



New York State  
Psychiatric Institute



COLUMBIA  
COLUMBIA UNIVERSITY  
DEPARTMENT OF PSYCHIATRY

ASSC 26

# Introspective inference counteracts perceptual distortion

06/25/2023

Andra Mihali, Marianne Broeker, Florian Ragalmuto and Guillermo Horga





bioRxiv posts many COVID19-related papers. A reminder: they have not been formally peer-reviewed and should not guide health-related behavior or be reported in the press as conclusive.

New Results

 [Follow this preprint](#)

## Introspective inference counteracts perceptual distortion

Andra Mihali, Marianne Broeker, Florian Ragalmuto, Guillermo Horga

**doi:** <https://doi.org/10.1101/2021.11.13.468497>

This article is a preprint and has not been certified by peer review [what does this mean?].

[Abstract](#)[Full Text](#)[Info/History](#)[Metrics](#) [Preview PDF](#)

### Abstract

Introspective agents can recognize the extent to which their internal perceptual experiences deviate from the actual states of the external world. This ability, also known as insight, is critically required for reality testing and is impaired in psychosis, yet very little is known about its cognitive underpinnings. We developed a Bayesian modeling framework and a novel psychophysics paradigm to quantitatively characterize this type of insight while participants experienced a motion after-effect illusion. Participants could compensate for the illusion when judging the actual direction of a motion stimulus. Furthermore, confidence, reaction time, and

 [Previous](#)

Posted September 30, 2022.

 [Download PDF](#) [Print/Save Options](#) [Revision Summary](#) [Tweet](#) [Like 0](#)

**COVID-19 SARS-  
medRxiv and bio**

### Subject Area

[Neuroscience](#)

**Subject Areas**

**All Articles**

[Animal Behavior](#)[Biochemistry](#)[Biotechnology](#)



New York State  
Psychiatric Institute



COLUMBIA  
COLUMBIA UNIVERSITY  
DEPARTMENT OF PSYCHIATRY

ASSC 26

Introspective inference counteracts  
perceptual distortion

Quantitative *measure* and  
candidate *process model* for **insight**  
with relevance to psychiatry  
(computational psychiatry)



New York State  
Psychiatric Institute



COLUMBIA  
COLUMBIA UNIVERSITY  
DEPARTMENT OF PSYCHIATRY

ASSC 26

# Introspective inference counteracts perceptual distortion

Quantitative *measure* and candidate *process model* for **insight** with relevance to psychiatry (computational psychiatry)

Hallucinations  
Illusions



# Introspective inference counteracts perceptual distortion

Quantitative *measure* and candidate *process model* for **insight** with relevance to psychiatry (computational psychiatry)

Hallucinations  
**Illusions**

# Perceptual insight

## *Operationalization*

- The ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

## *Experimental test*

- Inducing perceptual distortion and probing compensation... through Motion after-effect (MAE)

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bjp.2019.217

Analysis

Insight and psychosis: the next 30 years

Anthony S. David

JAMA Network Insights

Working With Patients With Impaired Insight

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henderson, MD

**no objective measures  
or process models**

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bjp.2019.217

Analysis

Insight and psychosis: the next 30 years

Anthony S. David

JAMA Network Insights

Working With Patients With Impaired Insight

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henderson, MD

**no objective measures  
or process models**

# Visual metacognition

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bjp.2019.217

Analysis

## Insight and psychosis: the next 30 years

Anthony S. David

JAMA Network Insights

## Working With Patients With Impaired Insight

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henderson, MD

**no objective measures or process models**

# Visual metacognition

the ability for one's subjective sense of confidence to track the objective accuracy

**several objective measures and process models**



© 2021 American Psychological Association  
ISSN: 0003-066X

American Psychologist

2021, Vol. 76, No. 9, 1445–1453  
<https://doi.org/10.1037/amp0000937>

## Visual Metacognition: Measures, Models, and Neural Correlates

Dobromir Rahnev

School of Psychology, Georgia Institute of Technology



RESEARCH ARTICLE

Comparing Bayesian and non-Bayesian accounts of human confidence reports

William T. Adler<sup>1\*</sup>, Wei Ji Ma<sup>1,2</sup>

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bjp.2019.217

Analysis

Insight and psychosis: the next 30 years

Anthony S. David

JAMA Network Insights

Working With Patients With Impaired Insight

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henderson, MD

**no objective measures or process models**

# Visual metacognition

the ability for one's subjective sense of confidence to track the objective accuracy

**several objective measures and process models**



© 2021 American Psychological Association  
ISSN: 0003-066X

American Psychologist

2021, Vol. 76, No. 9, 1445–1453  
<https://doi.org/10.1037/amp0000937>

Visual Metacognition: Measures, Models, and Neural Correlates

Dobromir Rahnev  
School of Psychology, Georgia Institute of Technology



RESEARCH ARTICLE

Comparing Bayesian and non-Bayesian accounts of human confidence reports

William T. Adler<sup>1\*</sup>, Wei Ji Ma<sup>1,2</sup>

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bj.2019.217

Analysis

## Insight and psychosis: the next frontier

Anthony S. David

JAMA Network Insights

## Working With Patients With Schizophrenia

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henc

**no objective measures or process models**



Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: [www.elsevier.com/locate/neubiorev](http://www.elsevier.com/locate/neubiorev)

Review article

Perceptual reality monitoring: Neural mechanisms dissociating imagination from reality

Nadine Dijkstra <sup>a,\*</sup>, Peter Kok <sup>a</sup>, Stephen M. Fleming <sup>a,b,c</sup>

<sup>a</sup> Wellcome Centre for Human Neuroimaging, University College London, United Kingdom

<sup>b</sup> Max Planck UCL Centre for Computational Psychiatry and Aging Research, University College London, United Kingdom

<sup>c</sup> Department of Experimental Psychology, University College London, United Kingdom



Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: [www.elsevier.com/locate/neubiorev](http://www.elsevier.com/locate/neubiorev)

Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders

Martin Rouy <sup>a,\*</sup>, Pauline Saliou <sup>a</sup>, Ladislas Nalborczyk <sup>b</sup>, Michael Pereira <sup>a</sup>, Paul Roux <sup>c,1</sup>, Nathan Faivre <sup>a,1</sup>

# Visual metacognition

the ability for one's subjective sense of confidence to track the objective accuracy

**several objective measures and process models**

American Psychologist

2021, Vol. 76, No. 9, 1445–1453  
<https://doi.org/10.1037/amp0000937>

**Measures, Models, and Neural Correlates**

Dobromir Rahnev  
Psychiatry, Georgia Institute of Technology

RESEARCH ARTICLE

Comparing Bayesian and non-Bayesian accounts of human confidence reports

William T. Adler <sup>1,\*</sup>, Wei Ji Ma <sup>1,2</sup>

# Perceptual insight

## *Operationalization*

- The ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

## *Experimental test*

- Inducing perceptual distortion and probing compensation... through Motion after-effect (MAE)

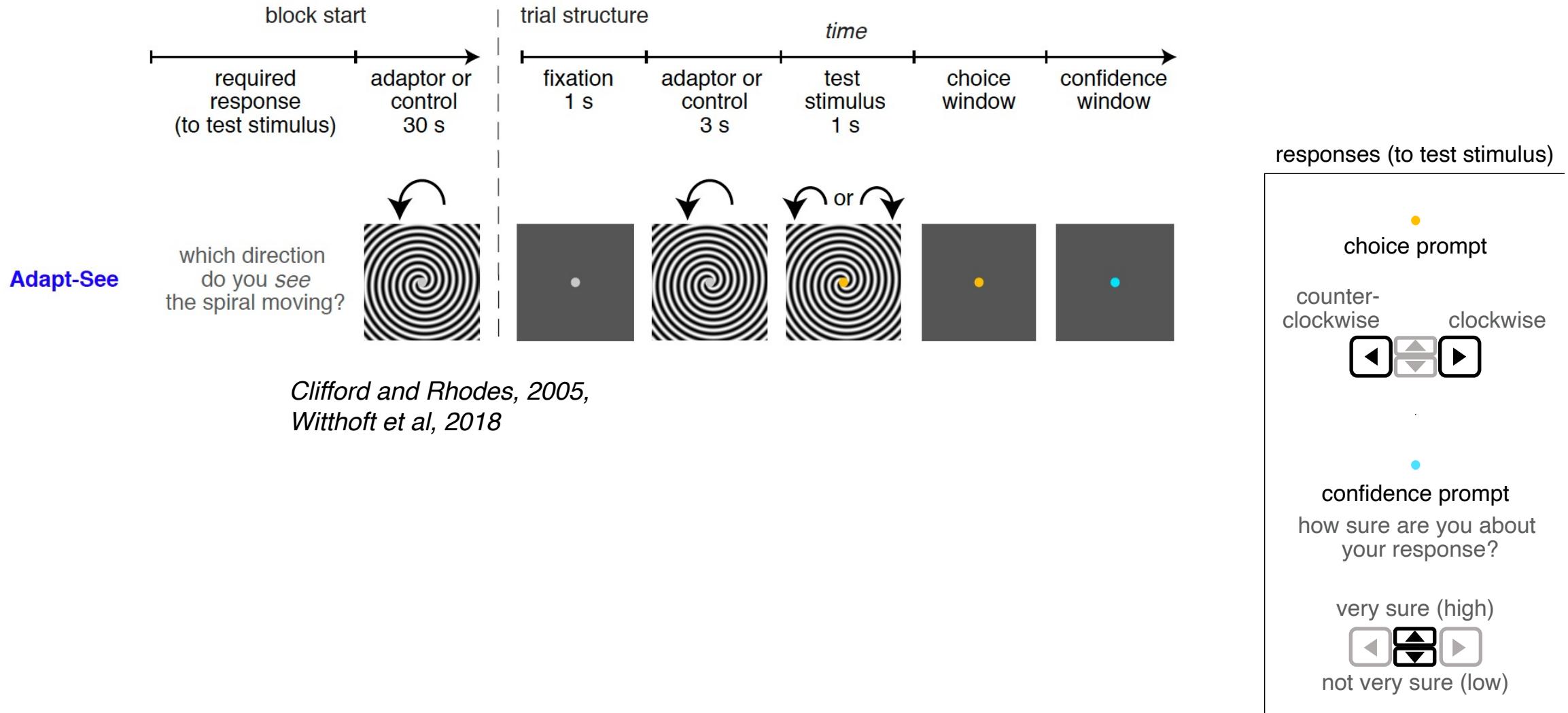
# Approach: Motion after-effect illusion

*Implementation of Archimedean spiral adapted from Peter Scarfe*  
<https://peterscarfe.com/spiraltexturedemo.html>

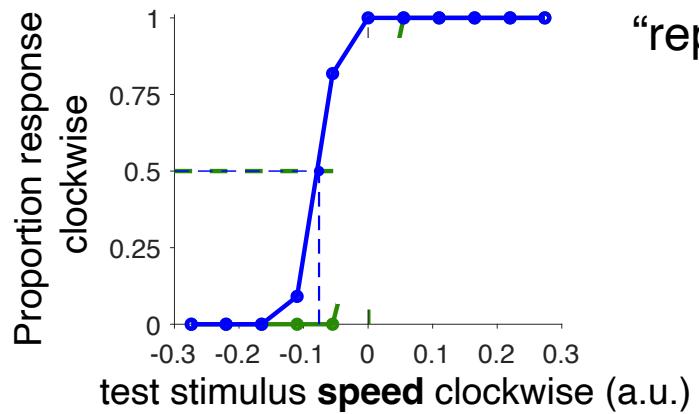
Adaptors with complex motion yield stronger MAE *Bex et al, 1999*  
Spiral after-effect *Pickersgill and Jeeves, 1958, Herrington and Claridge, 1965, Cavanaugh and Favreau, 1980, Kaunitz et al, 2011*



# Task



# How to measure the MAE strength / perceptual distortion?

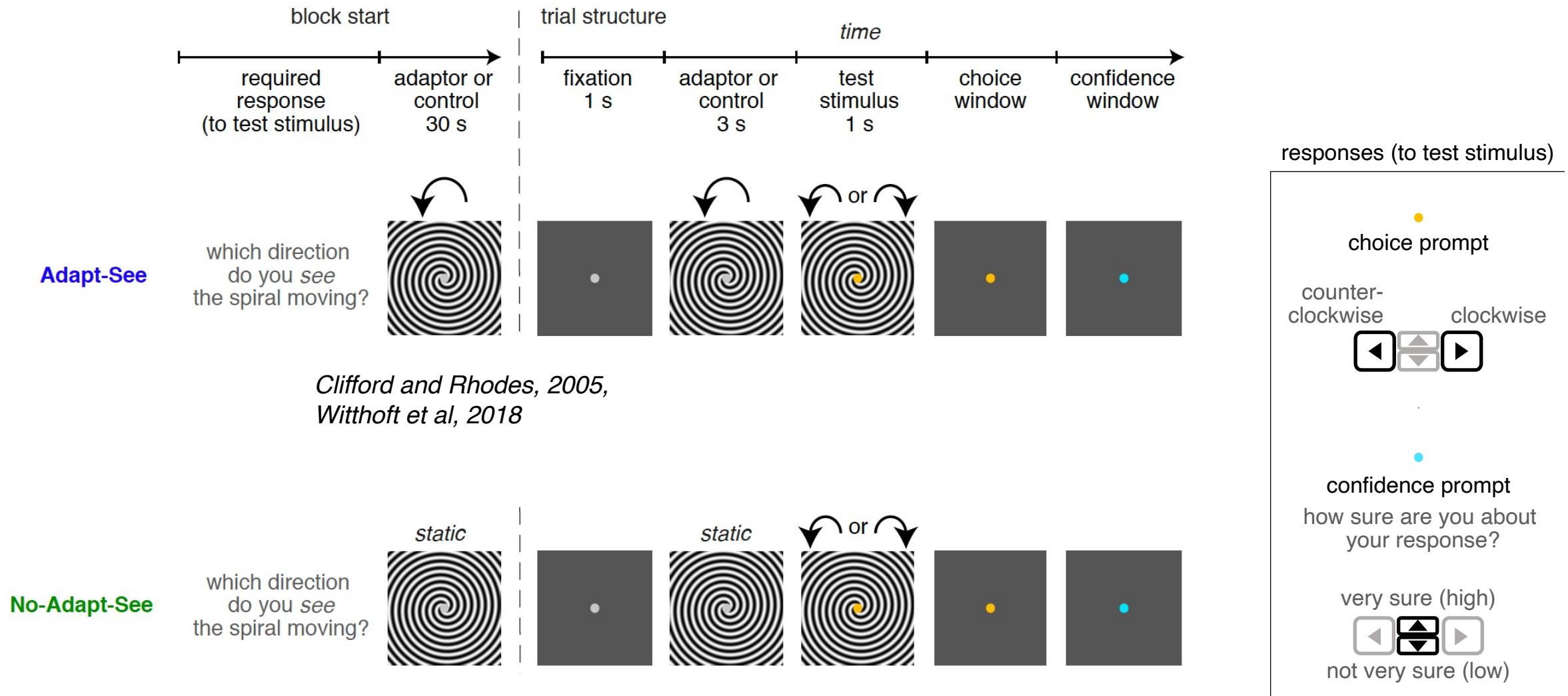


“repulsion effect” of adaptation  
*Levinson and Sekuler, 1976*

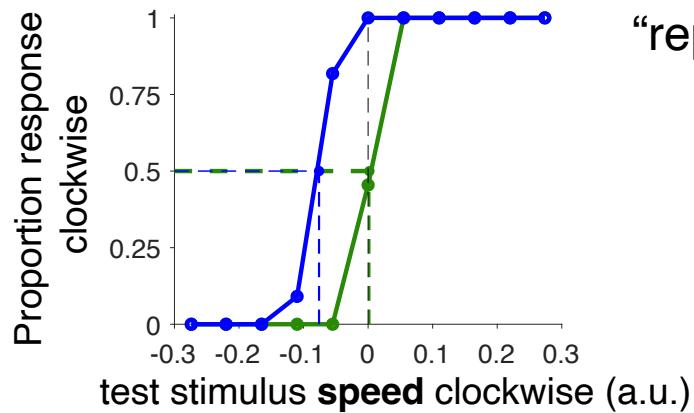
## Implicit nulling method

*Blake and Hiris, 1993*  
*Castet et al, 2002*  
*Thakkar et al, 2018, 2019*

# Task



# How to measure the MAE strength / perceptual distortion?



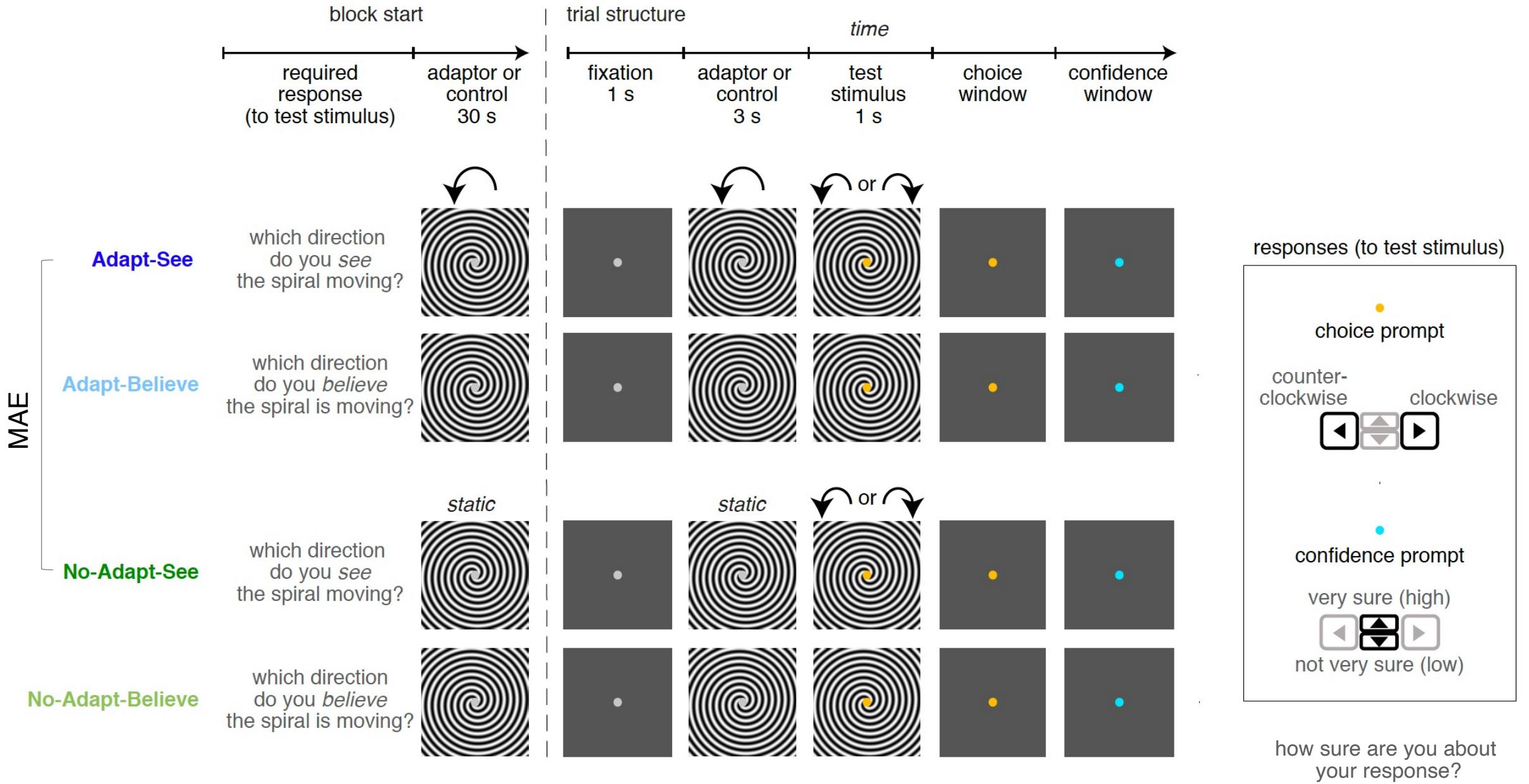
“repulsion effect” of adaptation  
*Levinson and Sekuler, 1976*

## Implicit nulling method

*Blake and Hirsh, 1993*  
*Castet et al, 2002*  
*Thakkar et al, 2018, 2019*

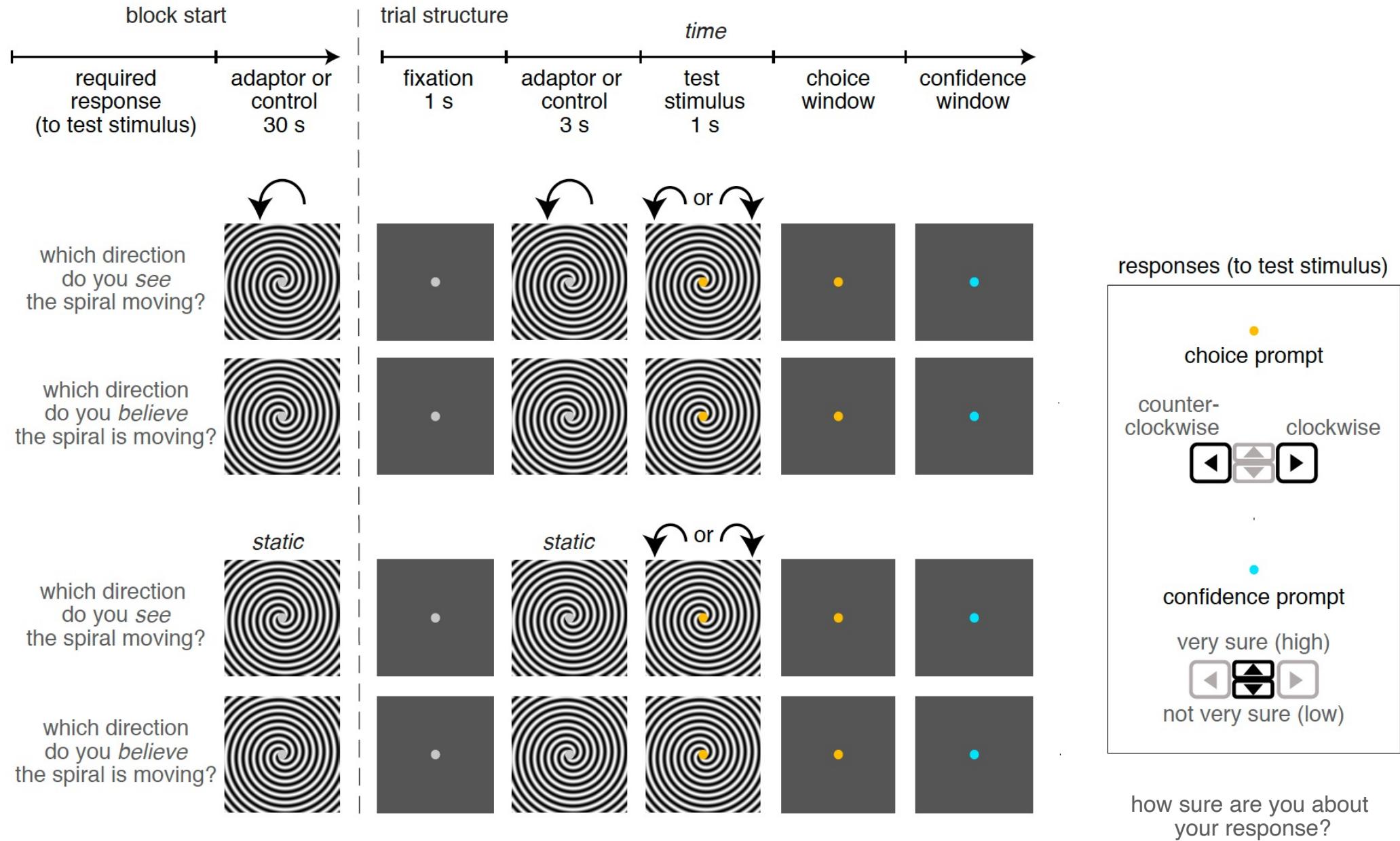
$$\text{MAE} = \mu_{\text{Adapt-See}} - \mu_{\text{No-Adapt-See}}$$

# Task design



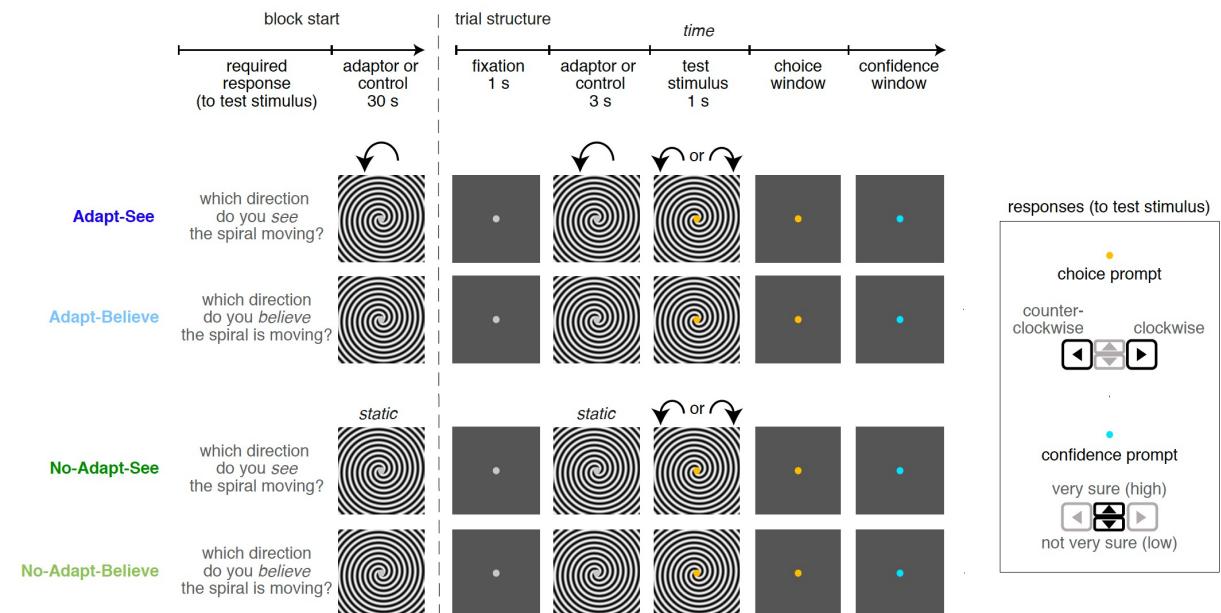
# Task design

MAE compensation

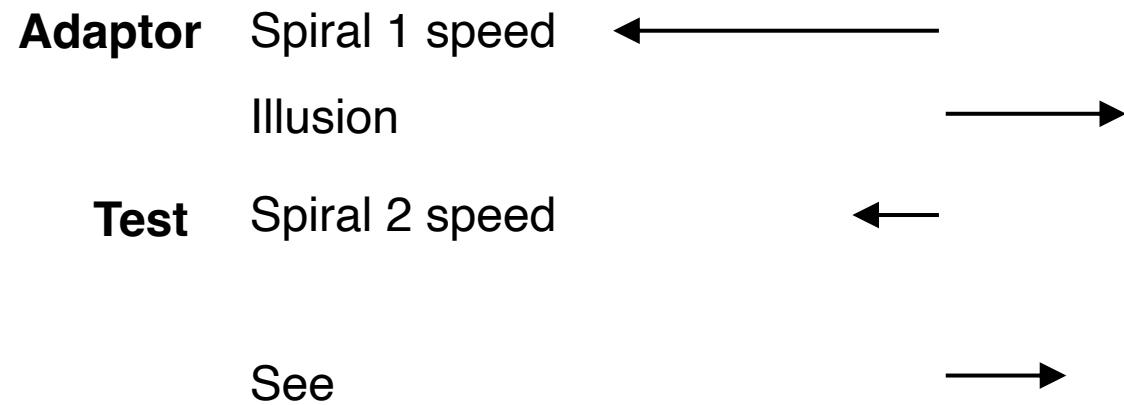


# Task outline

- Extensive training on the illusion and task components + quizzes
- 2 blocks of Adapt-See
- 2 blocks of Adapt-Believe
- 2 blocks of No-Adapt-See
- 2 blocks of No-Adapt-Believe



# Adapt-Believe condition training



# Adapt-Believe condition training

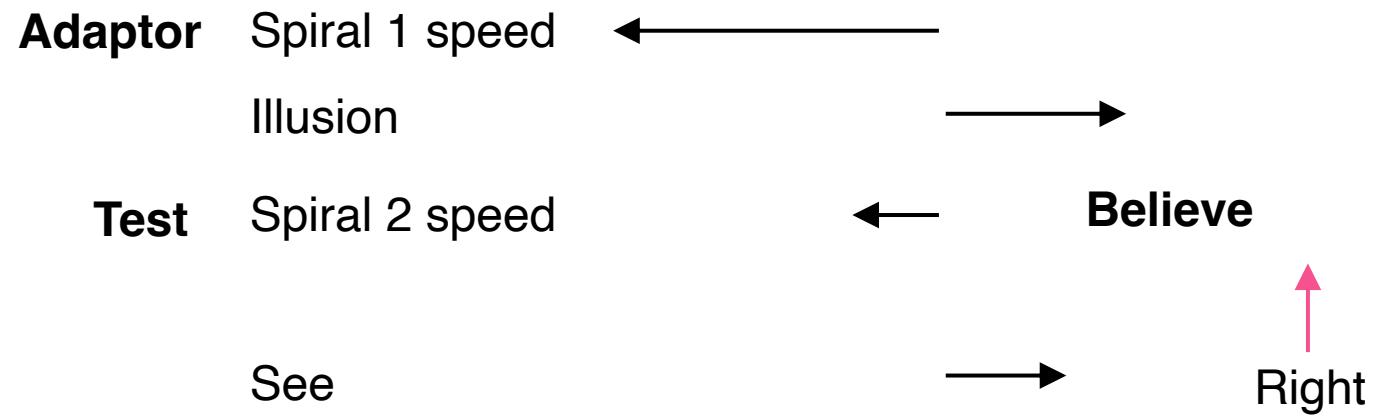
**Adaptor** Spiral 1 speed ←

Illusion →

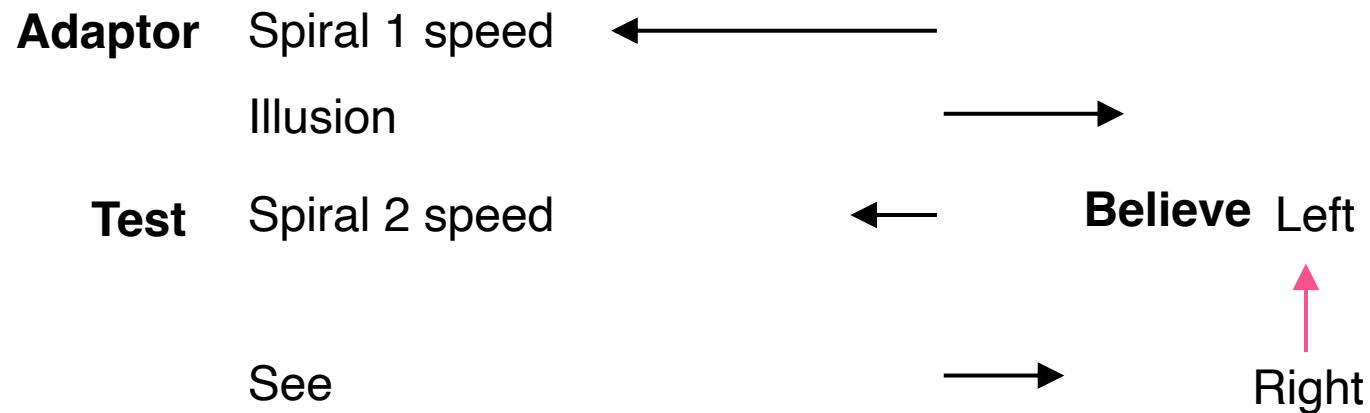
**Test** Spiral 2 speed ←

See → Right

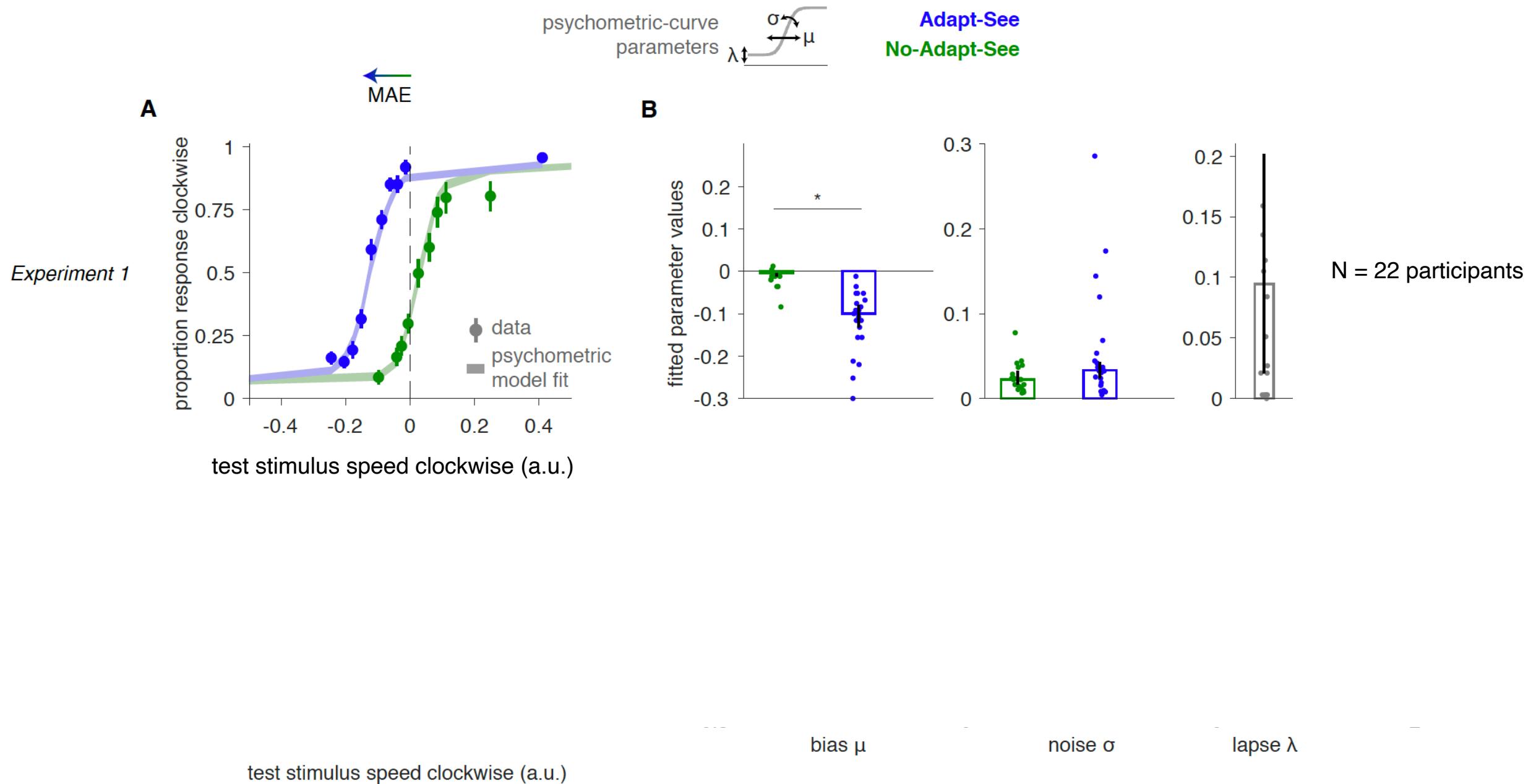
# Adapt-Believe condition training



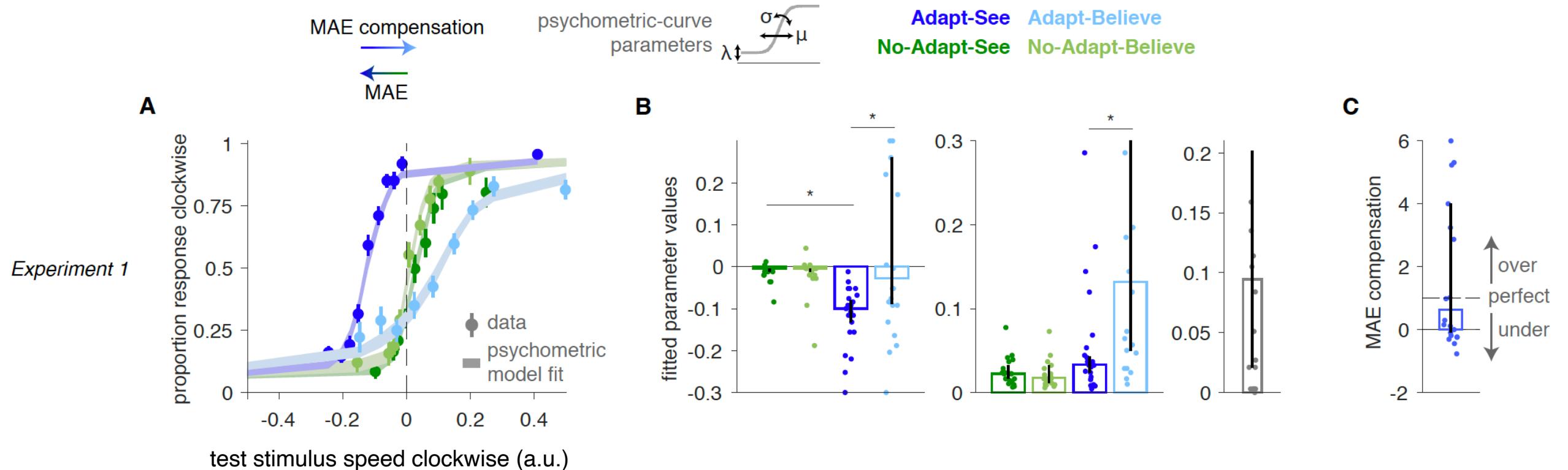
# Adapt-Believe condition training



# Results Adapt-See: MAE can induce a perceptual distortion (offset)



# Participants can compensate for perceptual distortions



$$\text{MAE compensation} = \frac{\mu_{\text{Adapt-Believe}} - \mu_{\text{Adapt-See}}}{|\mu_{\text{Adapt-See}}|}$$

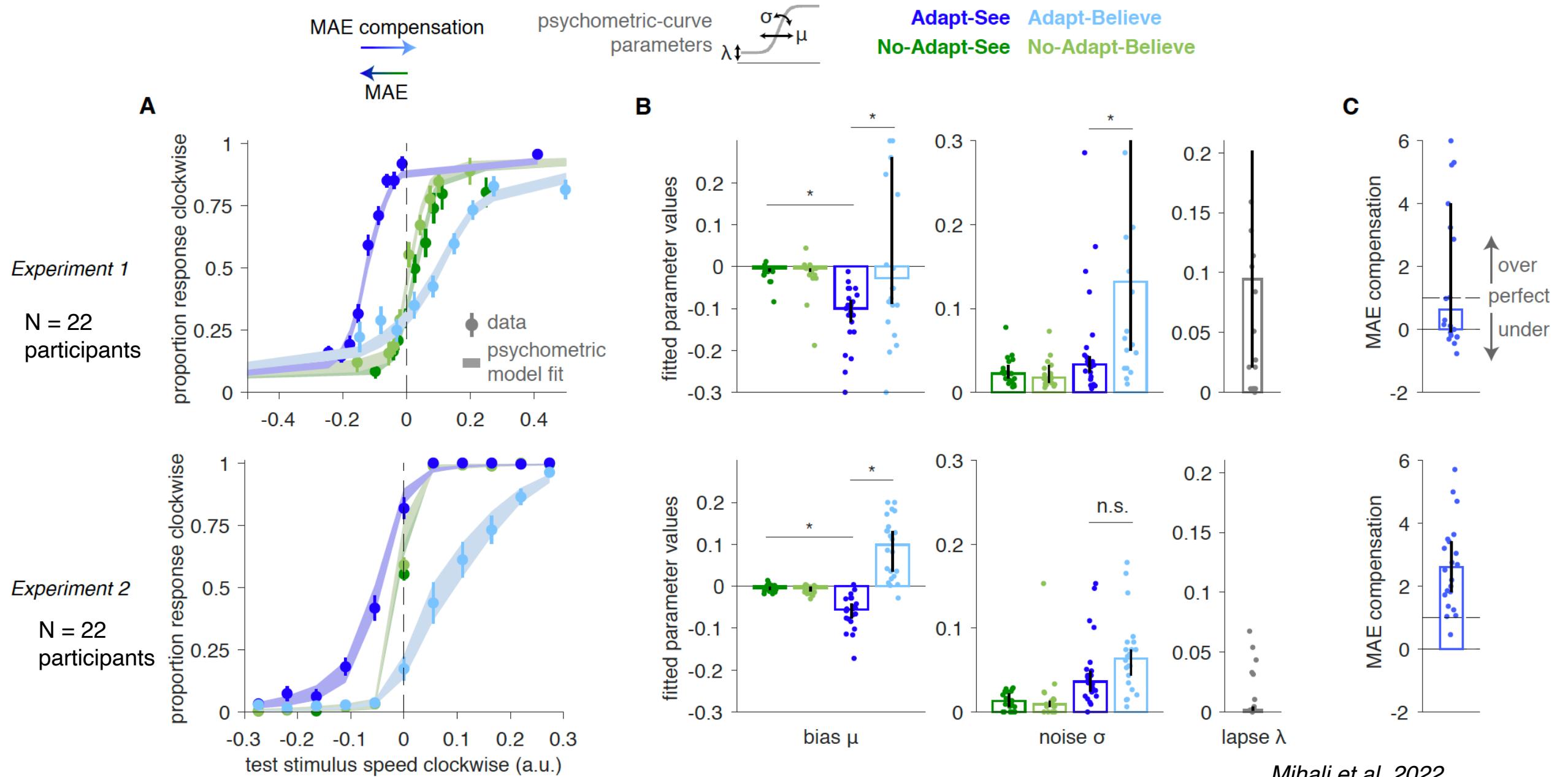
test stimulus speed clockwise (a.u.)

bias  $\mu$

noise  $\sigma$

lapse  $\lambda$

# Participants can (over)compensate for perceptual distortions



Healthy controls can compensate for the MAE, but is this *true insight*?

Is this compensation at an intermediate perceptual inference level?

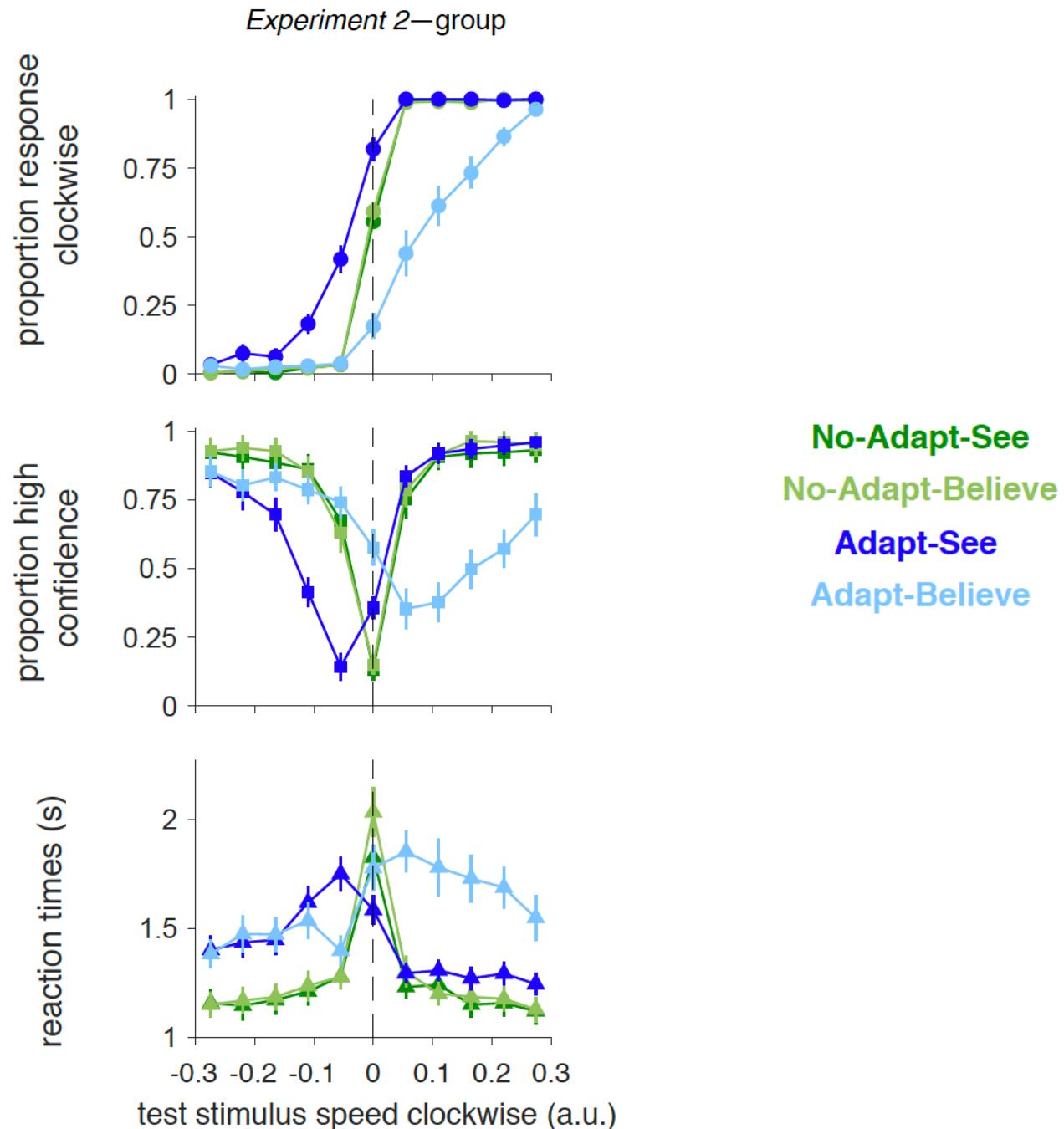
Healthy controls can compensate for the MAE, but is this *true insight*?

Is this compensation at an intermediate perceptual inference  
(and not a late response) level?

To gain clarity, we also looked at  
participants' confidence reports.

Maldonado Moscoso et al., *Proc R Soc B* 2020;  
Gallagher et al., *Sci Rep* 2019

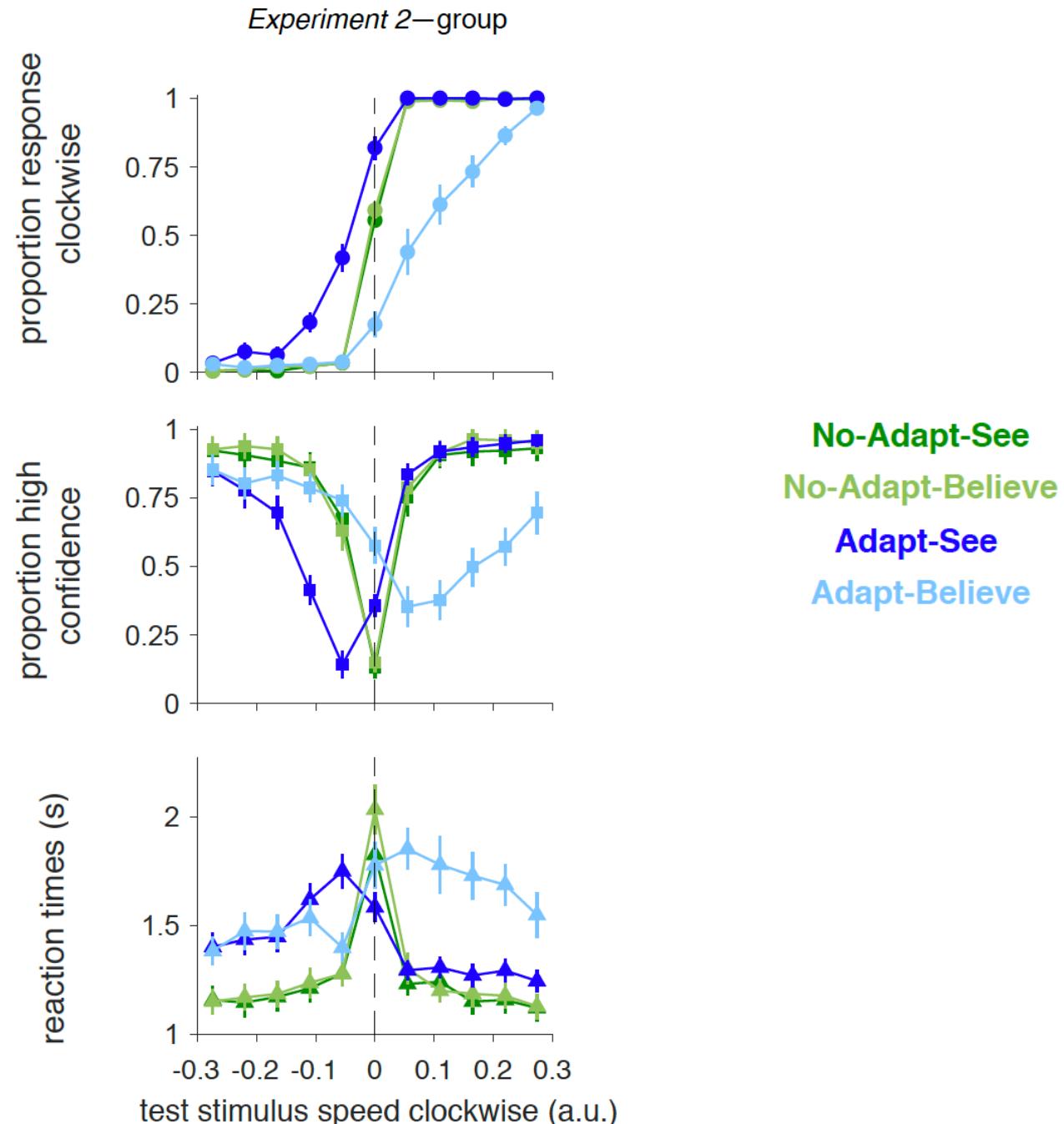
Psychometric curves  
for MAE and MAE  
**compensation** shift in  
tandem with  
confidence and RT  
curves



Psychometric curves  
for MAE and MAE  
**compensation** shift in  
tandem with  
confidence and RT  
curves

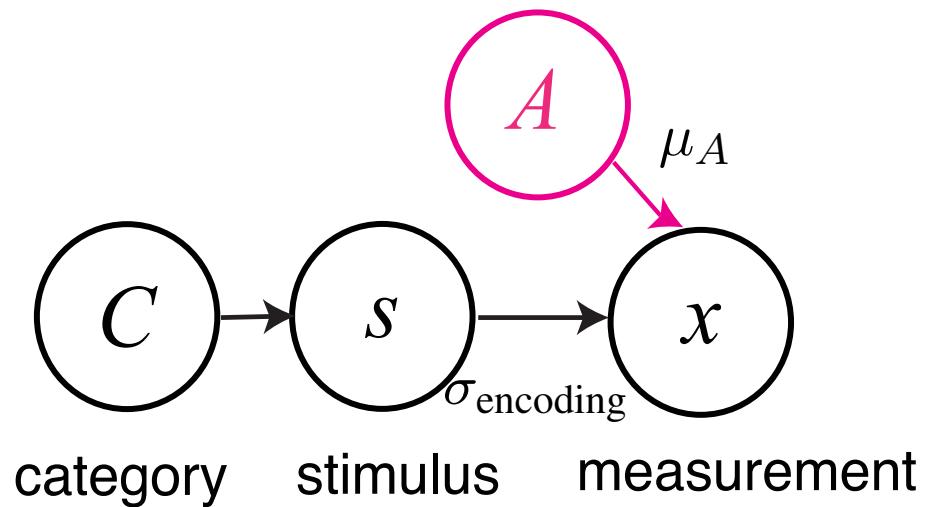
Bayesian process  
models of confidence

*Green and Swets, 1966*  
*Knill and Pouget, 2004,*  
*Kording and Wolpert, 2006,*  
*Adler and Ma, 2018*  
*Hsin and Ma, 2019*



# Bayesian model of perceptual insight

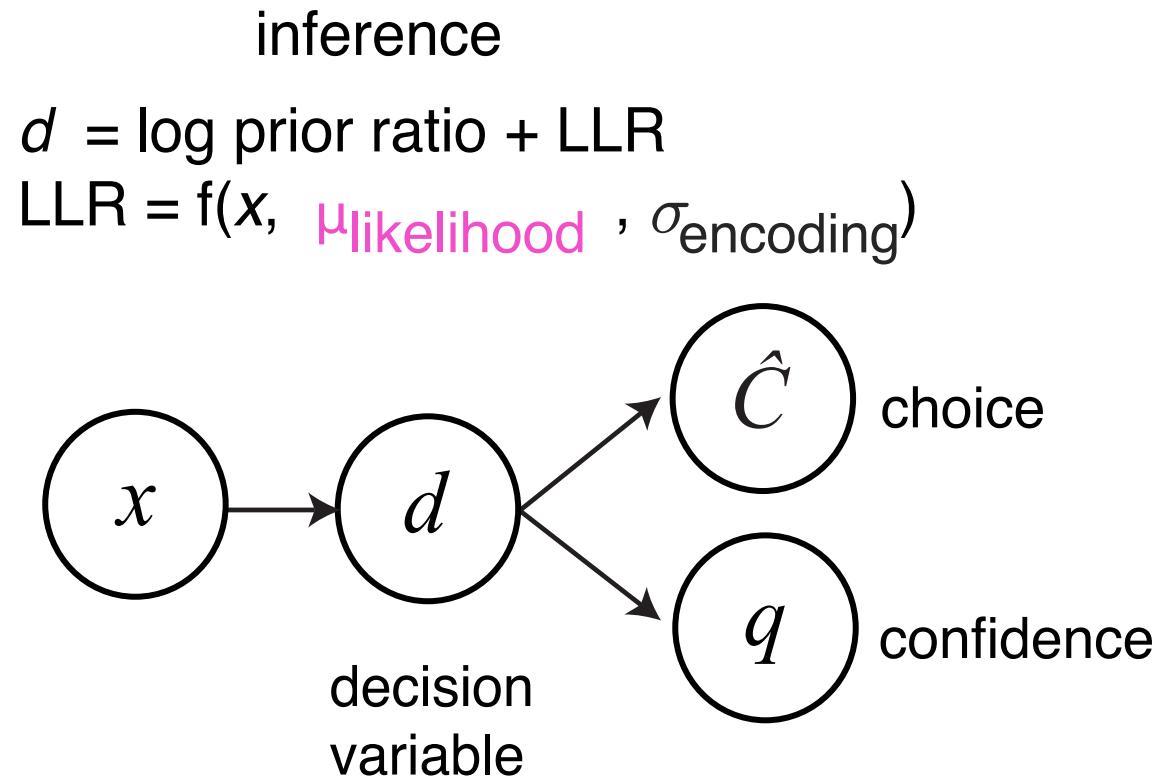
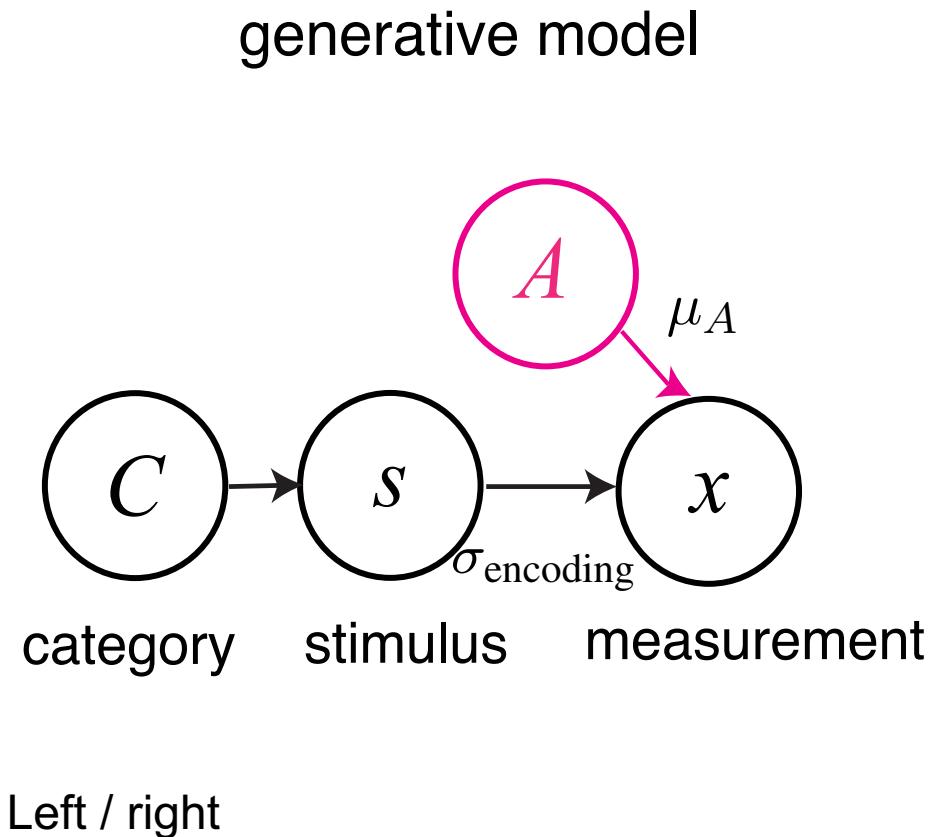
generative model



Left / right

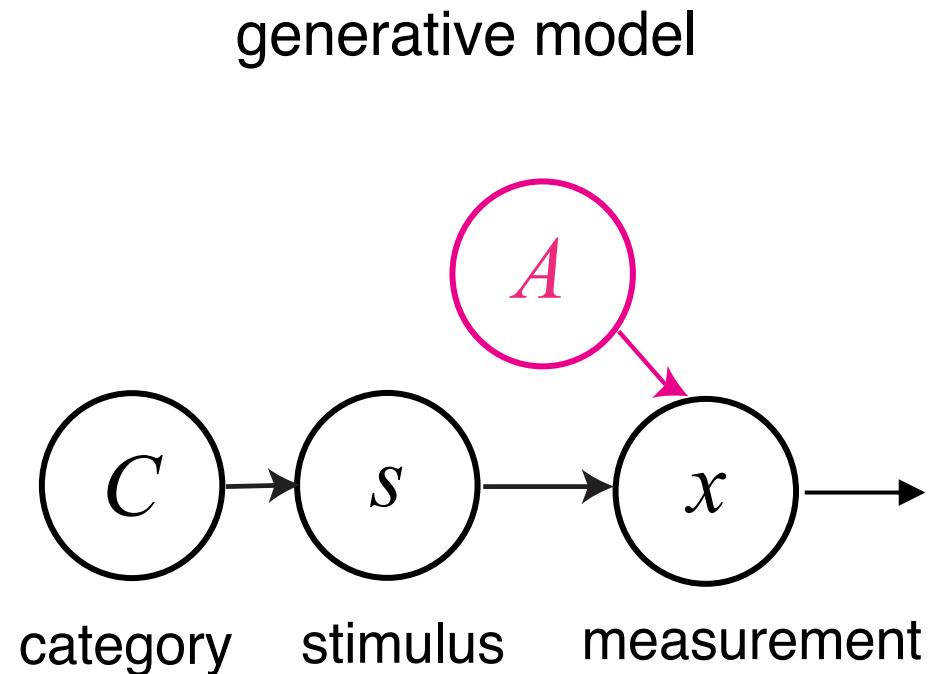
*Green and Swets, 1966  
Knill and Pouget, 2004,  
Kording and Wolpert, 2006,  
Adler and Ma, 2018  
Hsin and Ma, 2019  
Mihali et al, 2022*

# Bayesian model of perceptual insight

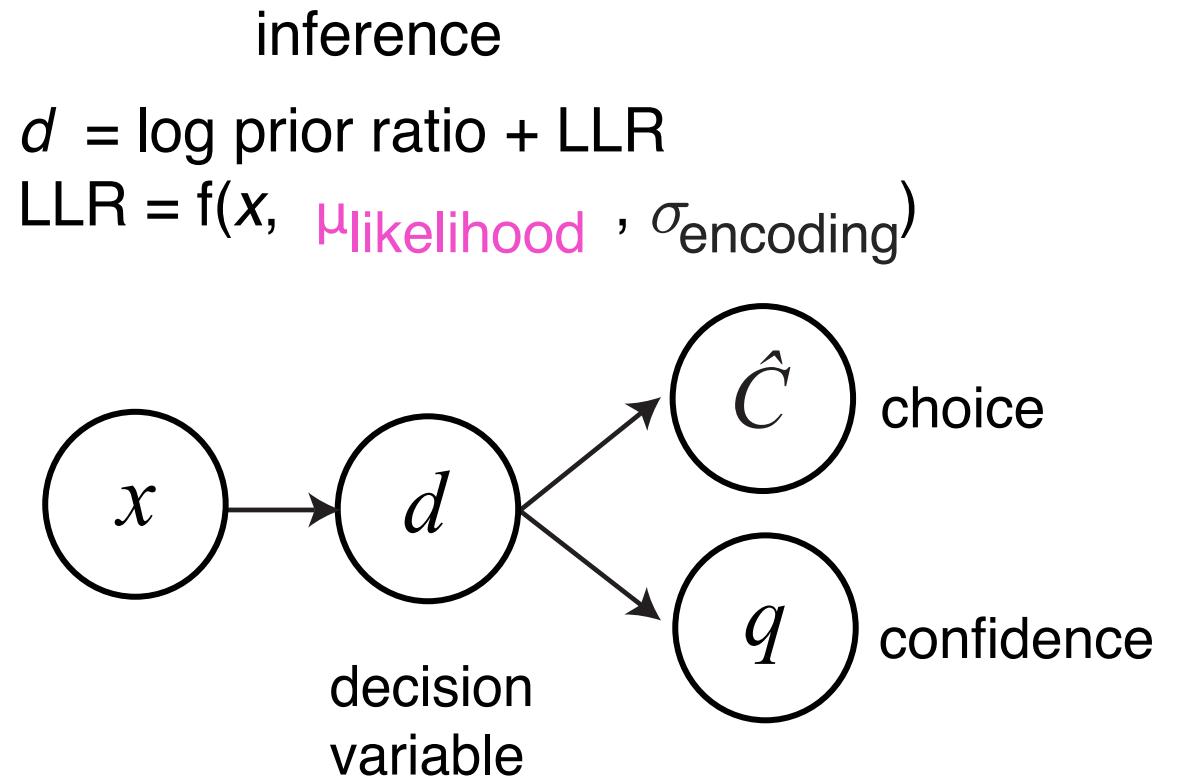


*Green and Swets, 1966  
Knill and Pouget, 2004,  
Kording and Wolpert, 2006,  
Adler and Ma, 2018  
Hsin and Ma, 2019  
Mihali et al, 2022*

# Bayesian model of perceptual insight



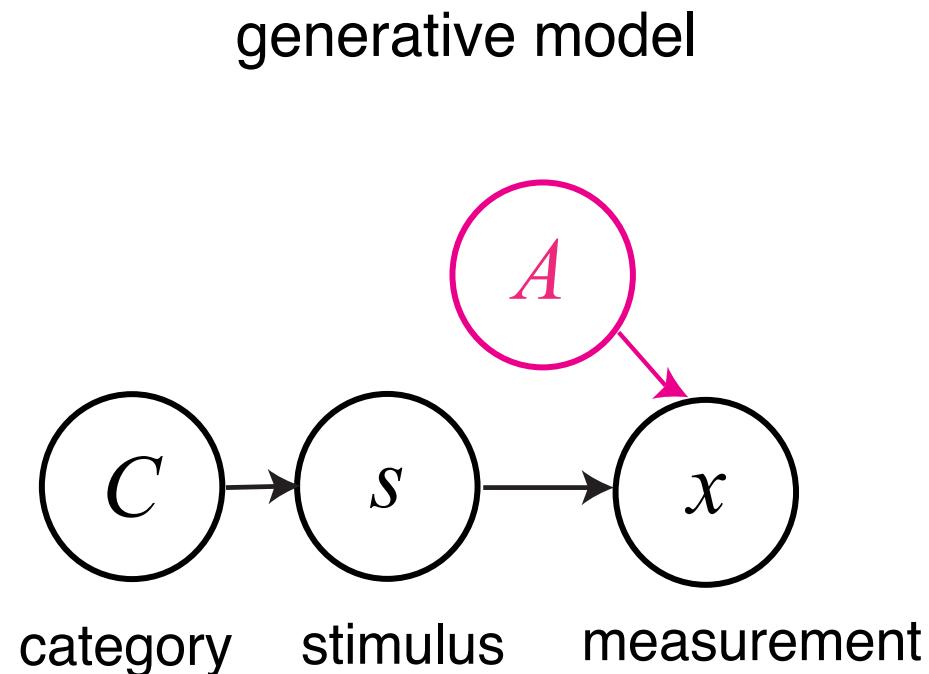
$$p(C = 1|x) = \frac{1}{1 + e^{-d}}$$
$$\hat{C} = \operatorname{argmax}_C p(C|x)$$



$$\text{confidence} = p(C = \hat{C}|x) = \frac{1}{1 + e^{|d|}}$$
$$q = \begin{cases} 1, & \text{if confidence} > k_{\text{confidence}} \\ 0, & \text{otherwise} \end{cases}$$

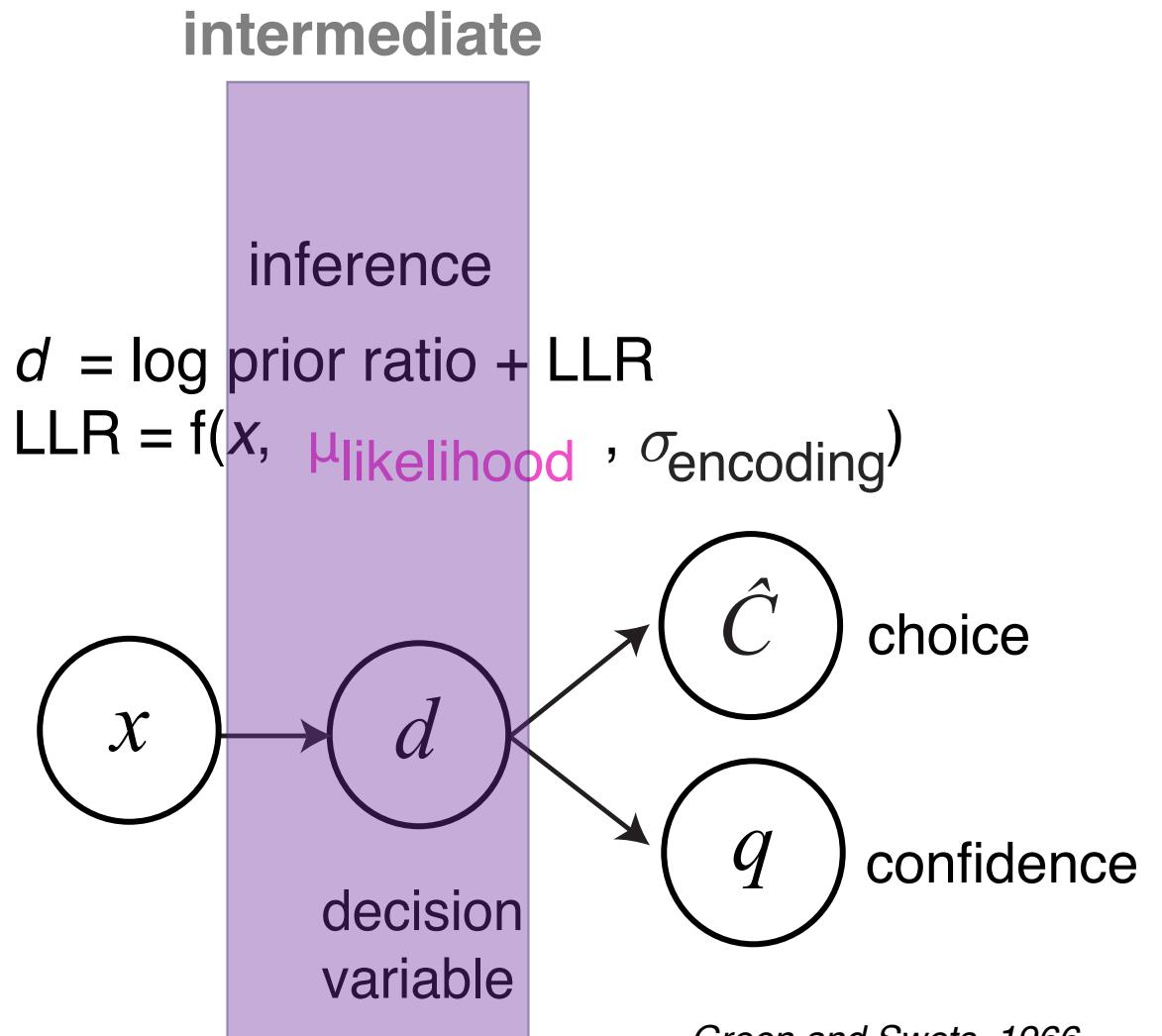
*Green and Swets, 1966  
Knill and Pouget, 2004,  
Kording and Wolpert, 2006,  
Adler and Ma, 2018  
Hsin and Ma, 2019  
Mihali et al, 2022*

# Bayesian model of perceptual insight



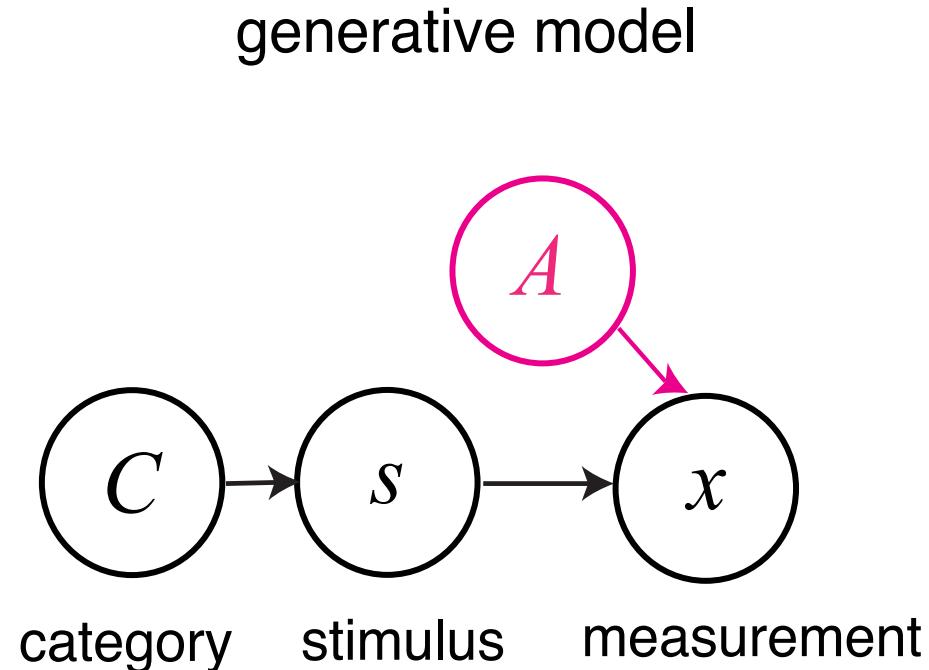
$$p(C = 1|x) = \frac{1}{1 + e^{-d}}$$
$$\hat{C} = \operatorname{argmax}_C p(C|x)$$

$$\text{confidence} = p(C = \hat{C}|x) = \frac{1}{1 + e^{|d|}}$$
$$q = \begin{cases} 1, & \text{if confidence} > k_{\text{confidence}} \\ 0, & \text{otherwise} \end{cases}$$



Green and Swets, 1966  
Knill and Pouget, 2004,  
Kording and Wolpert, 2006,  
Adler and Ma, 2018  
Hsin and Ma, 2019  
Mihali et al, 2022

# Bayesian model of perceptual insight

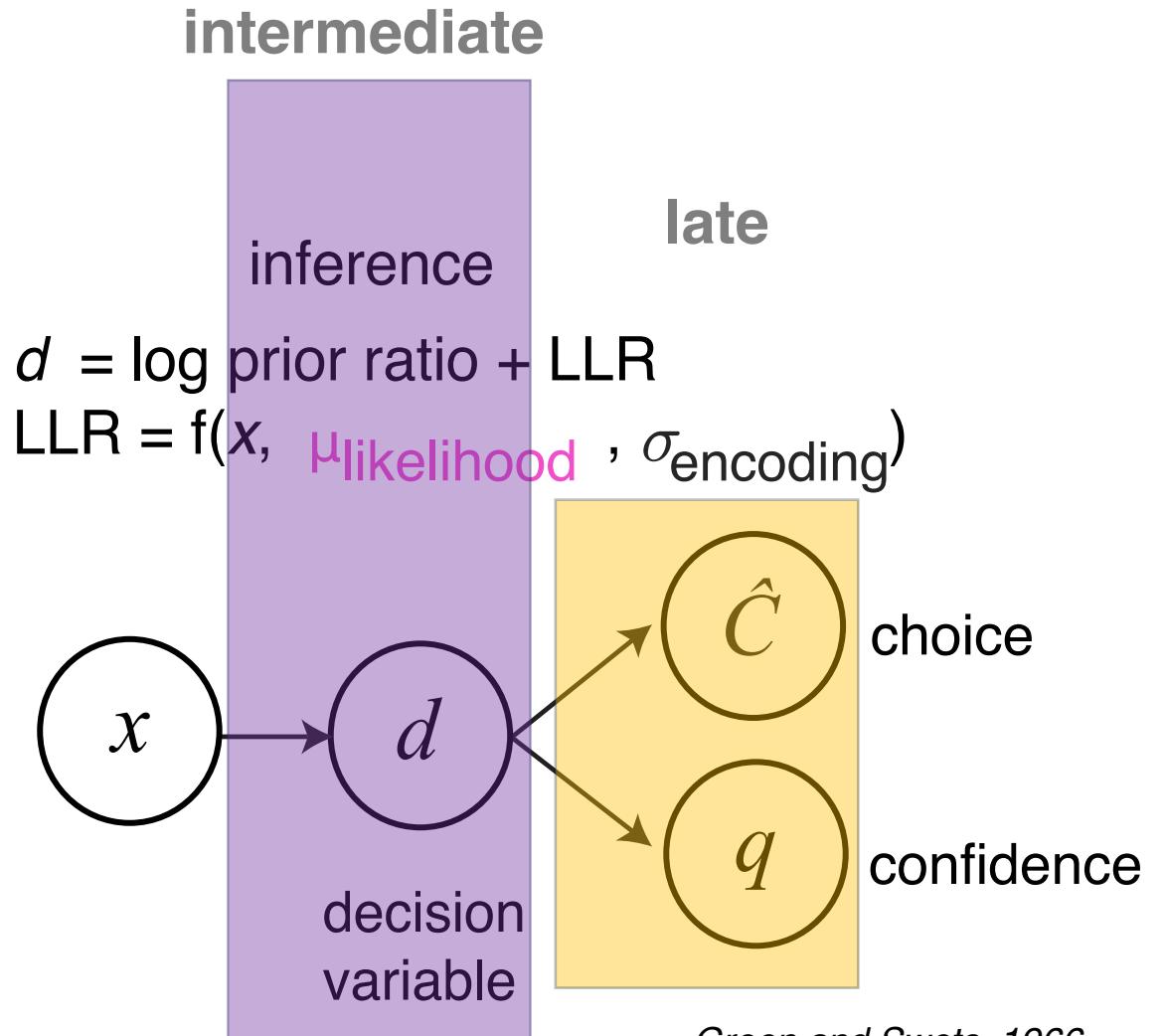


$$p(C = 1|x) = \frac{1}{1 + e^{-d}}$$

$$\hat{C} = \operatorname{argmax}_C p(C|x)$$

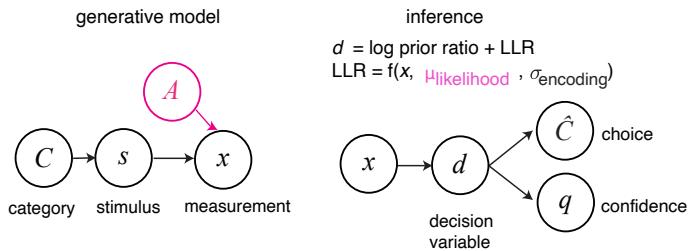
$$\text{confidence} = p(C = \hat{C}|x) = \frac{1}{1 + e^{|d|}}$$

$$q = \begin{cases} 1, & \text{if confidence} > k_{\text{confidence}} \\ 0, & \text{otherwise} \end{cases}$$

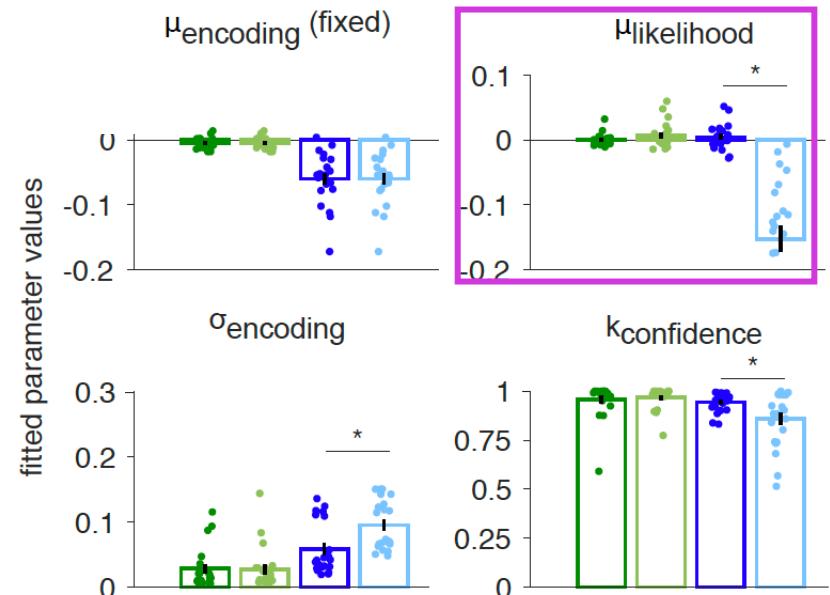
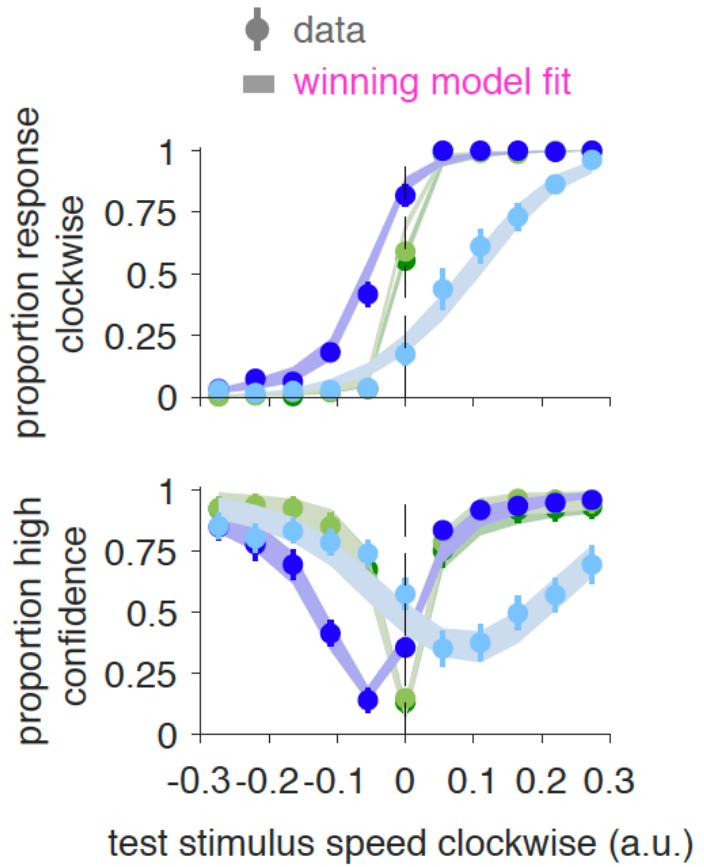


Green and Swets, 1966  
 Knill and Pouget, 2004,  
 Kording and Wolpert, 2006,  
 Adler and Ma, 2018  
 Hsin and Ma, 2019  
 Mihali et al, 2022

# Bayesian model of perceptual insight: model fits & model parameters



No-Adapt-See  
No-Adapt-Believe  
Adapt-See  
Adapt-Believe



# Conclusions

- People can experience perceptual distortions with the MAE illusion
- People with knowledge of the MAE illusion can (over)compensate for it
- A Bayesian inference model of perceptual insight with adjustments in the intermediate inference stage captured the data best and revealed differences in Adapt-See vs Adapt-Believe in a key parameter for perceptual insight,  $\mu_{\text{likelihood}}$
- Healthy participants demonstrated perceptual insight, specifically the ability to **incorporate introspective knowledge about distortions** in internal percepts to effectively infer the actual state of the external world

# Insight

the ability to incorporate introspective knowledge about distortions in internal percepts to effectively infer the actual state of the external world

BJPsych

The British Journal of Psychiatry (2020)  
217, 521–523. doi: 10.1192/bjp.2019.217

Analysis

Insight and psychosis: the next 30 years

Anthony S. David

JAMA Network Insights

Working With Patients With Impaired Insight

Hannah E. Brown, MD; Sydney Feinstein, BA; David C. Henderson, MD

**no objective measures or process models**

# Visual metacognition

the ability for one's subjective sense of confidence to track the objective accuracy

**several objective measures and process models**



© 2021 American Psychological Association  
ISSN: 0003-066X

American Psychologist

2021, Vol. 76, No. 9, 1445–1453  
<https://doi.org/10.1037/amp0000937>

Visual Metacognition: Measures, Models, and Neural Correlates

Dobromir Rahnev  
School of Psychology, Georgia Institute of Technology



RESEARCH ARTICLE

Comparing Bayesian and non-Bayesian accounts of human confidence reports

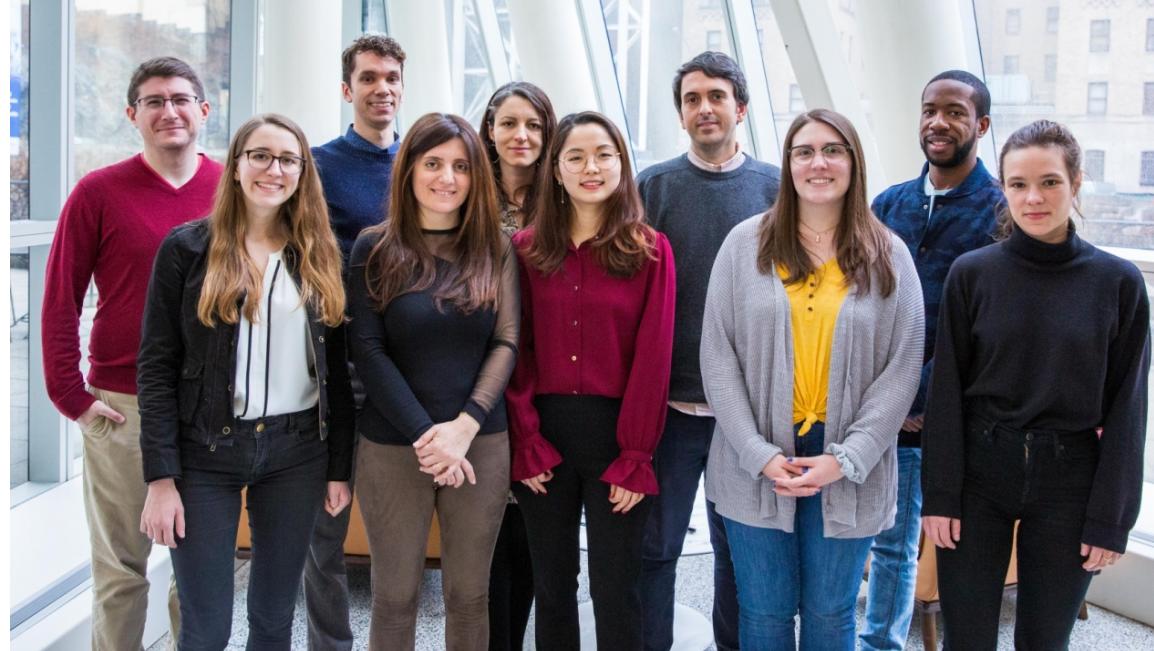
William T. Adler<sup>1\*</sup>, Wei Ji Ma<sup>1,2</sup>

# Acknowledgements

- Horga Lab: **Guillermo Horga**
- **Marianne Broeker**
- **Florian Ragalmuto**
- Alissa Fogelson
- Najate Ojeil
- Anastasia Velikovskaya
- Carmen Mendes de Leon
- Brandon Ashinoff
- Kenneth Wengler
- Justin Buck
- Nicholas Singletary
- Isabella Rosario
- Jocelyn Kim



COLUMBIA UNIVERSITY  
DEPARTMENT OF PSYCHIATRY



National Institute of Mental Health:  
awards R01MH117323 and R01MH114965  
to Guillermo Horga



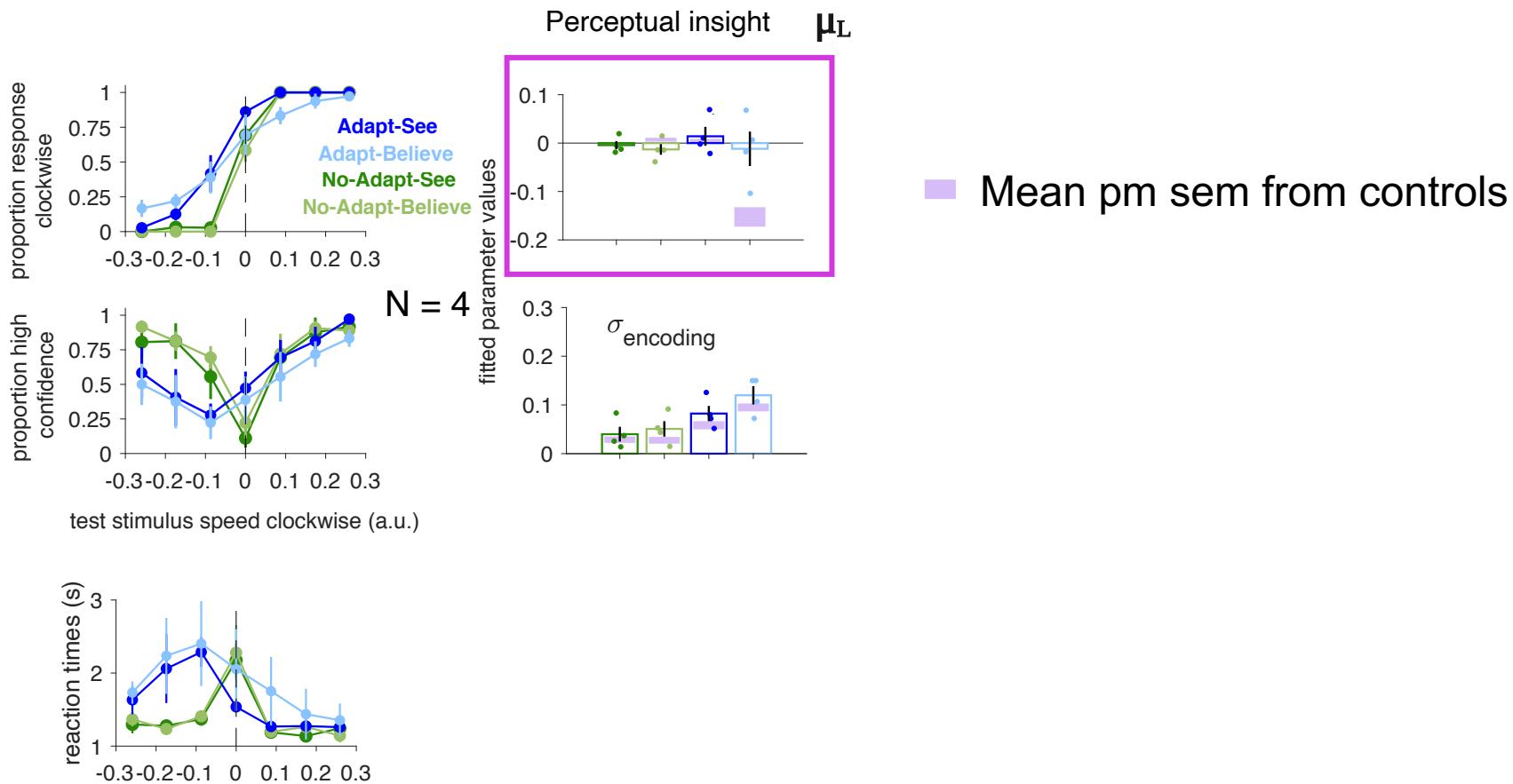
New York State  
Psychiatric Institute

- Nadine Dijkstra
- Kyo ligaya
- Daniel Wolpert
- Prady Sepulveda

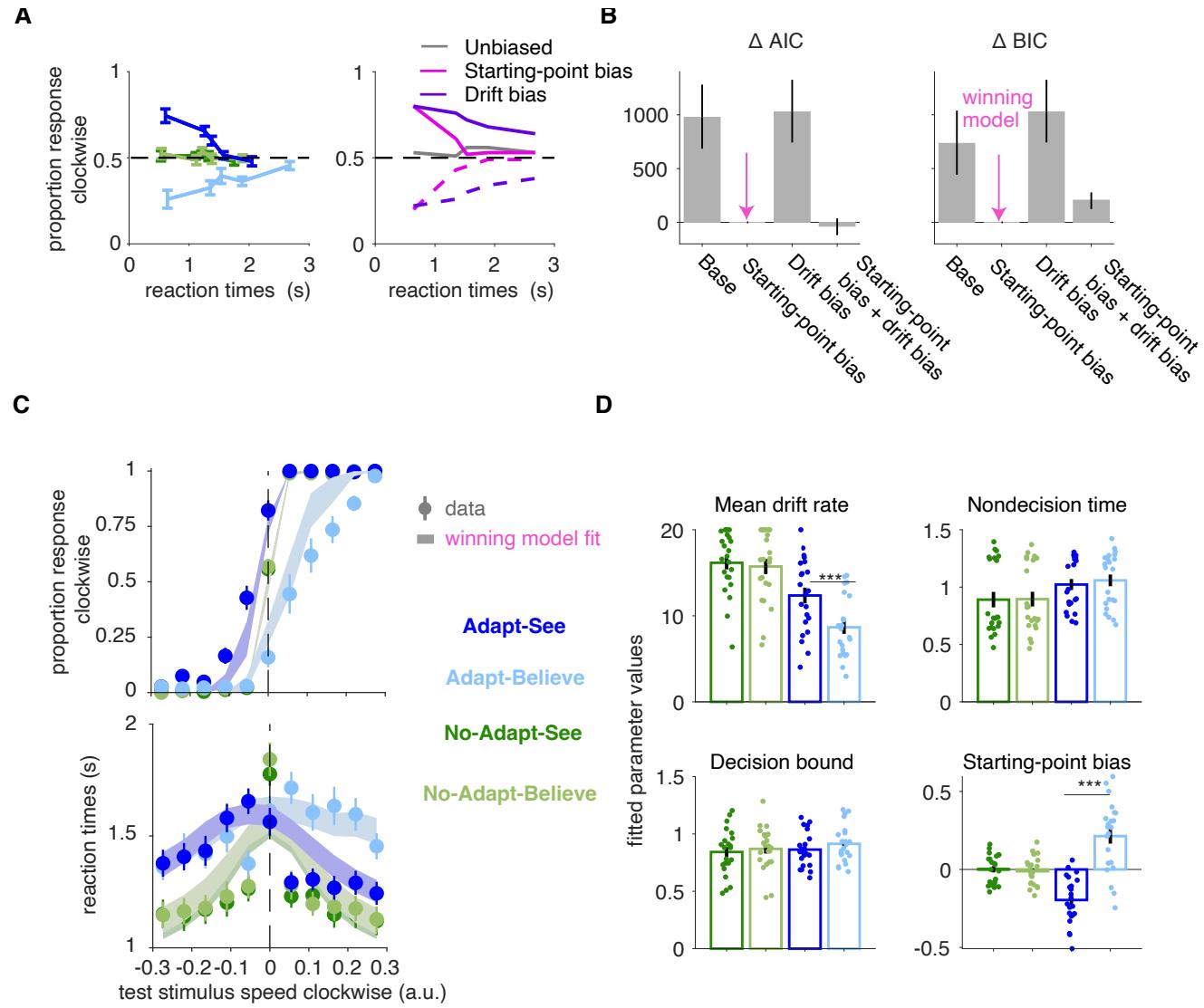
Thank you!

# Is the perceptual insight parameter meaningful? / Can it capture insight impairments in psychosis?

Preliminary data in patients with psychosis: N = 4 Patients may have decreased perceptual insight



# DDM



# Model fitting of the data from Experiment 2

## Strategy and parameters for the perceptual insight model

No Adapt

Adapt

No-Adapt-See

No-Adapt-Believe

1  $\mu_A$   
(fixed, shared)

2  $\sigma$

2  $\mu_L$

2  $k_{\text{confidence}}$

Adapt-See

Adapt-Believe

1  $\mu_A$   
(fixed, shared)

2  $\sigma$

2  $\mu_L$

2  $k_{\text{confidence}}$

14 parameters to fit total

Maximum likelihood estimation (MLE)

# Model fitting of the data from Experiment 2

## Strategy and parameters for the perceptual insight model

No Adapt

Adapt

No-Adapt-See

No-Adapt-Believe

1  $\mu_A$   
(fixed, shared)

2  $\sigma$

2  $\mu_L$

2  $k_{\text{confidence}}$

Adapt-See

Adapt-Believe

1  $\mu_A$   
(fixed, shared)

2  $\sigma$

2  $\mu_L$

2  $k_{\text{confidence}}$

14 parameters to fit total

Maximum likelihood estimation (MLE)

### Perceptual insight model

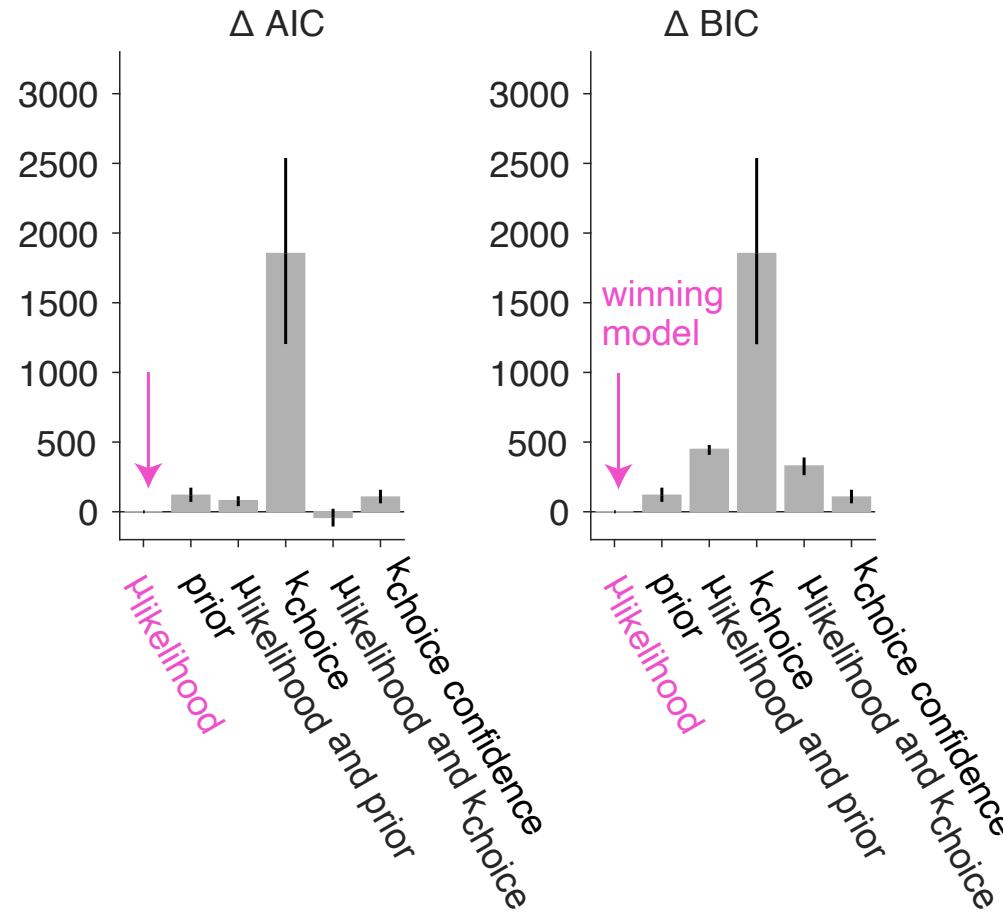
### Response bias model

Stage	Model					
	Perceptual insight " $\mu_{\text{likelihood}}$ "	"prior"	" $\mu_{\text{likelihood}} + \text{prior}$ "	Late compensation " $k_{\text{choice}}$ "	" $\mu_{\text{likelihood}} + k_{\text{choice}}$ "	" $k_{\text{choice}} \text{ confidence}$ "
Early	$\mu_{\text{encoding}} \times 2$ (fixed) $\sigma_{\text{encoding}} \times 4$					
Intermediate	$\mu_{\text{likelihood}} \times 4$	prior $\times 4$	$\mu_{\text{likelihood}} \times 4$ prior $\times 4$		$\mu_{\text{likelihood}} \times 4$	
Late	$k_{\text{confidence}} \times 4$	$k_{\text{confidence}} \times 4$	$k_{\text{confidence}} \times 4$	$k_{\text{confidence}} \times 4$ $k_{\text{choice}} \times 4$	$k_{\text{confidence}} \times 4$ $k_{\text{choice}} \times 4$	$k_{\text{confidence}} \times 4$ $k_{\text{choice}} \times 4$

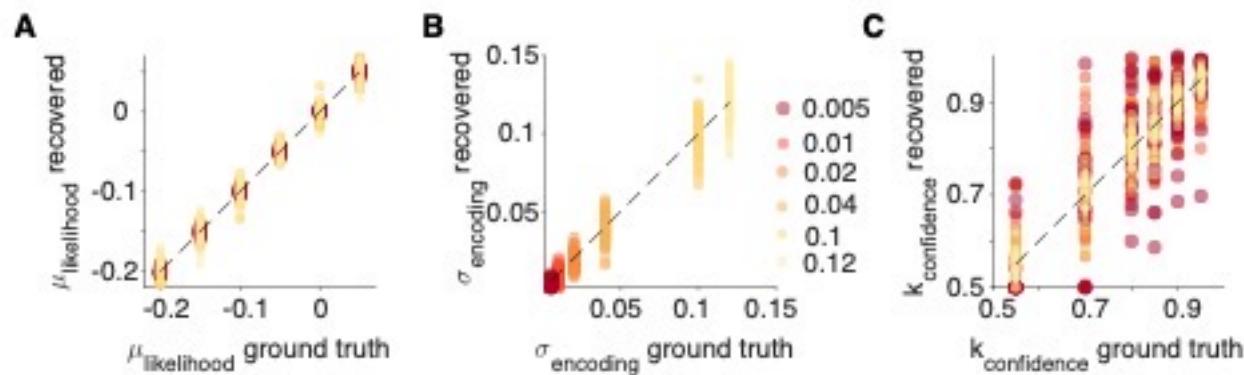
$$\hat{C} = \begin{cases} 1, & \text{if } d > 0 \\ -1, & \text{otherwise} \end{cases}$$

$$\hat{C} = \begin{cases} 1, & \text{if } d > k_{\text{choice}} \\ -1, & \text{otherwise} \end{cases}$$

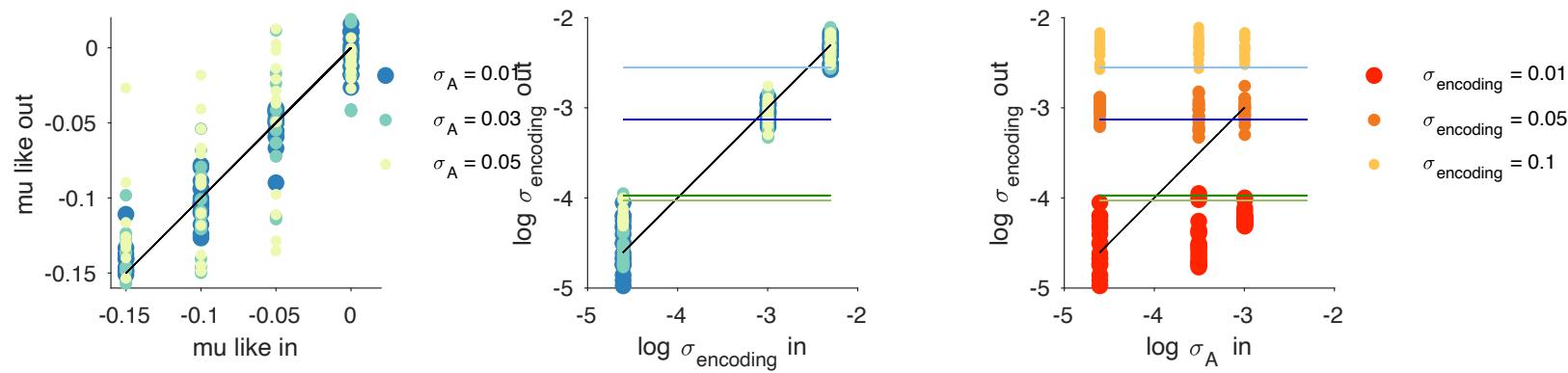
Model comparison:  
The best fitting  
model is the  
perceptual insight  
model



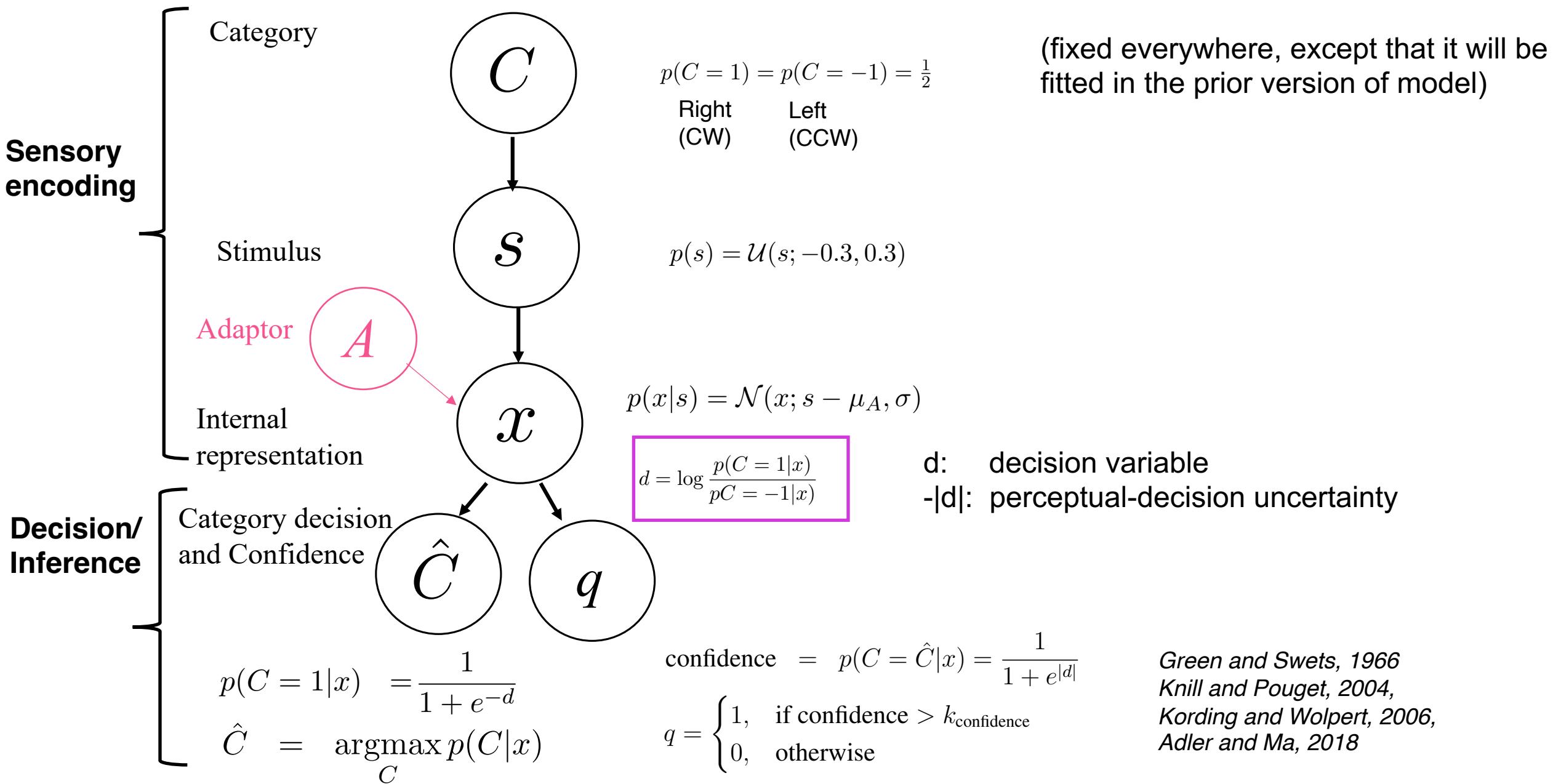
# Parameter recovery



**Figure S7: Parameter recovery in the  $\mu_{\text{likelihood}}$  model, based on simulated datasets with 121 trials each. A)  $\mu_{\text{likelihood}}$ . B)  $\sigma_{\text{encoding}}$ . C)  $k_{\text{confidence}}$ .** The Spearman correlations for parameter recovery were 0.98 for  $\mu_{\text{likelihood}}$ , 0.97 for  $\sigma_{\text{encoding}}$  and 0.85 for  $k_{\text{confidence}}$ .

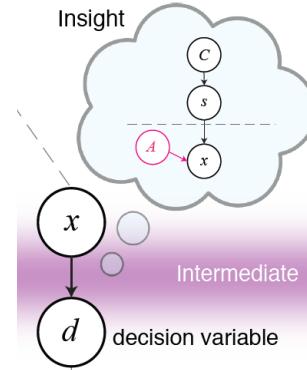


# Bayesian model and inference in this task



# Inference

$$d = \log \frac{p(C = 1|x)}{p(C = -1|x)}$$



Perceptual insight as incorporation of knowledge of the distortion  $-\mu_L$ - into the decision variable

$$d = \log \frac{p(C = 1)}{1 - p(C = 1)} + \log \left( \frac{\Phi(0.3 - \mu_L - x; 0, \sigma) - \Phi(-\mu_L - x; 0, \sigma)}{\Phi(-\mu_L - x; 0, \sigma) - \Phi(-0.3 - \mu_L - x; 0, \sigma)} \right) \text{ Scaled with the sensory noise } \sigma$$

$\Phi$ : Cumulative density function (cdf) of the Gaussian