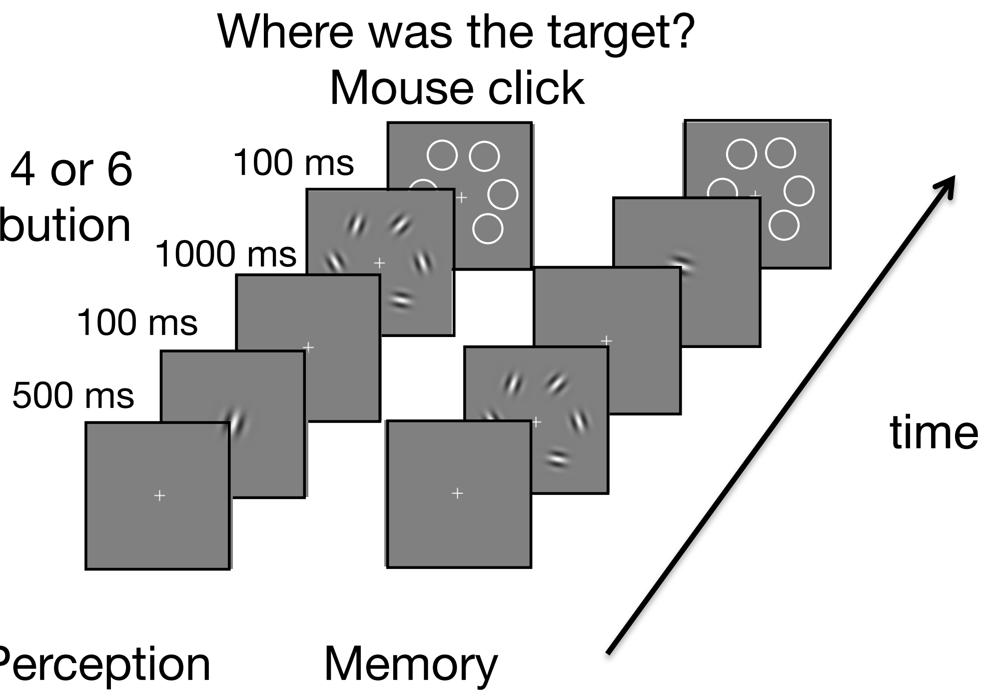
Set size: 2, 3, 4 or 6
Uniform distribution



Perception

n-AFC localization in perception and memory

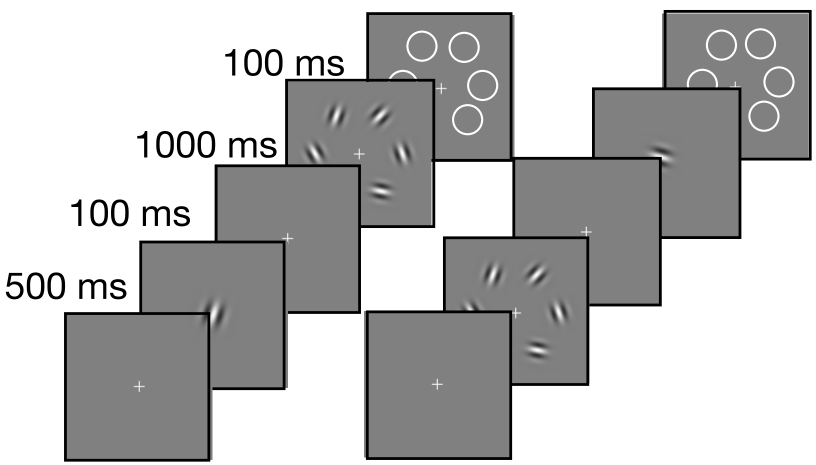
Set size: 2, 3, 4 or 6
Uniform distribution



n-AFC localization and detection

Localization

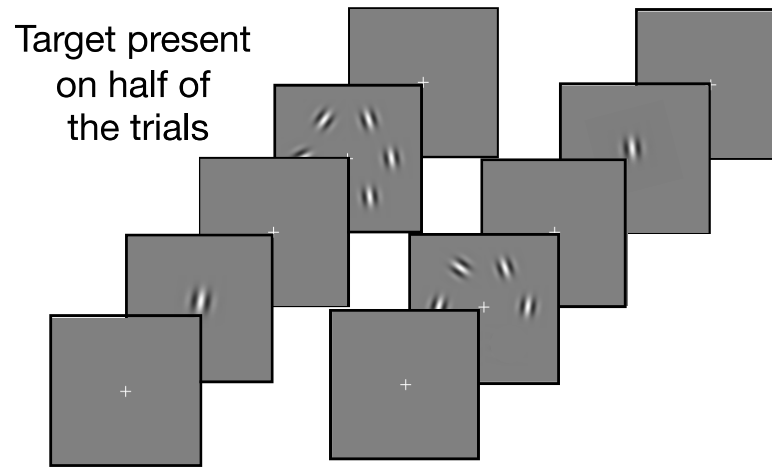
Where was the target?
Mouse click



Perception Memory

Detection

Target present or absent?
Button press

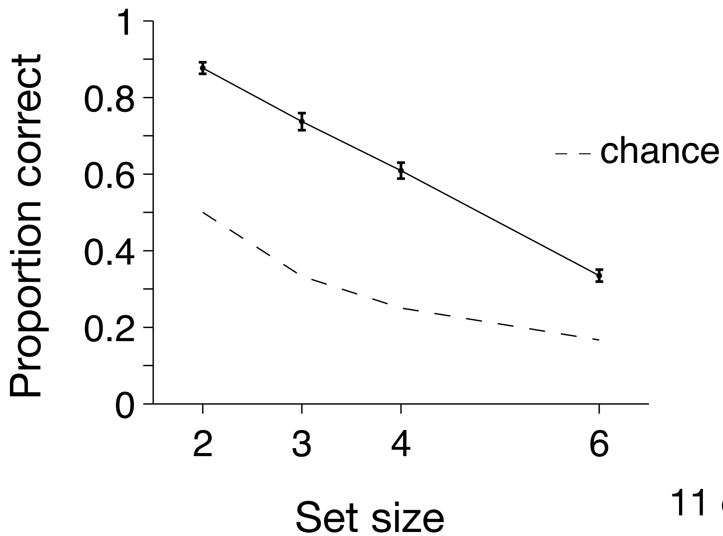


Perception Memory

Mazyar, van den Berg and Ma, 2012

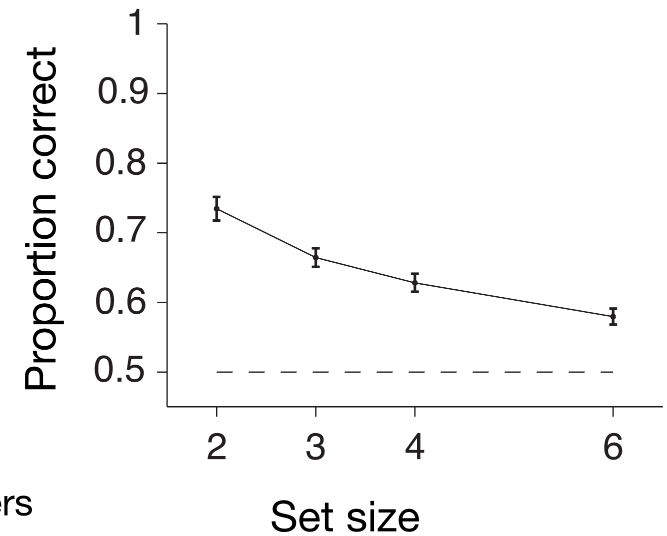
Summary statistic #1:
Accuracy as a function of set size

Localization

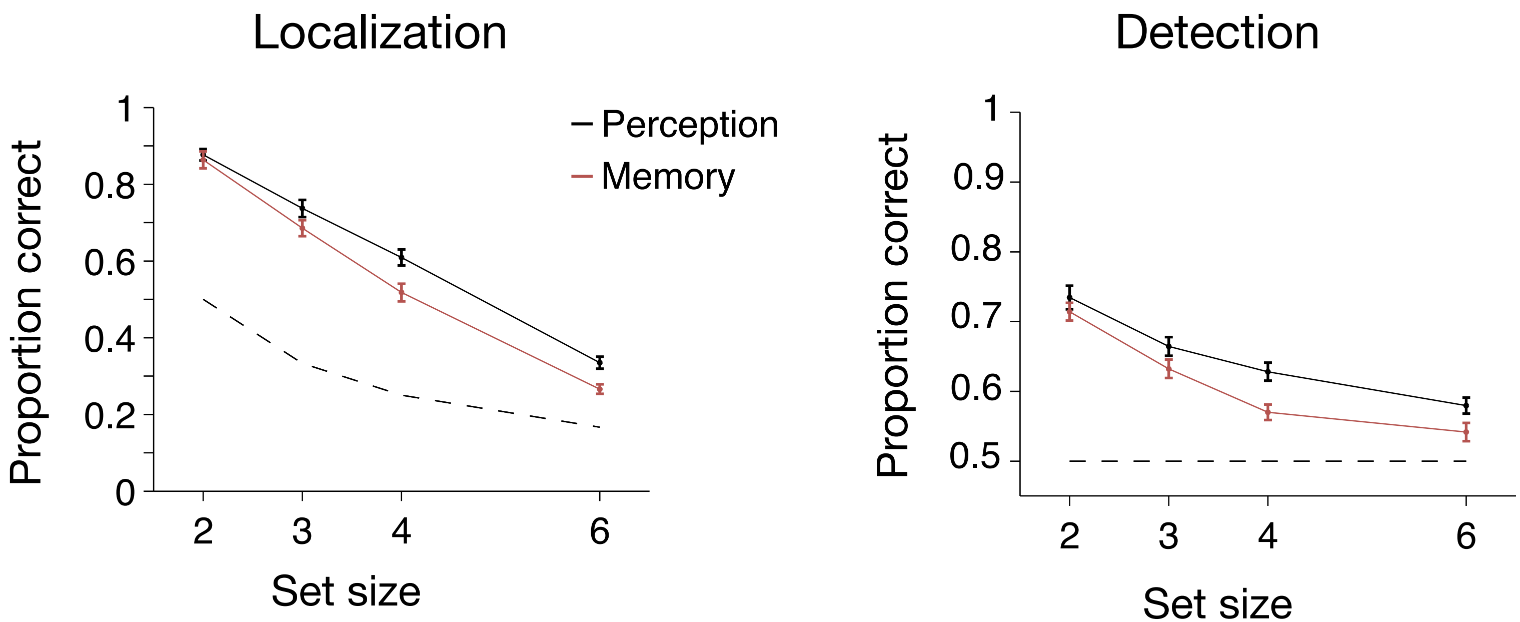


11 observers

Detection



Summary statistic #1: Accuracy as a function of set size



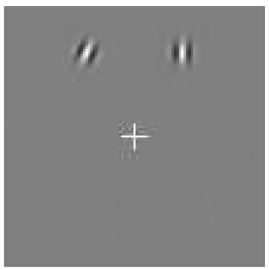
Effect of set size: $p < 0.001$
Effect of Perception/Memory: $p < 0.001$

Effect of set size: $p < 0.001$
Effect of Perception/Memory: $p < 0.001$

- Goal: understand the underlying process through quantitative modeling.
- We need a more detailed characterization of behavior.
- How can we better characterize the behavioral data given that distractors are heterogeneous?

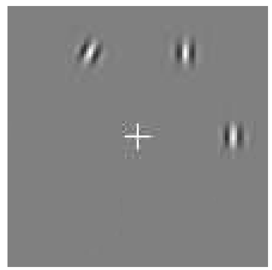
Challenge: Distractor space is (N-1)-dimensional

$N = 2$



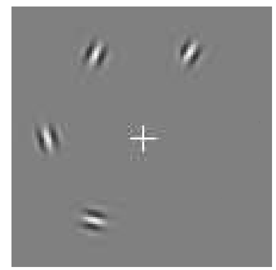
1 distractor

$N = 3$



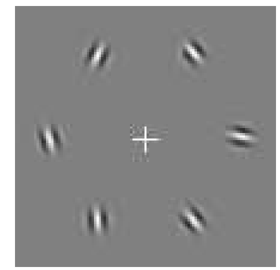
2 distractors

$N = 4$



3 distractors

$N = 6$



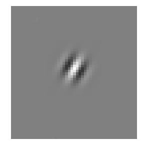
5 distractors

We need a smart “dimensionality reduction”.

One way to reduce N-1 distractors to a single number: minimum target-distractor (T-D) orientation distance

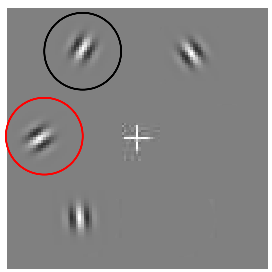
Target orientation

$s_T = 30°$

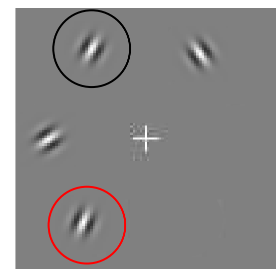


Easier?

$N = 4$



Harder?

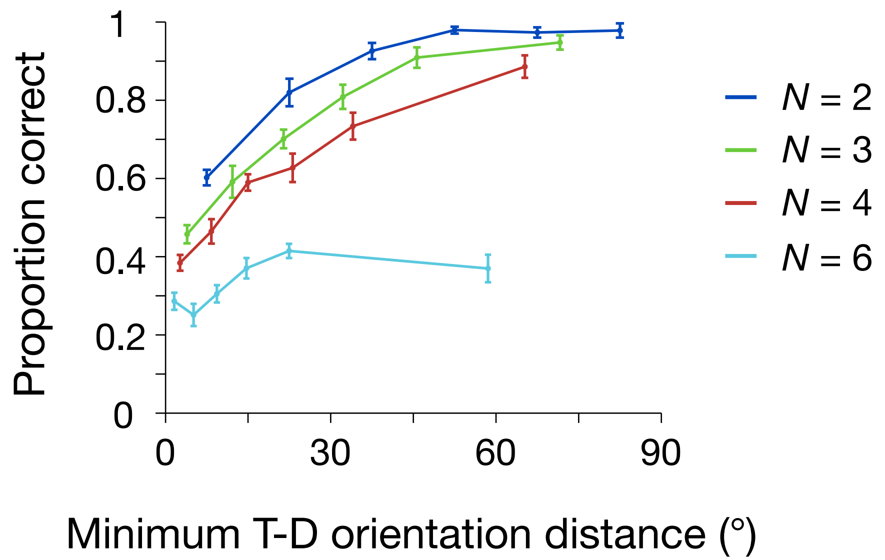


Distractor orientation (°)	Target-distractor orientation distance (°)
-35	65
55	25
-5	35

Distractor orientation (°)	Target-distractor orientation distance (°)
-35	65
55	25
24	**6**

Minimum T-D orientation distance

Summary statistic #2: Accuracy as a function of minimum target-distractor orientation distance



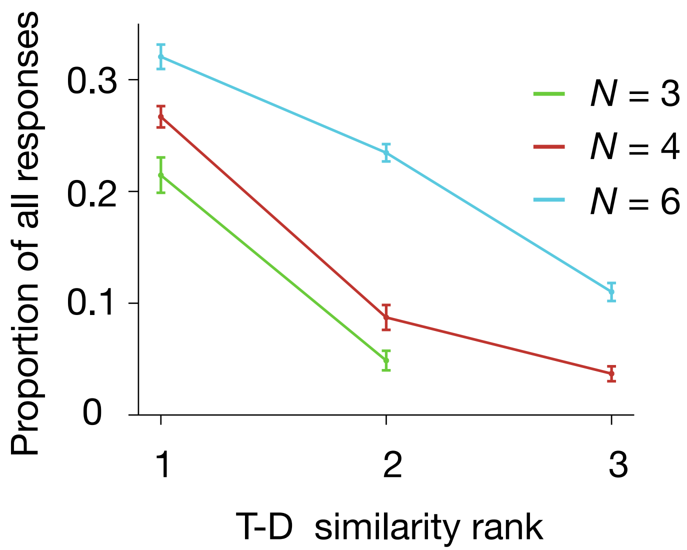
Effect of set size: $p < 0.001$
Effect of minimum T-D orientation distance: $p < 0.001$

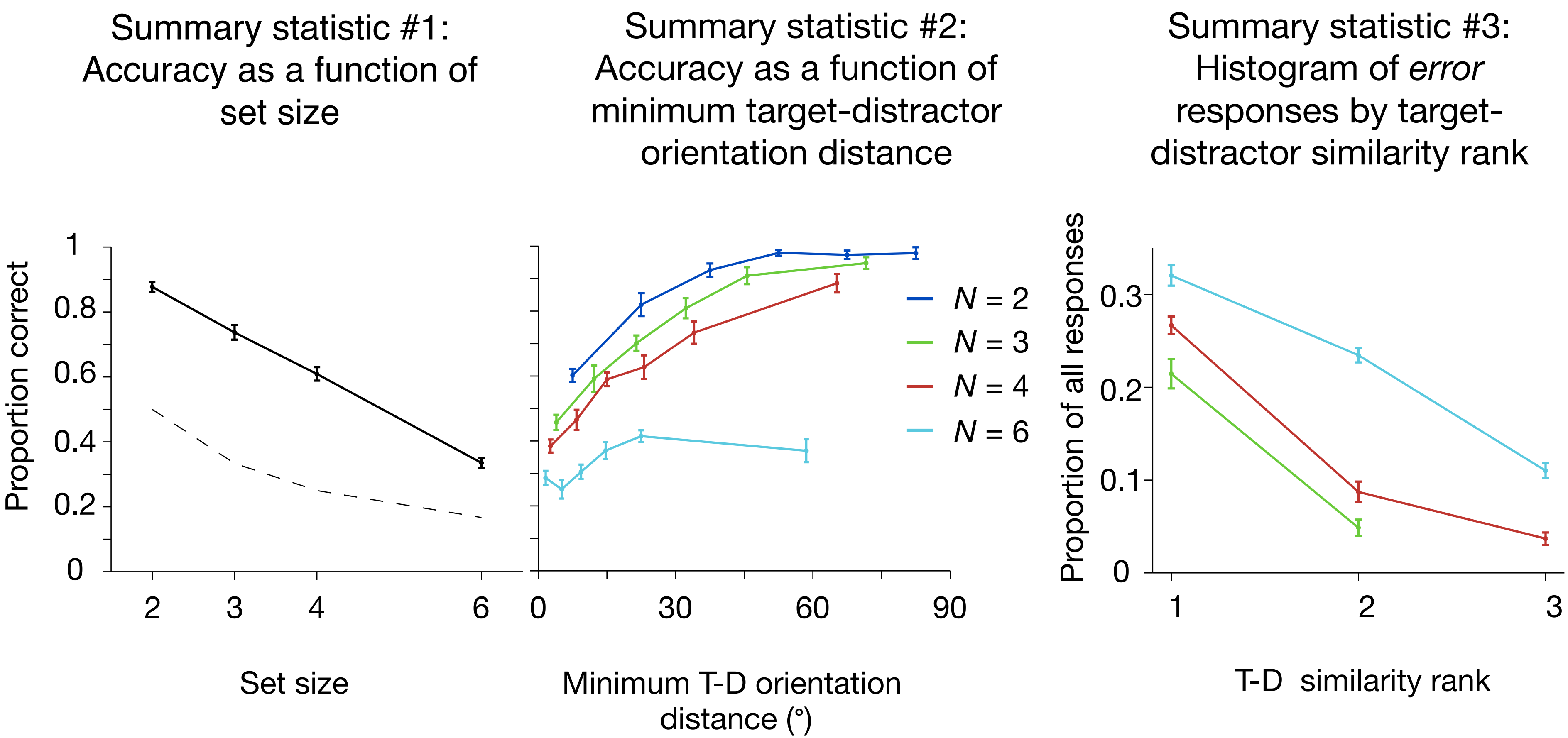
Not just qualitative trends. We will use the exact shapes of these curves to constrain the model.

So far, we have only examined proportion correct.
Next, we will break down the error responses.

Summary statistic #3: Histogram of *error* responses by target-distractor similarity rank

Distractor orientation (°)	Target-distractor orientation distance (°)	Target-distractor similarity rank
-35	65	3
55	25	2
11	19	1





Goal: fit all these curves together (on an individual-subject basis) with an optimal-observer model

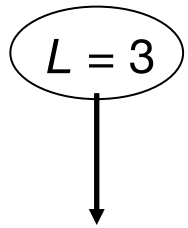
Optimal-observer modeling - two stages:

Encoding: stimuli ➔ noisy measurements

Decision:

- Calculate probabilities of possible target locations based on the set of noisy measurements and knowledge of the encoding process
- Choose the most probable location: this maximizes proportion correct

Optimal-observer model: encoding stage



Each location is equally likely to contain the target

Optimal-observer model: encoding stage

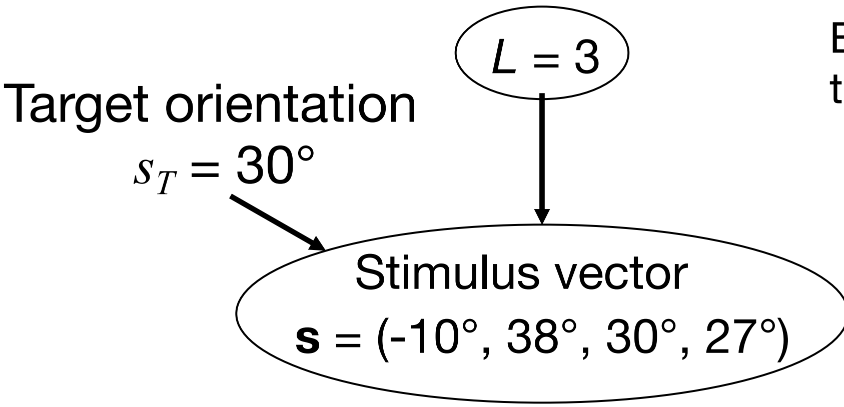
$L = 3$

Target orientation

$s_T = 30°$

Each location is equally likely to contain the target

Optimal-observer model: encoding stage



Each location is equally likely to contain the target

All orientations independently drawn from uniform

Optimal-observer model: encoding stage

$L = 3$

Each location is equally likely to contain the target

Target orientation

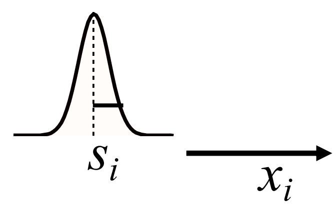
$s_T = 30°$

Stimulus vector

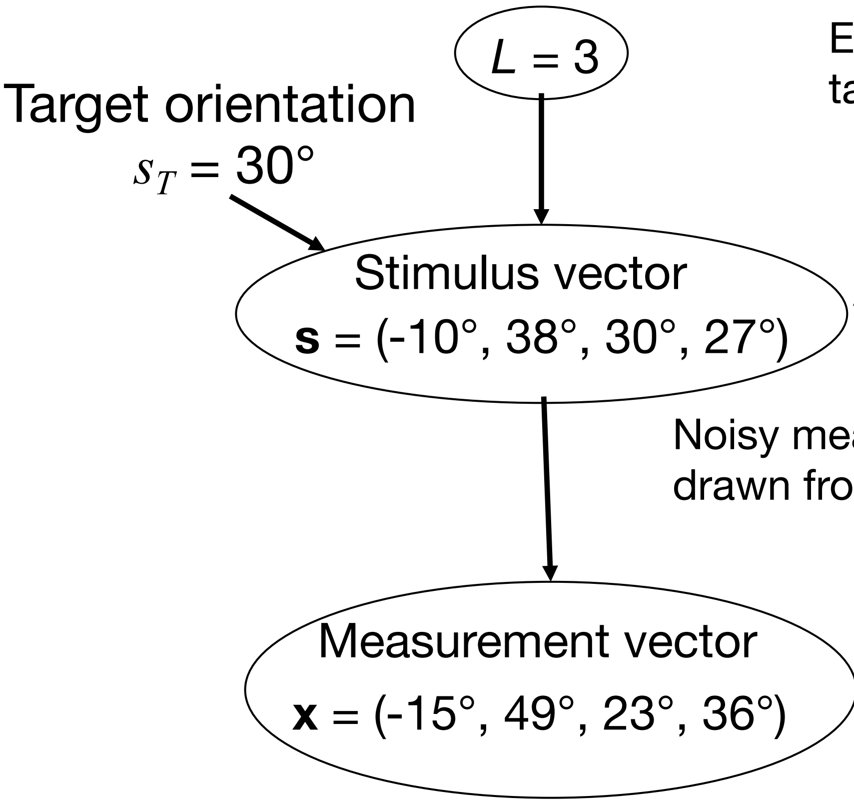
$\mathbf{s} = (-10°, 38°, 30°, 27°)$

All orientations independently drawn from uniform

Noisy measurements of these orientations, independently drawn from Von Mises centered on the true orientation s_i



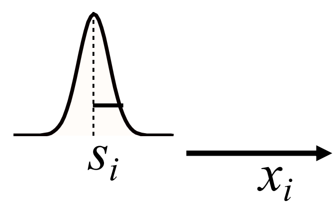
Optimal-observer model: encoding stage



Each location is equally likely to contain the target

All orientations independently drawn from uniform

Noisy measurements of these orientations, independently drawn from Von Mises centered on the true orientation s_i



Optimal-observer model: encoding stage

$L = 3$

Each location is equally likely to contain the target

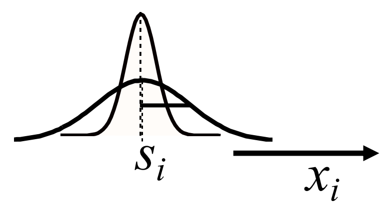
Target orientation
$s_T = 30°$

Stimulus vector
$\mathbf{s}$ = (-10°, 38°, 30°, 27°)

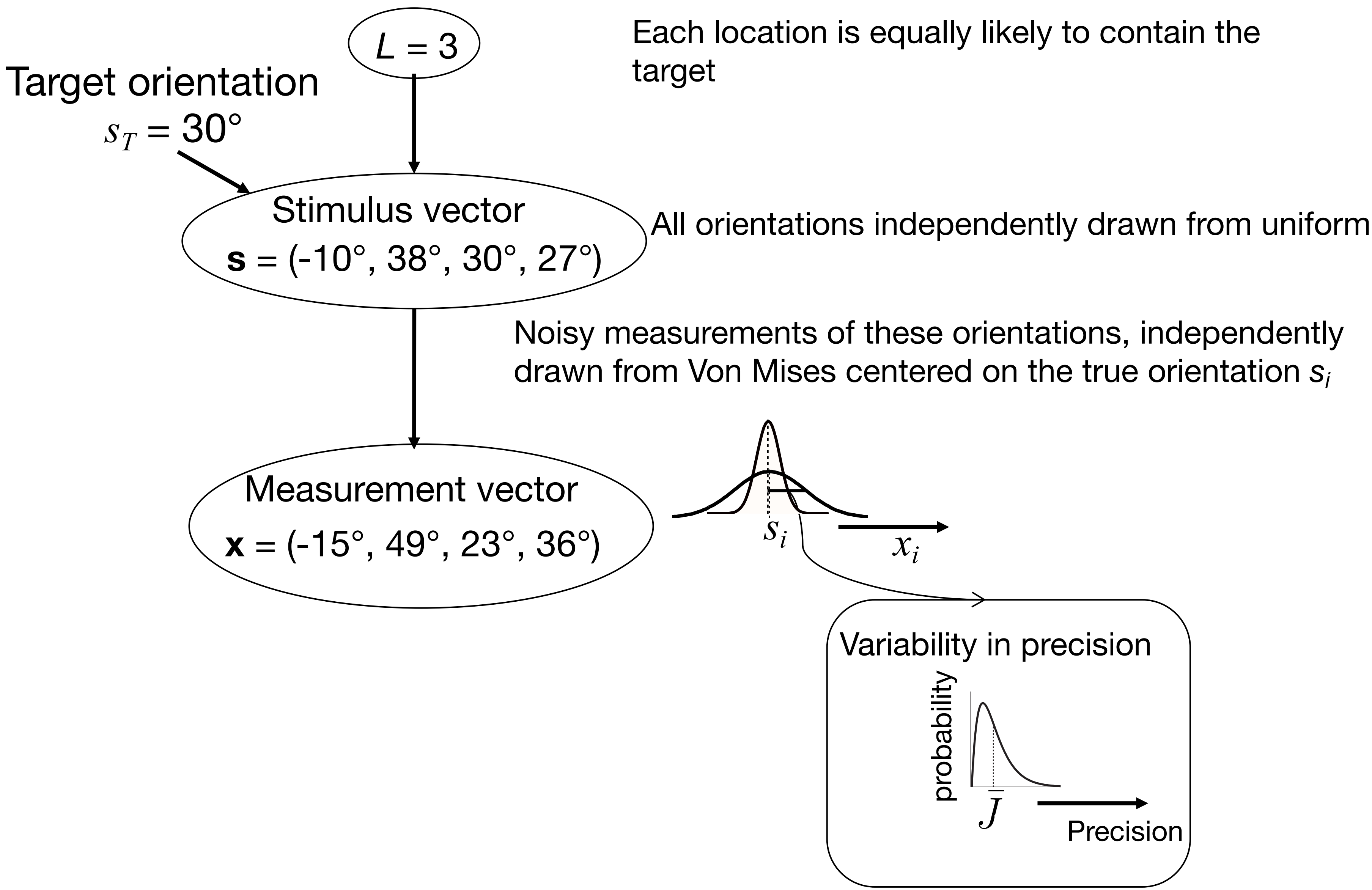
All orientations independently drawn from uniform

Noisy measurements of these orientations, independently drawn from Von Mises centered on the true orientation s_i

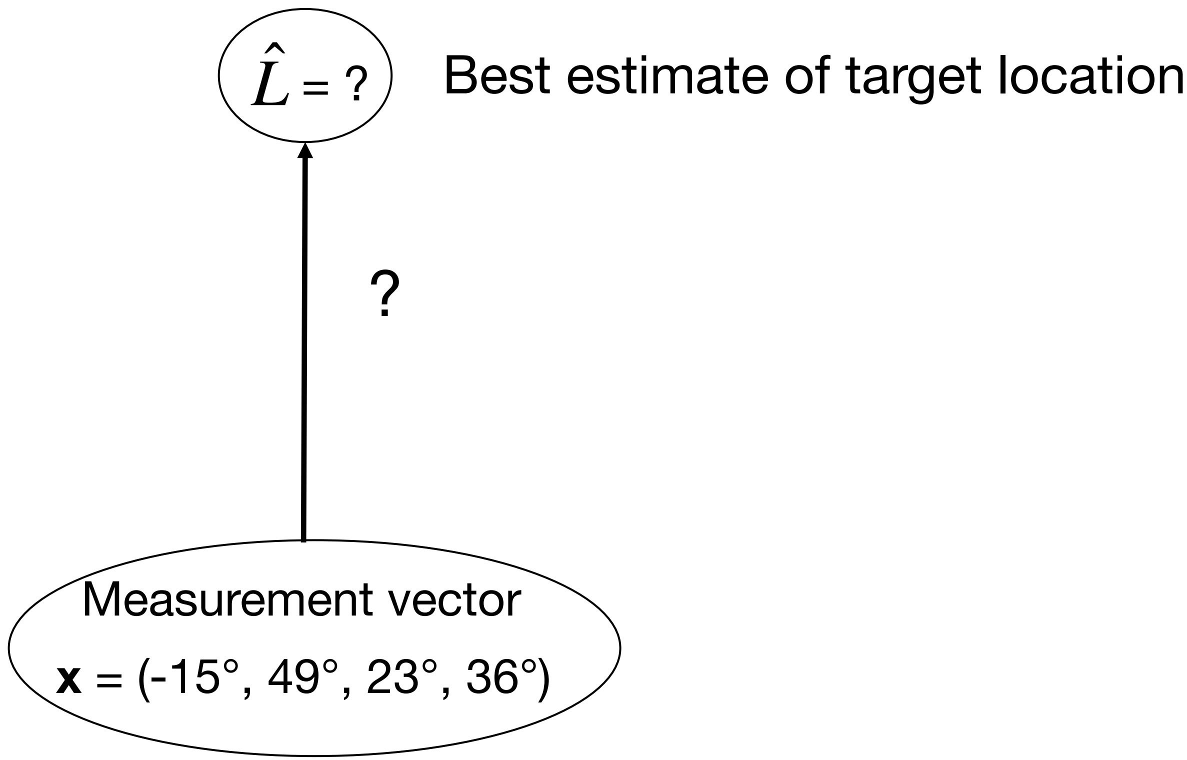
Measurement vector
$\mathbf{x}$ = (-15°, 49°, 23°, 36°)



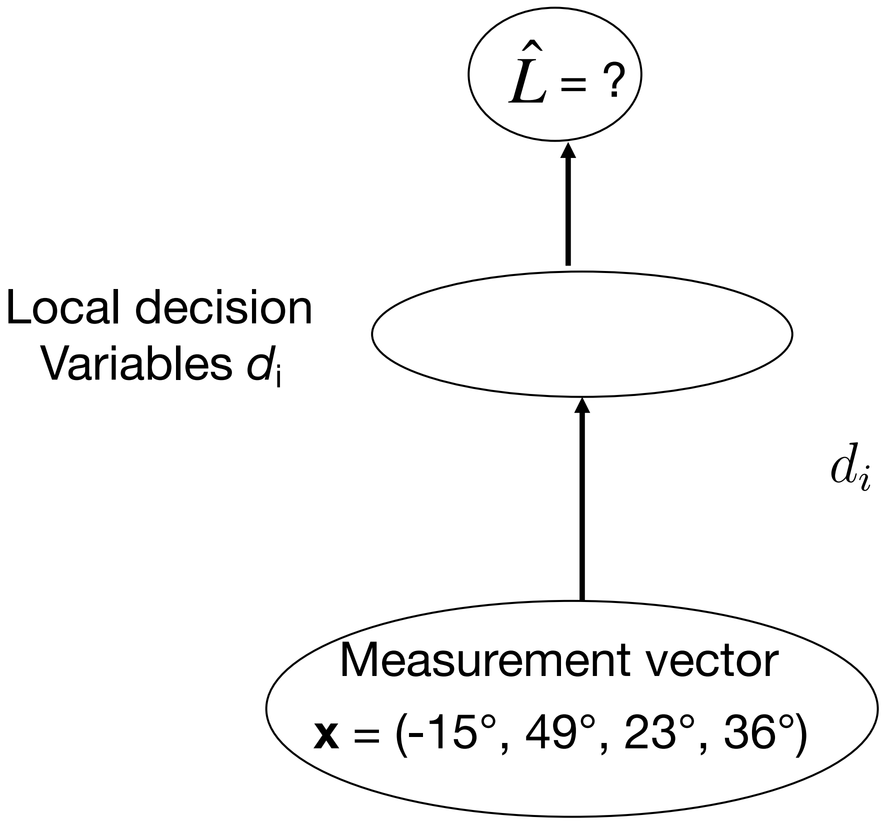
Optimal-observer model: encoding stage



Optimal-observer model: decision stage

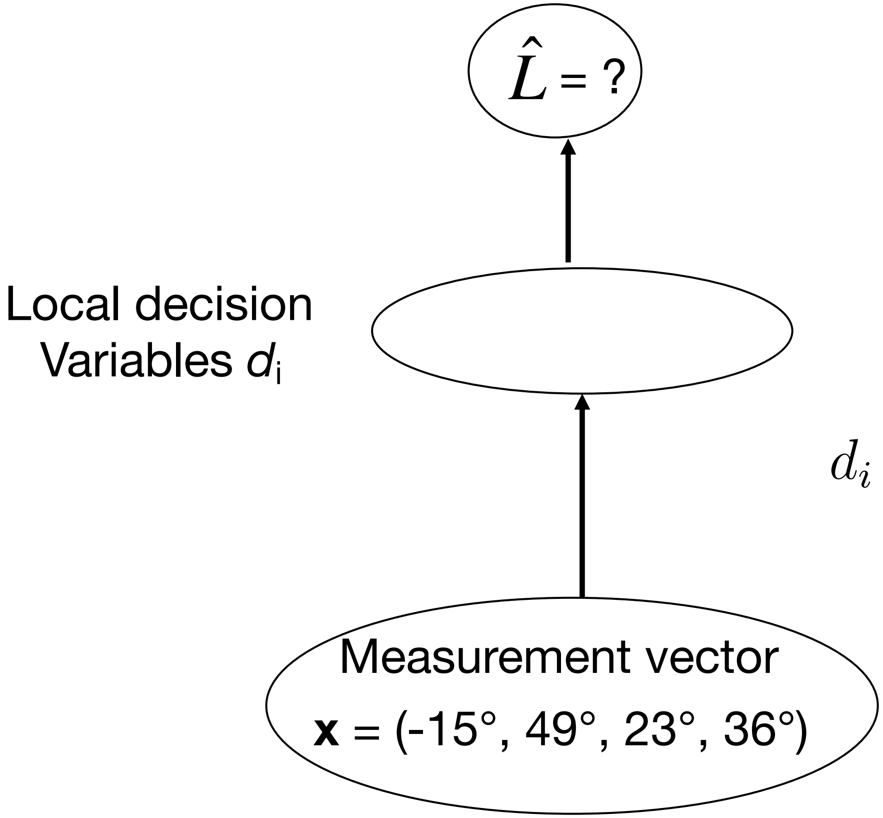


Optimal-observer model: decision stage



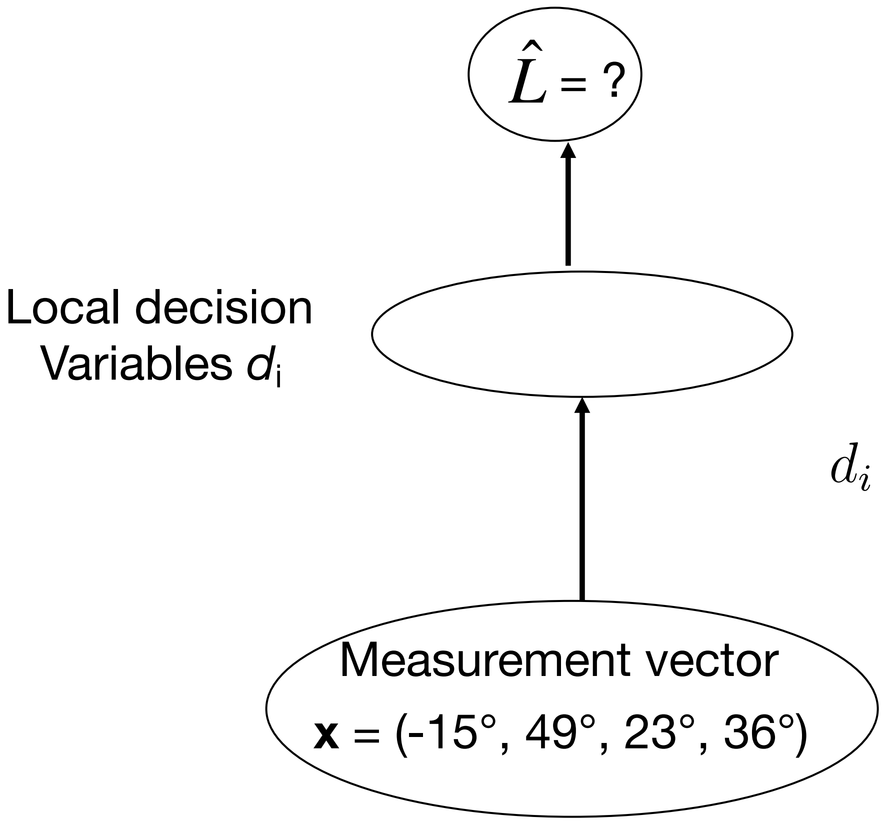
$$d_i \quad = \quad \log \frac{p(x_i|\text{target at location } i)}{p(x_i|\text{distractor at location } i)}$$

Optimal-observer model: decision stage



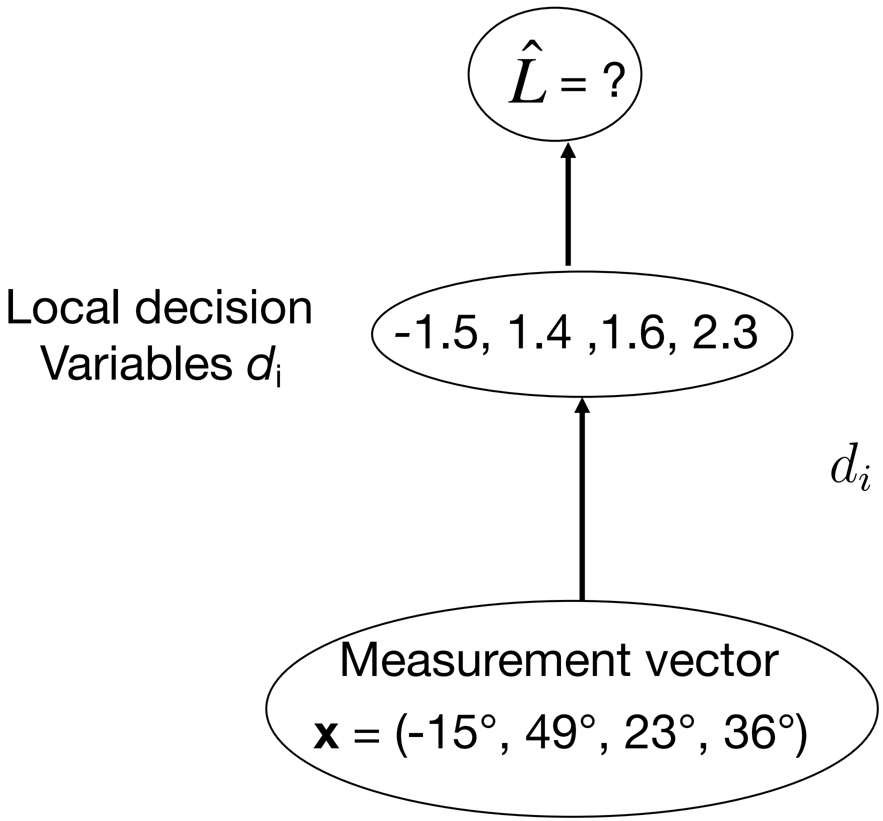
$$
\begin{aligned}
d_i &= \log \frac{p(x_i|\text{target at location } i)}{p(x_i|\text{distractor at location } i)} \\
&= -\log I_0(\kappa_i) + \kappa_i \cos(x_i - s_T)
\end{aligned}
$$

Optimal-observer model: decision stage



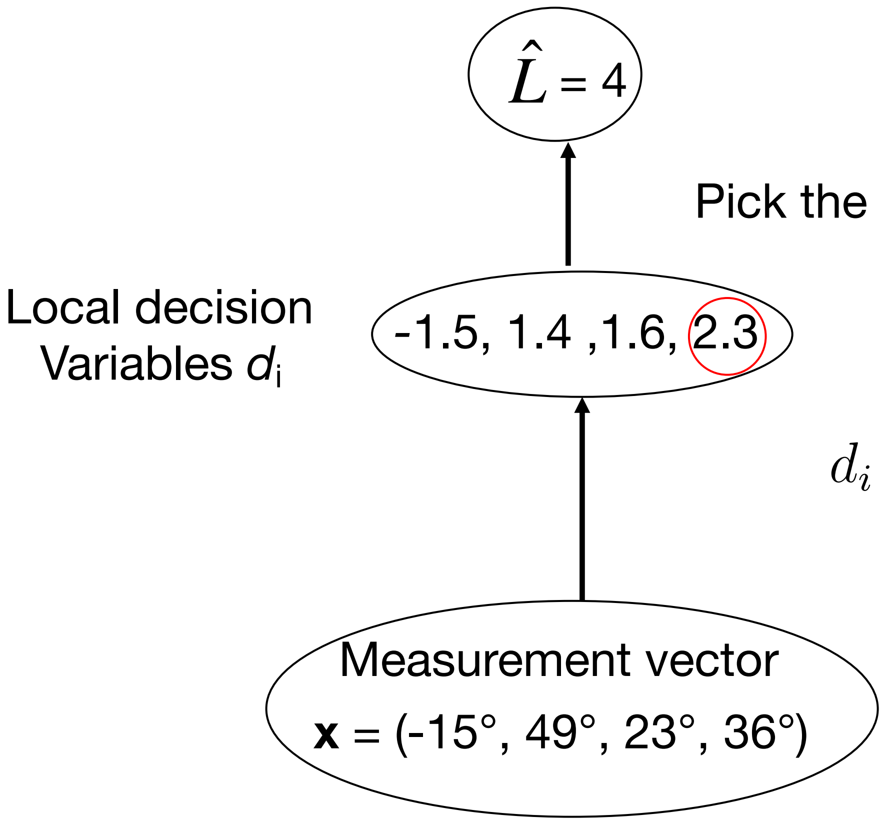
$$
\begin{aligned}
d_i &= \log \frac{p(x_i|\text{target at location } i)}{p(x_i|\text{distractor at location } i)} \\
&= -\log I_0(\kappa_i) + \kappa_i \cos(x_i - s_T)
\end{aligned}
$$

Optimal-observer model: decision stage



$$
\begin{aligned}
d_i &= \log \frac{p(x_i|\text{target at location } i)}{p(x_i|\text{distractor at location } i)} \\
&= -\log I_0(\kappa_i) + \kappa_i \cos(x_i - s_T)
\end{aligned}
$$

Optimal-observer model: decision stage



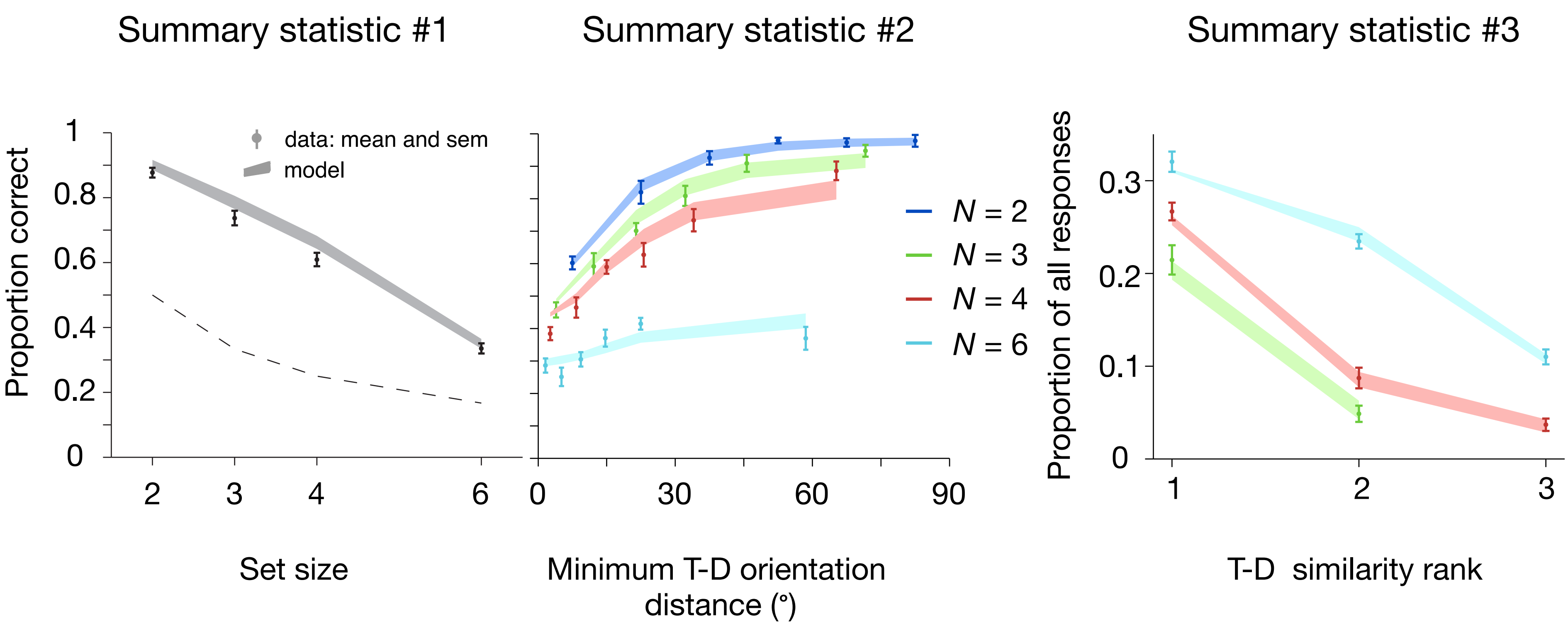
$$
\begin{aligned}
d_i &= \log \frac{p(x_i | \text{target at location } i)}{p(x_i | \text{distractor at location } i)} \\
&= -\log I_0(\kappa_i) + \kappa_i \cos(x_i - s_T)
\end{aligned}
$$

Parameters and model fitting

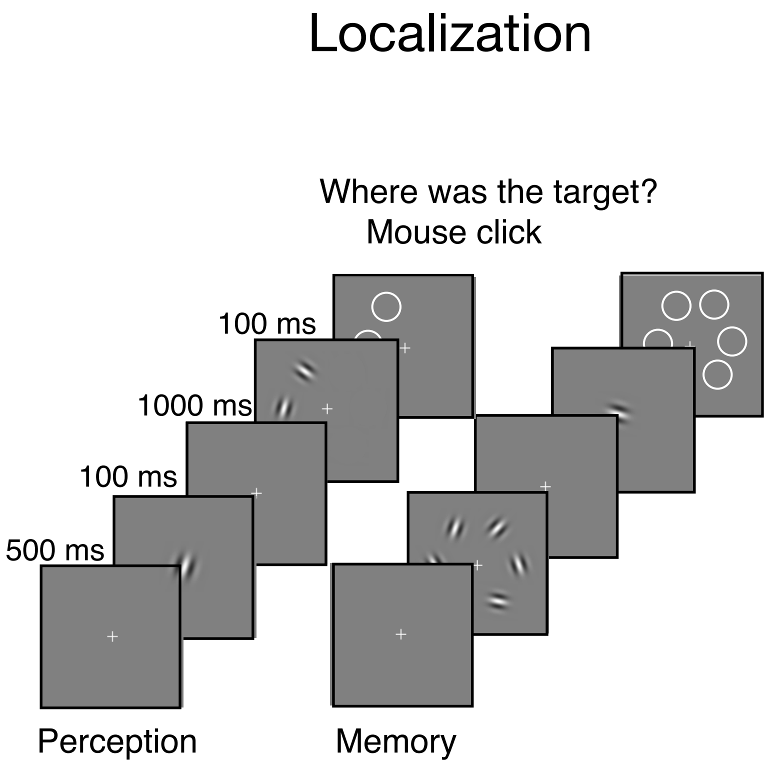
- Five free parameters:
  - Mean precisions at set sizes 2, 3, 4, 6
  - Variability in precision

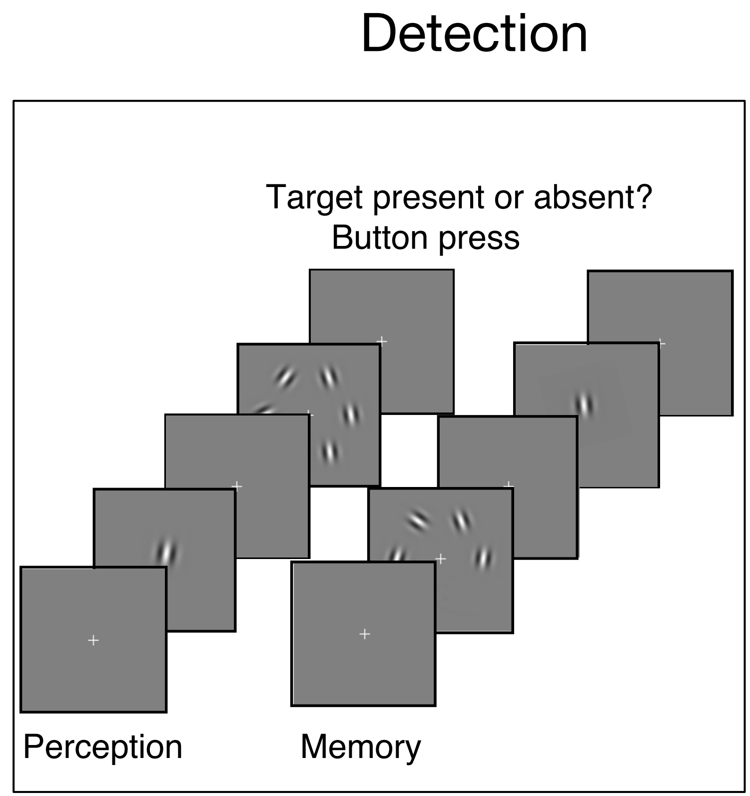
- Simulate predictions $p(\hat{L} \mid \mathbf{s})$ (2000 samples)
- Model fitting:
  - Fit trial-to-trial responses to find parameter estimates that maximize the likelihood, *p*(data | model parameters)
  - Use Bayesian Adaptive Direct Search, *Acerbi and Ma, 2017*
- Model checking: simulate summary statistics with fitted parameters (per subject)

Optimal-observer model captures the localization data very well

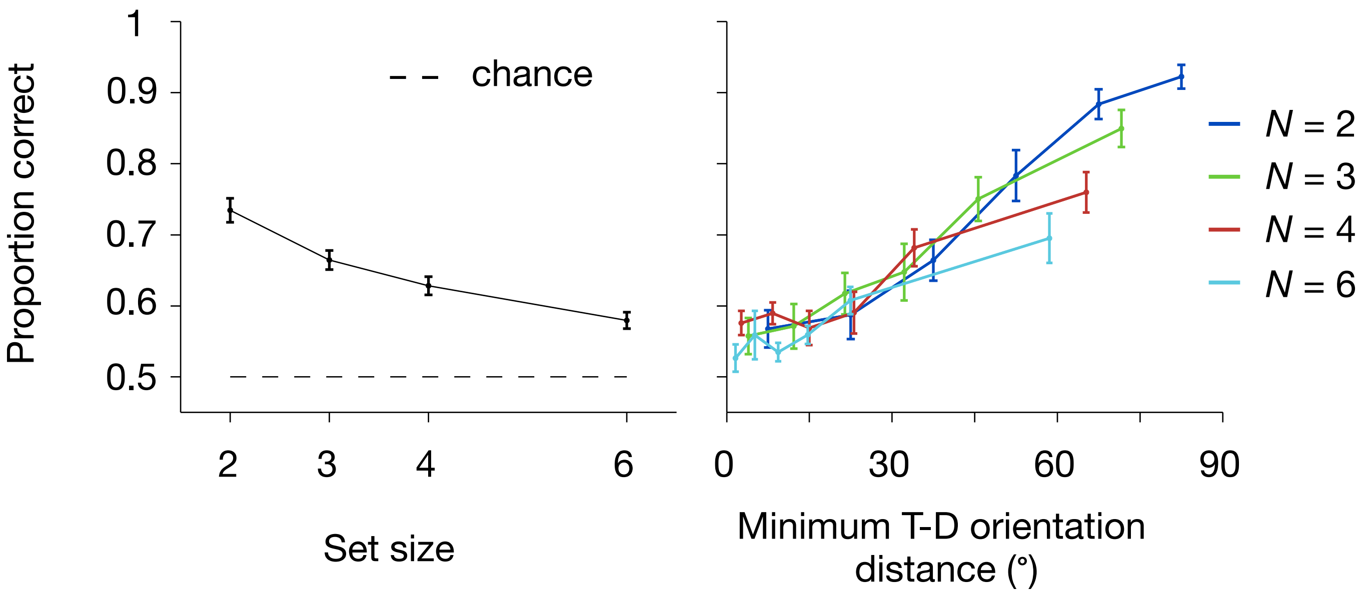


Every subject also did a second task, detection:





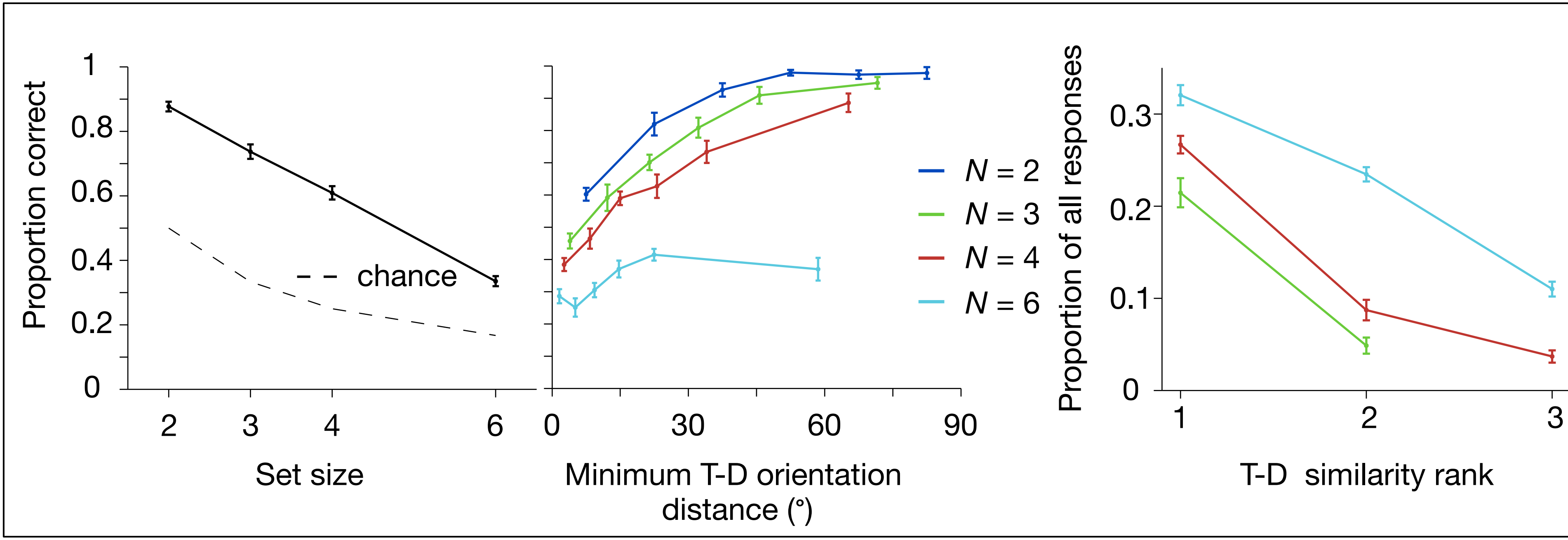
Summary statistics #1 and #2 for detection



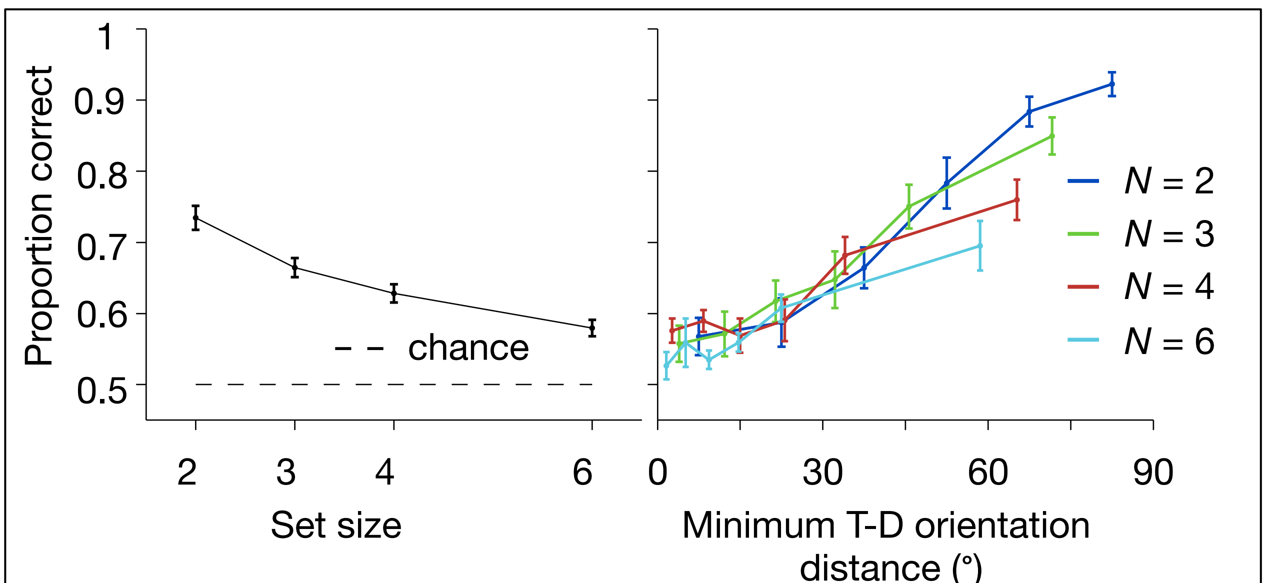
Effect of set size: $p < 0.001$
Effect of time: $p < 0.001$

Effect of set size: $p < 0.001$
Effect of time: $p < 0.001$

Localization



Detection



Goal: fit all these curves together within the same subject *with the same parameters*

Fitting the optimal-observer model *jointly* to localization and detection

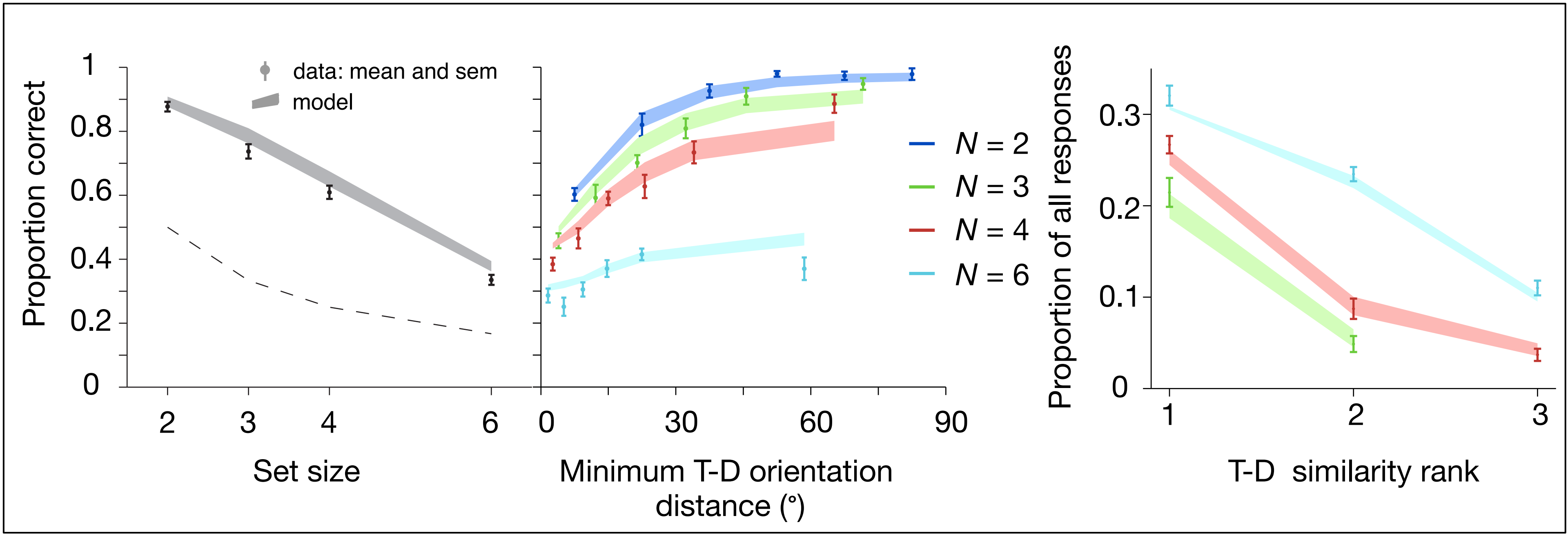
- Same encoding model for both localization and detection
- Slightly different decision rule (in view of task difference)

Ma et al. 2011, Mazyar, van den Berg and Ma 2012

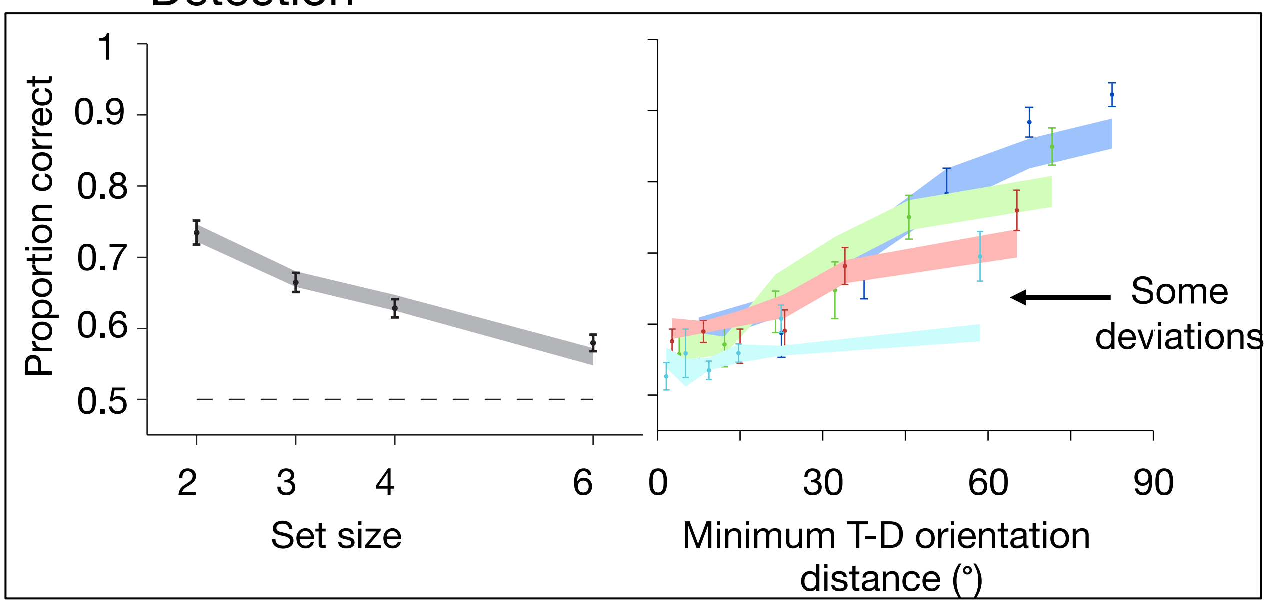
- Per subject, *parameters are shared between tasks*, with one additional parameter for Detection ($p_{present}$): one more parameter, double the data!

Joint fits for perception

Localization



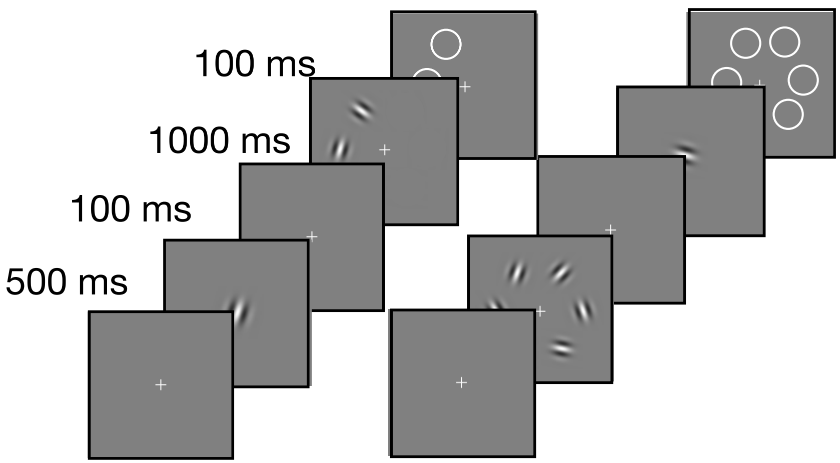
Detection



n-AFC localization and detection

Localization

Where was the target?
Mouse click

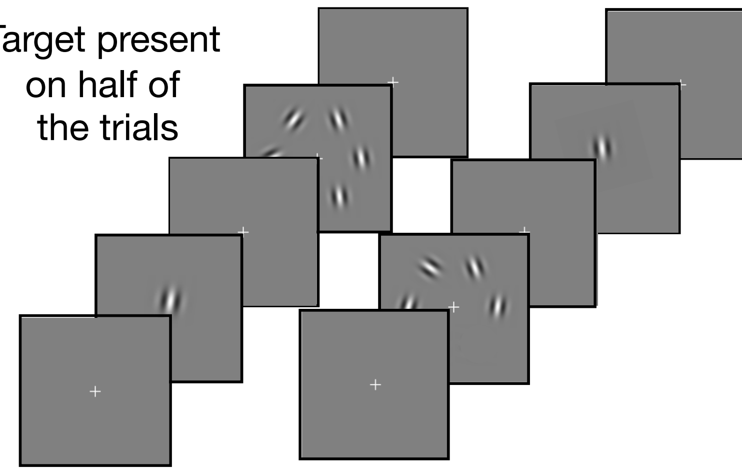


Perception | Memory

Detection

Target present or absent?
Button press

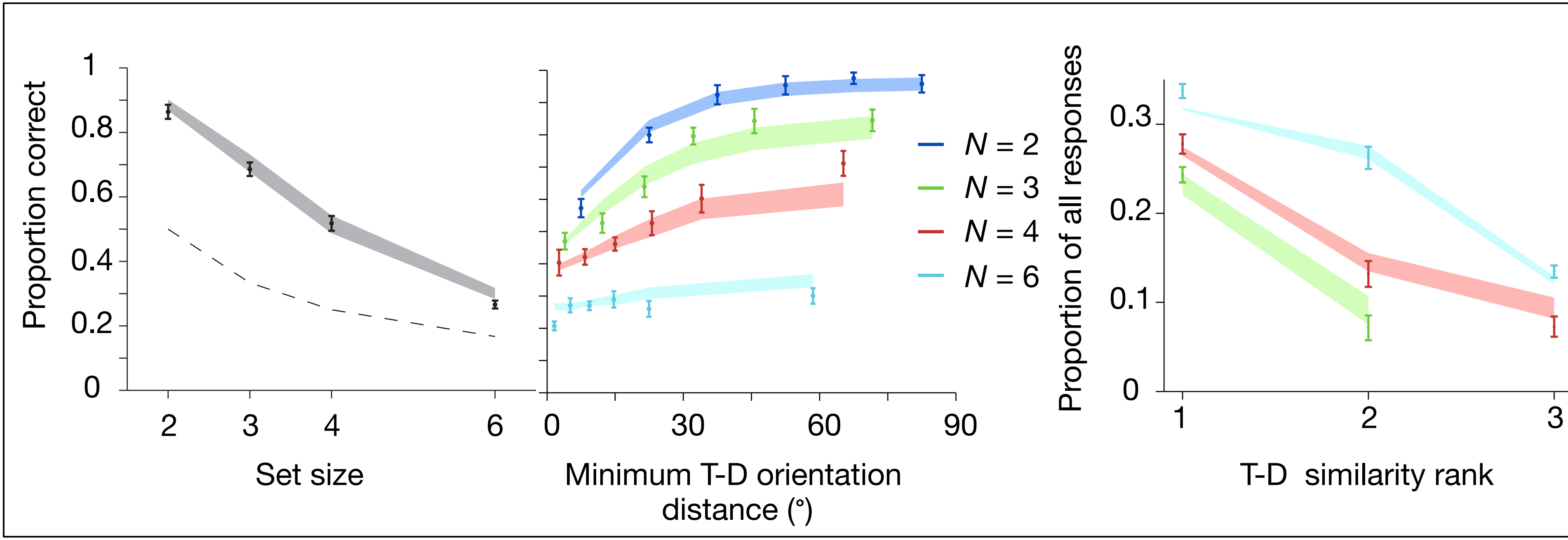
Target present on half of the trials



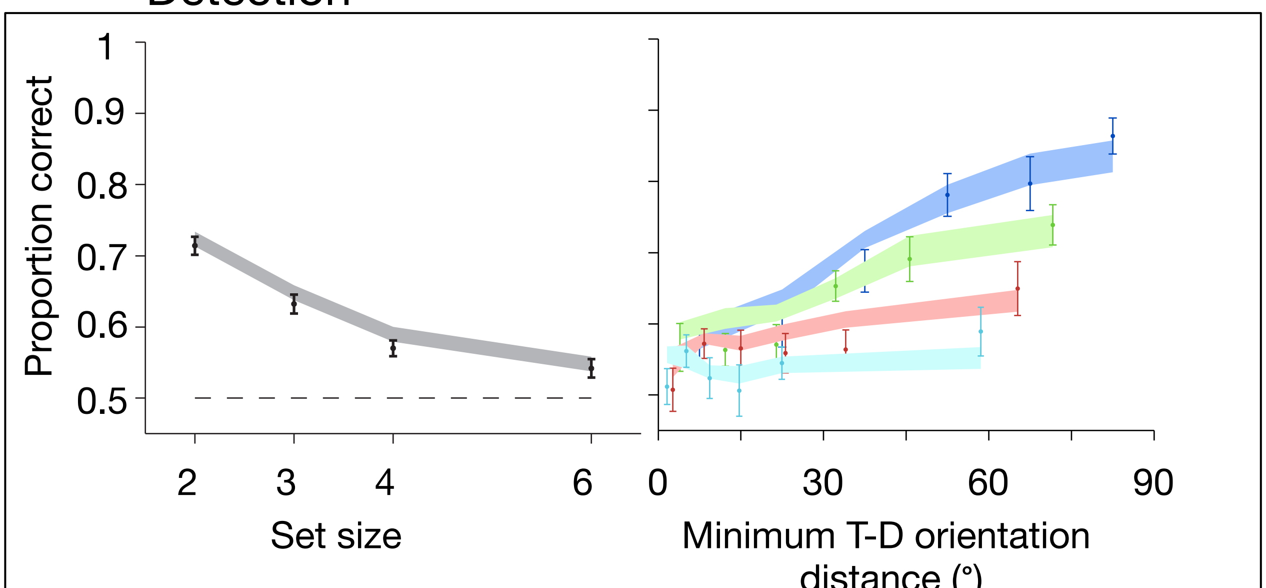
Perception | Memory

Joint fits for memory

Localization

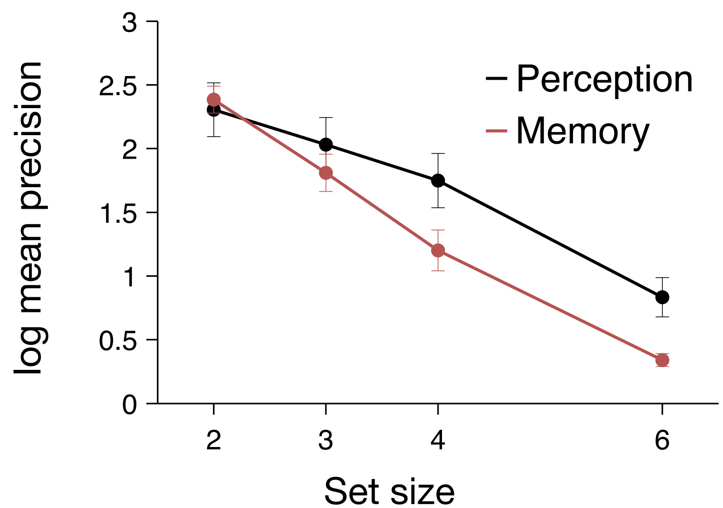


Detection



The same process might underlie perception- and memory-based visual search.

Mean precision parameters decrease with set size in perception and memory



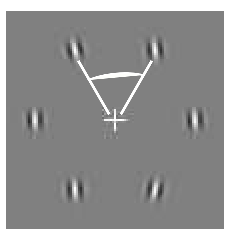
Effect of set size: $p < 0.001$
Effect of Perception/Memory: $p = 0.009$

Does stimulus proximity affect this pattern of results?

- Are the results in Experiment 1 conditional on the far spacing of stimuli?
  - There is some reason to believe that observers might swap items across locations *Bays 2009, 2016*
- Experiment 2 repeats Experiment 1 with a smaller spacing, still outside the Bouma distance for crowding *Bouma 1971*

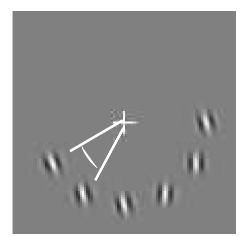
Experiment 1

60 deg



Experiment 2

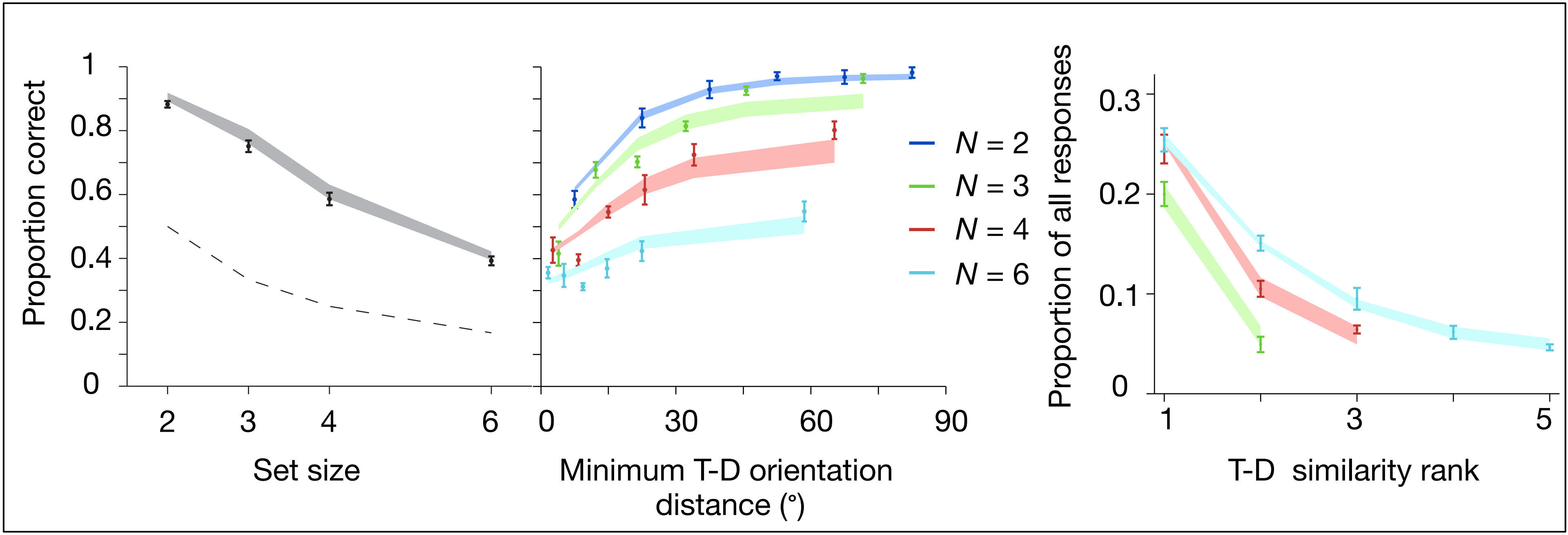
30 deg



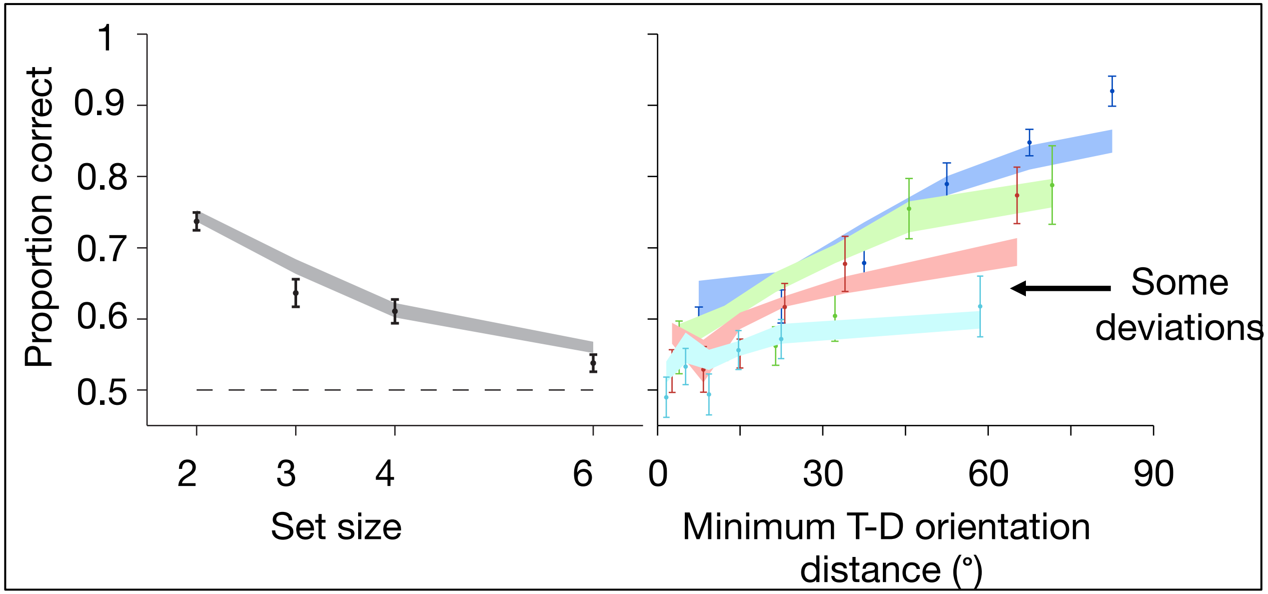
- Does the optimal-observer model still capture the data?

Smaller spacing: Joint fits for perception

Localization

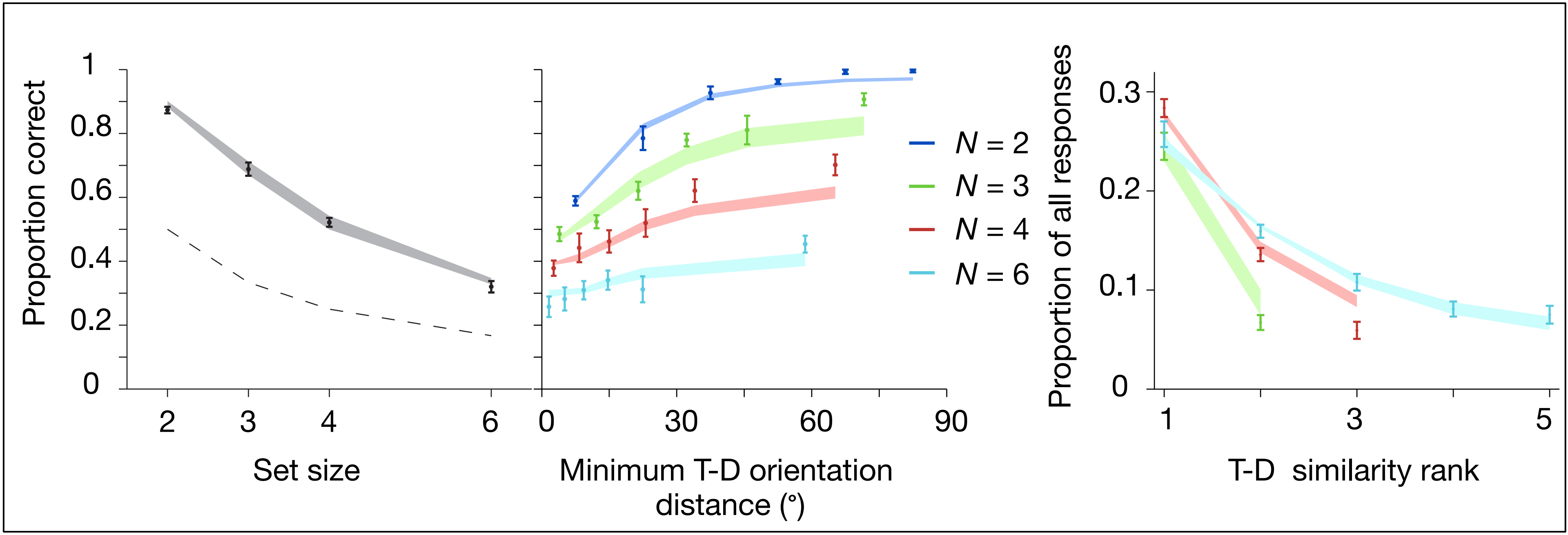


Detection

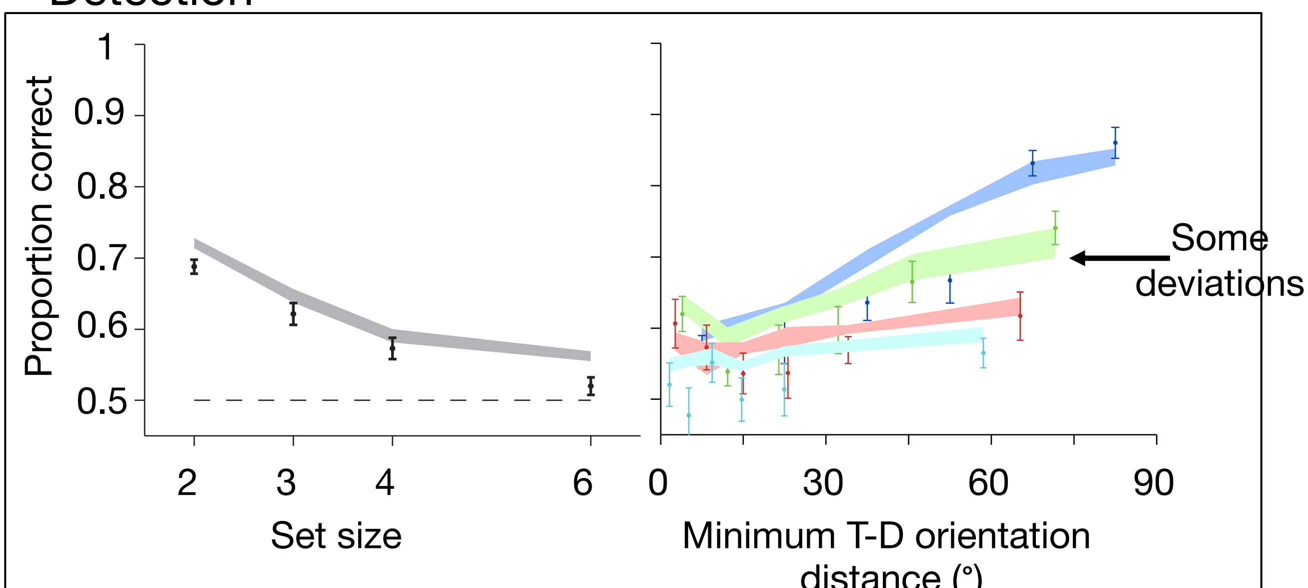


Smaller spacing: Joint fits for memory

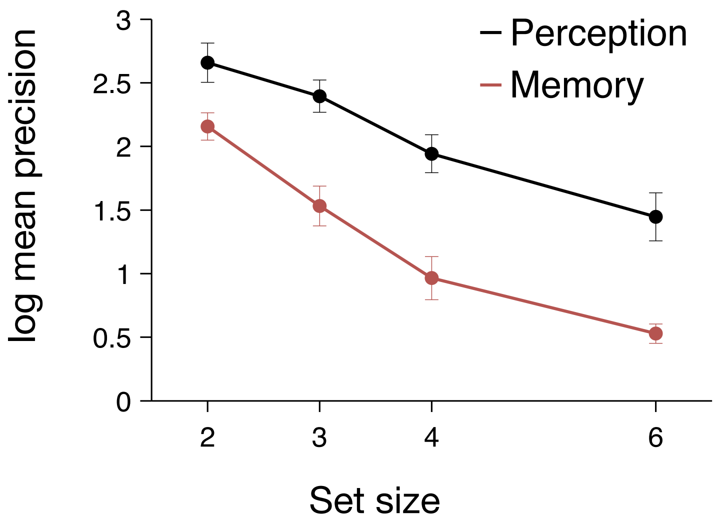
Localization



Detection



Smaller spacing: Mean precision parameters decrease with set size in perception and memory



Effect of set size: *p* < 0.001
Effect of Perception/Memory: *p* = 0.003

Conclusions part 1

- Goal: study visual search with heterogeneous distractors (towards real-world search), but maintain modelability
- Varied set size, task, timing condition (perception/memory) and spacing
- Accuracy decreases with increasing set size
- Accuracy increases with increasing minimum target-distractor orientation distance
- In localization, distractors more similar to the target were chosen more often
- Optimal-observer model captured the localization data well
  - Same model captured the joint localization and detection data, suggesting shared encoding processes
  - However, some deviations in the detection task
- Same process model for perception and memory-based search
  - In both Perception and Memory, precision decreases with set size
  - Precision Perception > Precision Memory
- Similar results with higher stimulus proximity

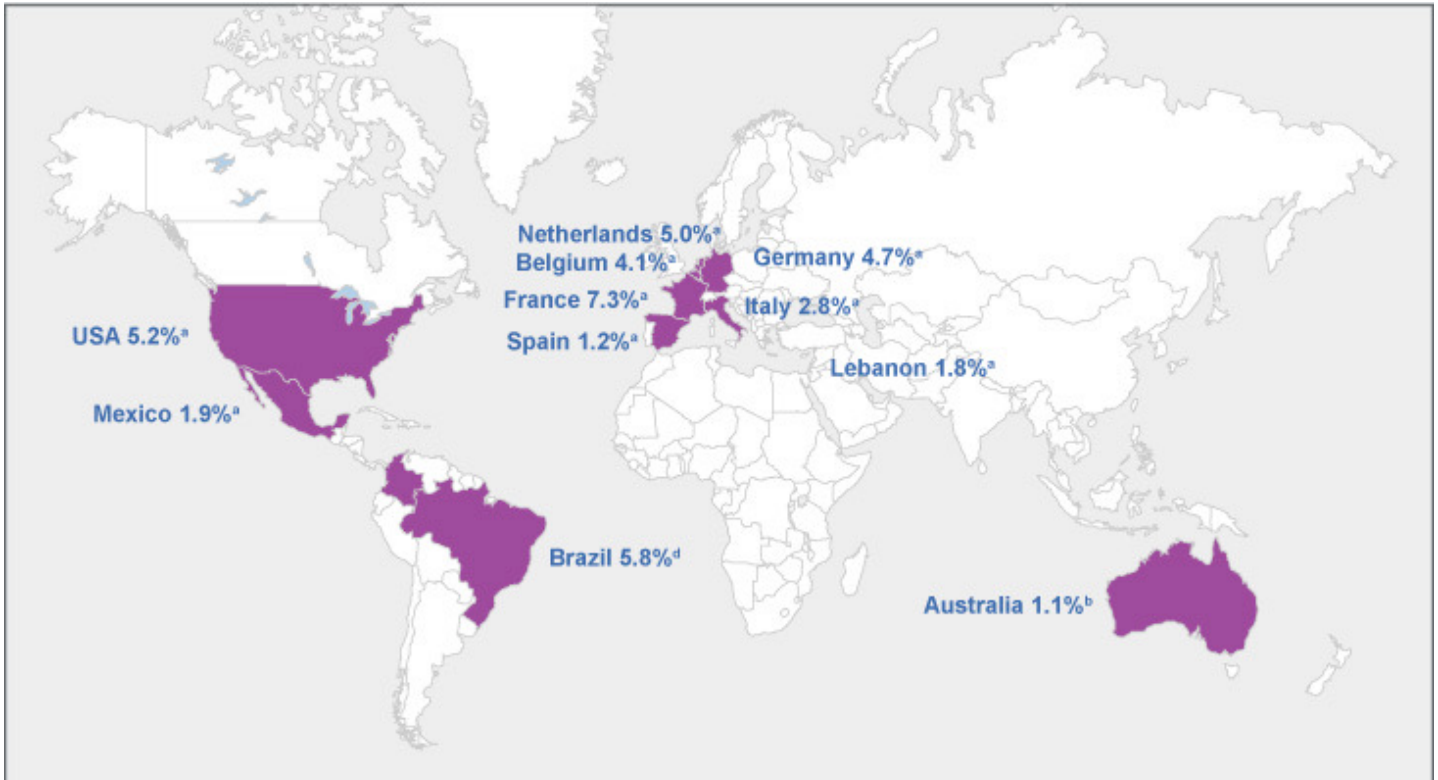
Part 2: Spatial and feature cueing and task-switching in neurotypicals and ADHD

*Allison Young**

Lenard Adler

Michael Halassa

Attention-deficit/hyperactive disorder (ADHD) is highly prevalent around the world



Data taken from (a) Fayyad J, de Graaf R, Kessler R, et al. Br J Psychiatry 2007; 190: 402-409. (b) Ebejer JL, Medland SE, van der Werf J et al. PLoS One 2012; 7: e47404. (c) de Zwaan M, Gruss B, Muller A, et al. Eur Arch Psychiatry Clin Neurosci 2012; 262: 79-86. (d) Polanczyk G, Laranjeira R, Zaleski M, et al. Int J Methods Psychiatr Res 2010; 19: 177-84.

Source: http://adhd-institute.com/burden-of-adhd/epidemiology/

ADHD diagnosis example questions

Adult ADHD Self-Report Scale

	Never	Rarely	Sometimes	Often	Very Often
1. How often do you have trouble wrapping up the final details of a project, once the challenging parts have been done?					

2. How often do you have difficulty getting things in order when you have to do a task that requires organization?

Inattention

.

.

.

15. How often do you find yourself talking too much when you are in social situations?

Hyperactivity/
Impulsivity

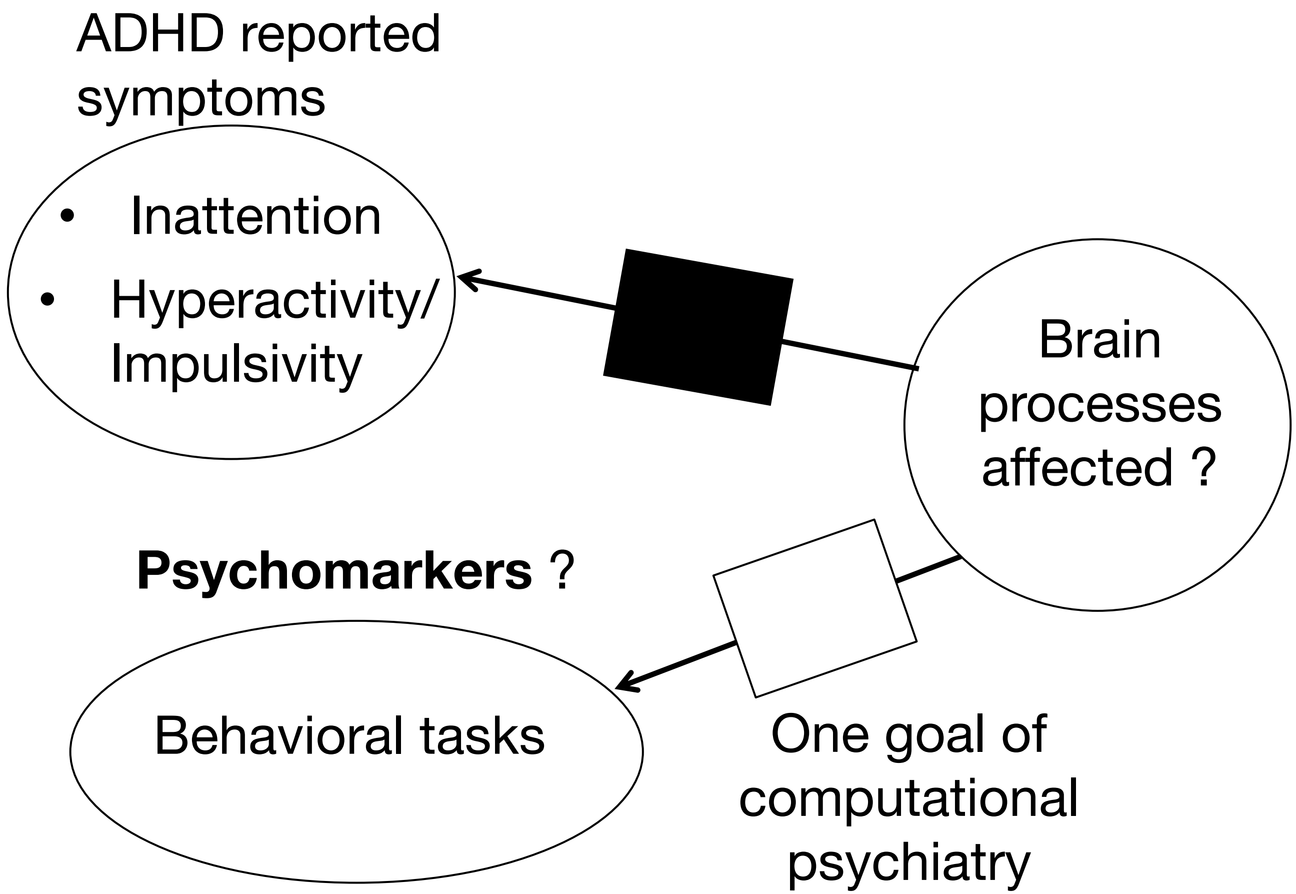
.

.

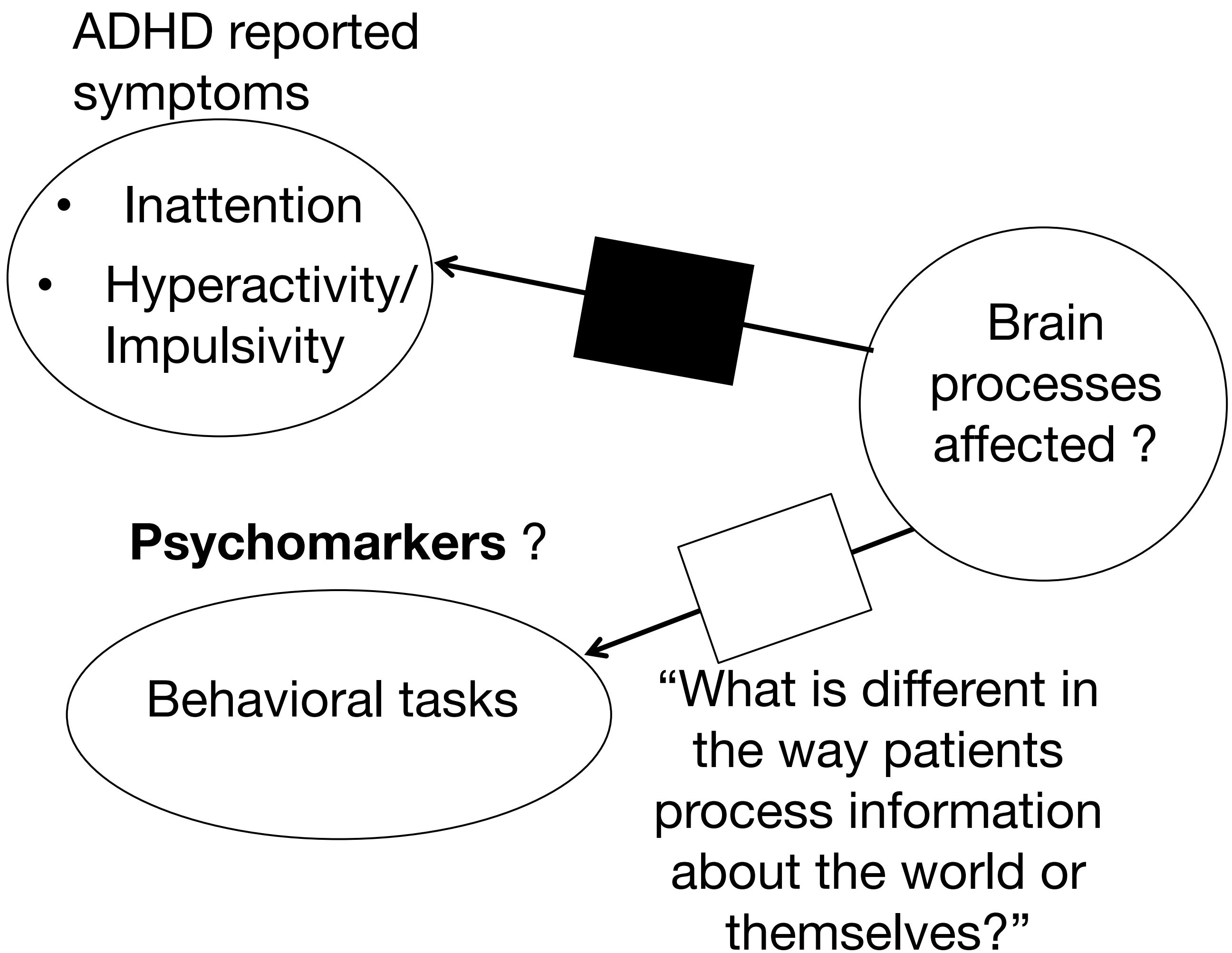
.

These survey questions might contain some inherent subjectivity and give insufficient insight into the underlying processes.

Searching for new psychomarkers in ADHD



Searching for new psychomarkers in ADHD

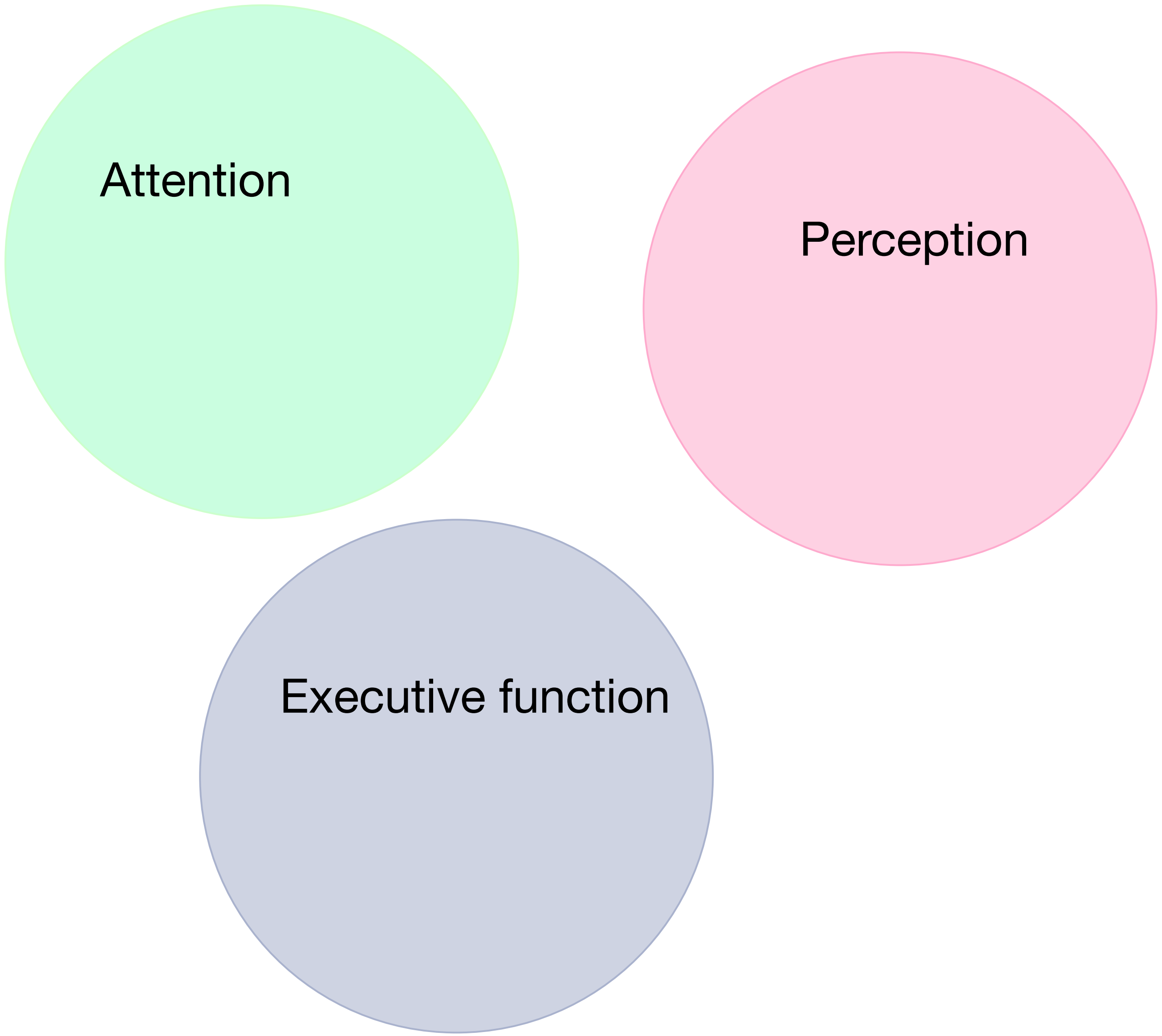


Redish and Gordon 2016

Participants

- 20 ADHD and 20 Controls
- Recruited through local advertisements, including flyers, newspaper and radio advertisements – part of a larger study at NYUMC
- Controls matched to ADHD in age and gender distributions
  - ADHD: 11F, 9M. Control: 12F, 8M
  - ADHD: 32.5 ± 6.1 yo. Control: 35.3 ± 10.0 yo
- A trained clinician assessed every participant using the following scales:
  - Adult ADHD Clinician Diagnostic Scale (ACDS)
  - Adult ADHD Investigator Symptom Rating Scale (AISRS)
  - Clinical Global Impressions-Severity of Illness (CGI-S) Scale
  - M.I.N.I International Neuropsychiatric Interview
- Participants also completed self-report scales:
  - Adult ADHD Self-Report Scale (ASRS v.1.1.)
  - Adult ADHD Quality of Life (AAQoL) Scale
  - World Health Organization Disability Assessment Schedule (WHODAS- II)
  - Behavior Rating Inventory of Executive Function Adult Version (BRIEF-A)

What is impaired in ADHD?



What is impaired in ADHD?

- Divided attention ?
- Selective attention ?
- Sustained attention/vigilance?
- Flexibility?

Van Zomeren and Brouwer, 1994
Huang-Pollock and Nigg 2003
Roberts et al. 2017

Perception

Executive function

What is impaired in ADHD?

- Divided attention ?
- Selective attention ?
- Sustained attention/vigilance?
- Flexibility?

~No

Van Zomeren and Brouwer, 1994
Huang-Pollock and Nigg 2003
Roberts et al. 2017

Perception

Executive function

What is impaired in ADHD?

- Divided attention ?
- Selective attention ?
- Sustained attention/vigilance?
- Flexibility?

~No

Van Zomeren and Brouwer, 1994
Huang-Pollock and Nigg 2003
Roberts et al. 2017

Perception

~No

Fuermaier et al. 2017

Executive function

What is impaired in ADHD?

- Divided attention ?
- Selective attention ?
- Sustained attention/vigilance?
- Flexibility?

Van Zomeren and Brouwer, 1994
Huang-Pollock and Nigg 2003
Roberts et al. 2017

~No

Perception

Fuermaier et al. 2017

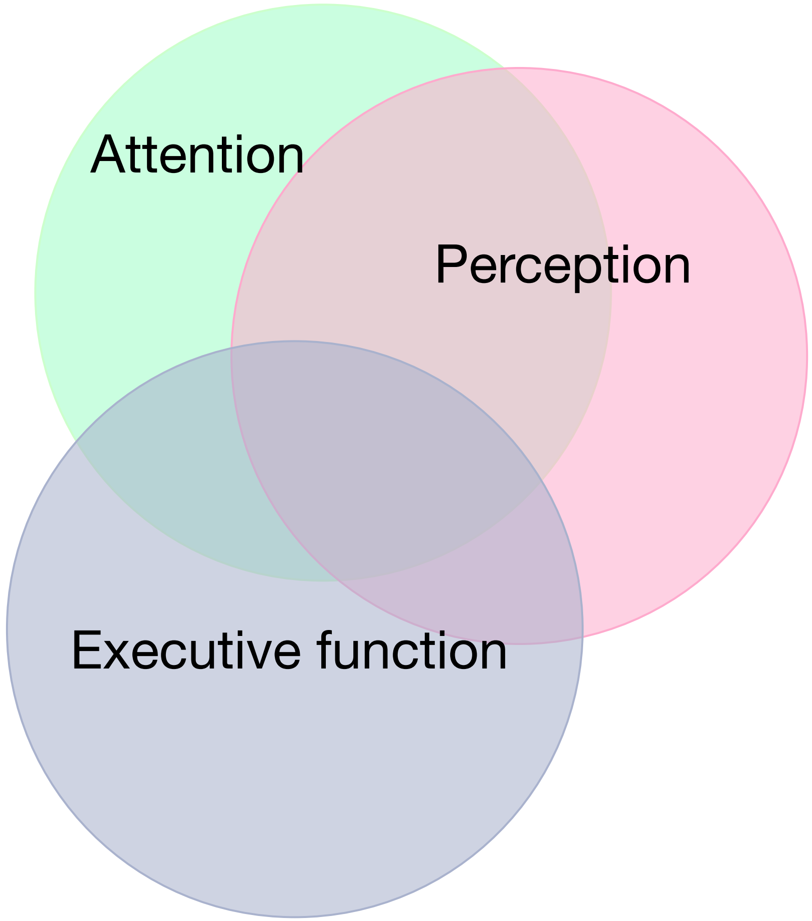
~No

- Cognitive control / set shifting / flexibility
- Response inhibition
- Working memory

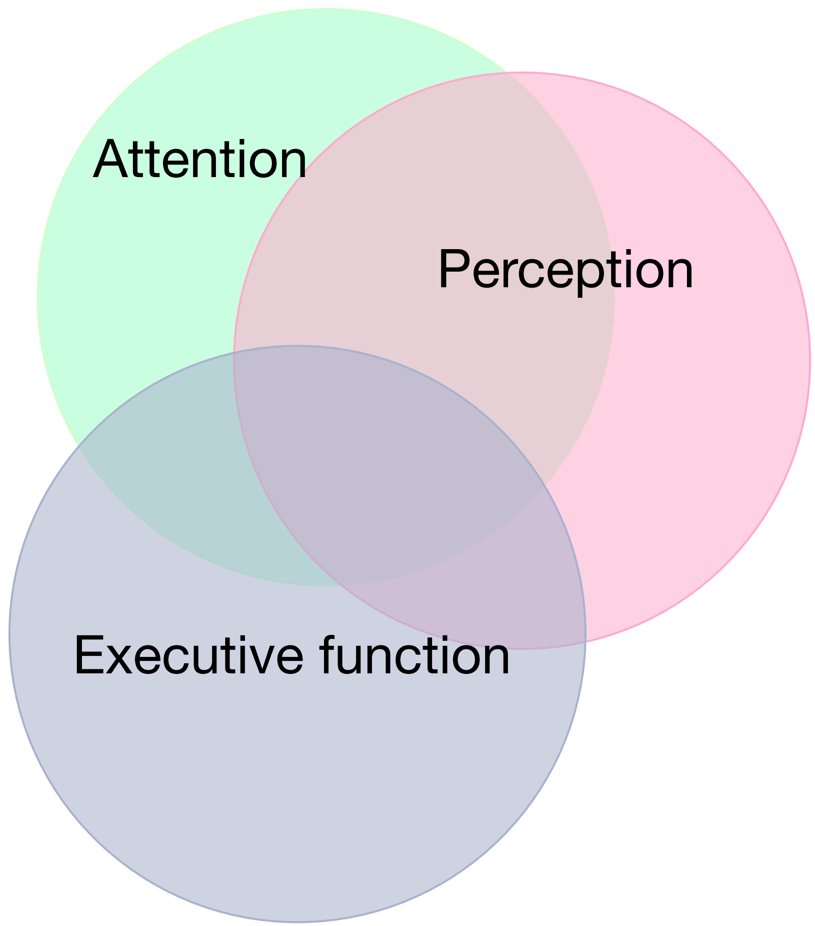
Diamond 2013
Willcutt 2005
de Zeeuw and Durston 2017

~Yes

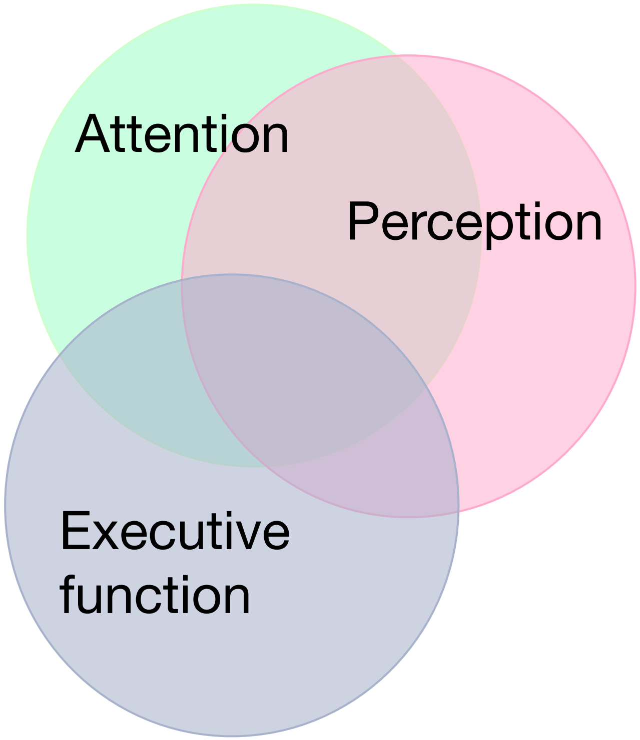
Hypothesis: Perceptual deficits emerge if attention and executive function are simultaneously taxed.



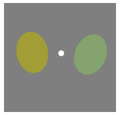
Goal: get separate measures of perceptual encoding precision and executive function



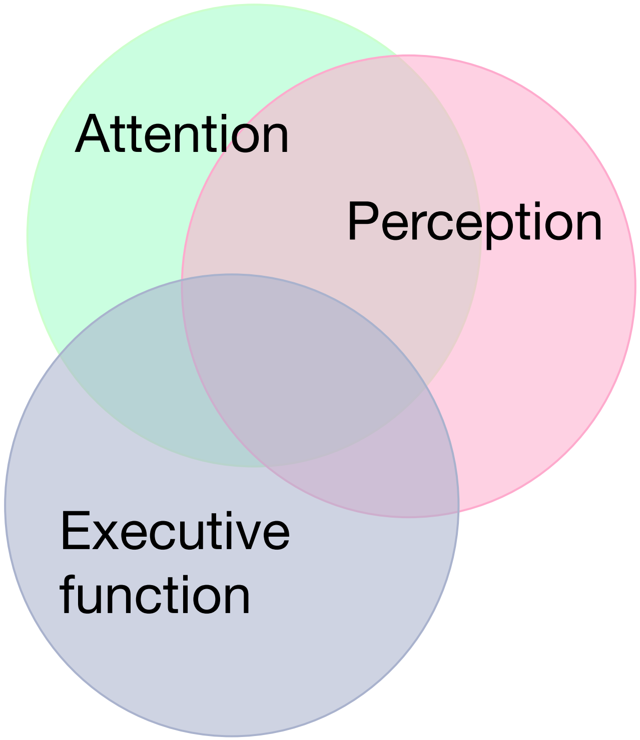
Our approach



Stimuli



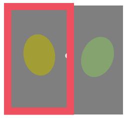
Our approach



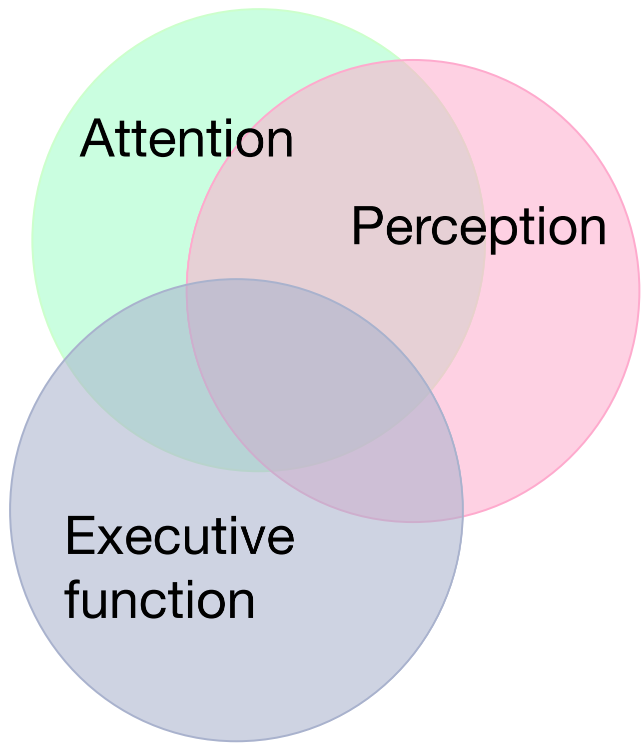
Stimuli

yes-no

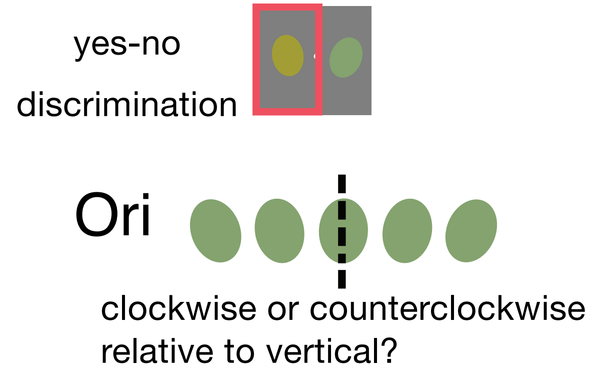
discrimination



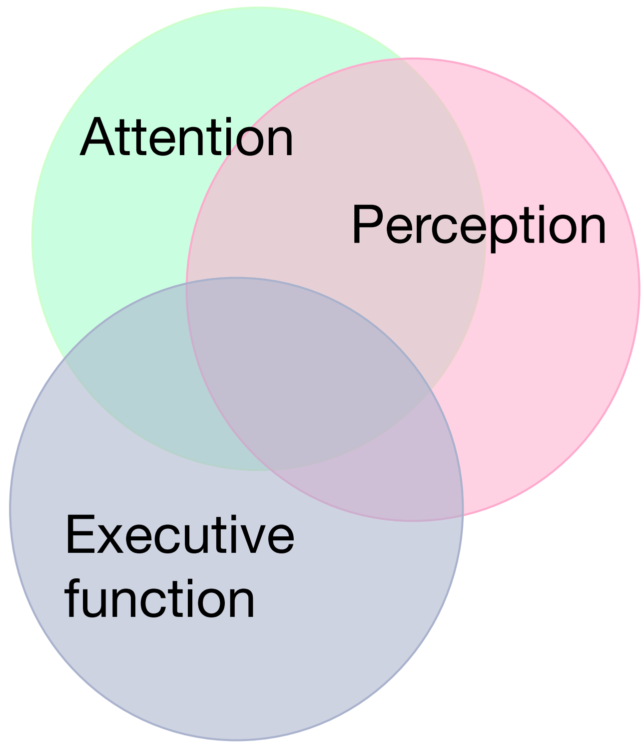
Our approach



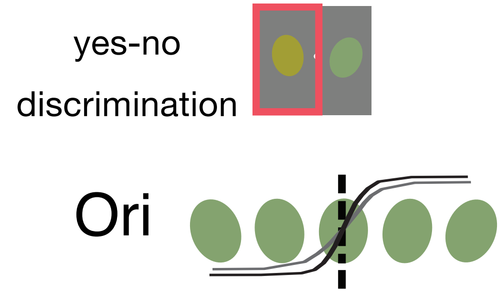
Stimuli



Our approach



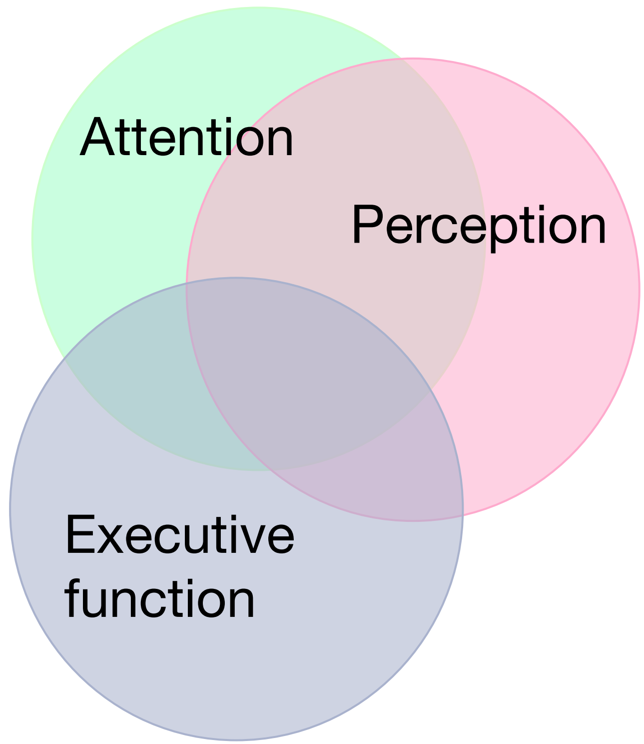
Stimuli



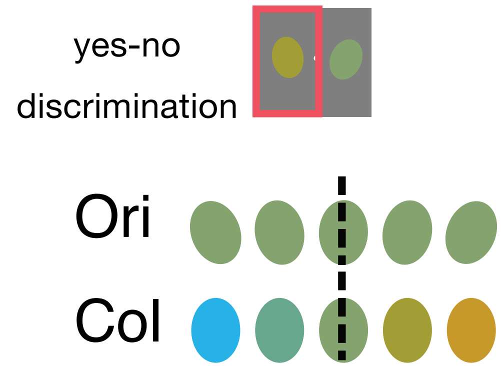
clockwise or counterclockwise relative to vertical?

Fit psychometric curves

Our approach

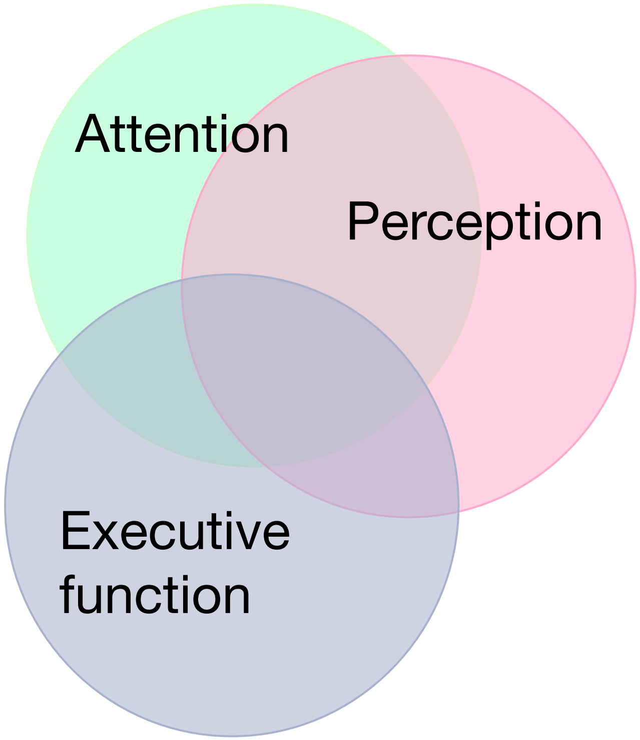


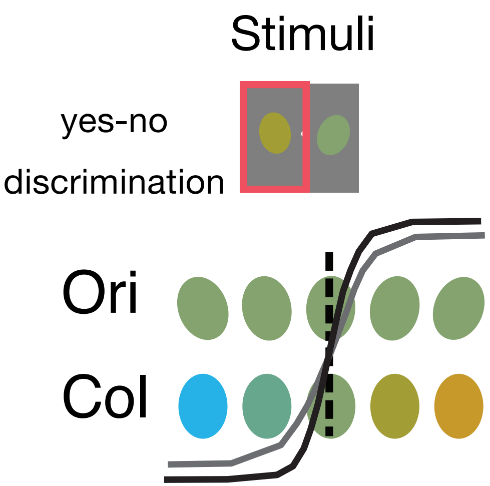
Stimuli



more yellow or more blue
relative to middle green?

Our approach



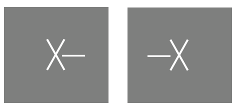


Our approach

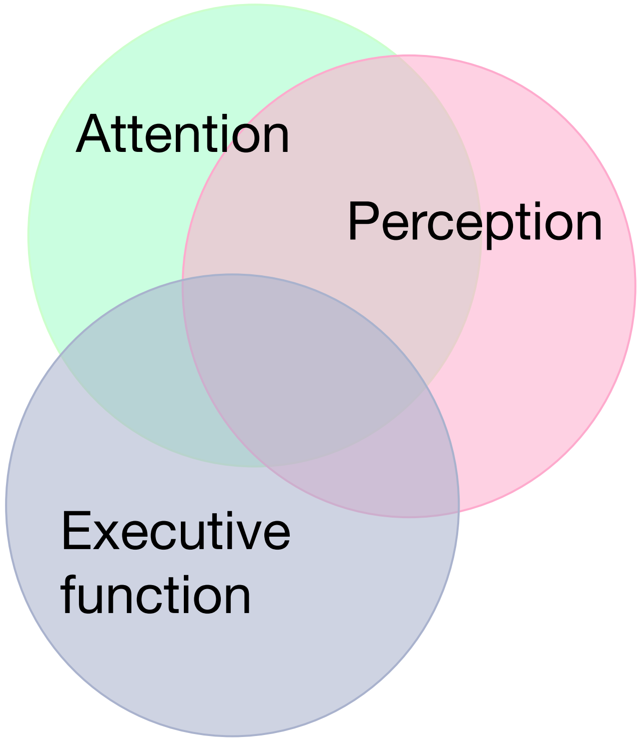
Selective attention:
Spatial cues

Feature cues

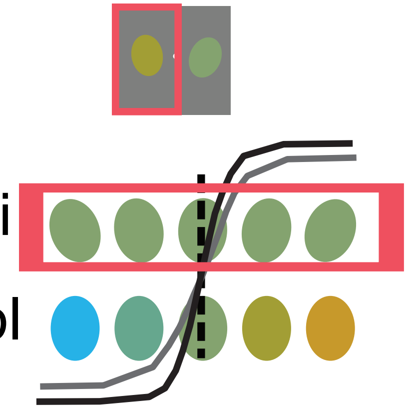
Ori



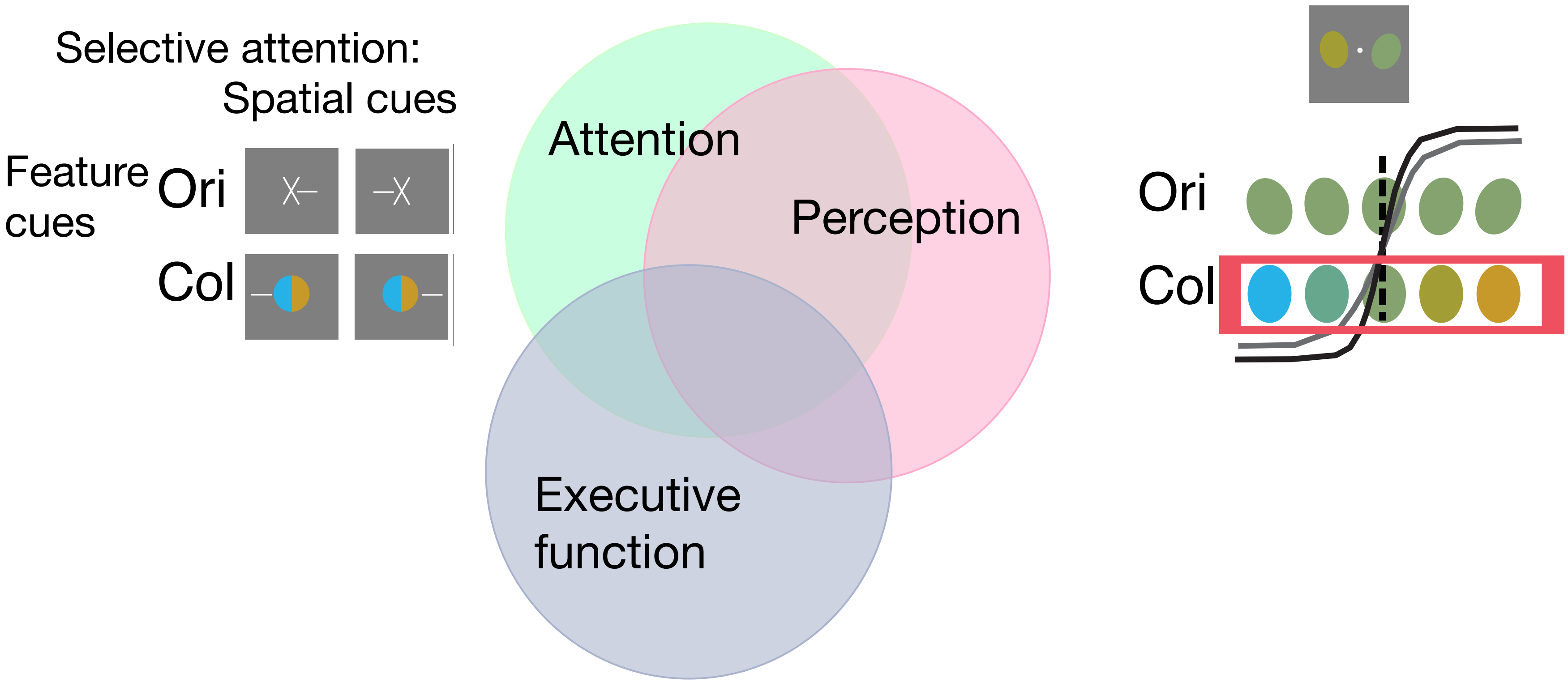
Col

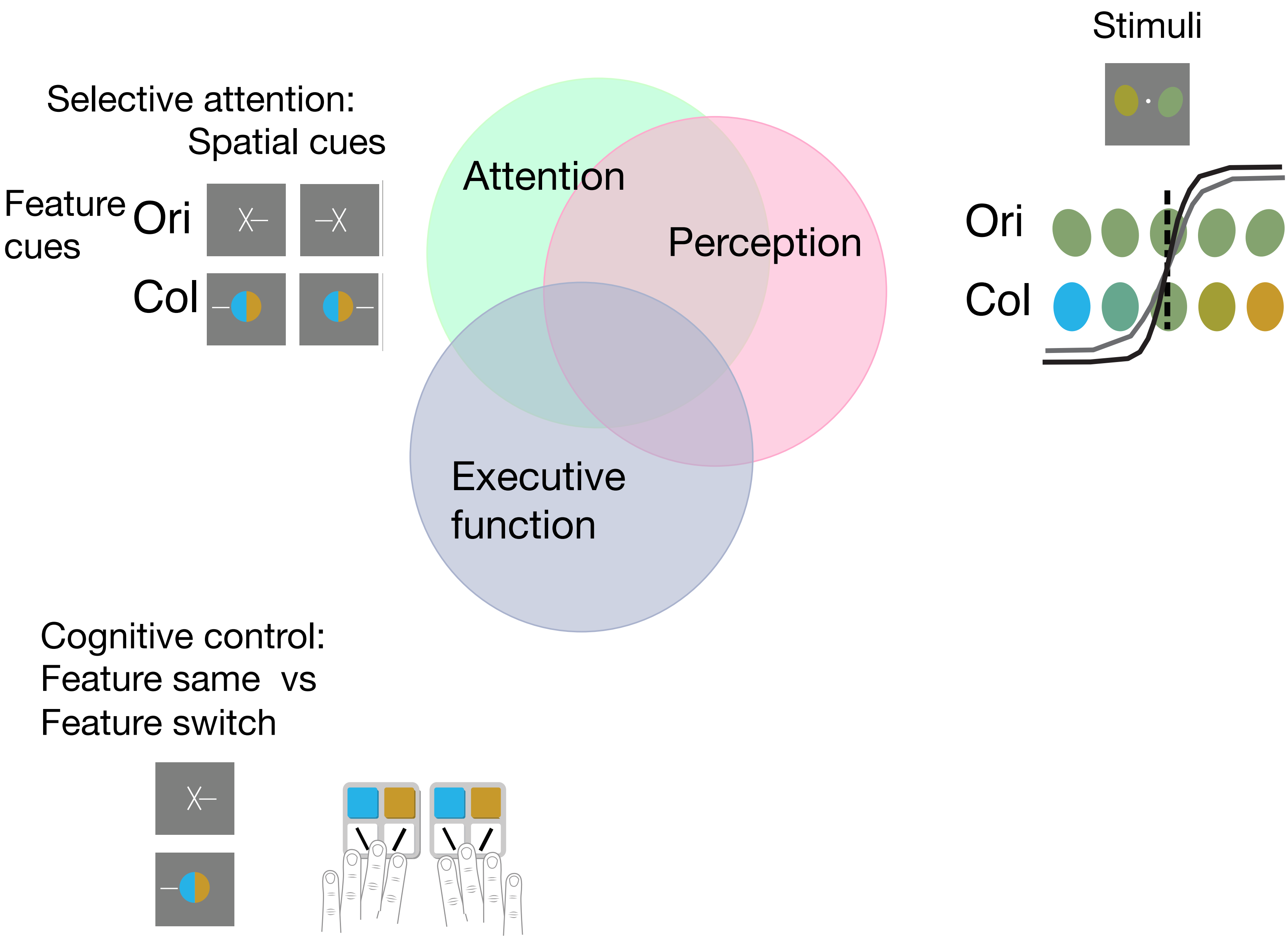


Stimuli



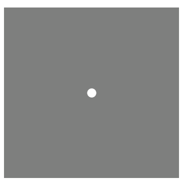
Our approach




Selective attention:
Spatial cues
Feature cues
Ori
Col
Attention
Perception
Executive function
Stimuli
Ori
Col
Cognitive control:
Feature same vs
Feature switch

Perceptual decision-making with spatial and feature dimension task switching

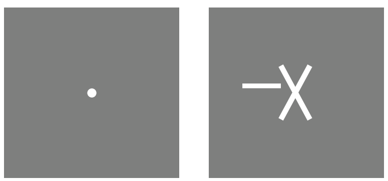
Trial time course



100 ms

Perceptual decision-making with spatial and feature dimension task switching

Trial time course

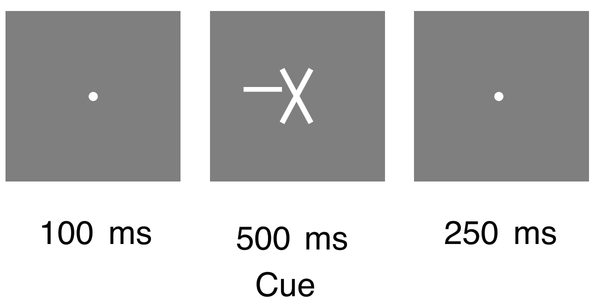


100 ms

500 ms
Cue

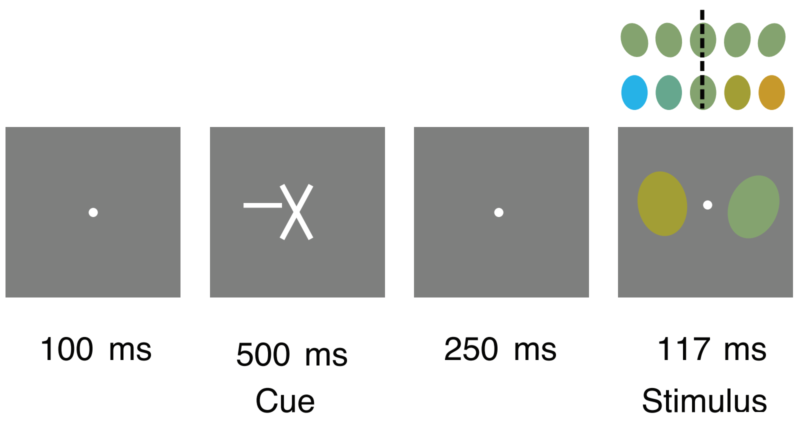
Perceptual decision-making with spatial and feature dimension task switching

Trial time course



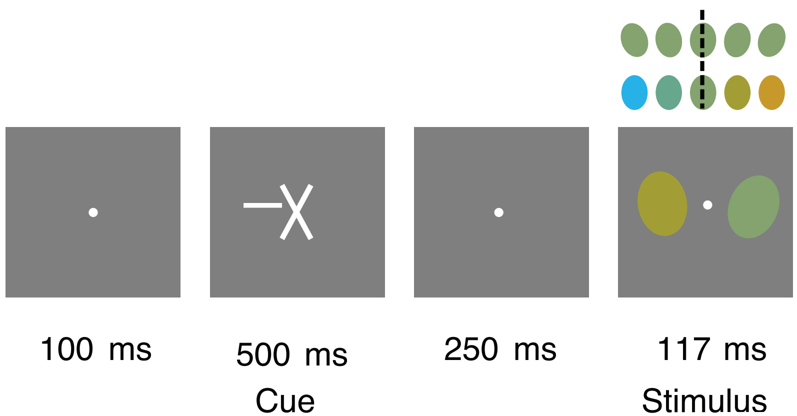
Perceptual decision-making with spatial and feature dimension task switching

Trial time course



Perceptual decision-making with spatial and feature dimension task switching

Trial time course



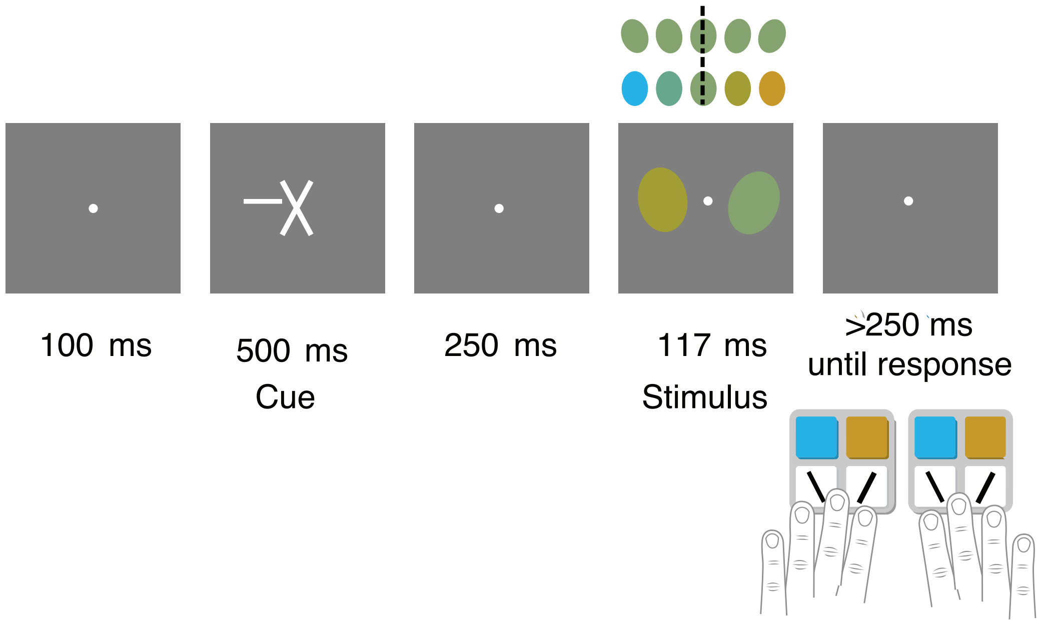
-Adaptive (efficient) stimulus selection method

(Kontsevich and Tyler 1999, Acerbi 2016)

-Especially important for running ADHD participants

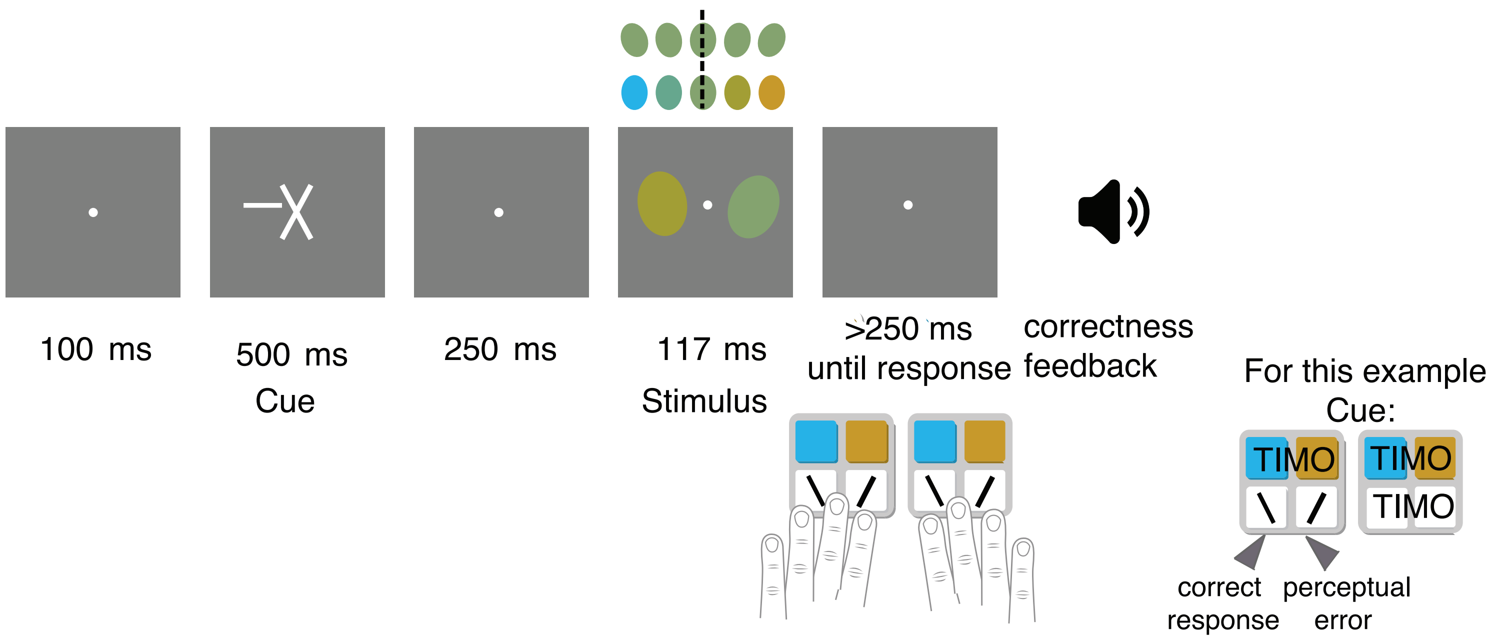
Perceptual decision-making with spatial and feature dimension task switching

Trial time course



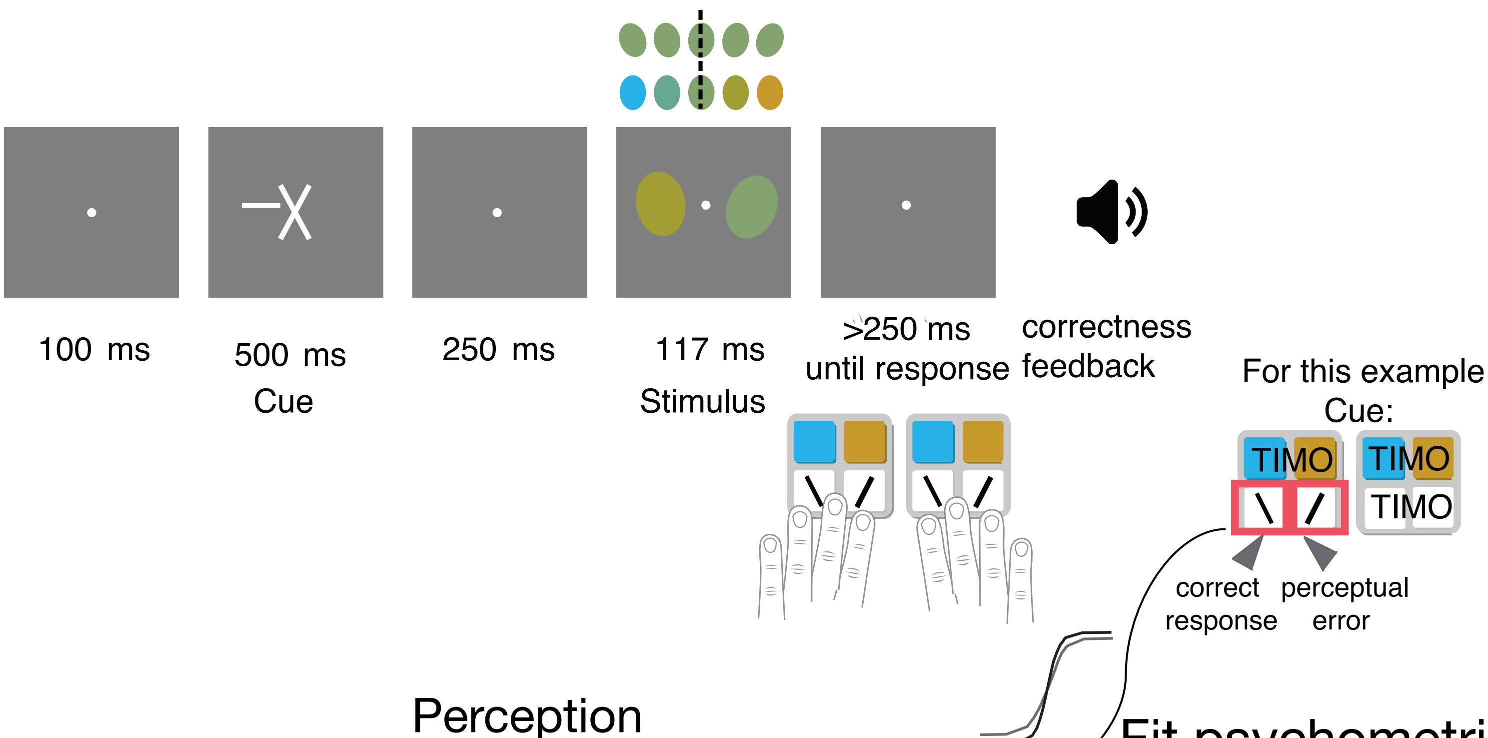
Perceptual decision-making with spatial and feature dimension task switching

Trial time course



Perceptual decision-making with spatial and feature dimension task switching

Trial time course



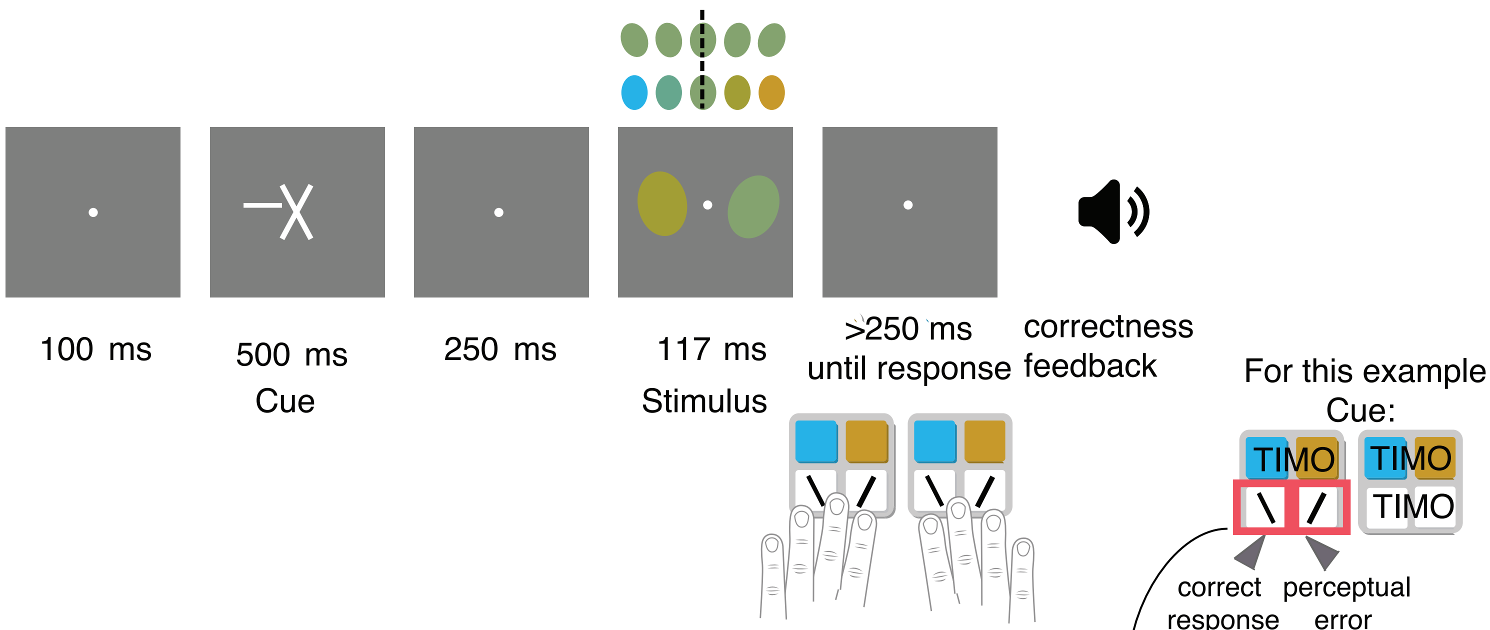
Fit psychometric curves

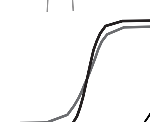
Perception

Measure of perceptual precision:
Perceptual variability parameter

Perceptual decision-making with spatial and feature dimension task switching

Trial time course





Perception/Attention

Measure of perceptual precision:
Perceptual variability parameter

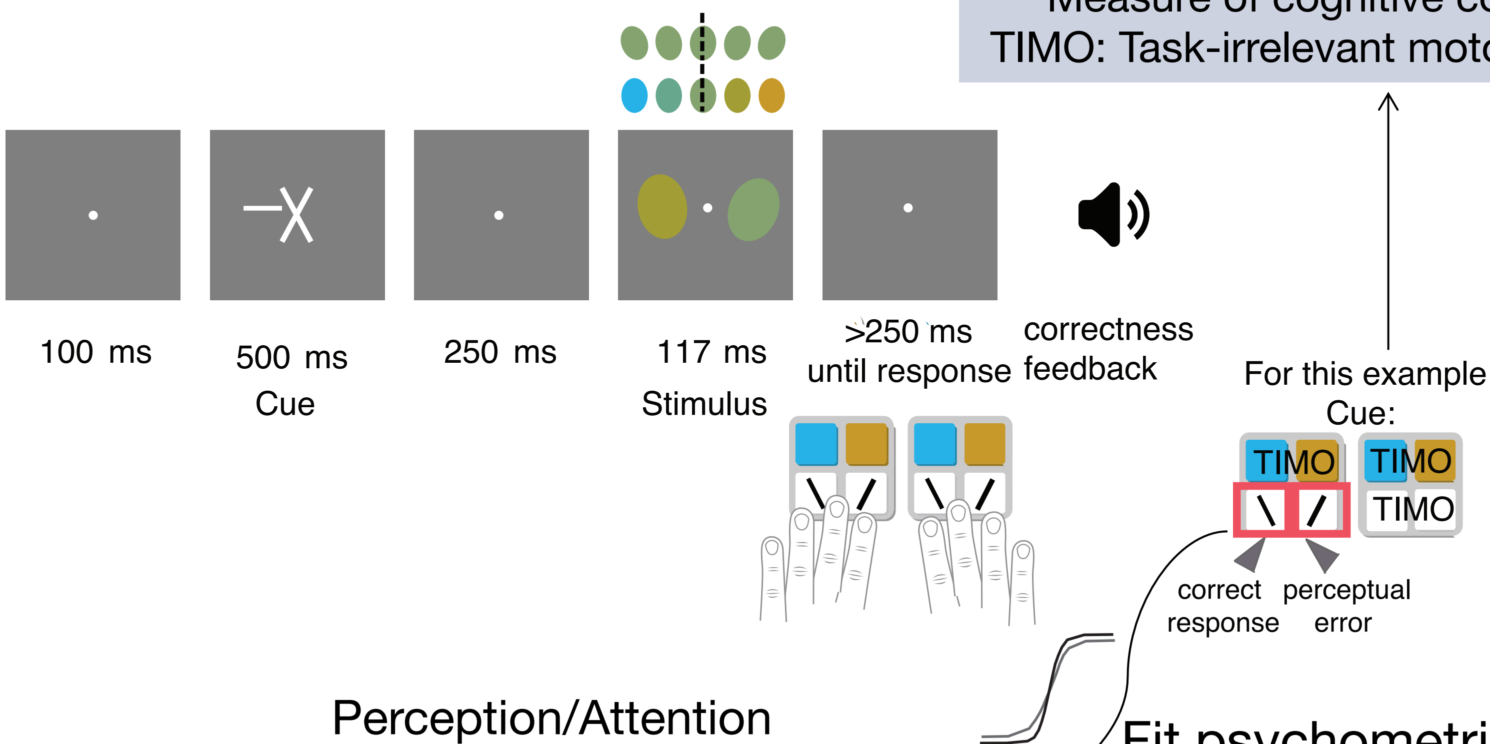
Fit psychometric curves

Perceptual decision-making with spatial and feature dimension task switching

Trial time course

Executive function

Measure of cognitive control:
TIMO: Task-irrelevant motor output



Fit psychometric curves

Perception/Attention

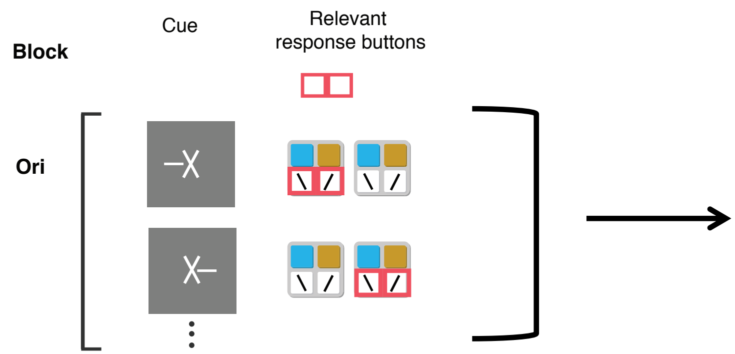
Measure of perceptual precision:
Perceptual variability parameter

Perceptual decision-making with spatial and feature dimension task switching

3 types of blocks:

Trials in 4 conditions:

Feature same

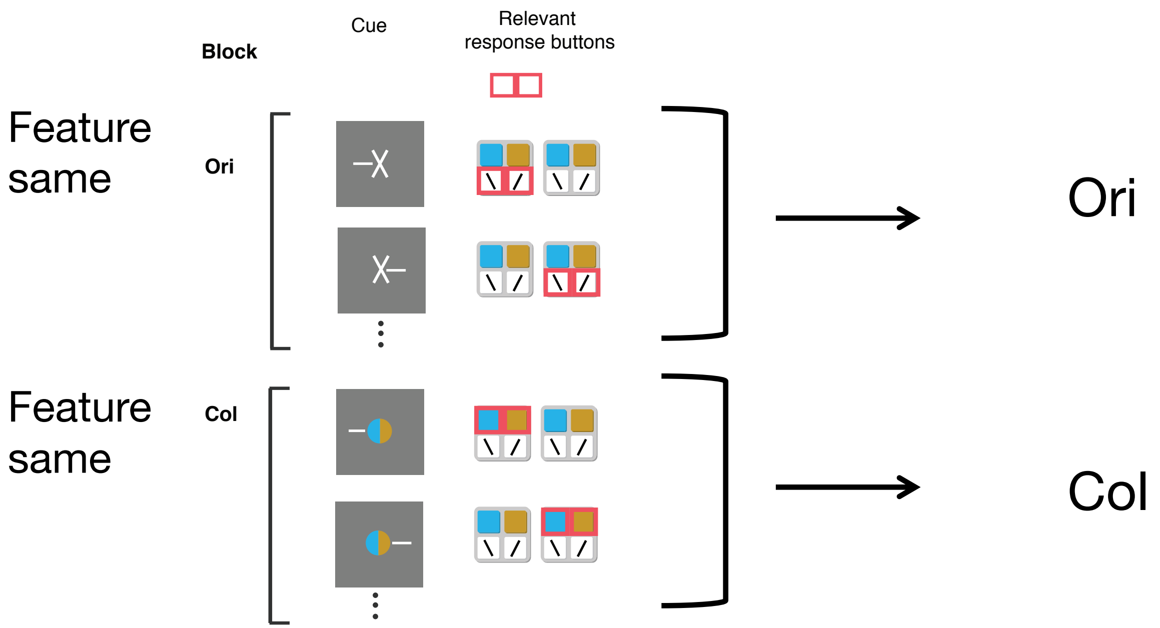


Ori

Perceptual decision-making with spatial and feature dimension task switching

3 types of blocks:

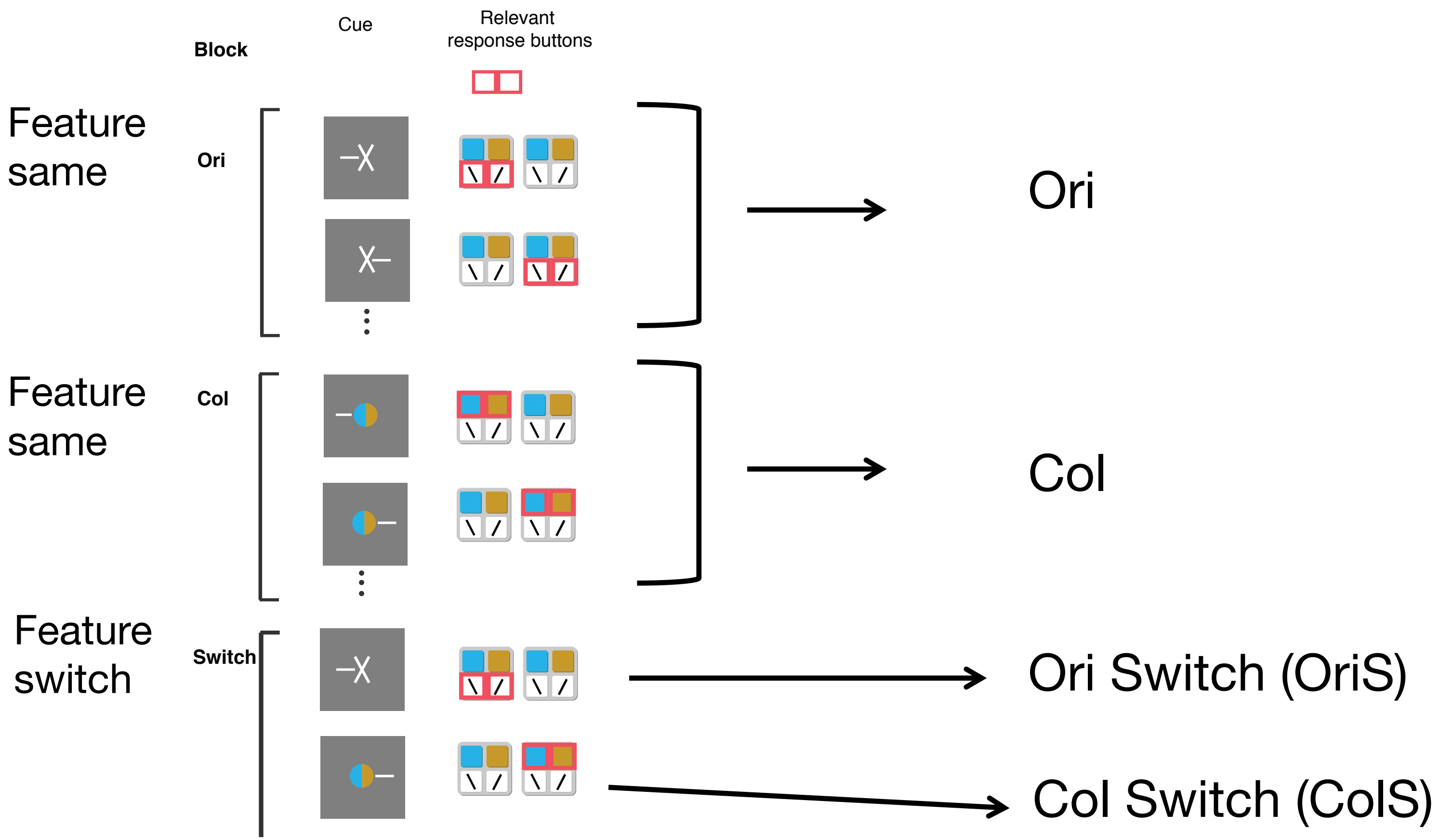
Trials in 4 conditions:



Perceptual decision-making with spatial and feature dimension task switching

3 types of blocks:

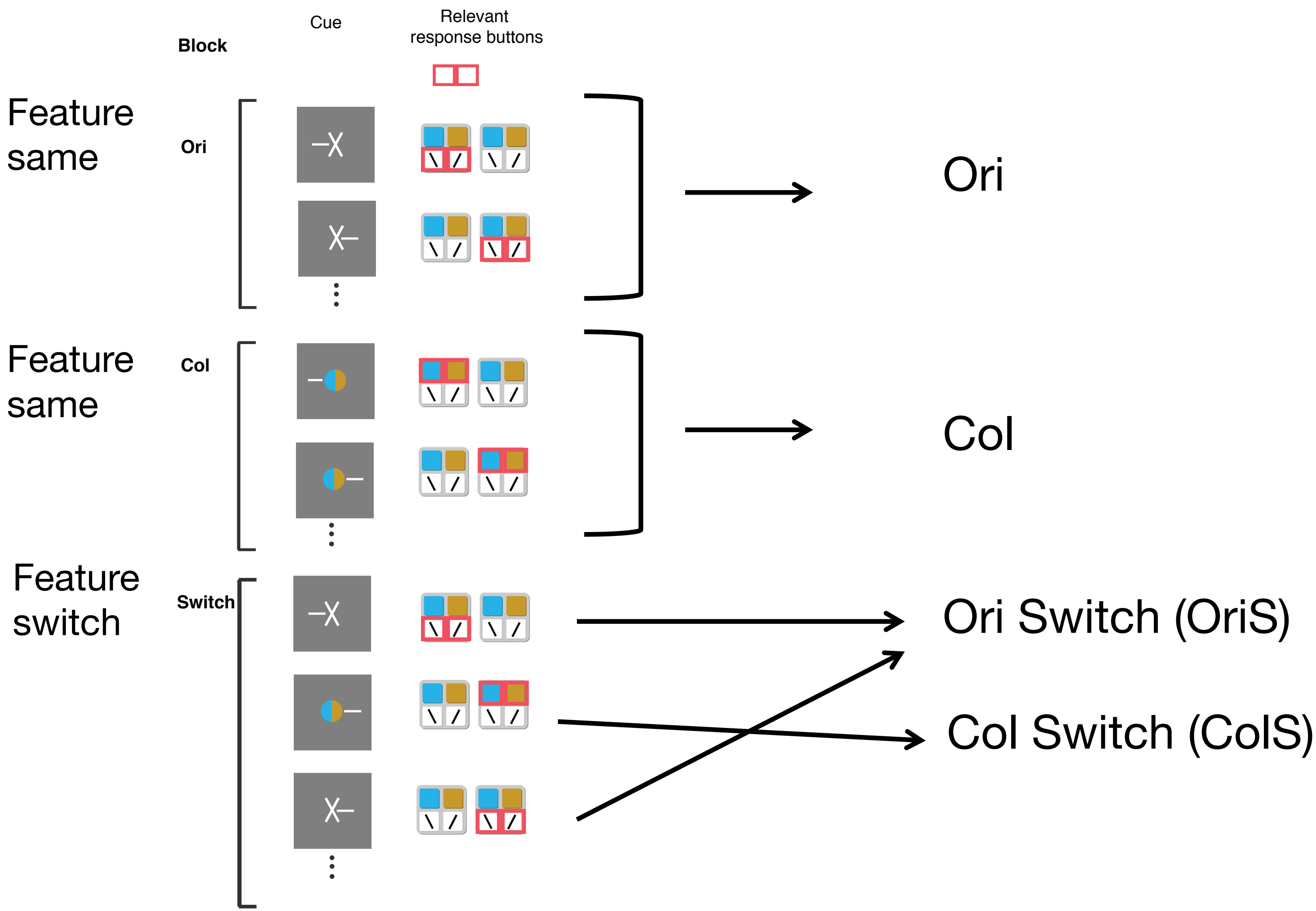
Trials in 4 conditions:



Perceptual decision-making with spatial and feature dimension task switching

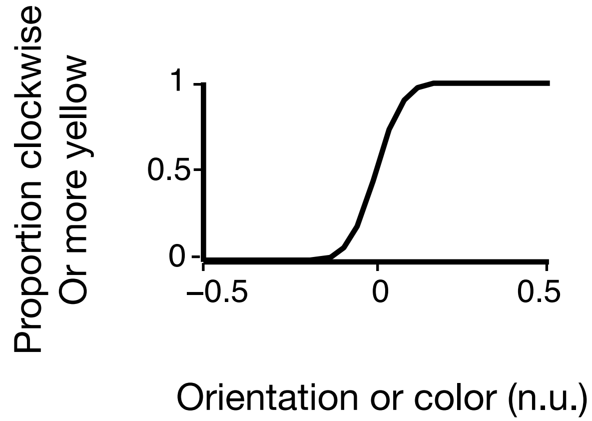
3 types of blocks:

Trials in 4 conditions:



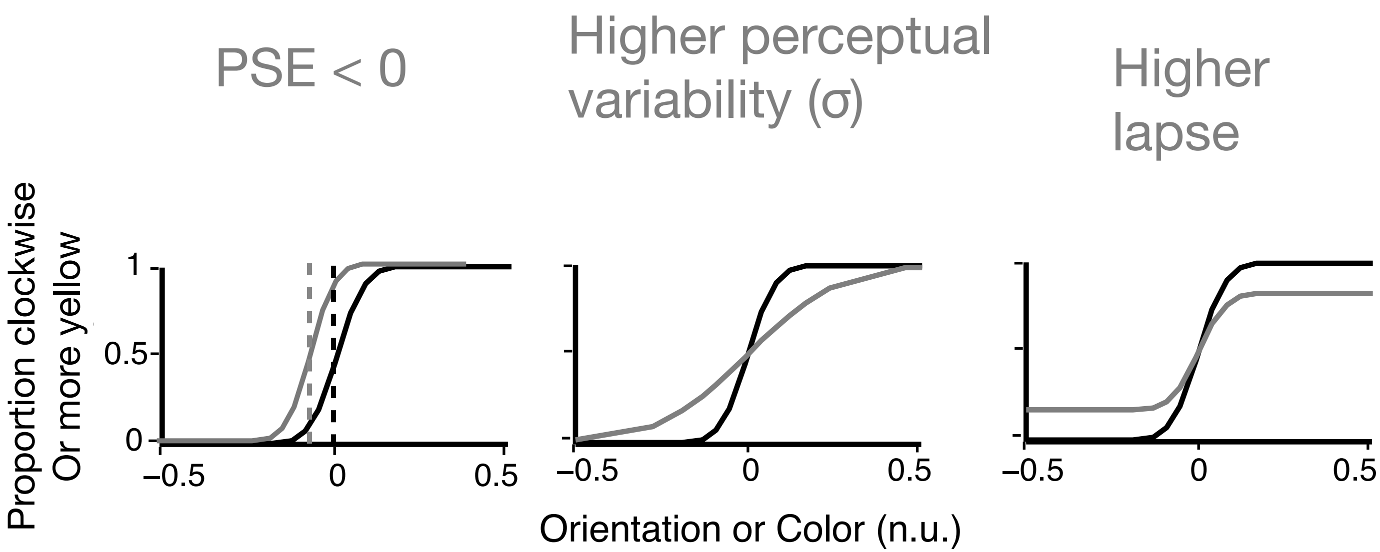
Psychometric curve fitting

- On non-TIMO trials, we fit psychometric curves with parameters PSE, perceptual variability (σ), and lapse



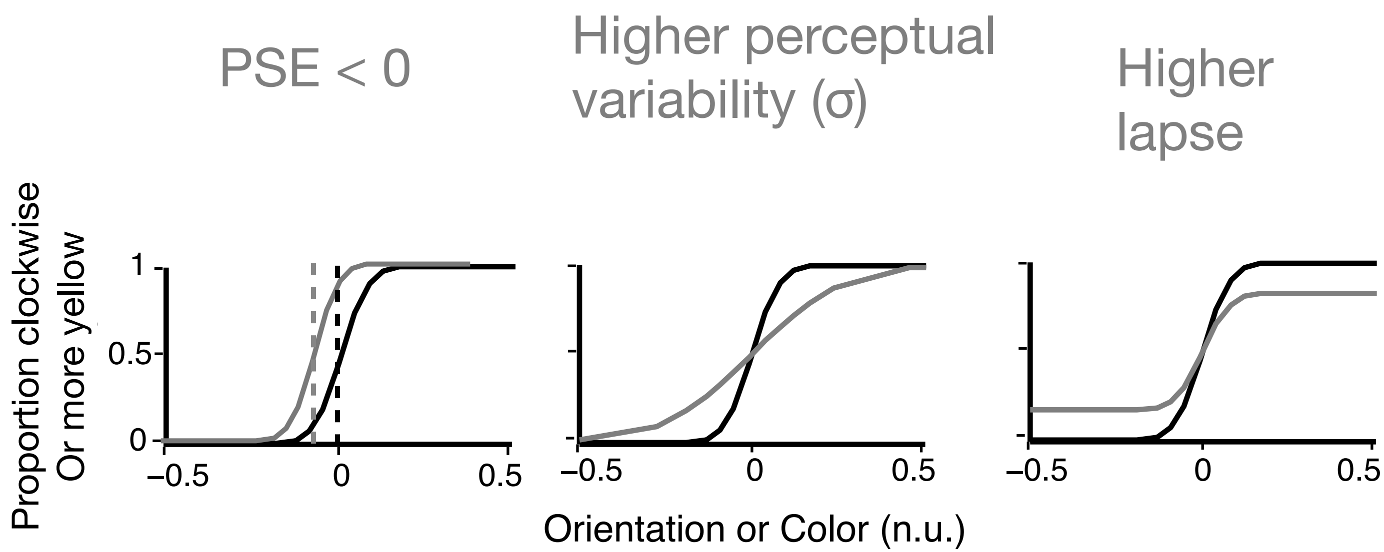
Psychometric curve fitting

- On non-TIMO trials, we fit psychometric curves with parameters PSE, perceptual variability (σ), and lapse



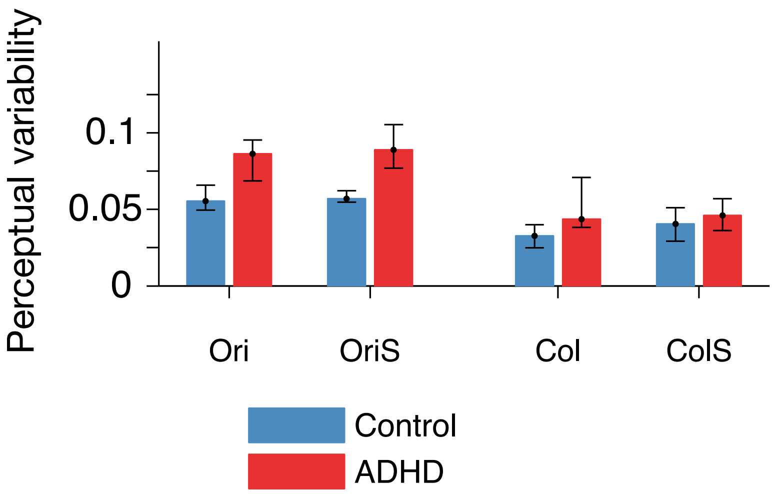
Psychometric curve fitting

- On non-TIMO trials, we fit psychometric curves with parameters PSE, perceptual variability (σ), and lapse



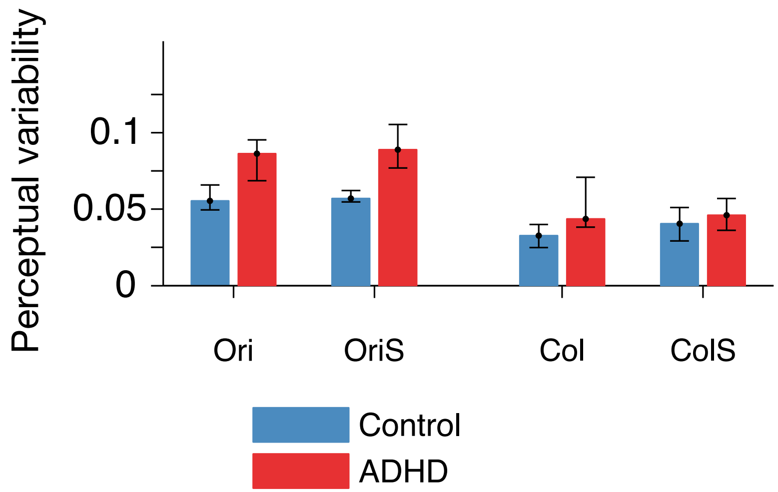
- We assume that PSE and lapse are shared between Switch/No-Switch conditions for a given feature, while perceptual variability might vary.

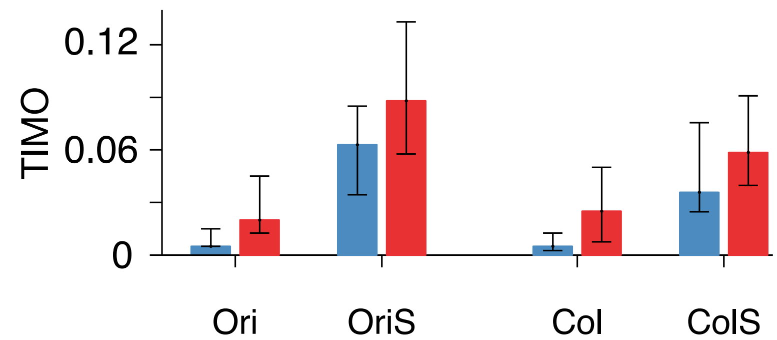
ADHD participants have higher perceptual variability



Effect of group: *p* = 0.002
Effect of feature Ori/Col: p < 0.001
Effect of load no-switch/Switch: *p* = 0.33

ADHD participants have higher perceptual variability and higher TIMO

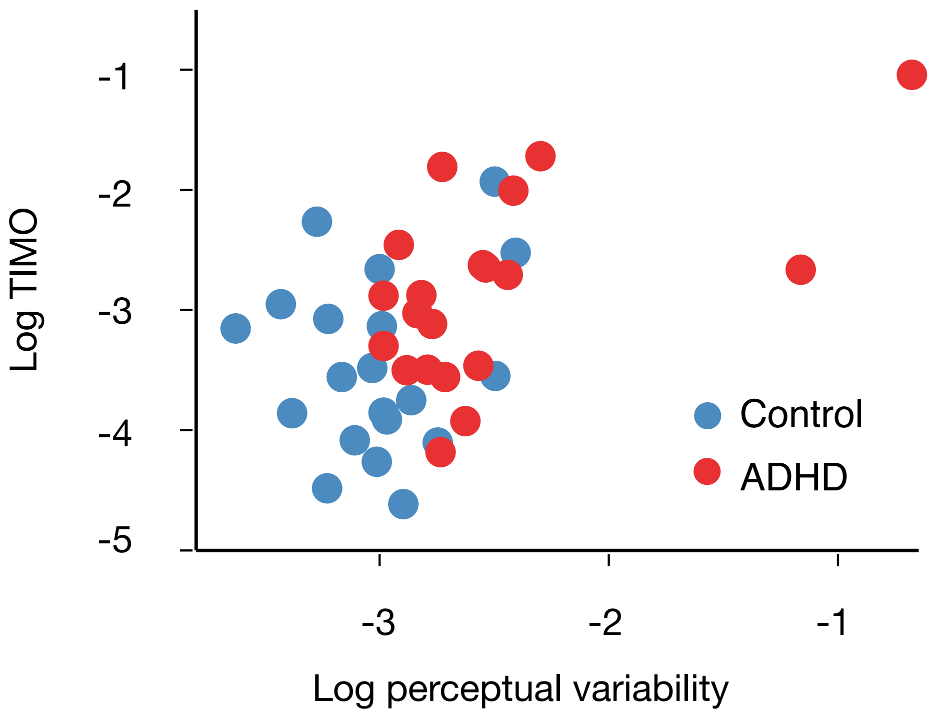




Effect of group: p = 0.002
Effect of feature Ori/Col: p < 0.001
Effect of No-switch/Switch: p = 0.33

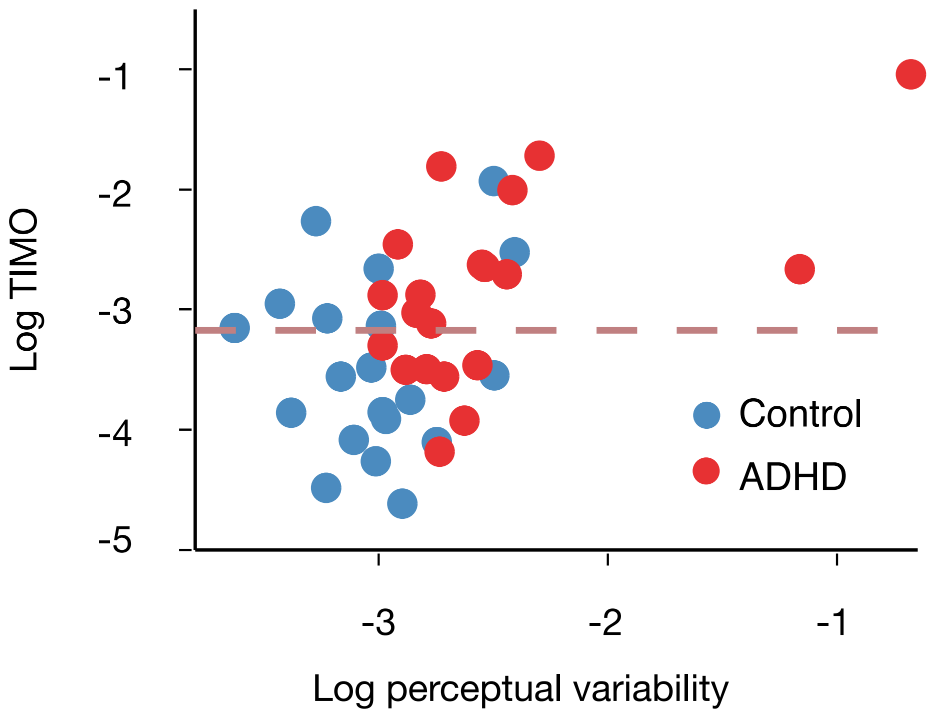
Effect of group: p = 0.005
Effect of feature Ori/Col: p = 0.21
Effect of No-switch/Switch: p < 0.001

Perceptual variability and TIMO are correlated



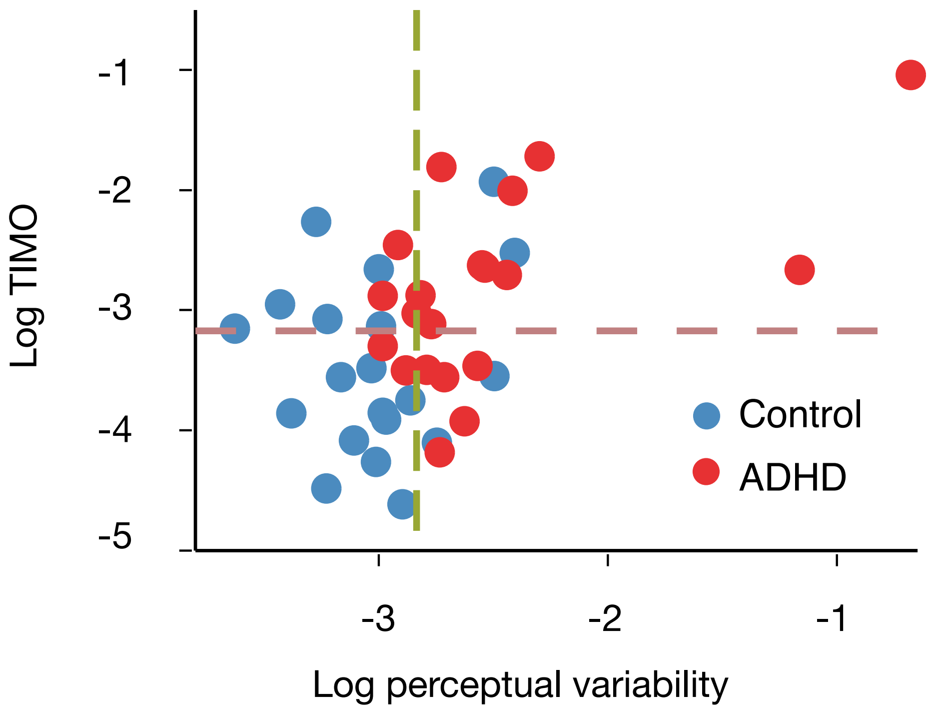
$$\boldsymbol{\rho = 0.41}$$
$$p = 0.0085$$

Logistic regression classifier based on TIMO



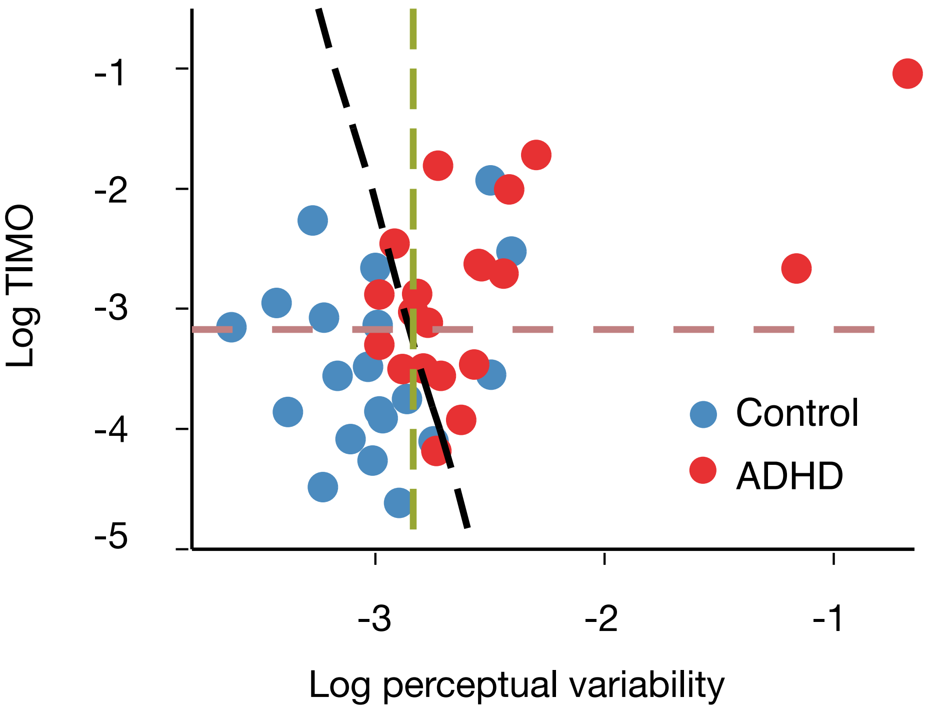
Classifier	accuracy	mean cross-validated accuracy
TIMO	62.5%	63.1%

Logistic regression classifier based on perceptual variability



Classifier	accuracy	mean cross-validated accuracy
TIMO	62.5%	63.1%
σ	77.5%	77.1%

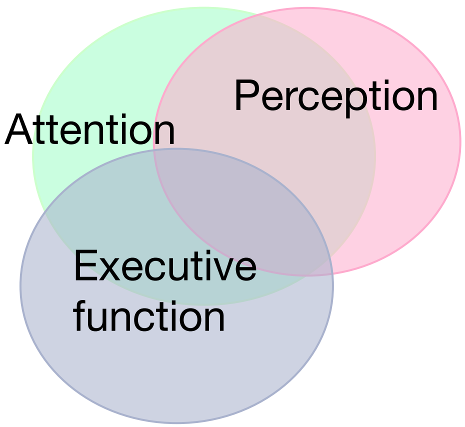
Logistic regression classifiers including perceptual variability yield high diagnosis accuracy



Classifier	accuracy	mean cross-validated accuracy
TIMO	62.5%	63.1%
σ	77.5%	77.1%
σ and TIMO	82.5%	77.8%

Conclusions part 2

- Goal: get separate measures of perceptual encoding precision and cognitive control
  - Taxed attention through spatial and feature cueing
  - Taxed executive function through stimulus-response rules



- ADHD participants had worse cognitive control
  - In line with previous studies
- ADHD participants had higher perceptual variability
  - Contrast with previous studies
  - Perhaps perceptual differences in ADHD only manifest when attention and executive function are simultaneously taxed

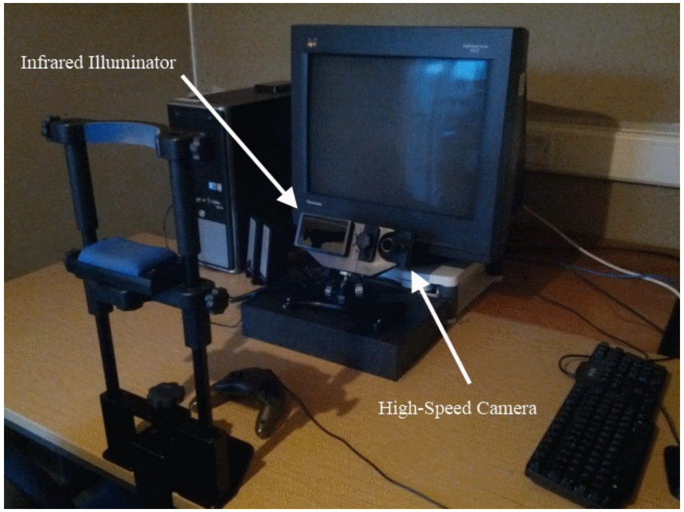
- Classifier based on perceptual variability parameter yielded good diagnosis accuracy ➔ candidate psychomarker.

Part 3: Bayesian microsaccade detection, with an application to ADHD

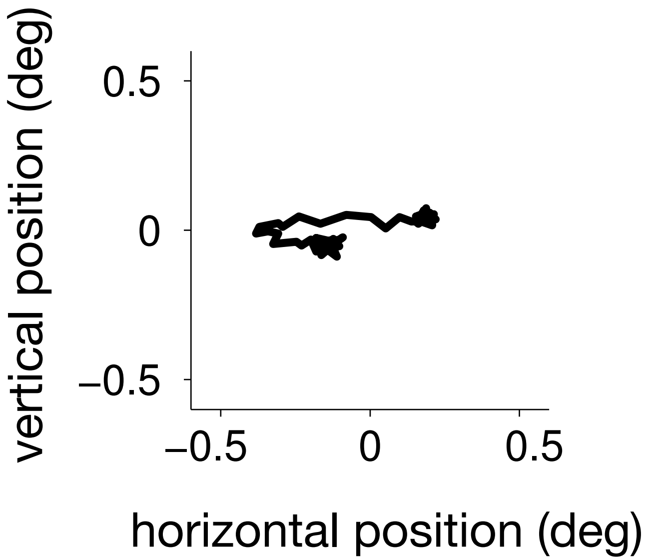
*with Bas van Opheusden**

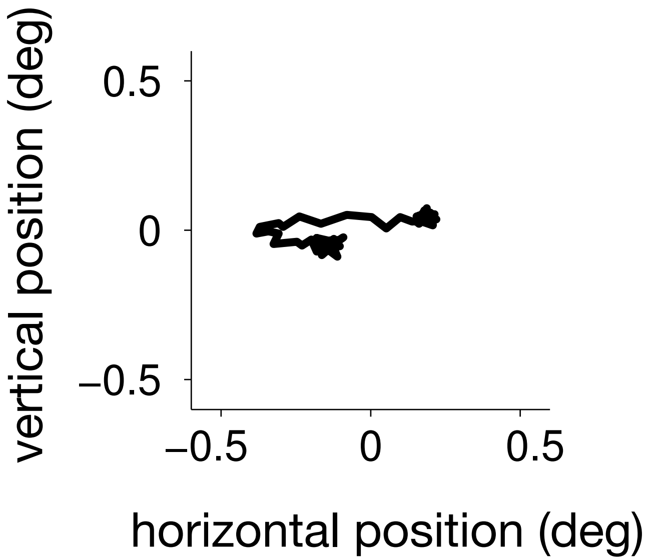
Microsaccades

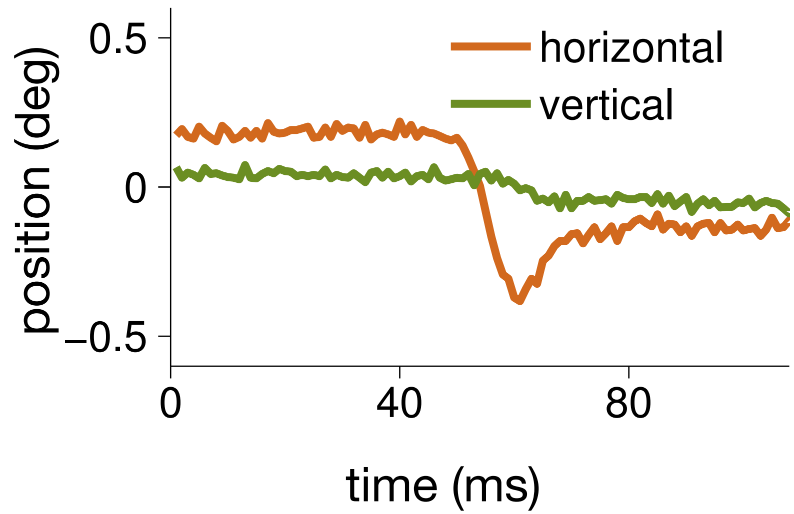
- High velocity fixational eye movements
- Several proposed perceptual and cognitive roles, also in indexing covert attention *Engbert and Kliegl 2003, Martinez-Conde et al. 2004, 2013, Rolfs 2009, Lara and Wallis 2012, Yuval-Greenberg et al. 2014*
- Detection is complicated partly due to the eye tracker noise → important to have a good detection algorithm



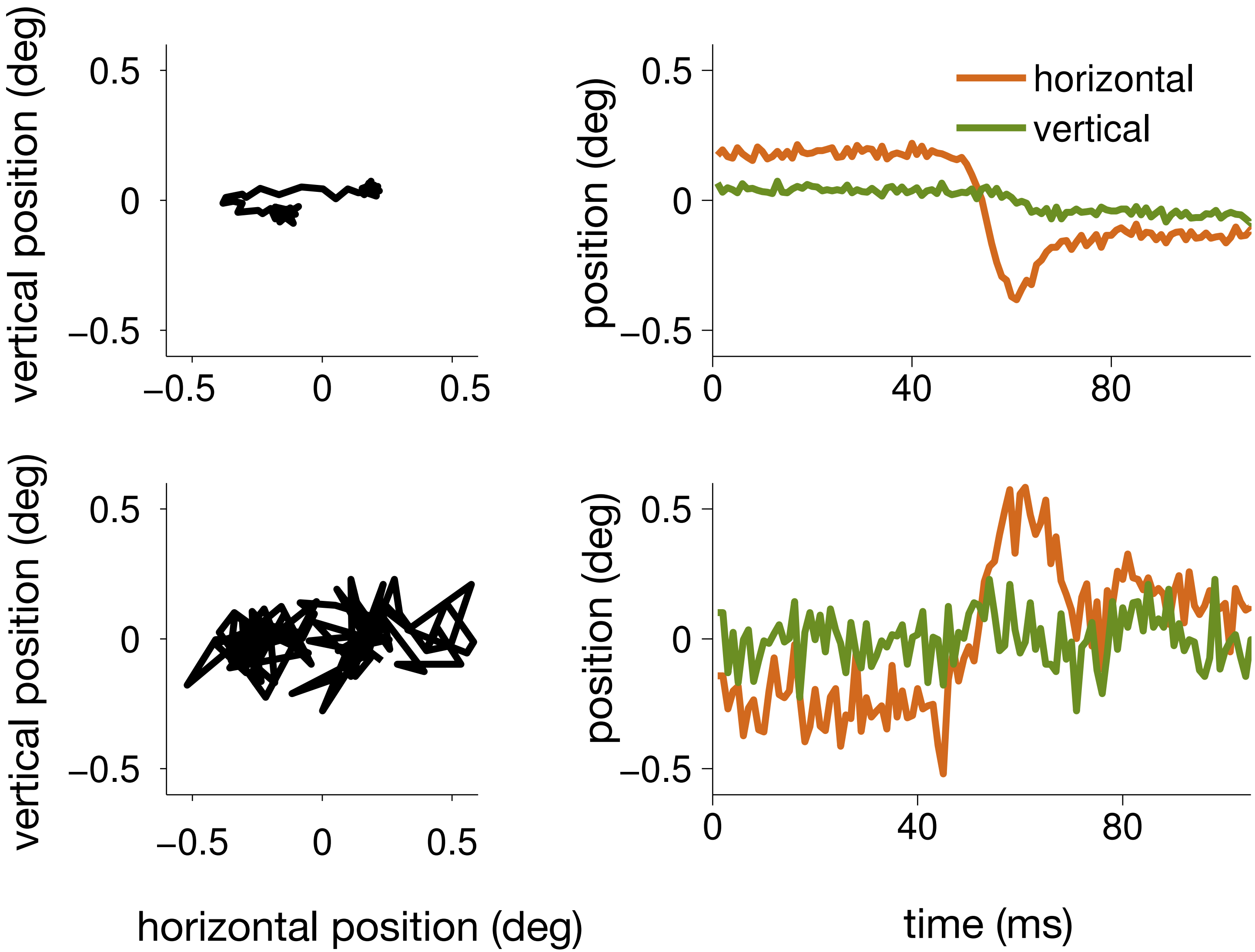
Source: https://wiki.psychwire.co.uk/?page_id=180


vertical position (deg)
0.5
0
–0.5
–0.5
0
0.5
horizontal position (deg)


vertical position (deg)
0.5
0
–0.5
–0.5
0
0.5
horizontal position (deg)

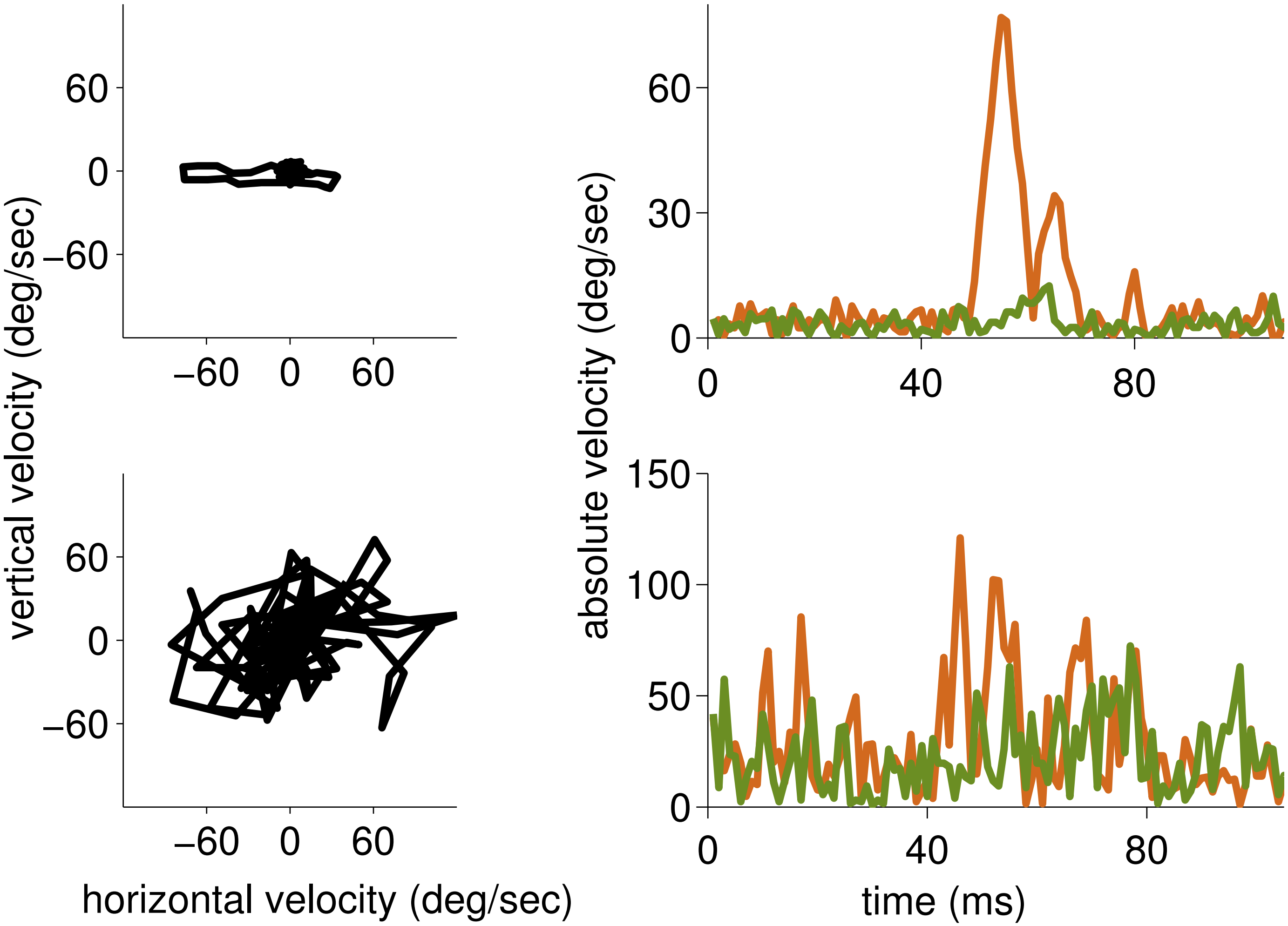

horizontal
vertical
position (deg)
0.5
0
–0.5
0
40
80
time (ms)

Microsaccades can be hard to detect under high noise



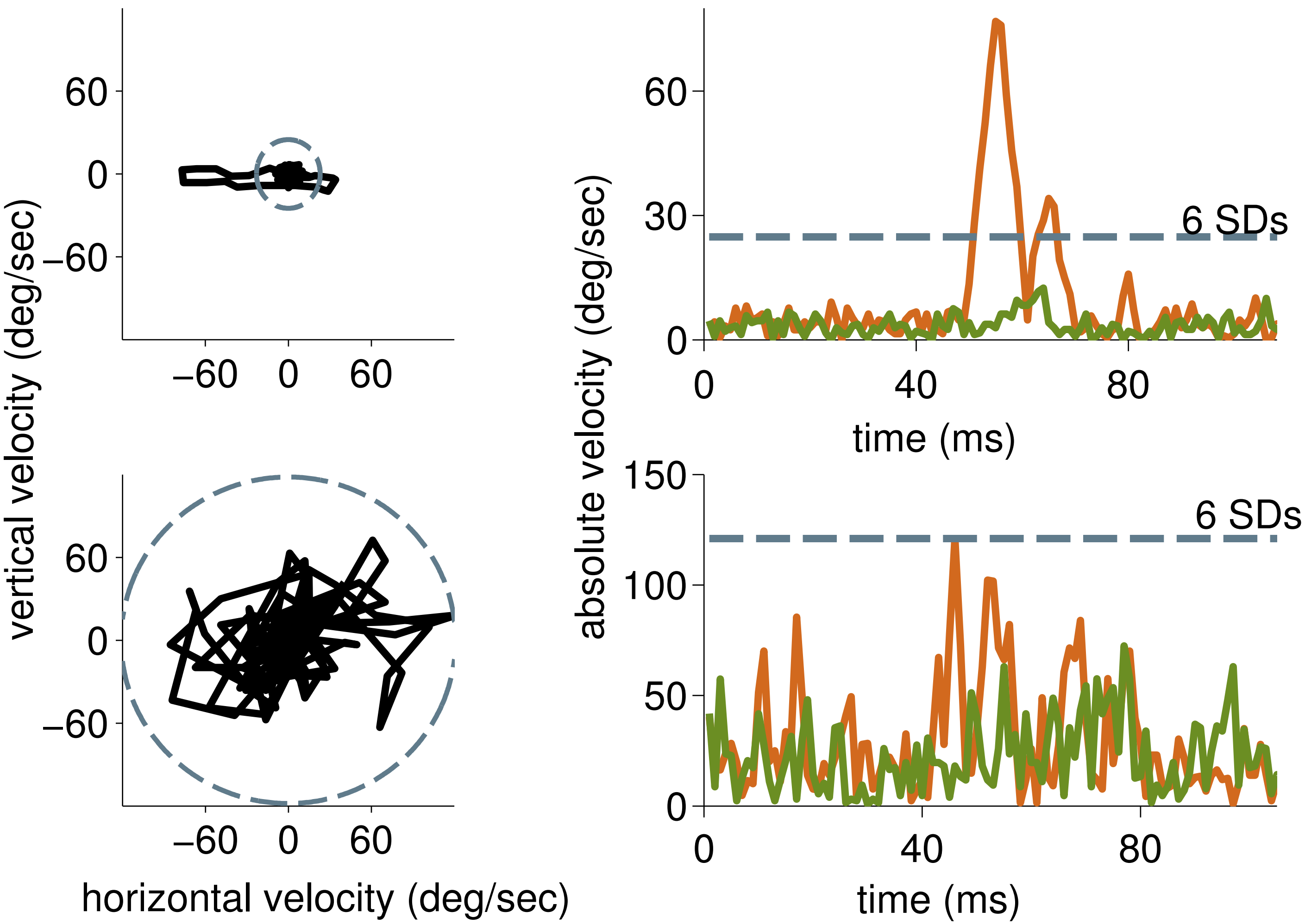
Engbert & Kliegl (EK) velocity threshold algorithm

Engbert and Kliegl, 2003



Engbert & Kliegl (EK) velocity threshold algorithm

Engbert and Kliegl, 2003

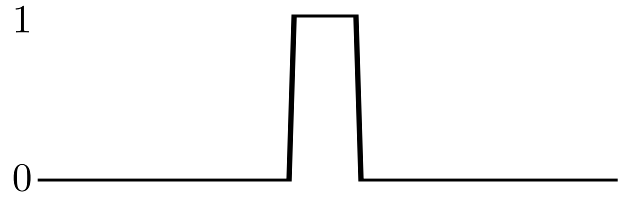


Bayesian microsaccade detection (BMD)

- Explicit assumptions about the process by which measured eye positions are generated
  - Probabilistic, not mechanistic
- Returns at each time point a probability instead of a binary answer

Generative model of fixational eye movements

eye state **C**

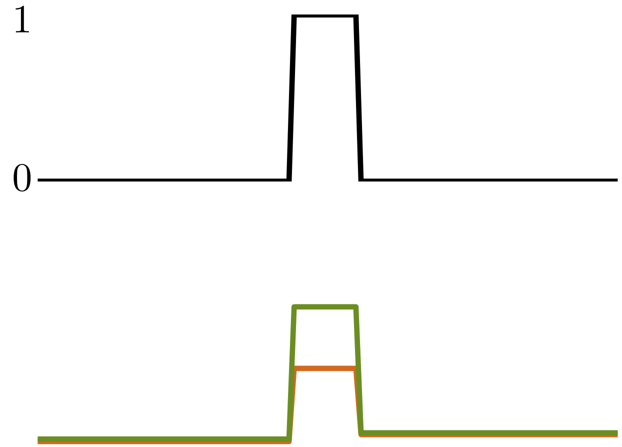


time

Generative model of fixational eye movements

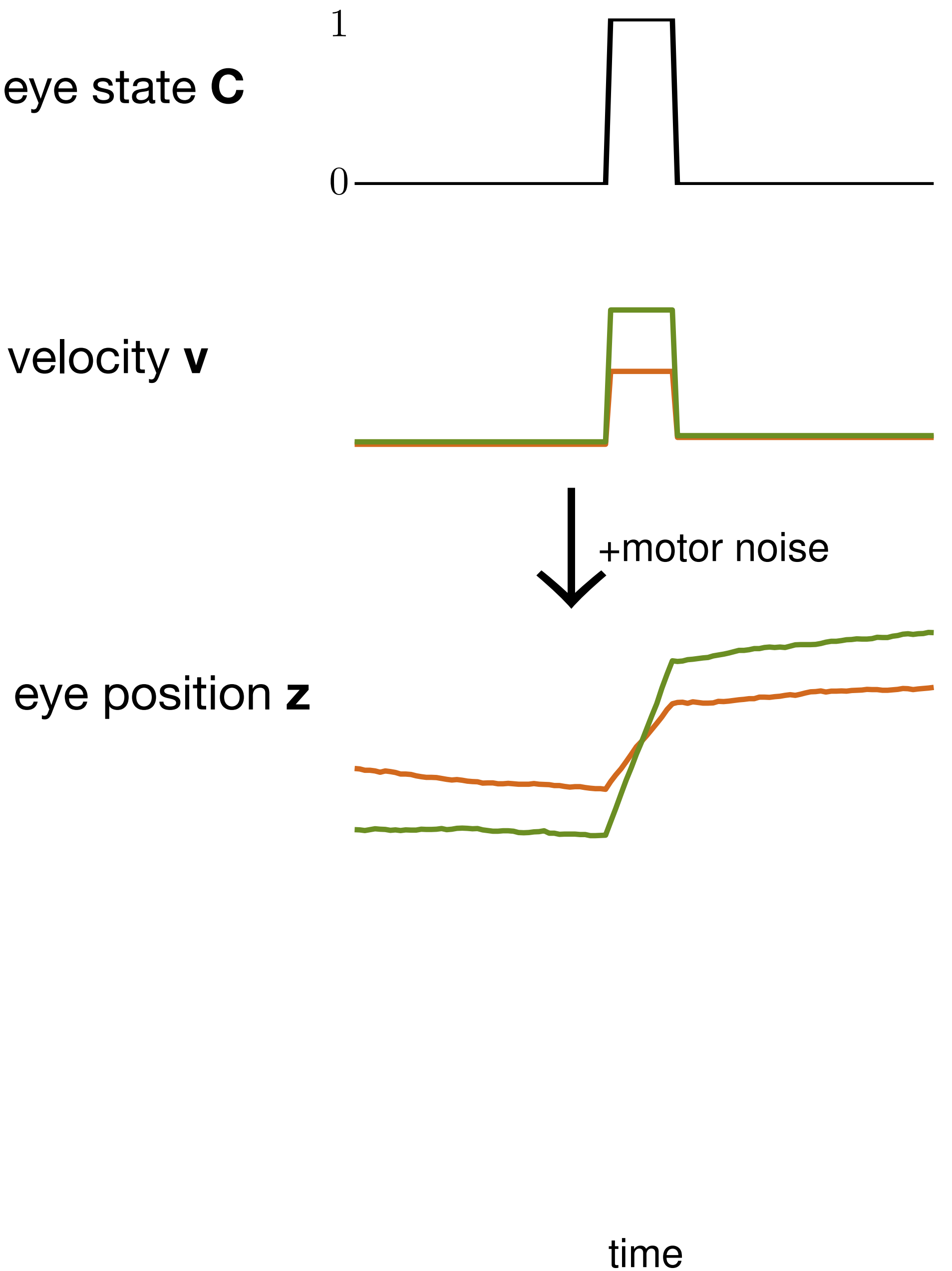
eye state **C**

velocity **v**

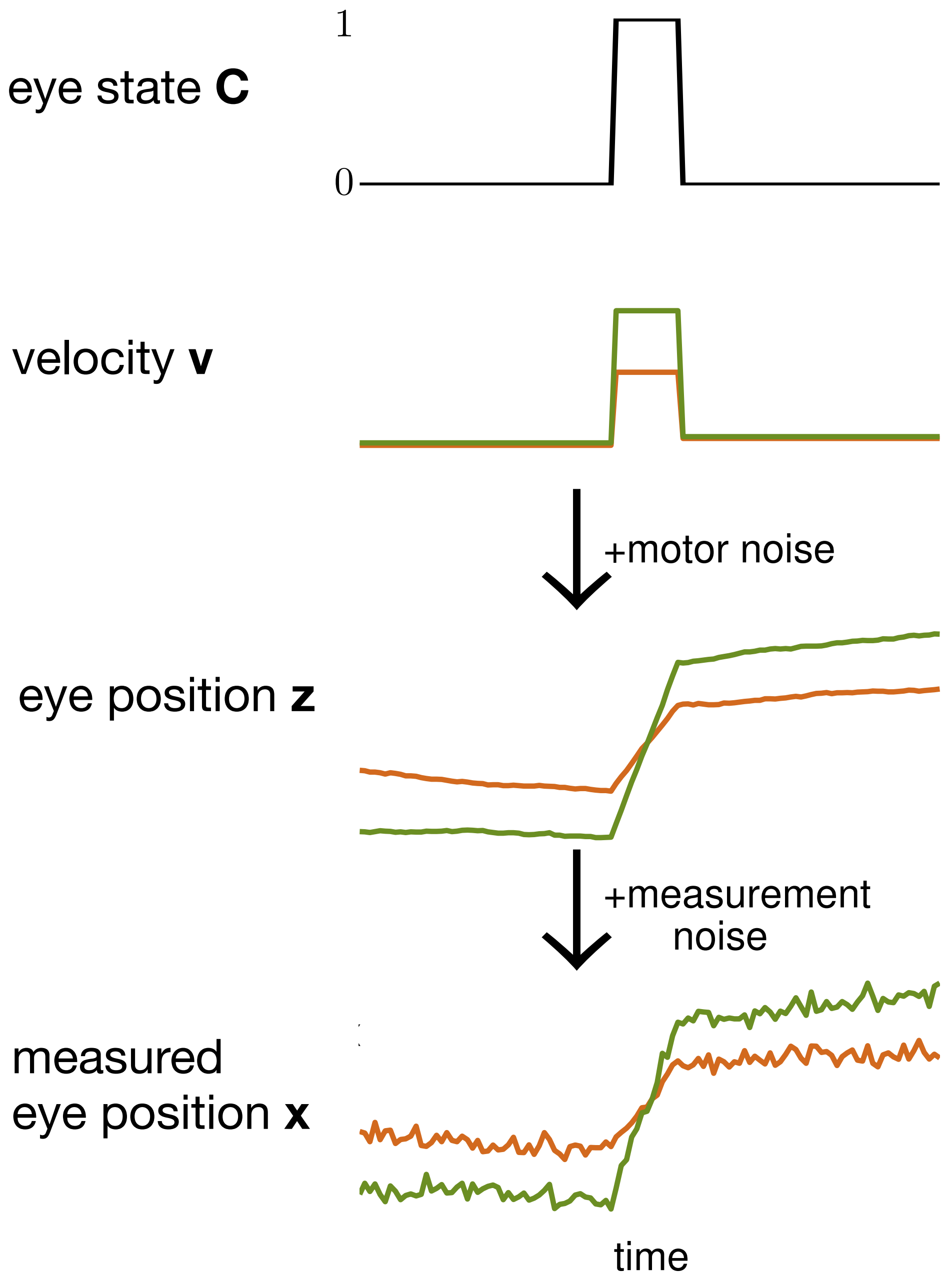


time

Generative model of fixational eye movements



Generative model of fixational eye movements



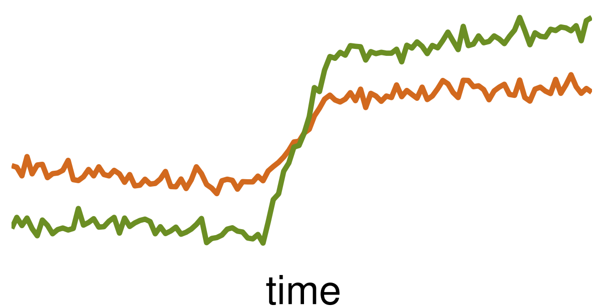
eye state **C**

?

Bayesian inference

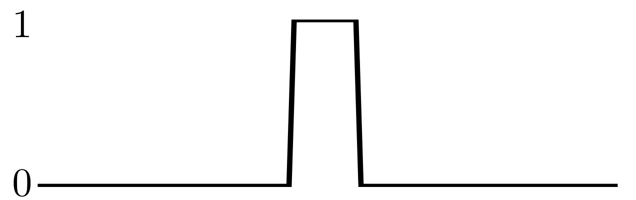
$$p(\mathbf{C}|\mathbf{x})$$

measured
eye position **x**



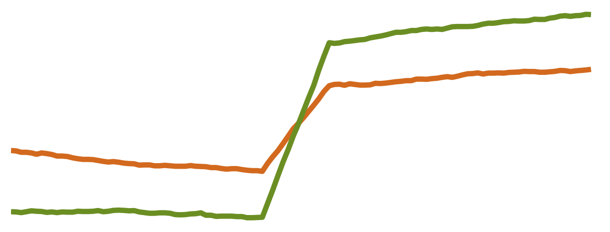
Approximate inference: BMD algorithm

$p(\mathbf{C}|\hat{\mathbf{z}})$ eye state **C**



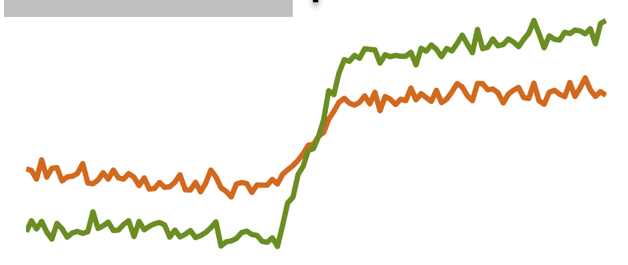
-Numerical approximation of the likelihood $p(\hat{\mathbf{z}}|\mathbf{C})$
-MCMC sampling

estimated eye position $\hat{\mathbf{z}}$



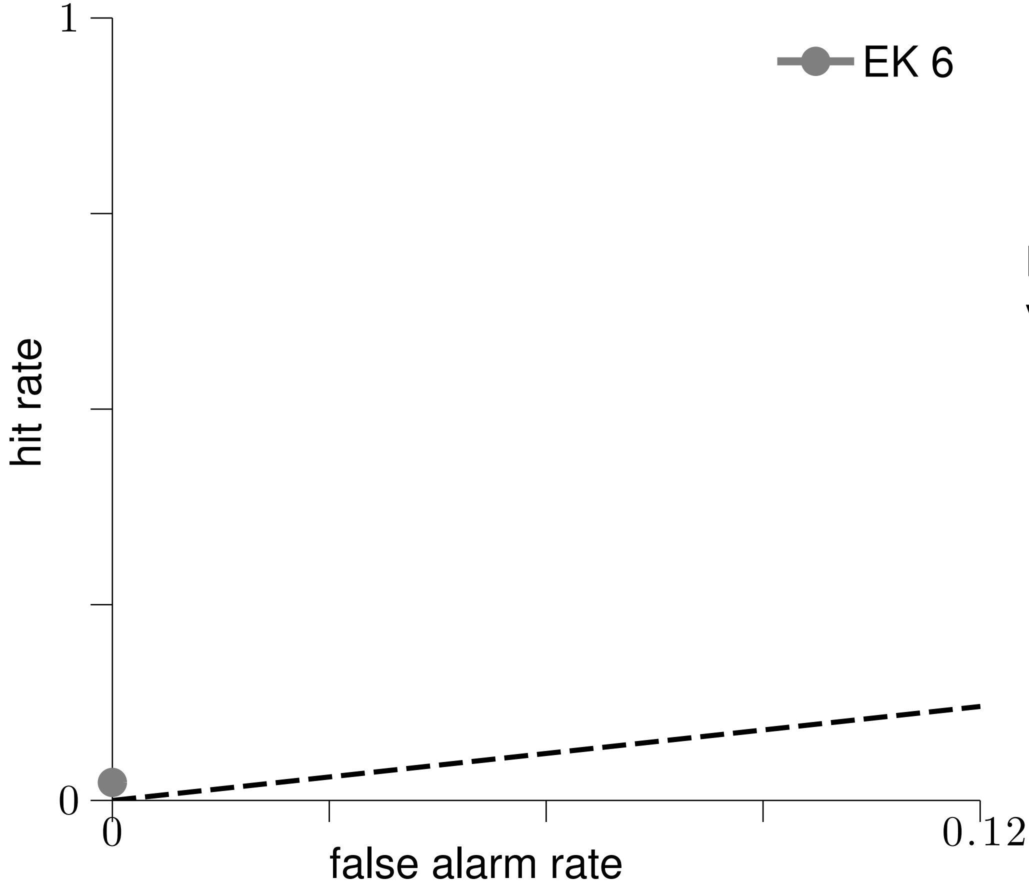
Kalman smoother

measured eye position **x**

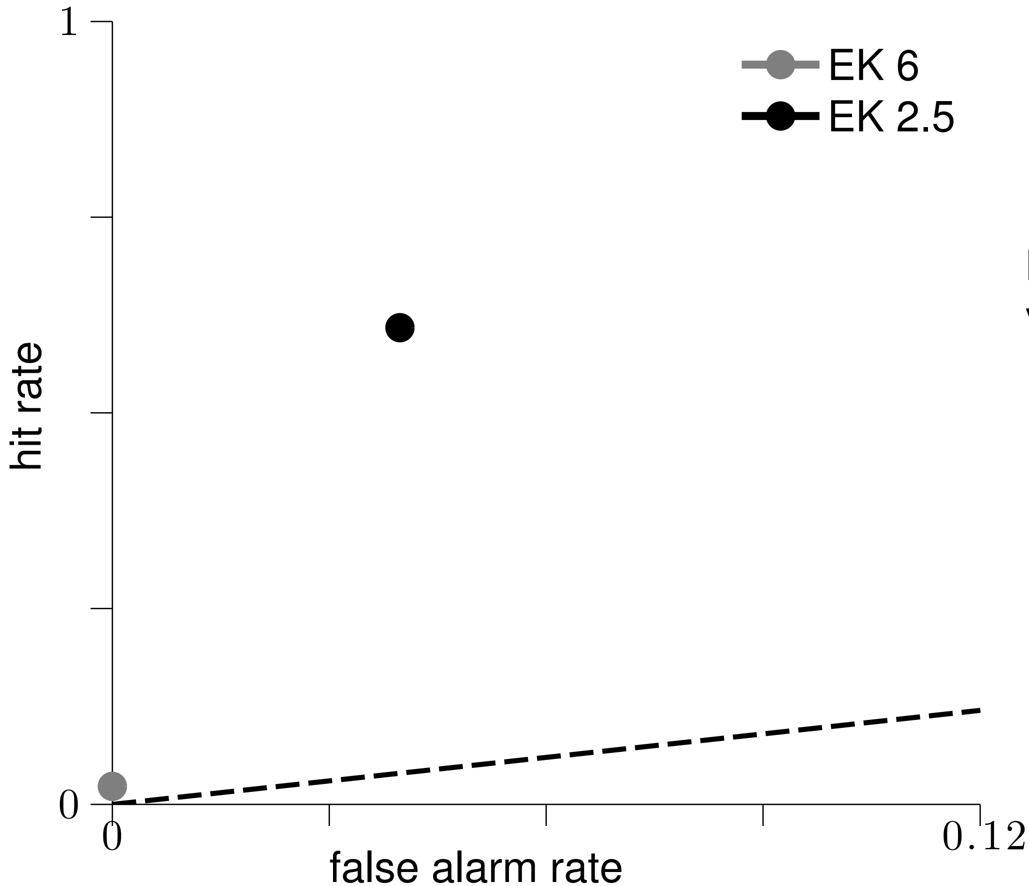


Test BMD algorithm: simulated data

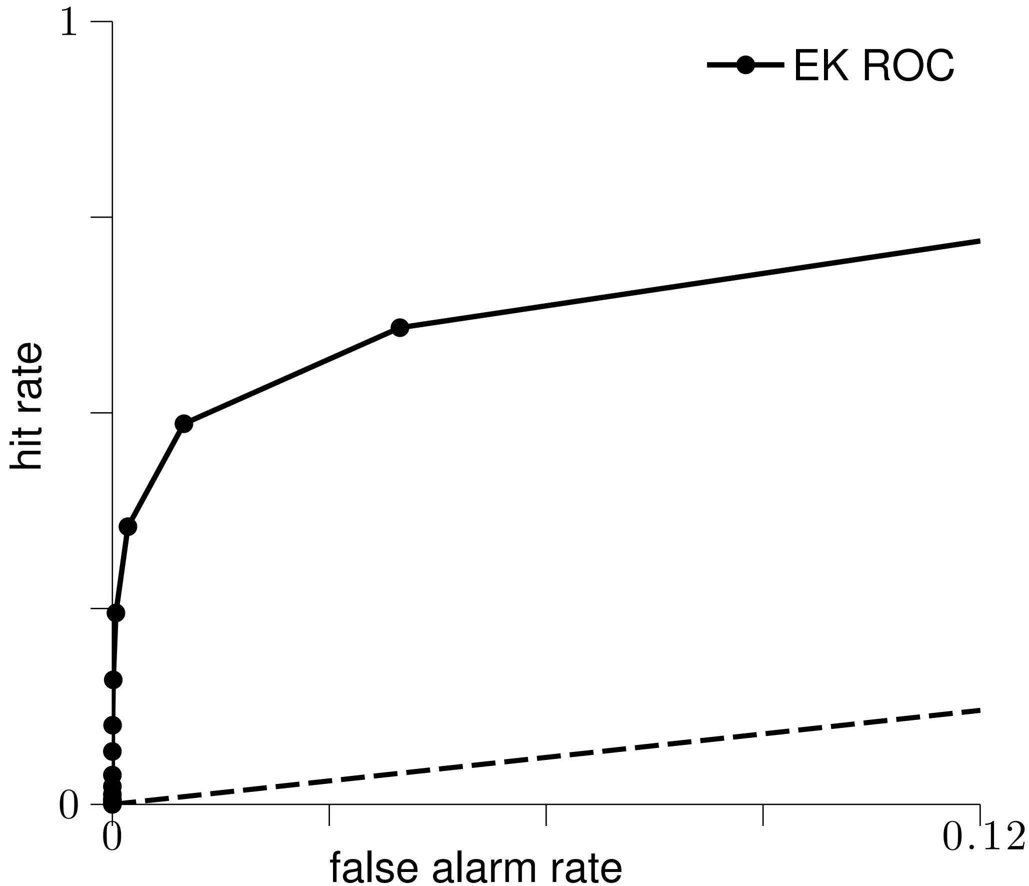
1. Simulate data from generative model with increasing levels of motor and measurement noise
2. Test BMD against alternative algorithms: variants of EK velocity threshold, unsupervised clustering (Otero-Milan, 2014)
3. Use hit rate and false alarm rate as metric for performance



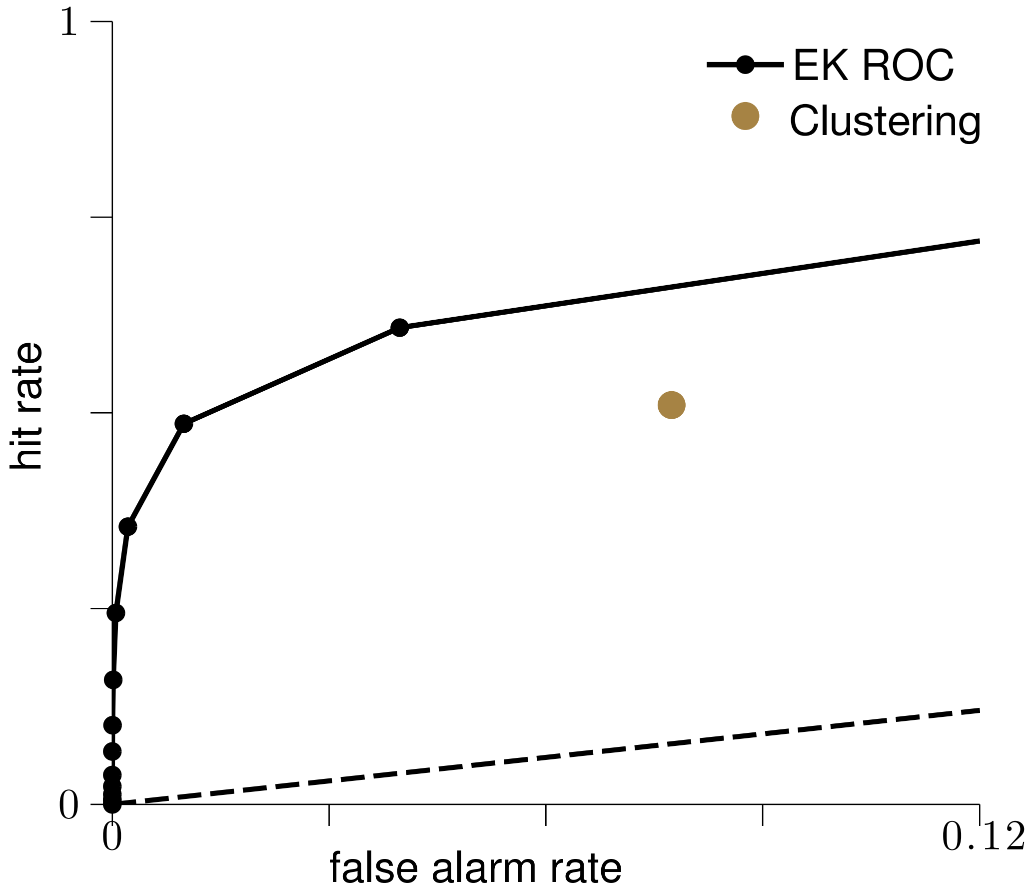
EK: Engbert and Kliegl
velocity threshold



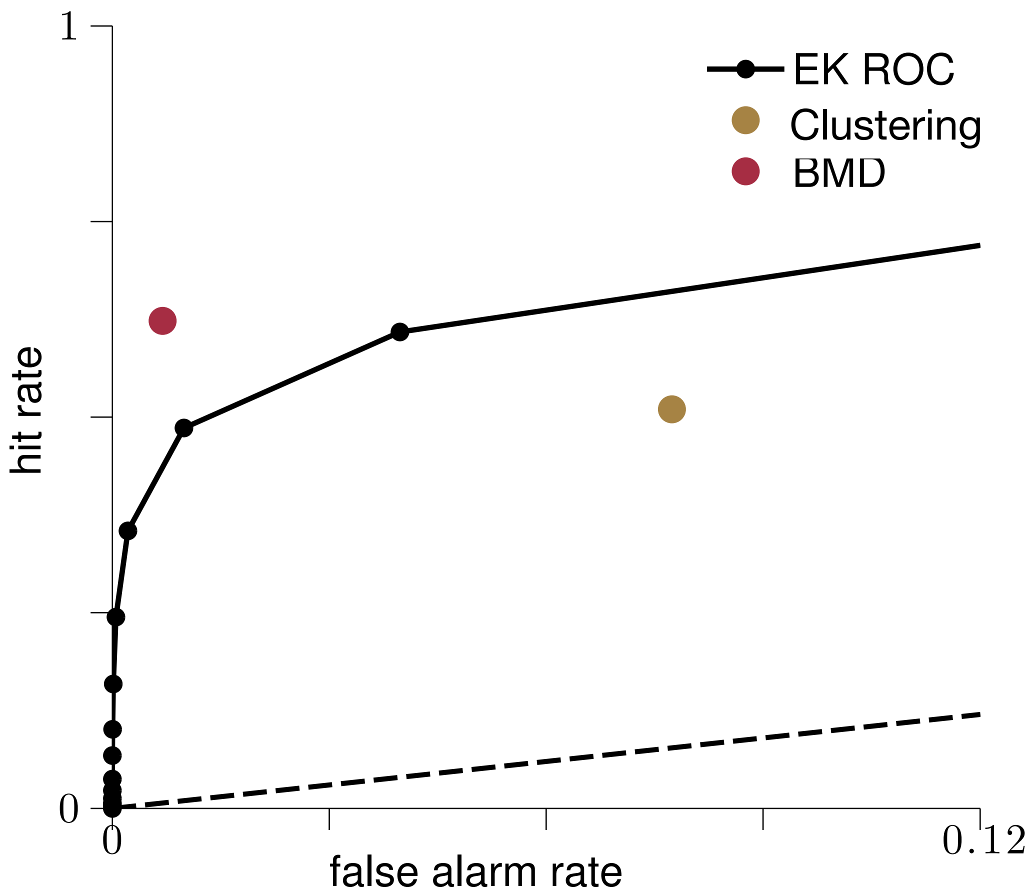
EK: Engbert and Kliegl
velocity threshold


1
0
0
0.12
hit rate
false alarm rate
EK ROC

Unsupervised clustering is less robust to noise than EK on simulated data



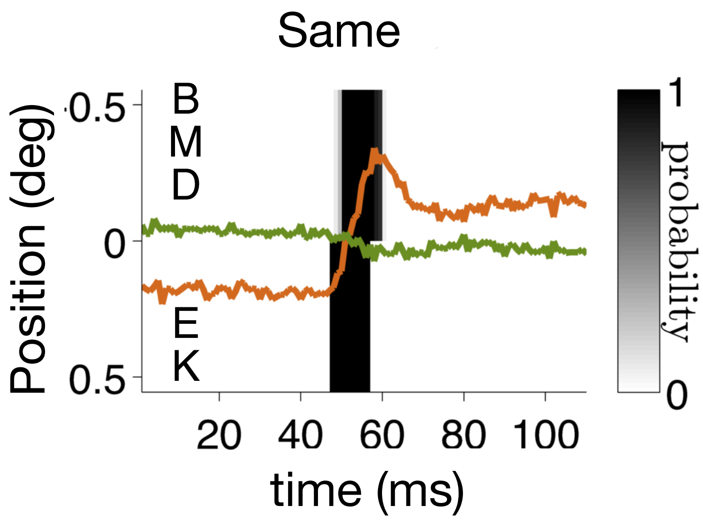
BMD is more robust to noise than EK on simulated data



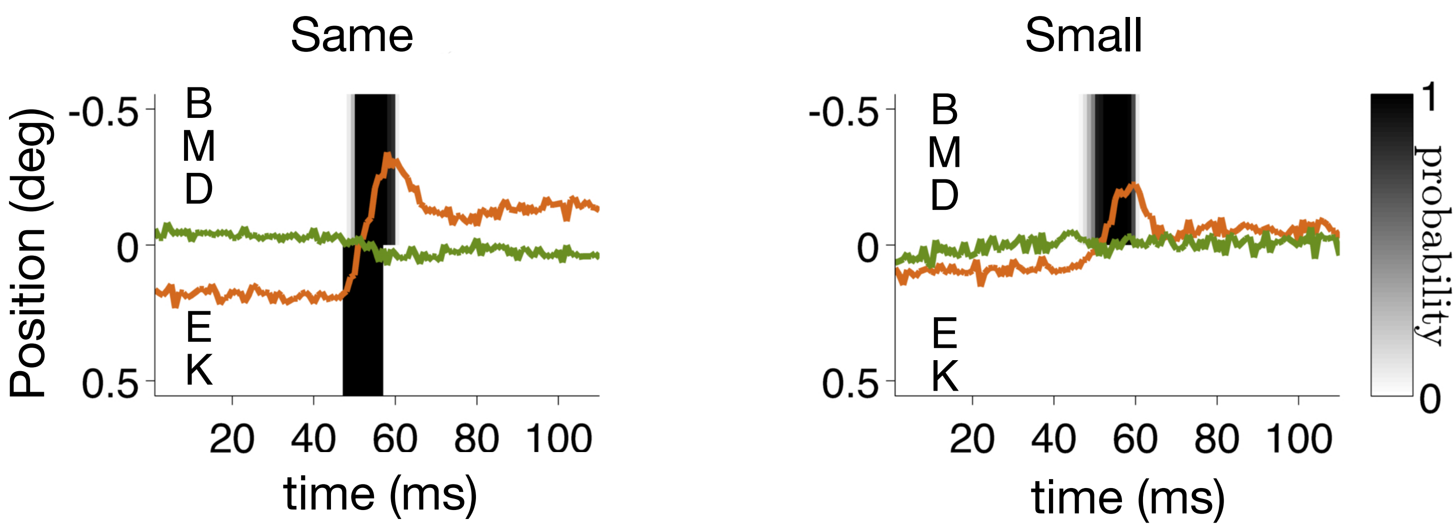
Tests on real data

1. EyeLink data
    - High measurement noise

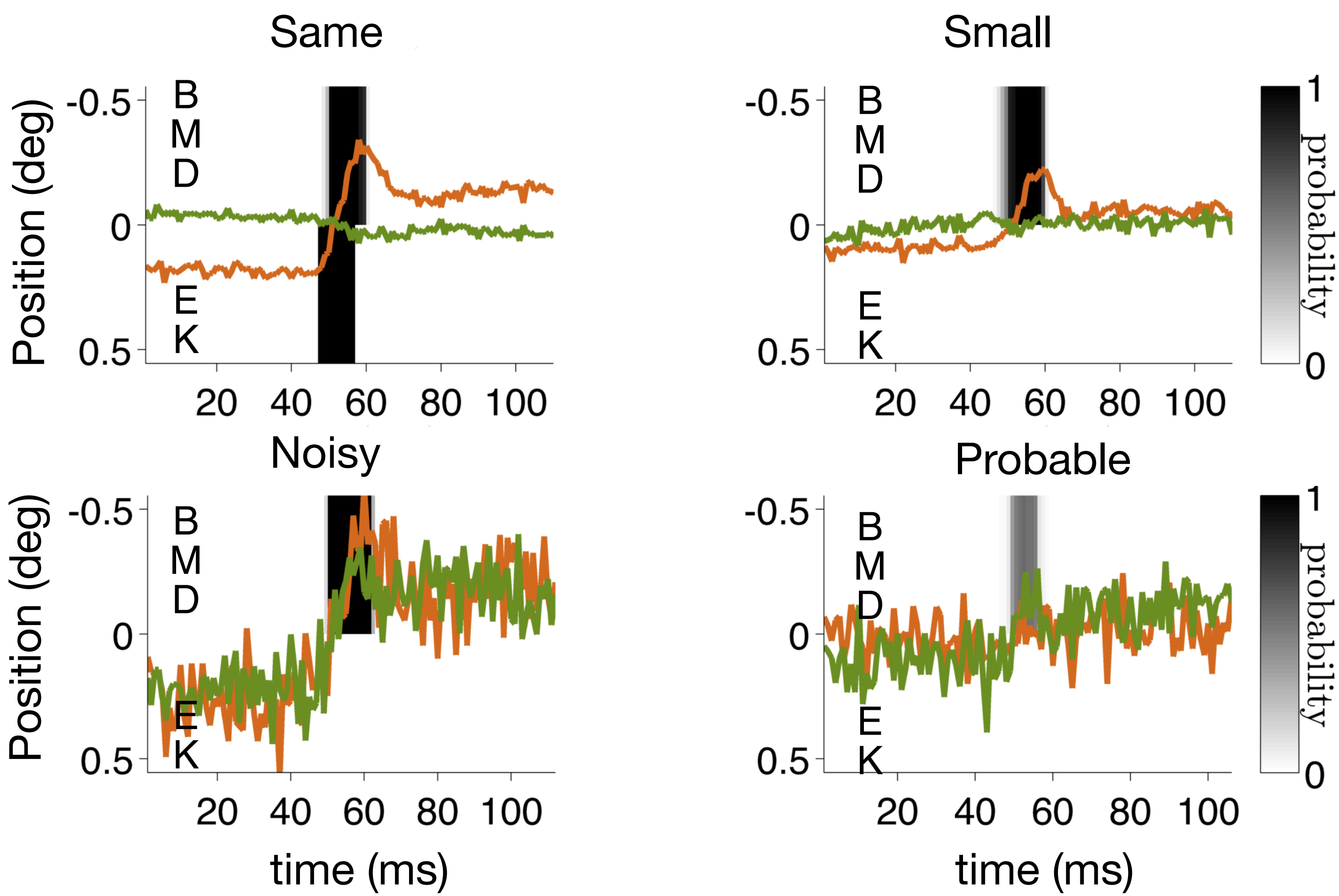
In most cases, BMD infers almost the same microsaccade as EK



BMD infers a **small** microsaccade missed by EK



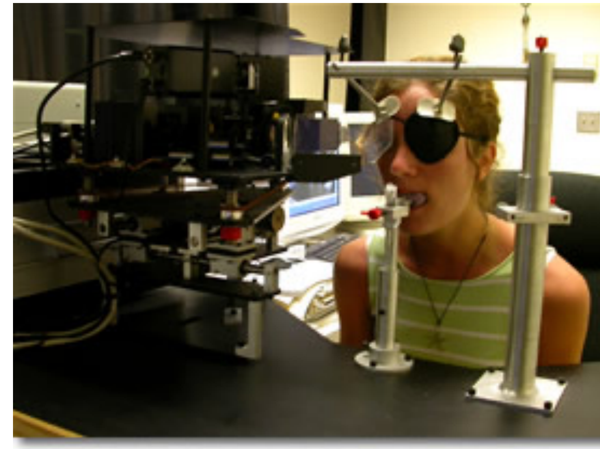
BMD infers probable microsaccades embedded in noise, undetected by EK



Tests on real data

1. EyeLink data
   - High measurement noise

2. Dual Purkinje Image data
   - Low measurement noise
   - Ko, Poletti and Rucci, 2010 and Poletti and Rucci, 2015

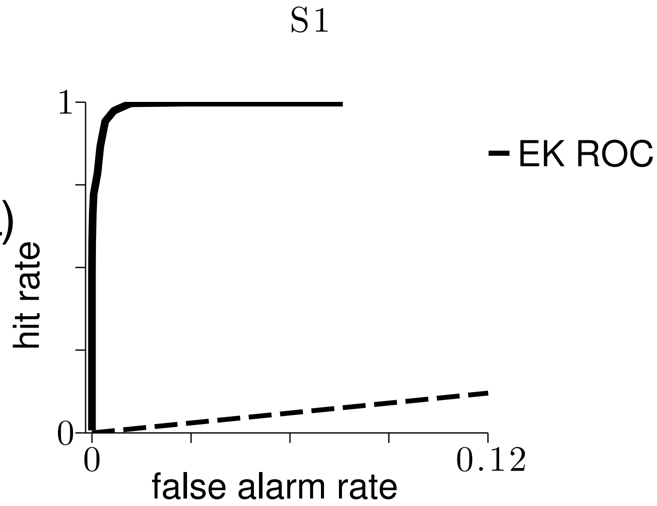


Poletti and Rucci lab

Strategy

- Set ground truth as the inferred microsaccades on low-measurement noise DPI data
- Artificially add measurement noise
- Compare robustness of EK, clustering and BMD

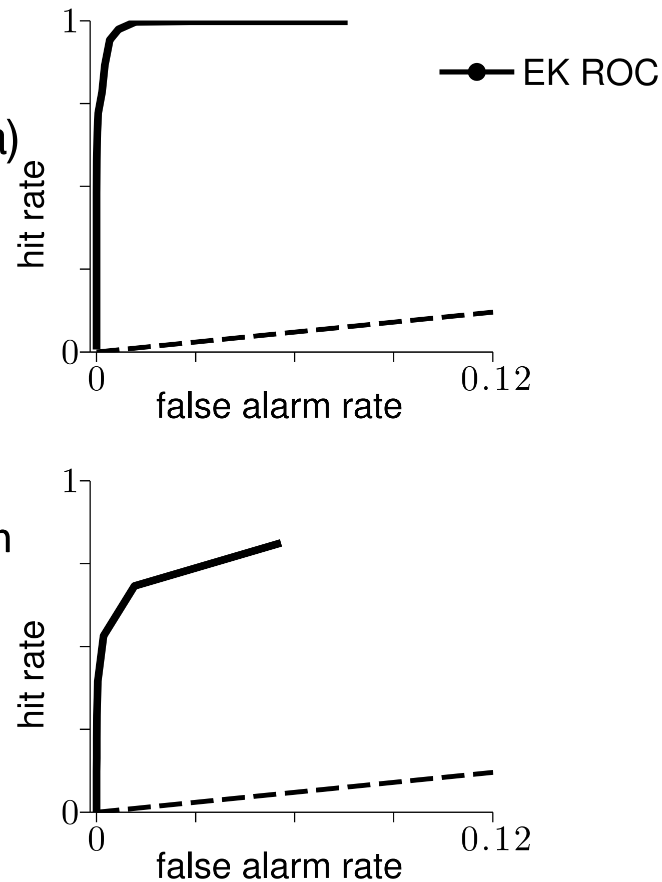
Low noise
(actual data)



Low noise
(actual data)

add
measurement
noise

Medium
noise



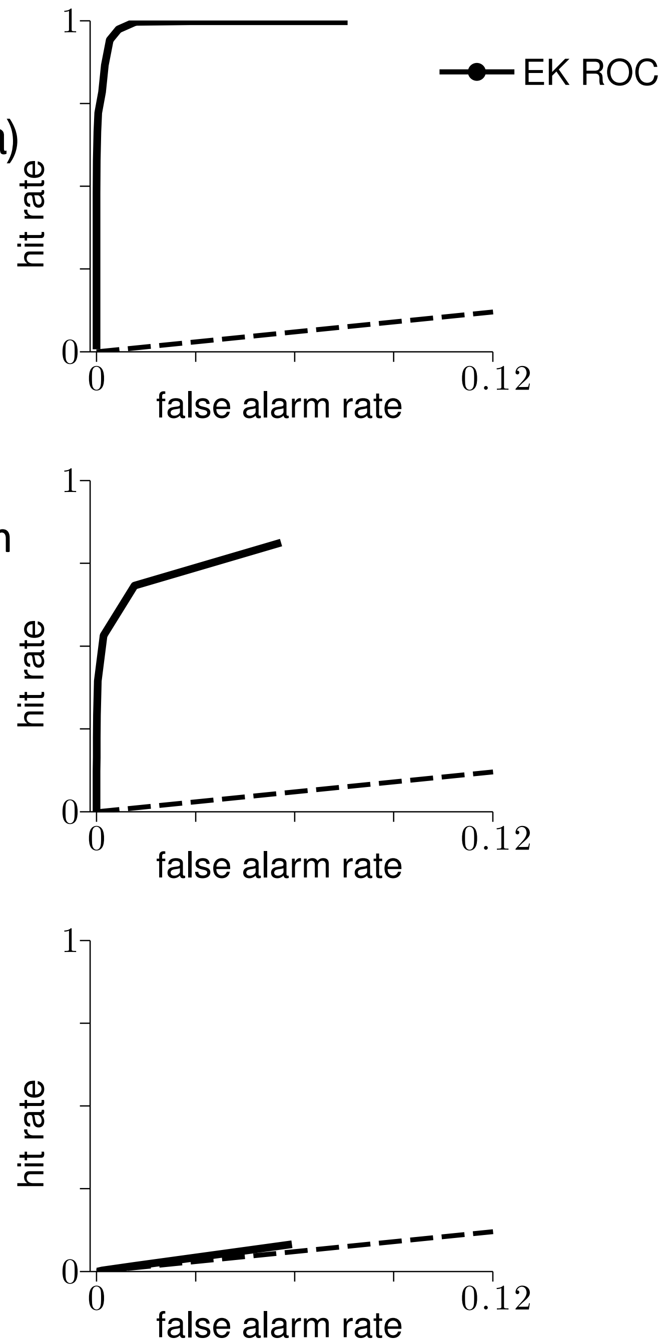
S1

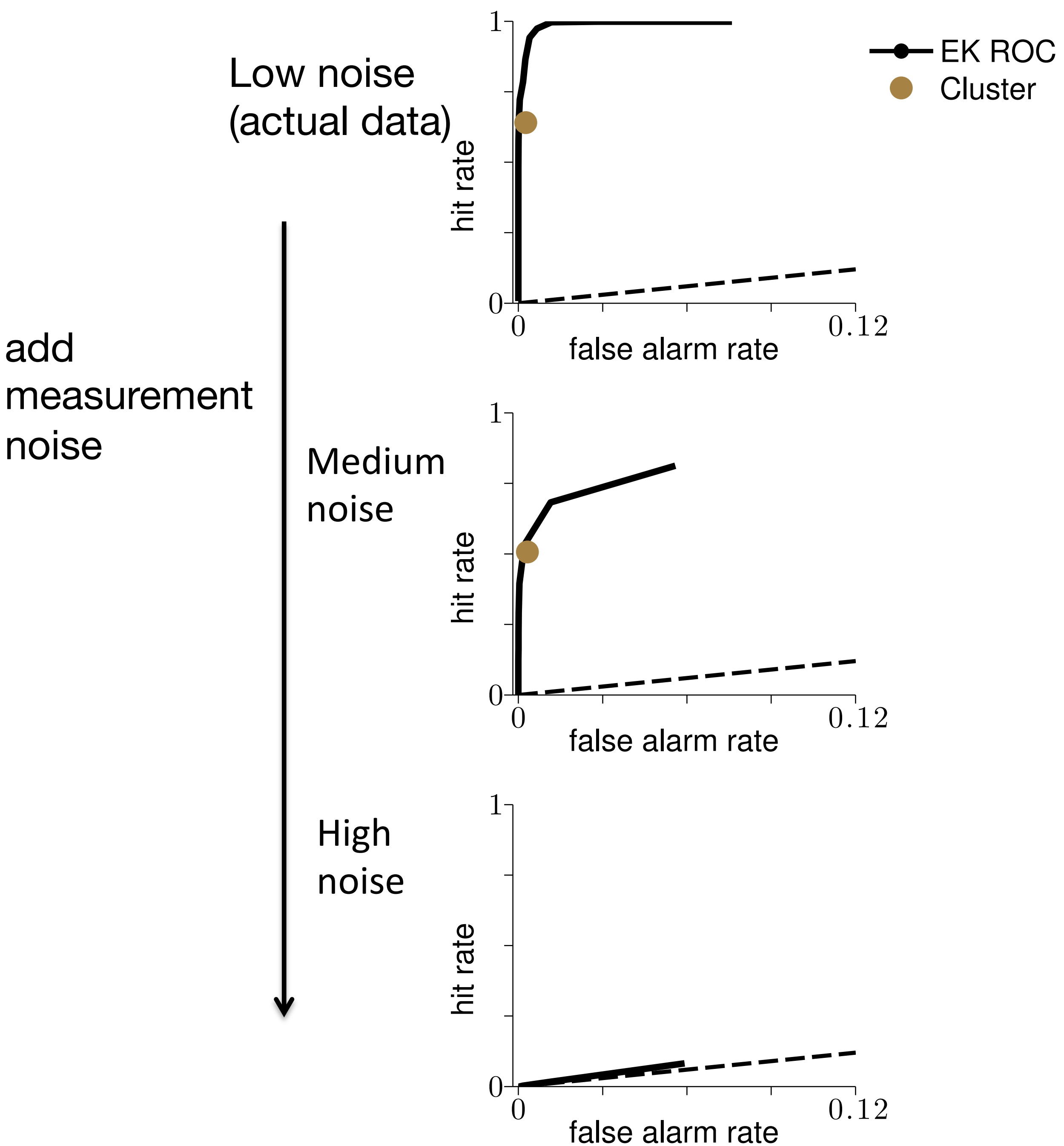
Low noise (actual data)

add measurement noise

Medium noise

High noise



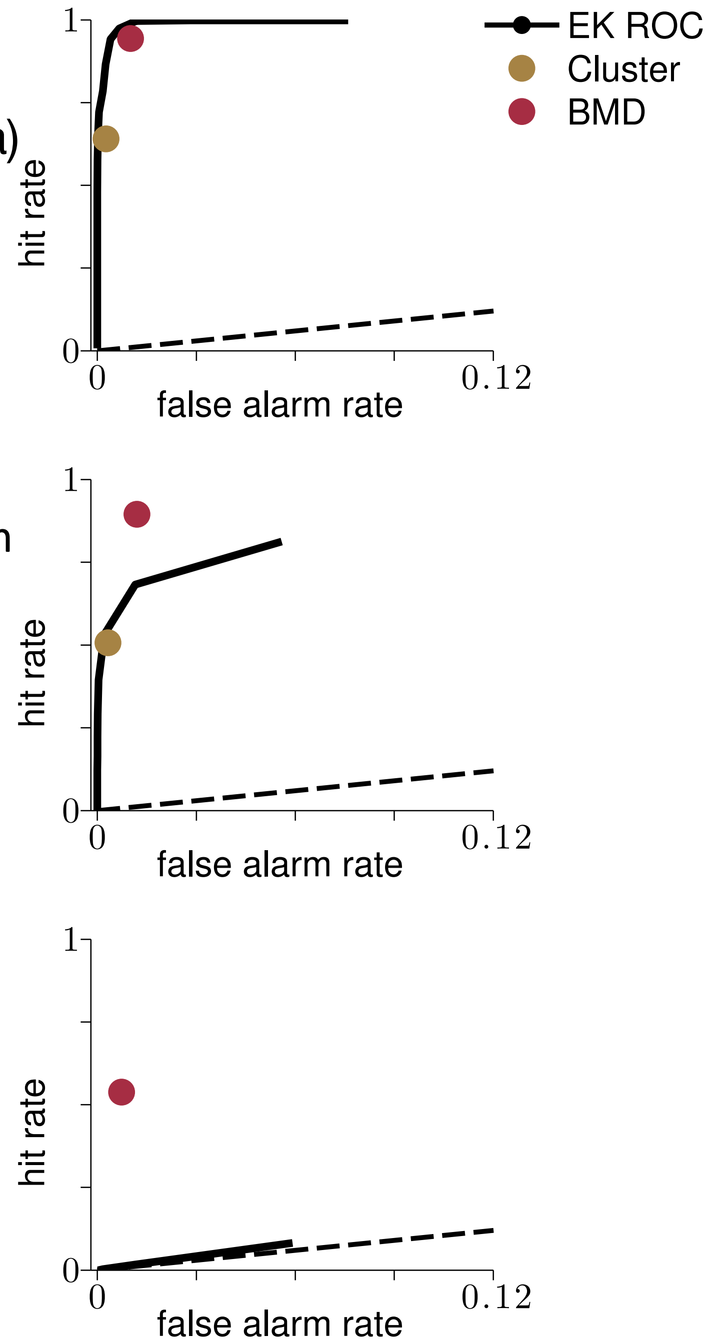

Low noise
(actual data)
add
measurement
noise
Medium
noise
High
noise
EK ROC
Cluster
hit rate
false alarm rate
1
0
0
0.12

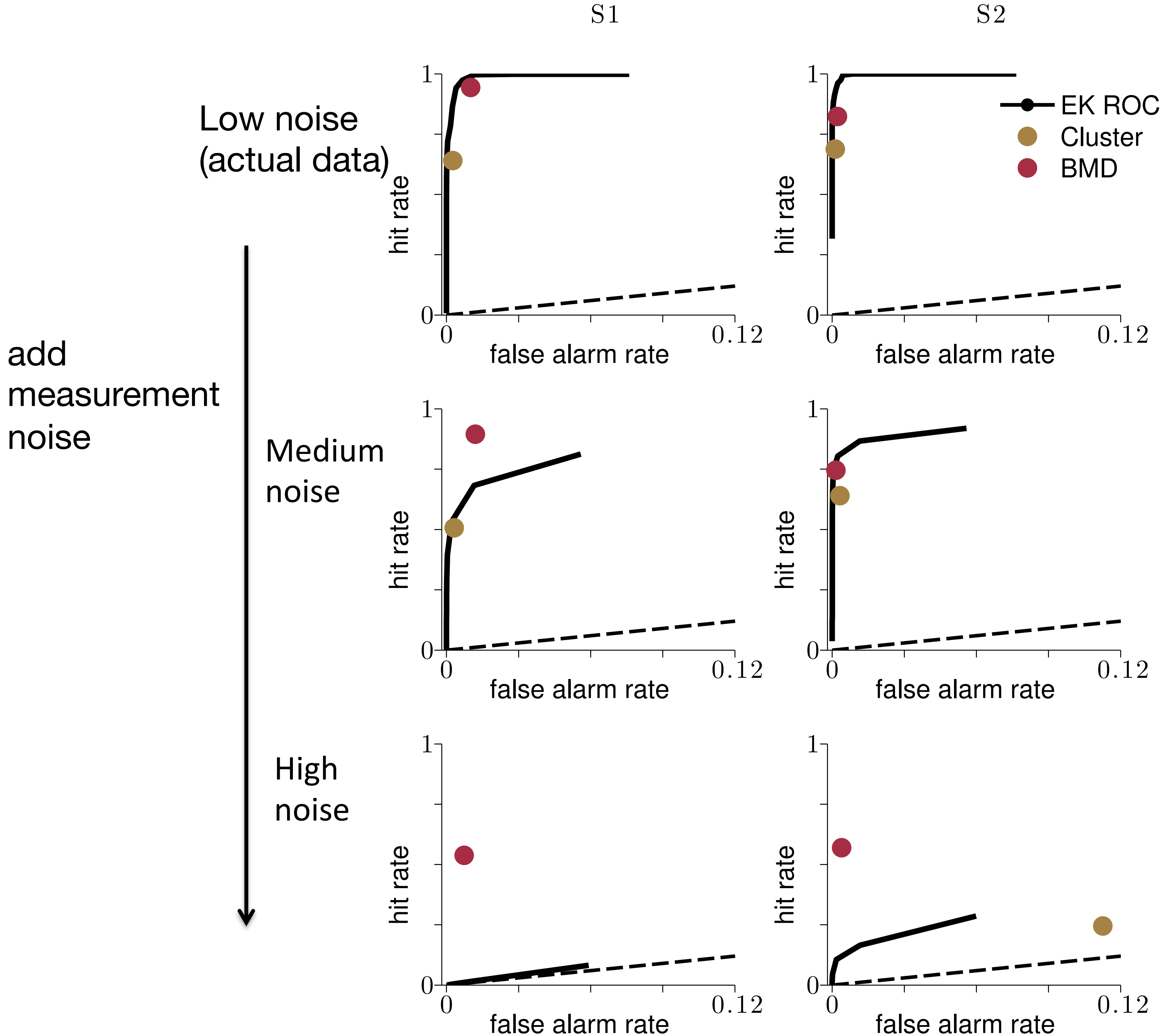
Low noise
(actual data)

add
measurement
noise

Medium
noise

High
noise




S1
S2
Low noise
(actual data)
add
measurement
noise
Medium
noise
High
noise
EK ROC
Cluster
BMD
hit rate
false alarm rate
1
0
0
0.12

- New Bayesian microsaccade detection (BMD) algorithm is more robust to measurement noise from the eye tracker

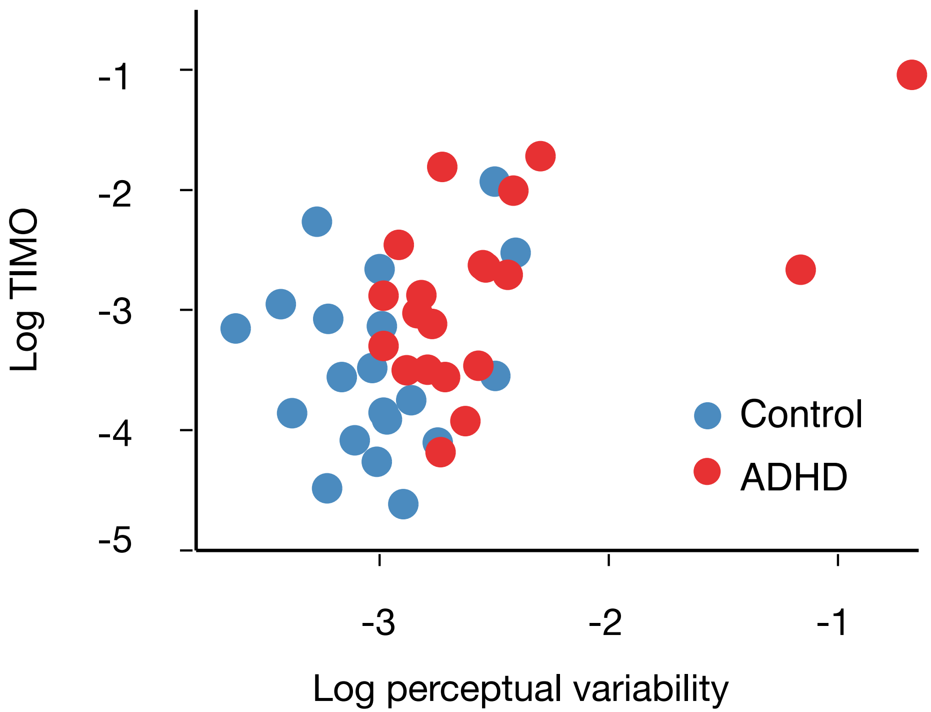
- What if I have data that was already filtered with EyeLink’s online filter and has super low noise?

  -Use unsupervised clustering or EK velocity threshold

BMD application to ADHD

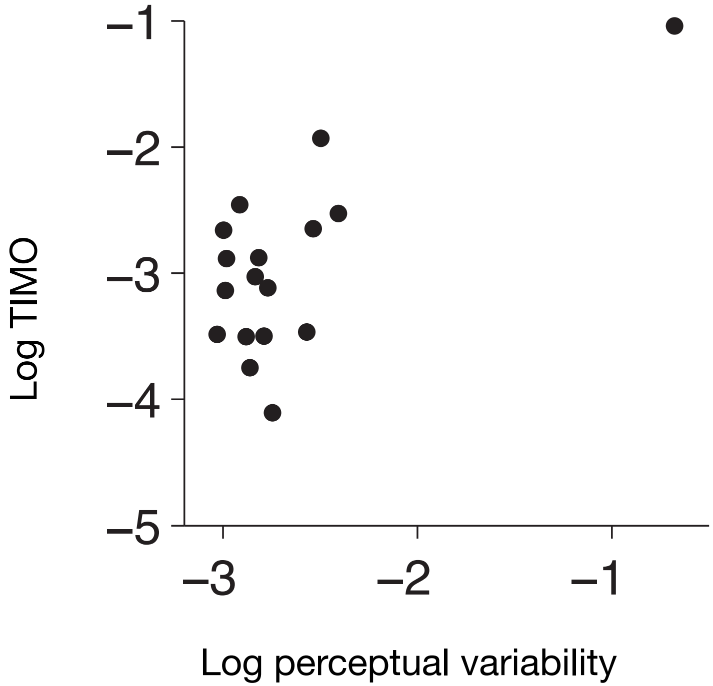
- Microsaccade suppression around stimulus onset could facilitate stimulus encoding *Martinez-Conde et al. 2013*
- Some evidence for differences in oculomotor control in ADHD *(Munoz et al. 2003)*, more recently microsaccades:
    - ADHD participants made more microsaccades during tasks that require stable fixation *Fried et al. 2014, Panagiotidi et al. 2017*
    - Less effective inhibition of microsaccade rate around expected stimulus onset *Dankner et al. 2017*

Part 2 data: Perceptual variability and TIMO are correlated



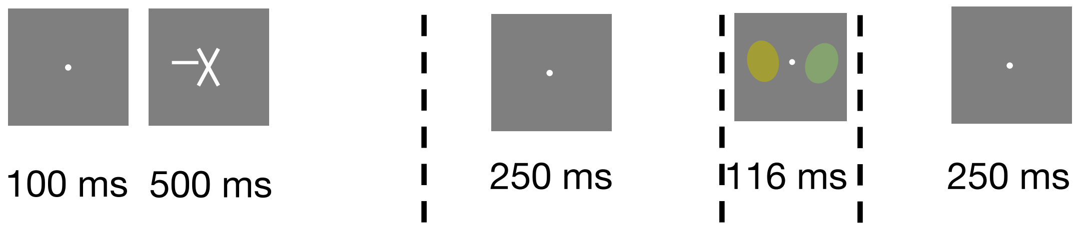
$$\boldsymbol{\rho = 0.41}$$
$$p = 0.0085$$

Perceptual variability and TIMO are correlated

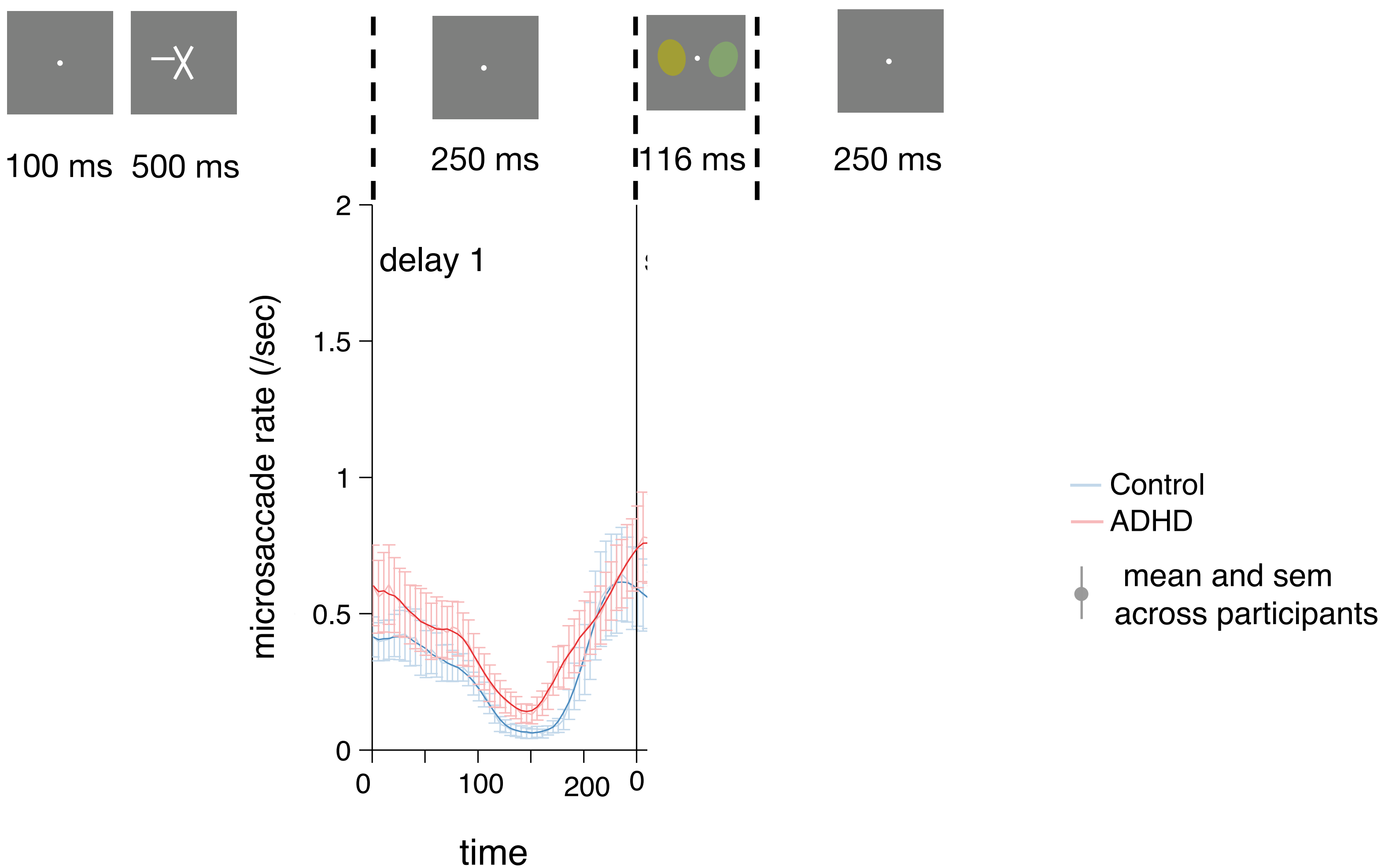


$\rho = 0.49$
$p = 0.03$

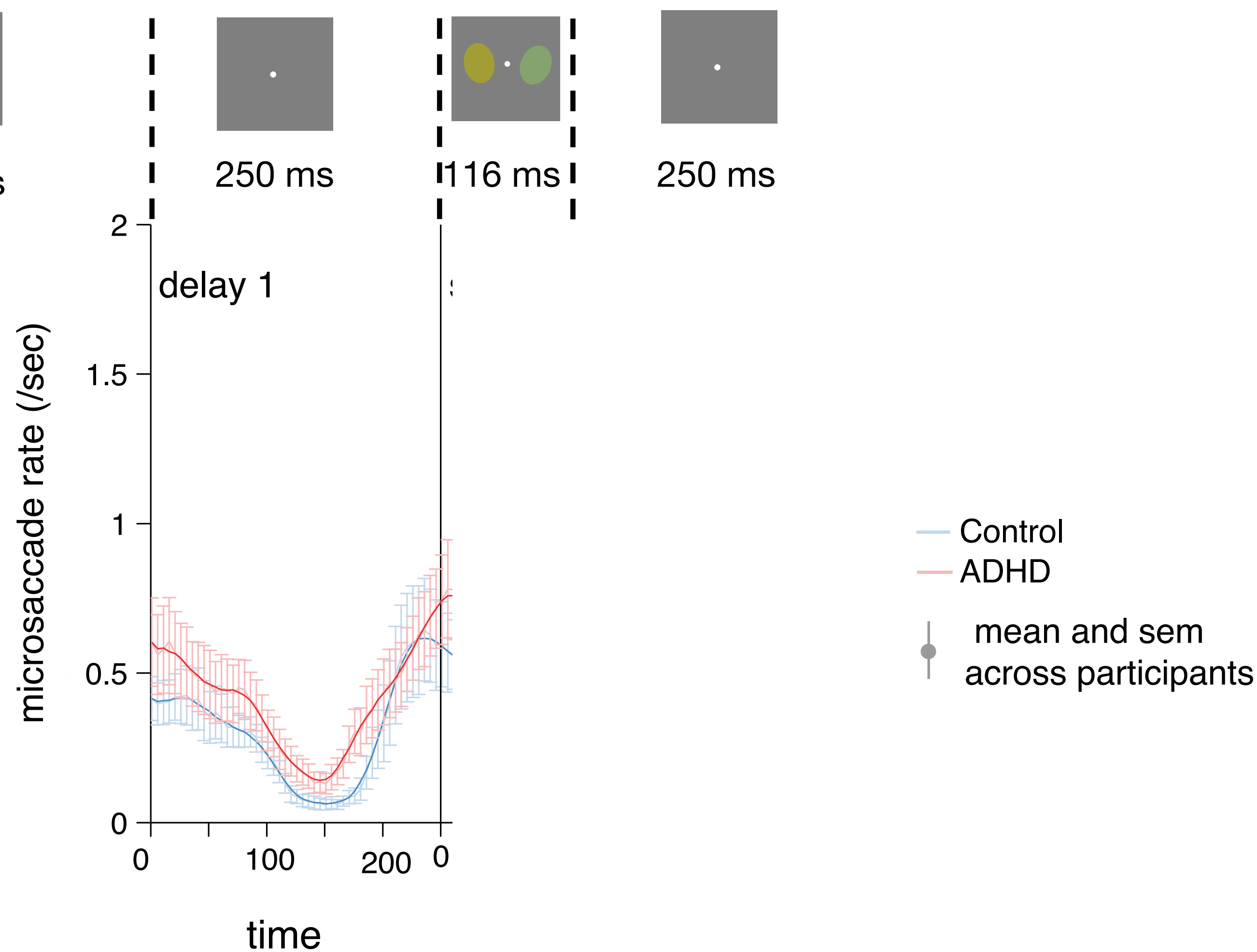
Microsaccade rate time courses



Microsaccade rate time courses around stimulus onset



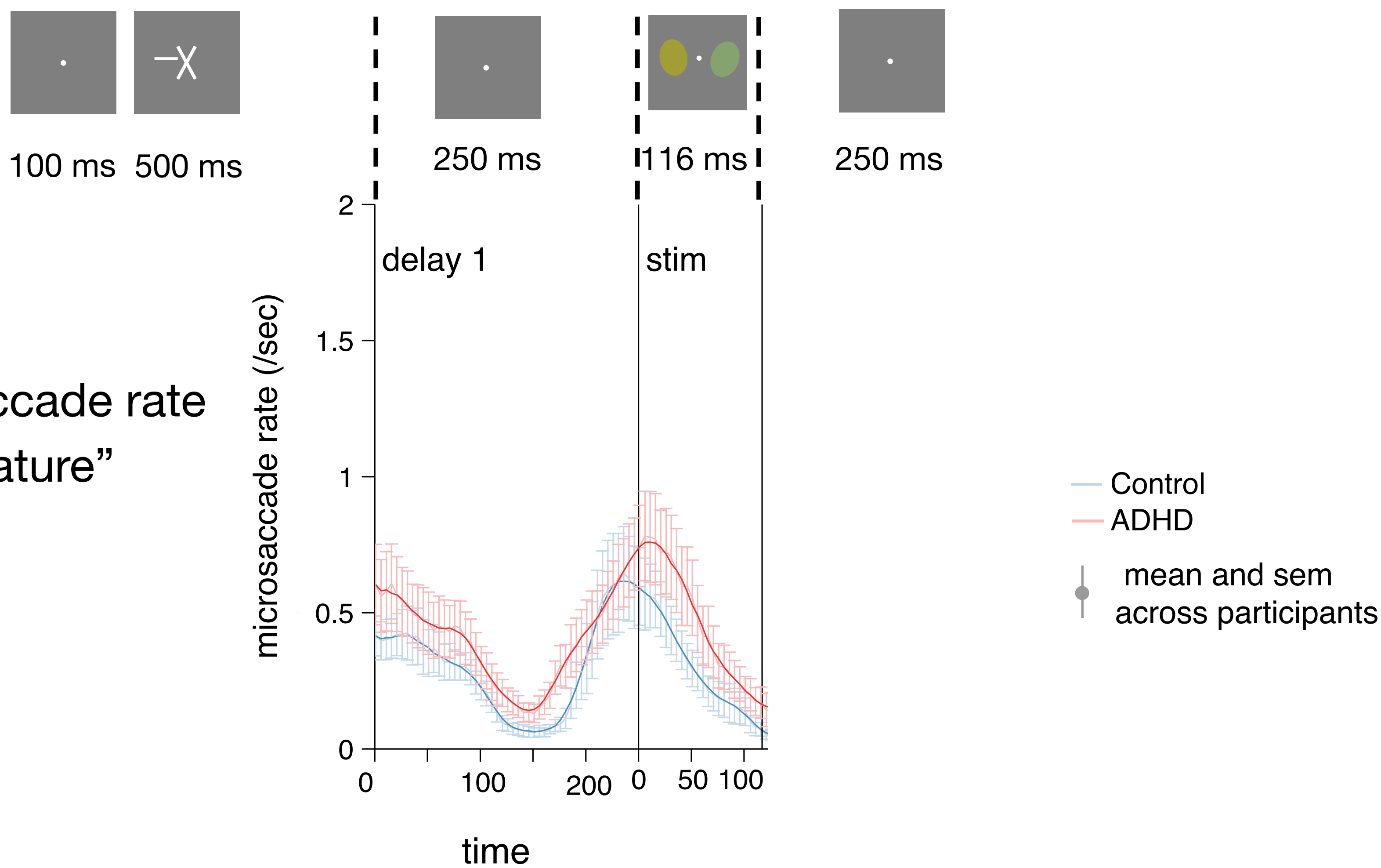
Microsaccade rate time courses around stimulus onset



"Microsaccade rate signature"

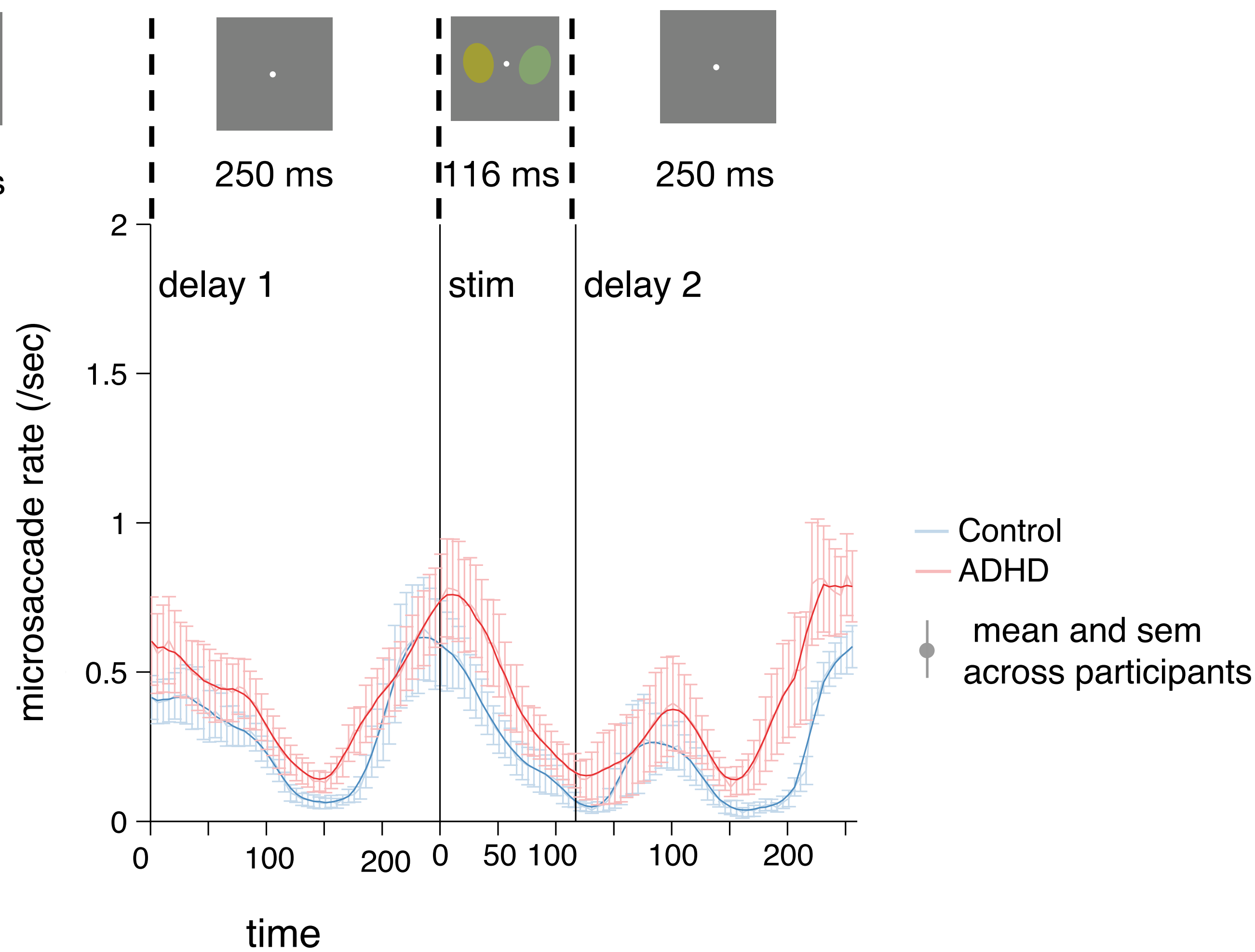
Microsaccade rate time courses around stimulus onset

“Microsaccade rate signature”



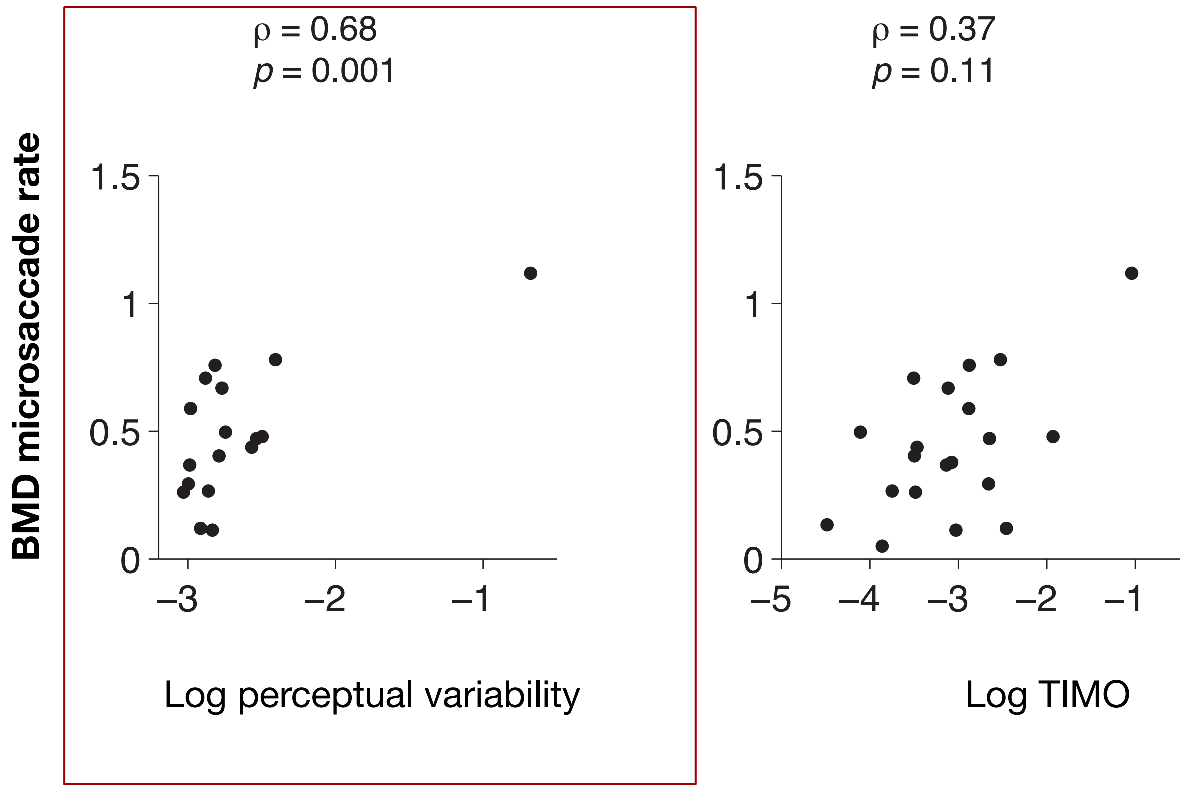
Microsaccade rate time courses around stimulus onset

"Microsaccade rate signature"



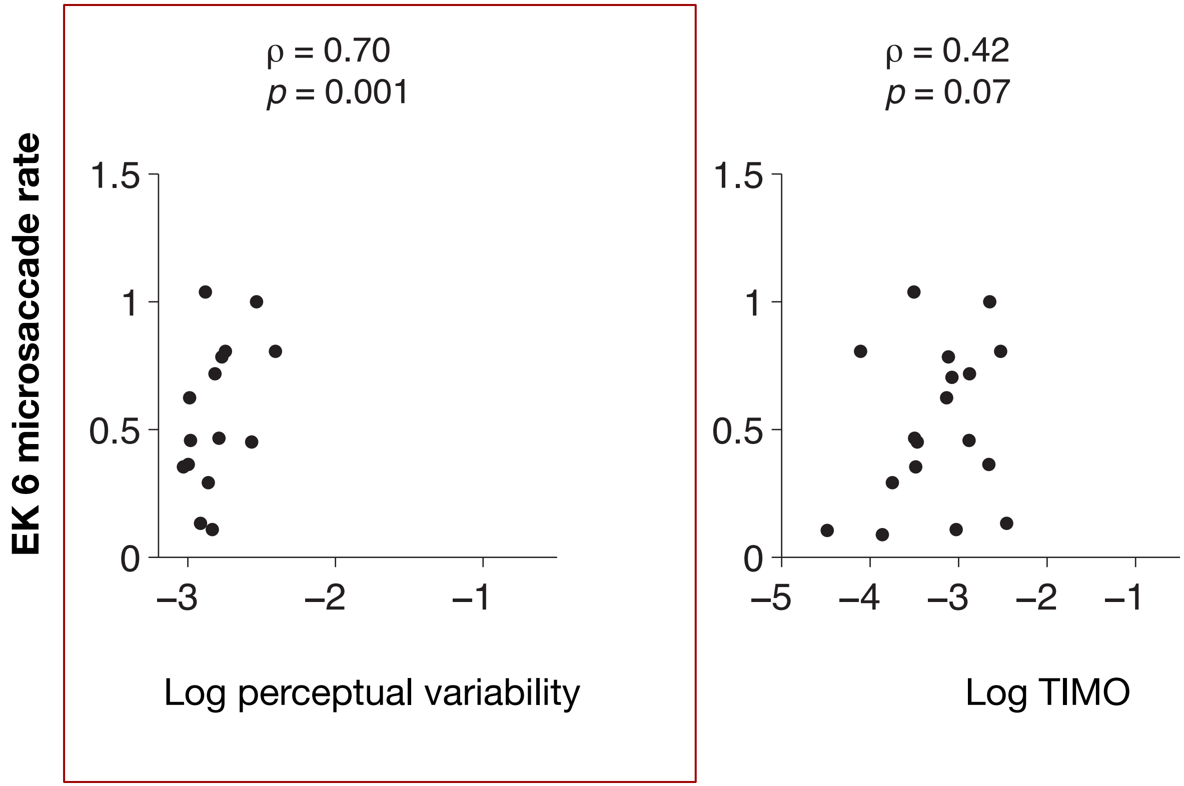
Perceptual variability and TIMO: differential correlation with microsaccade rate around stimulus onset

BMD algorithm



Perceptual variability and TIMO: differential correlation with microsaccade rate around stimulus onset

EK6 algorithm



Conclusions part 3

- Goal: develop a principled, Bayesian method for microsaccade detection and apply it to the ADHD and neurotypicals dataset in Part 2
- BMD recovers previously undetected microsaccades as it is more robust to eye tracker noise
- BMD algorithm brings a principled, Bayesian framework to microsaccade detection
- BMD has limitations:
  - It does not work well on low-noise or already filtered data
  - Speed
  - Assumptions, but improvements can be flexibly included
- Microsaccade rate around stimulus onset is correlated with the perceptual variability parameter, but not with TIMO
- Future: thorough exploration of microsaccade direction, velocity and amplitude distributions across all task periods

General conclusions

1. An **optimal-observer model** with a variable-precision encoding stage can capture observers' data in visual search with heterogeneous distractors.

2. We found **higher perceptual variability** in ADHD during a demanding task, along with worse cognitive control. Behavioral metrics alone yielded high diagnosis accuracy.

3. **Bayesian microsaccade detection** yields more robust inferences, especially under high measurement noise from the eye tracker.

4. Higher perceptual variability in ADHD could be associated with **less effective microsaccade suppression** around stimulus onset.

Acknowledgements

Ma Lab

- **Wei Ji Ma**
- **Bas van Opheusden**
- **Aditi Singh**
- Aspen Yoo
- Maija Honig
- Yanli Zhou
- Will Adler
- Luigi Acerbi
- Emin Orhan
- Qijia Zou
- Gianni Galbiati
- Zahy Bnaya
- Hongsup Shin

Outside collaborators:

- **Allison Young**
- Lenard Adler
- Michael Halassa

Committee:

- Marisa Carrasco
- Eero Simoncelli
- Roozbeh Kiani
- Susana Martinez-Conde

CNS, Psych and beyond

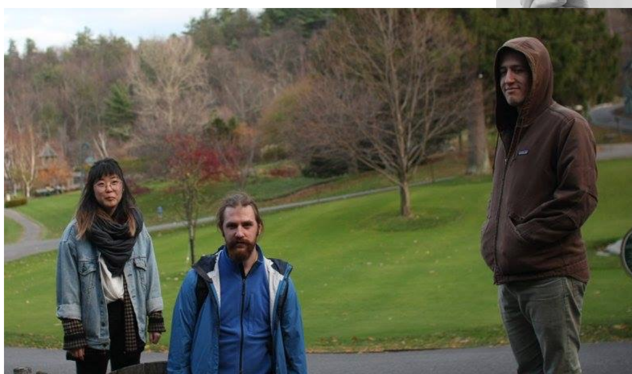
- Alex Reyes
- Mike Hawken
- Classmates
- Mariel Roberts
- Theory suite people
- Jess Holman
- Amala Ankolekar
- Paul Fan
- Erik Cruz
- Shenglong Wang

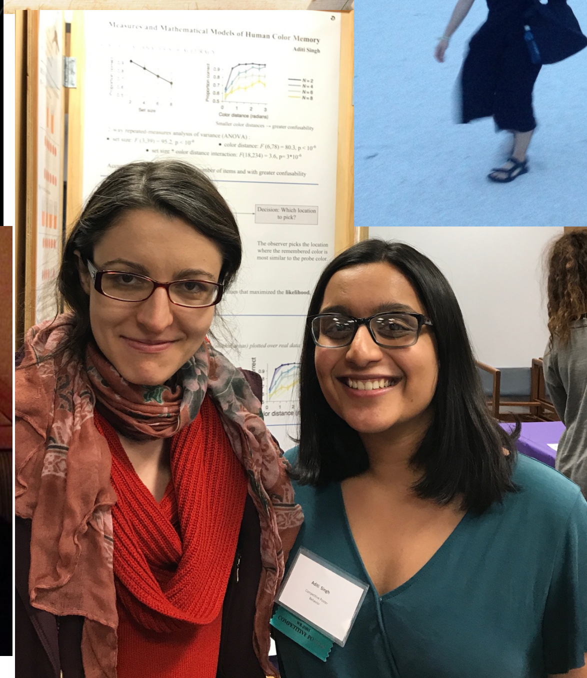
Talk practice & support

- Weiji ^n
- Ma lab ^2
- Christina Hatch
- Zhiwei Li
- Peyuan Zhang
- Jorge Jaramillo
- Lyuda Kovalchuke
- Emily Cowan
- Cristina Popa
- Maddy Joglekar

⋮



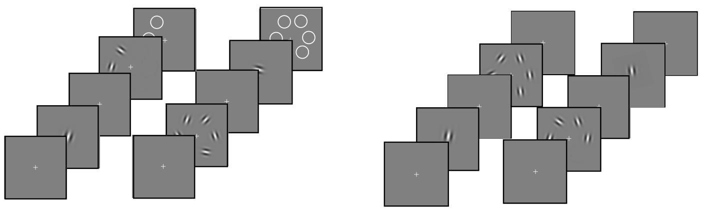


Measures and Mathematical Models of Human Color Memory
Aditi Singh

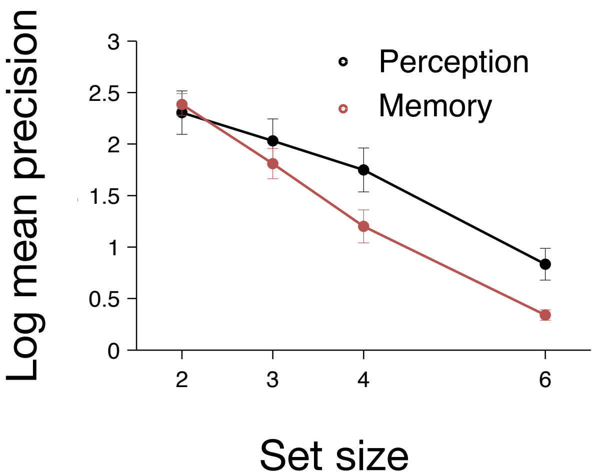

Recap measures related to visual covert attention

Part 1

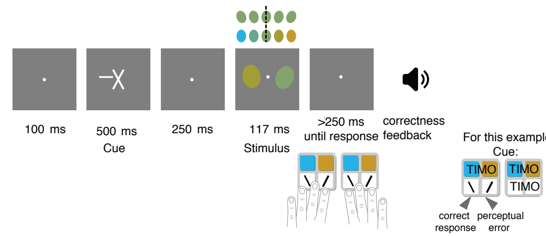
Localization

Detection

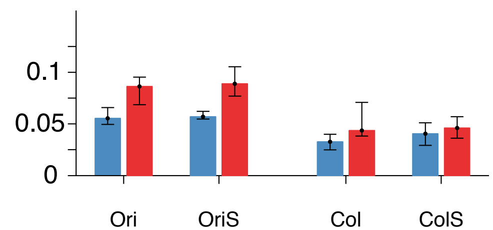




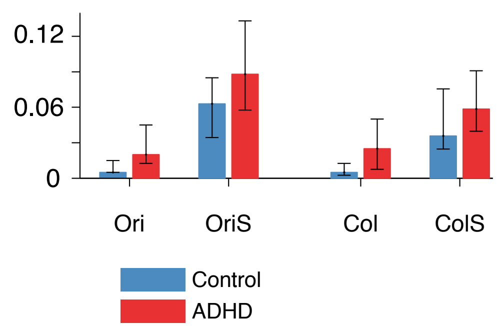
Part 2



Perception/Attention

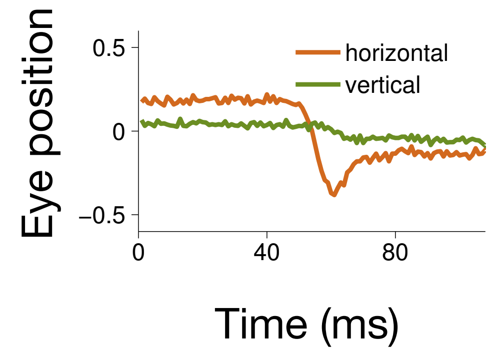


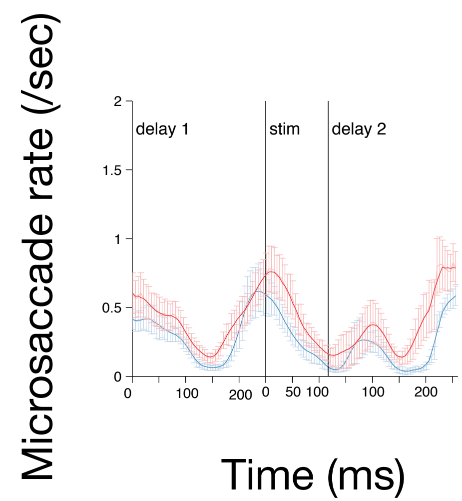
Cognitive control



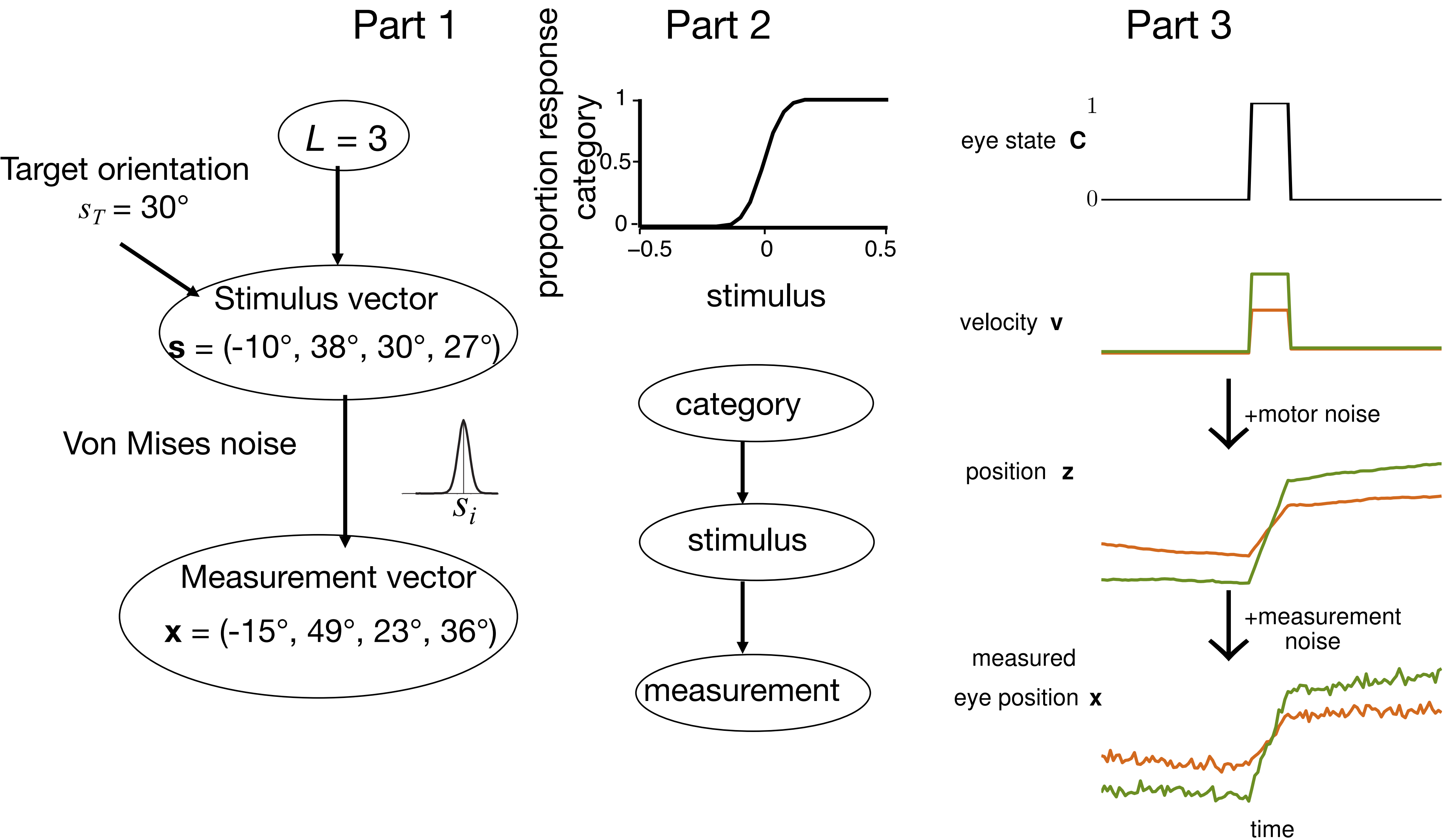
Part 3

Microsaccades

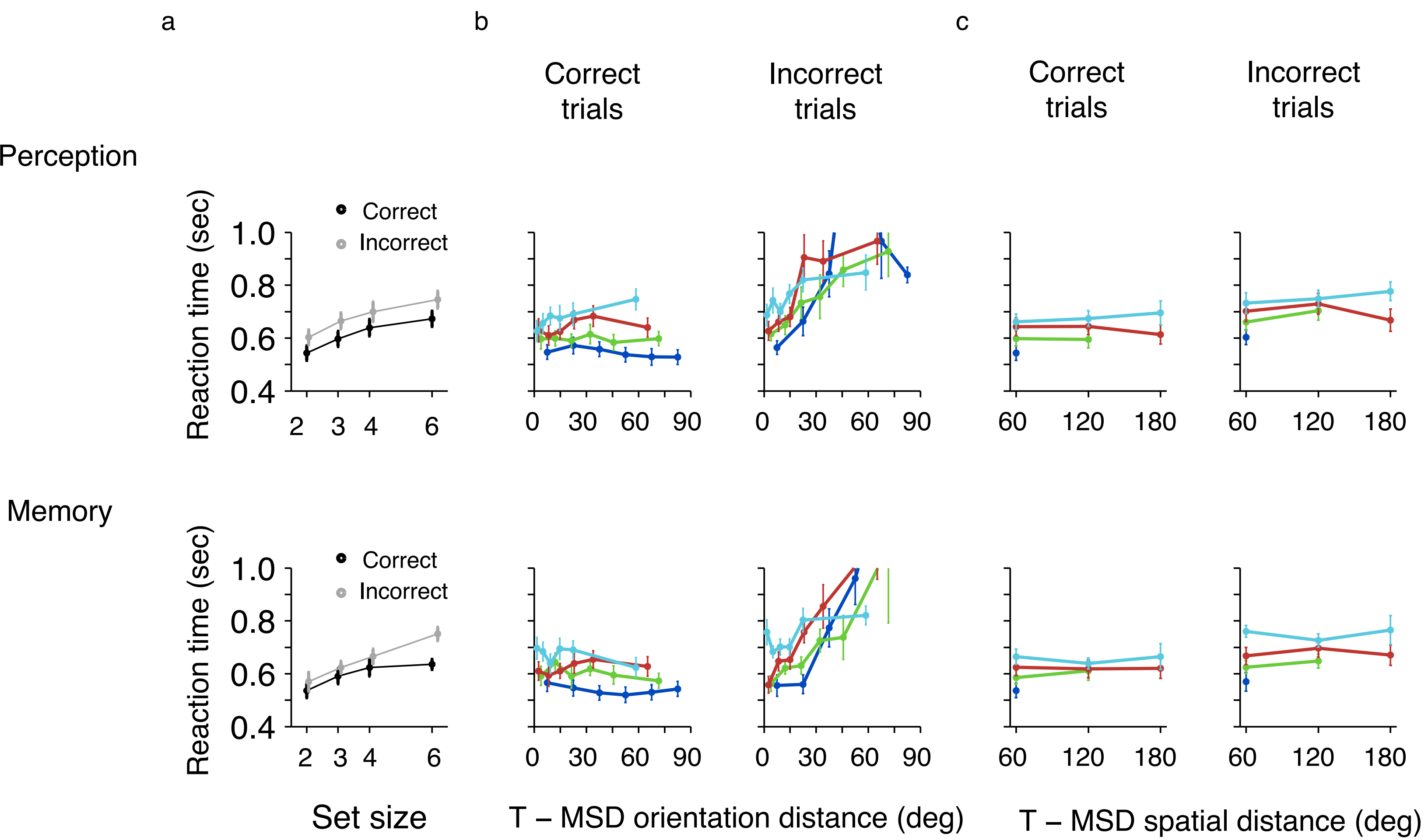




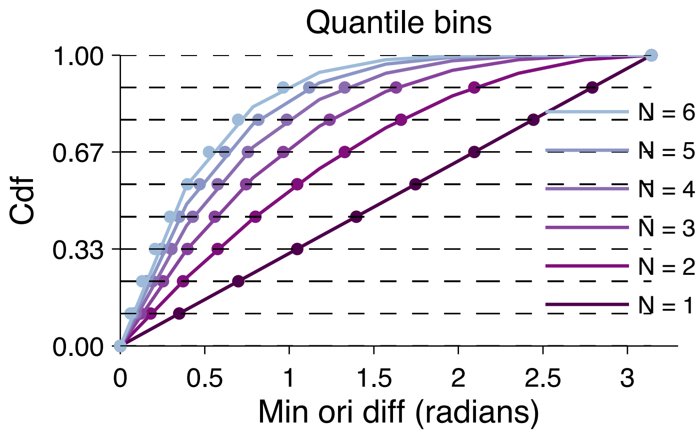
Recap models related to visual covert attention



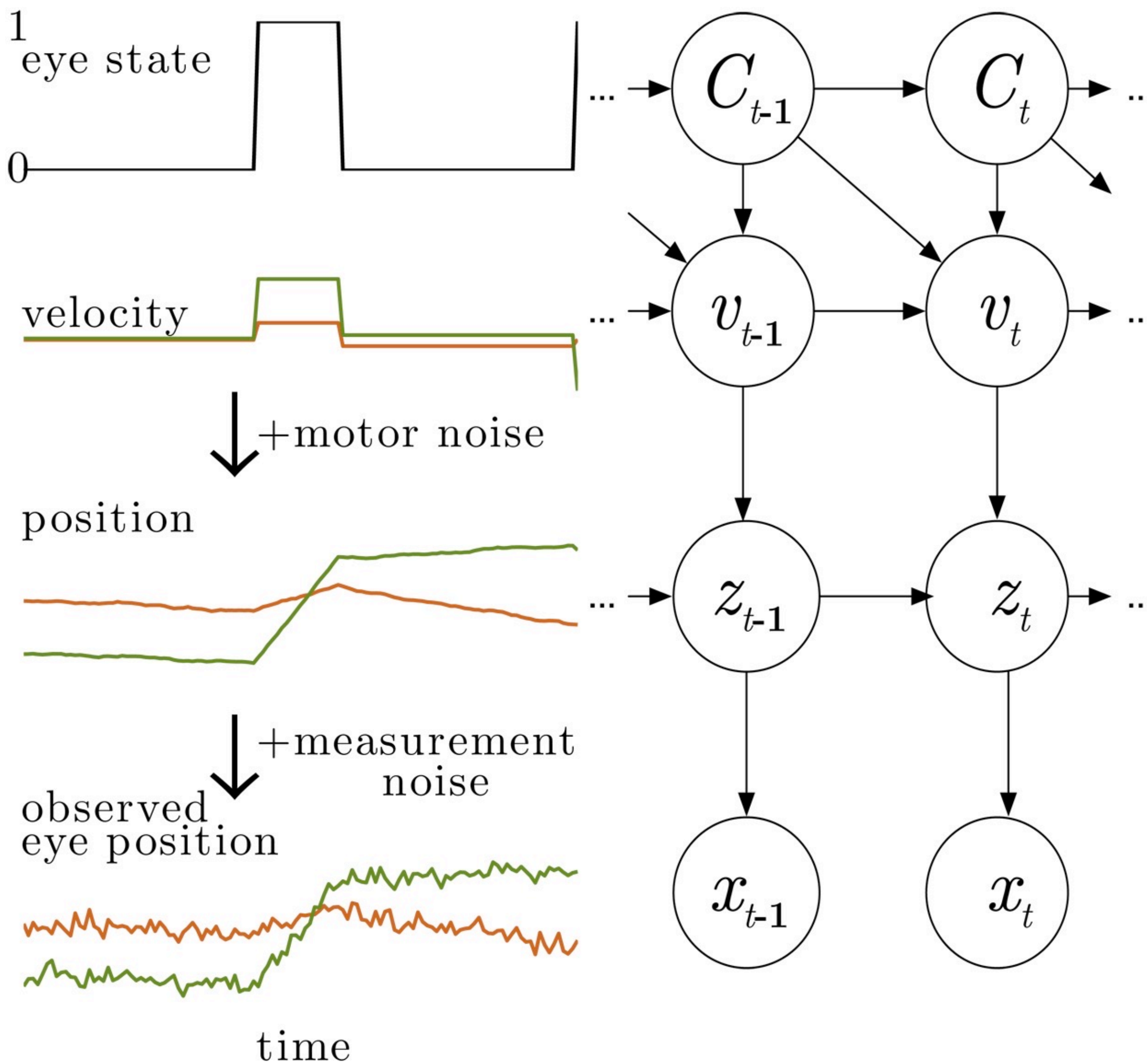
Reaction times Exp 1



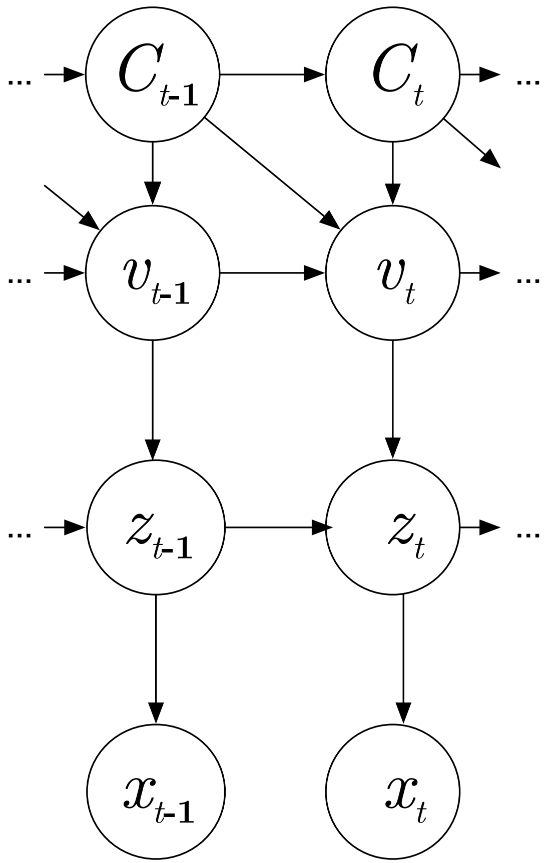
Quantile binning



Generative model of fixational eye movements



Generative model of fixational eye movements



$$p(\mathbf{C}) \quad \propto \quad \prod_{k=1}^{2n-1} Gamma\left(\Delta\tau_k; 2, \frac{1}{\lambda_{k \pmod 2}}\right)$$

$$p(v_t|C_t = C_{t-1}) = \delta(v_t - v_{t-1})$$

$$p(v_t|C_t = 1, C_{t-1} = 0) \quad \propto \quad Gen\ Gamma\left(||v_t||; \sigma_1, d_1, 2\right)$$

$$p(v_t|C_t = 0, C_{t-1} = 1) \quad \propto \quad Gen\ Gamma\left(||v_t||; \sigma_0, d_0, 2\right)$$

$$p(z_t|z_{t-1}, v_t) \quad = \quad \mathcal{N}(z_t; z_{t-1} + v_t, \Sigma_z)$$

$$p(x_t|z_t) \quad = \quad \mathcal{N}(x_t; z_t, \Sigma_x)$$

Step	Operation
0	Initialize **C**, $\widehat{\sigma_1}$, $\widehat{d_1}$, $\widehat{\sigma_0}$
1	Estimate the motor and measurement noise: $\hat{\sigma}_z$, $\hat{\sigma}_x$
2	Estimate $\hat{\mathbf{z}}$ from observations **x**: Kalman smoother
3	Sample from the posterior over **C**: MCMC
4	Estimate the velocity distribution parameters: maximum likelihood estimation (MLE)
	Return to Step 1

Table 1. BMD algorithm.