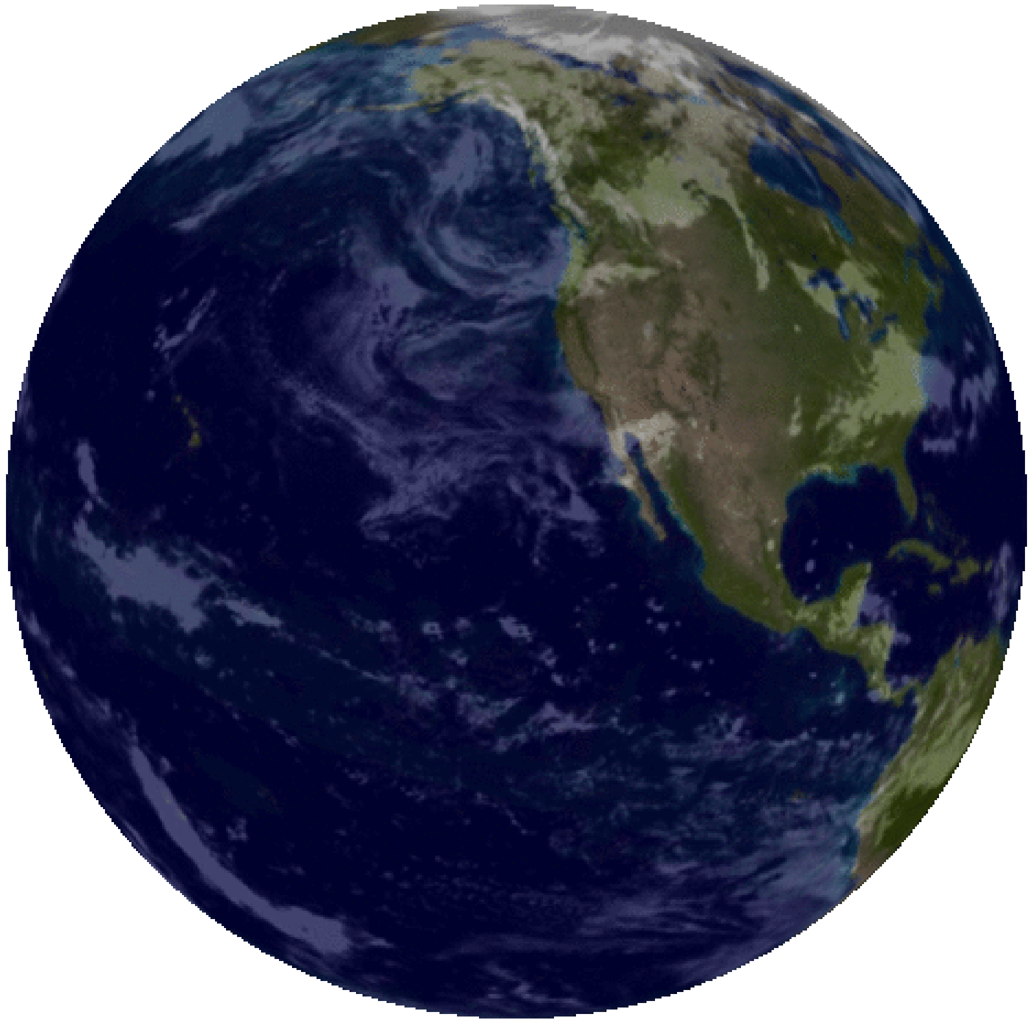
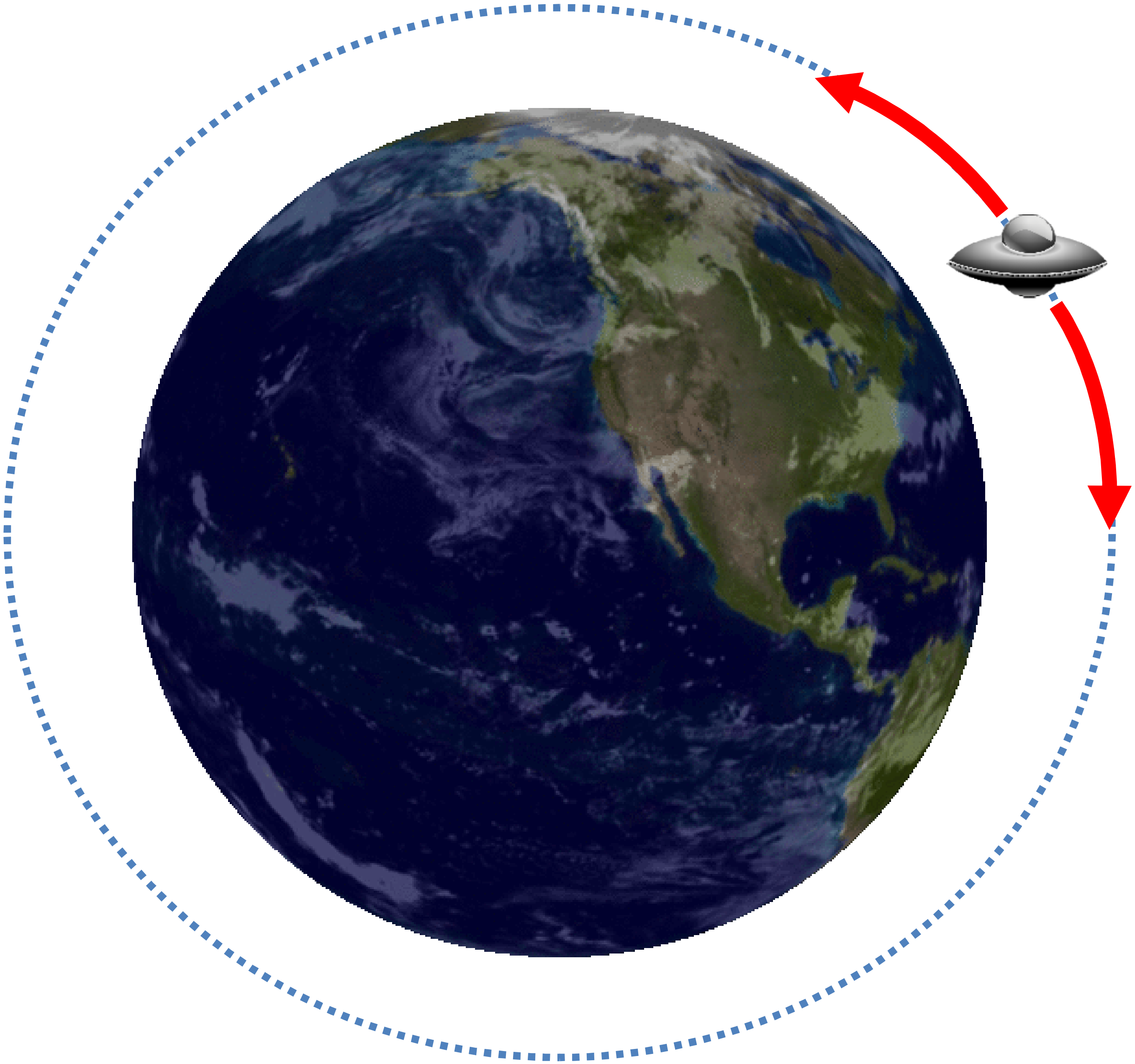
Velocity of change in the environment in the delta-rule model of reinforcement.

Carlos Velázquez, Arturo Bouzas.

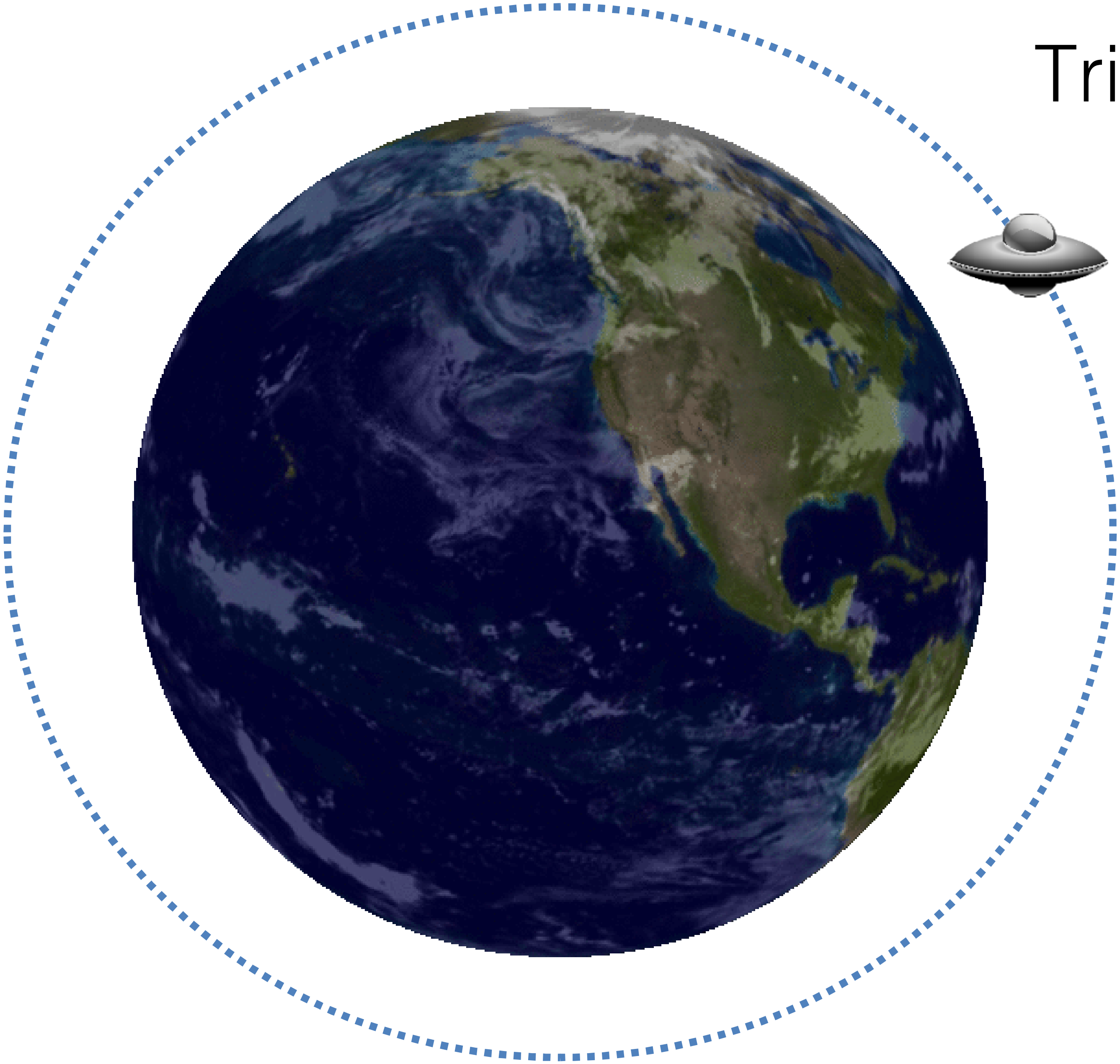
School of Psychology.

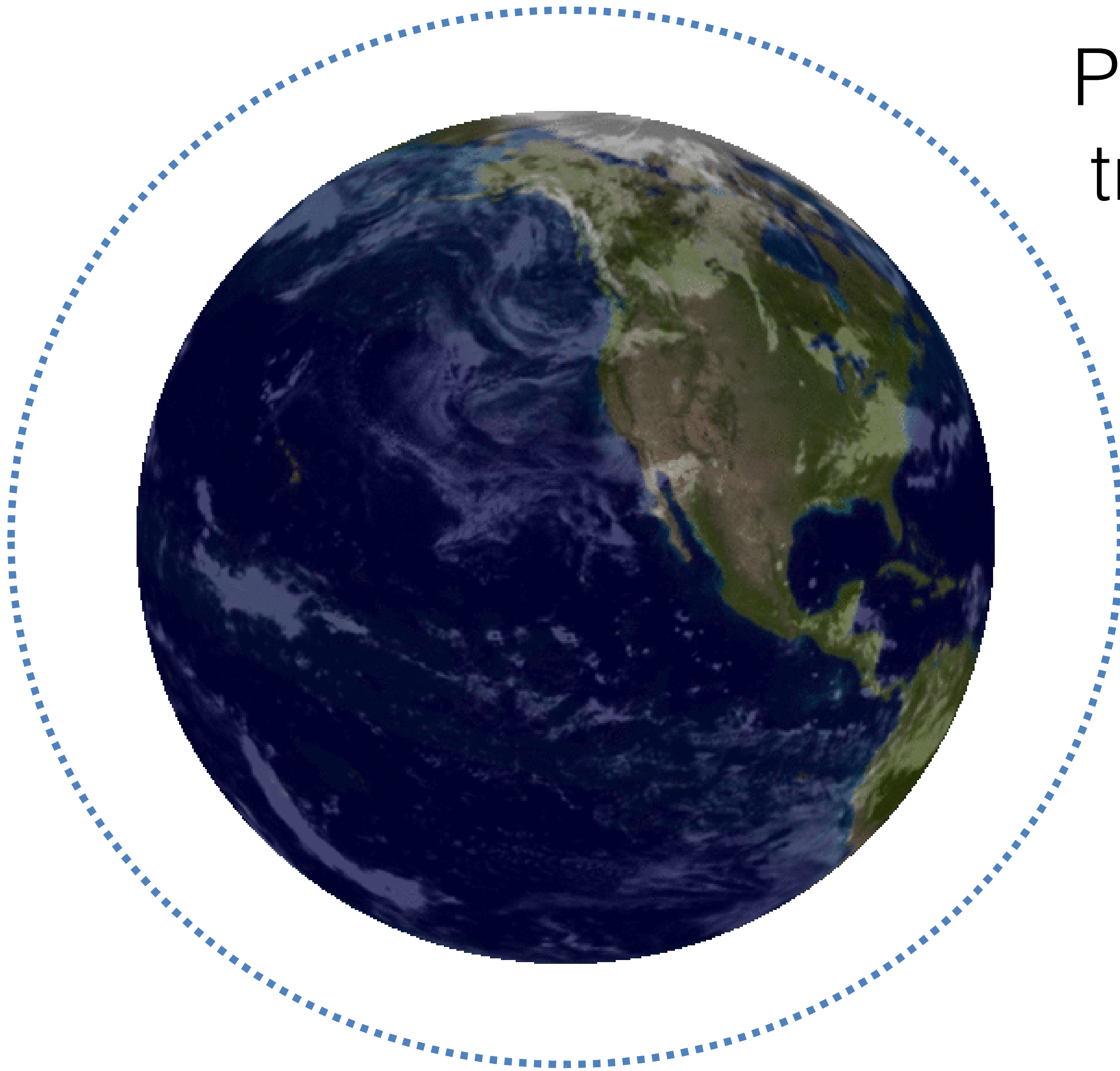
National Autonomous University of Mexico.



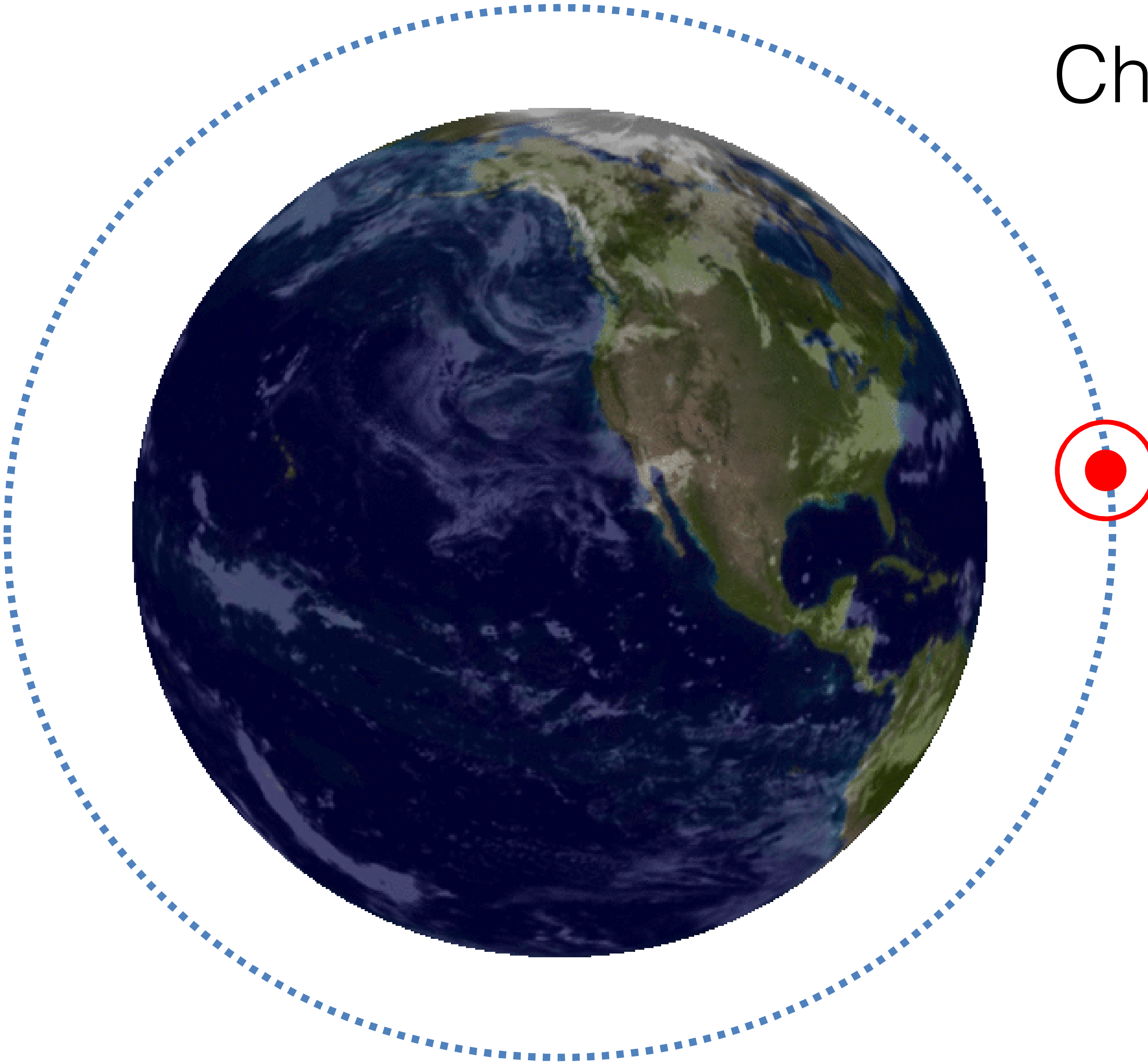


Trial 1.

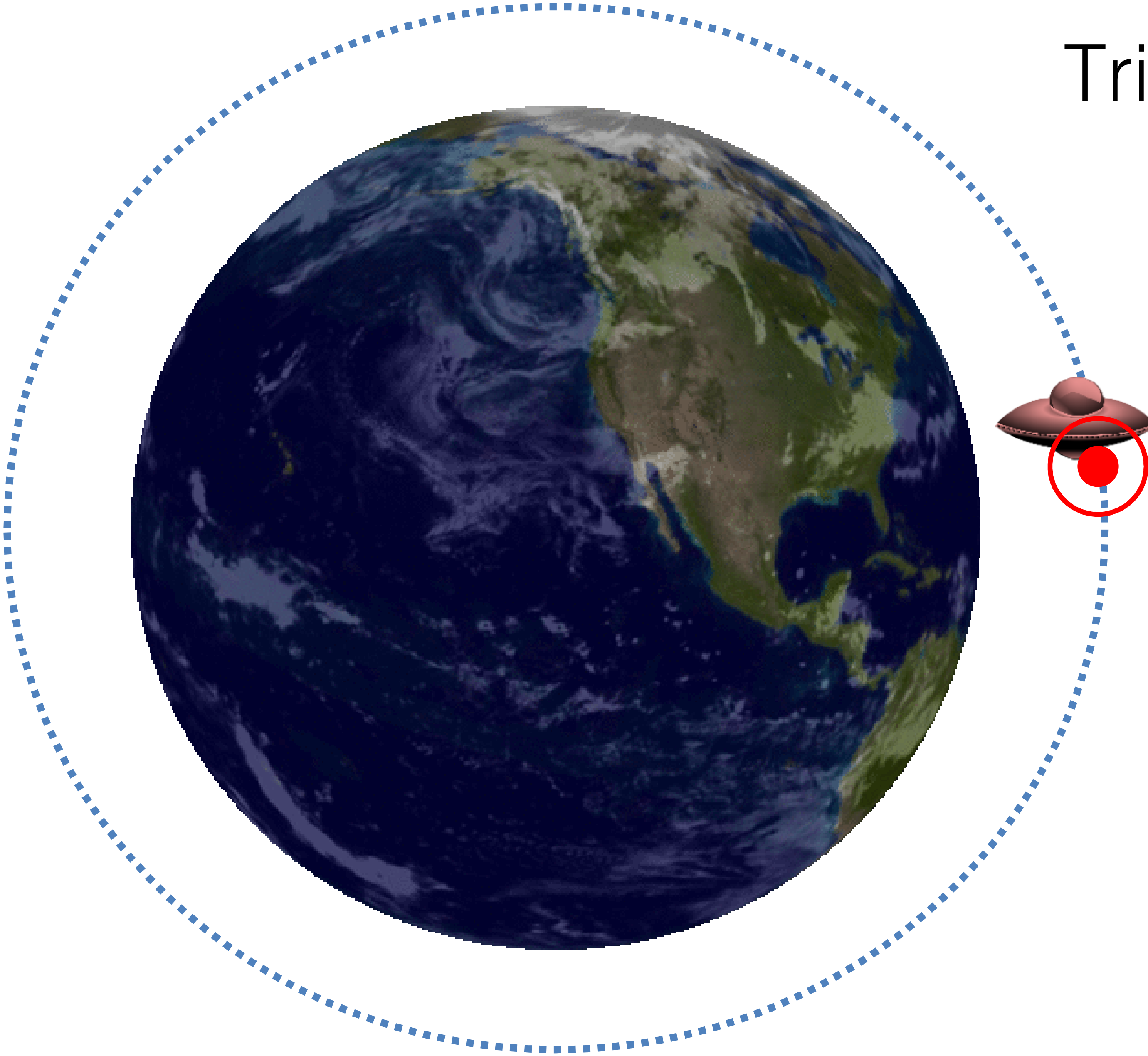




Predict trial 2.

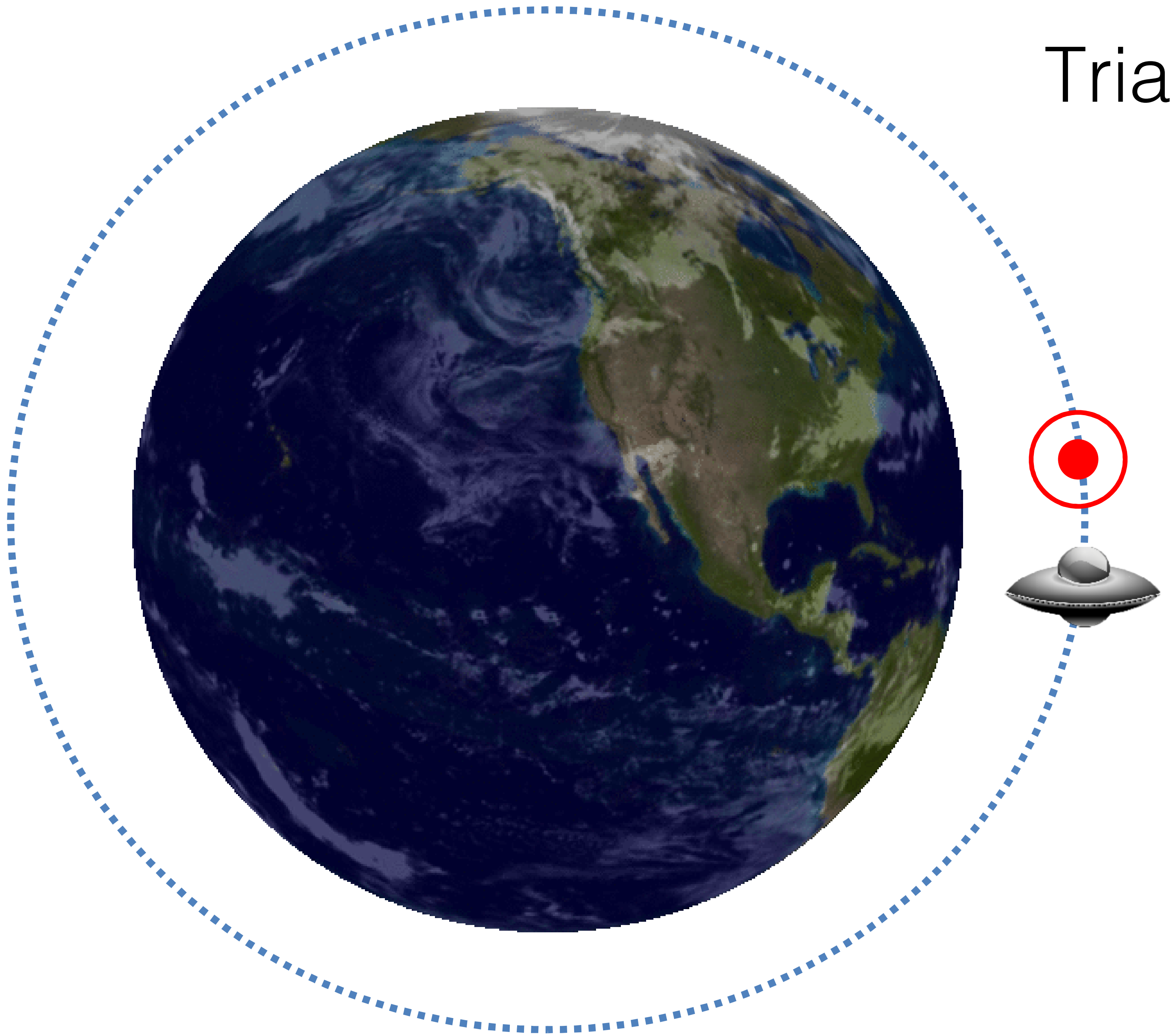


Choice.



Trial 2.

Trial 2.



Generation process.

- Consider the situation where the spaceship moves at a variable velocity and is corrupted with Gaussian noise.

Generation process.

- Consider the situation where the spaceship moves at a variable velocity and is corrupted with Gaussian noise.

$$\theta_{t+1} \sim N(x_{t+1}, \sigma_\theta^2)$$

$$x_{t+1} = x_t + v_t$$

$$v_{t+1} \sim N(v_t, \sigma_v^2)$$

Reinforcement learning model.

•Commonly used to model behavior in prediction tasks but provides an inaccurate description in changing environments.

$$V_{t+1} = V_t + \alpha\delta_t$$

Reinforcement learning model.

•Commonly used to model behavior in prediction tasks but provides an inaccurate description in changing environments.

$$V_{t+1} = V_t + \alpha\delta_t$$

• The prediction problem in our experiment can be solved by incorporating the estimation of the velocity component into the prediction.

Reinforcement learning model with a velocity term.

$$V_{t+1} = V_t + V'_{t+1} + \alpha\delta_t$$

$$V'_{t+1} = V'_t + \beta\delta_t$$

$$y_{t+1} \sim N(V_{t+1}, \eta)$$

where:

δ: Prediction error.

α: Learning rate for position

β : Learning rate for velocity

η: Variance for decision noise.

Reinforcement learning model with a velocity term.

$$V_{t+1} = V_t + V'_{t+1} + \alpha\delta_t$$

$$V'_{t+1} = V'_t + \beta\delta_t$$

$$y_{t+1} \sim N(V_{t+1}, \eta)$$

- This model assumes the generation process (position and velocity) can be recovered by taking noisy samples from it.

This work aims to answer two main questions:

- Is this a good model of people predictions in an environment that changes at a variable velocity?
- What is the relation between learning rates for position and velocity with the level of Gaussian noise?

Experimental design.

- 72 subjects performed the prediction task in four conditions defined by the signal-to-noise ratio (S/N):

Experimental design.

- 72 subjects performed the prediction task in four conditions defined by the signal-to-noise ratio (S/N):

$$\theta_{t+1} \sim N(x_{t+1}, \underline{\sigma_\theta^2})$$

$$x_{t+1} = x_t + v_t$$

$$v_{t+1} \sim N(v_t, \underline{\sigma_v^2})$$

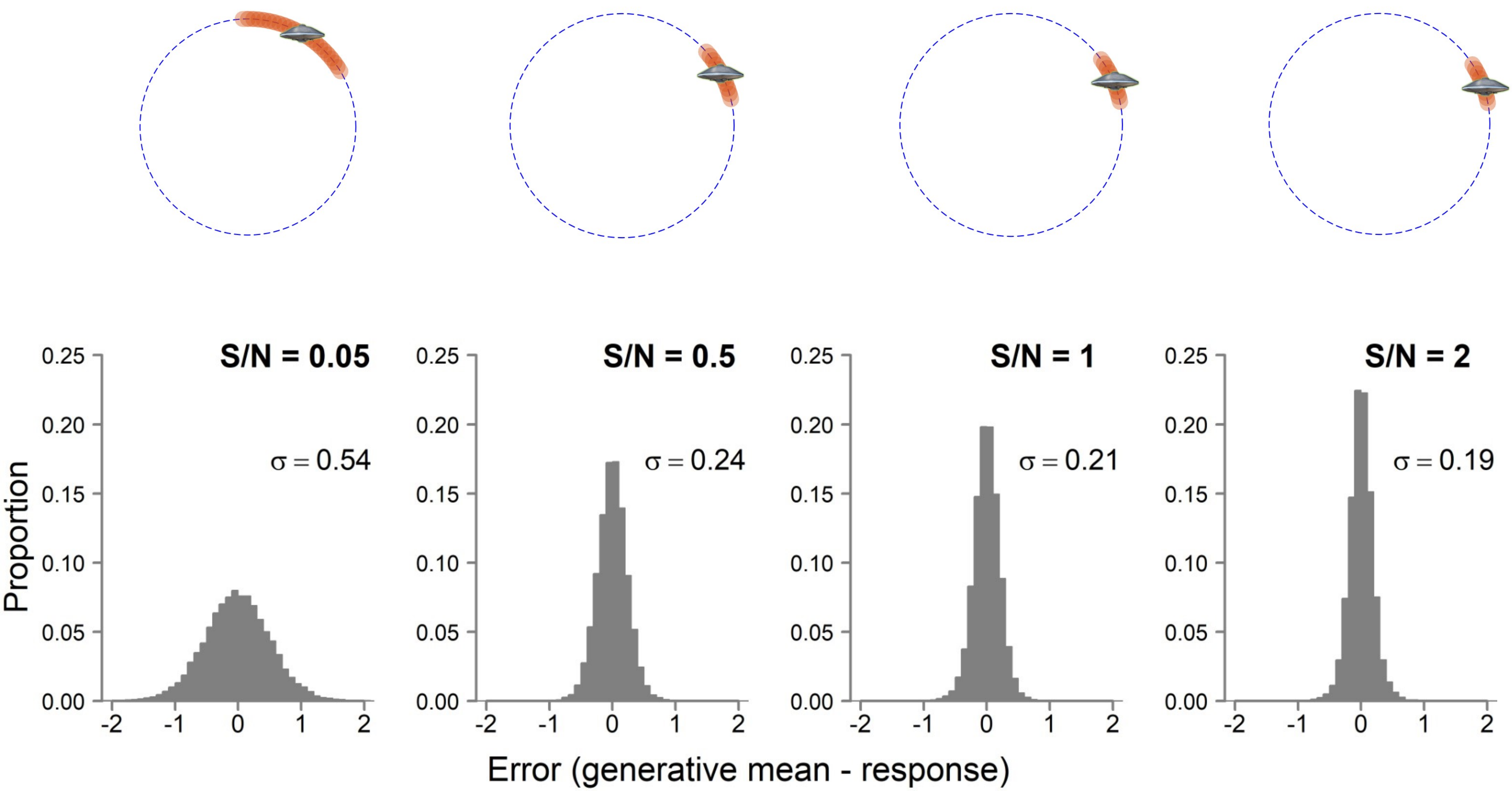
$$\frac{\sigma_v^2 \quad \leftarrow \text{Fixed}}{\sigma_\theta^2 \quad \leftarrow \text{Variable}}$$

Experimental design.

Condition	$\sigma_v^2 / \sigma_\theta^2$	Trials
1	0.05	300
2	0.5	300
3	1	300
4	2	300

- Order of conditions was randomized.
- Training phase with minimum of 50 trials.
- Units of observations and responses are in radians.

Results.

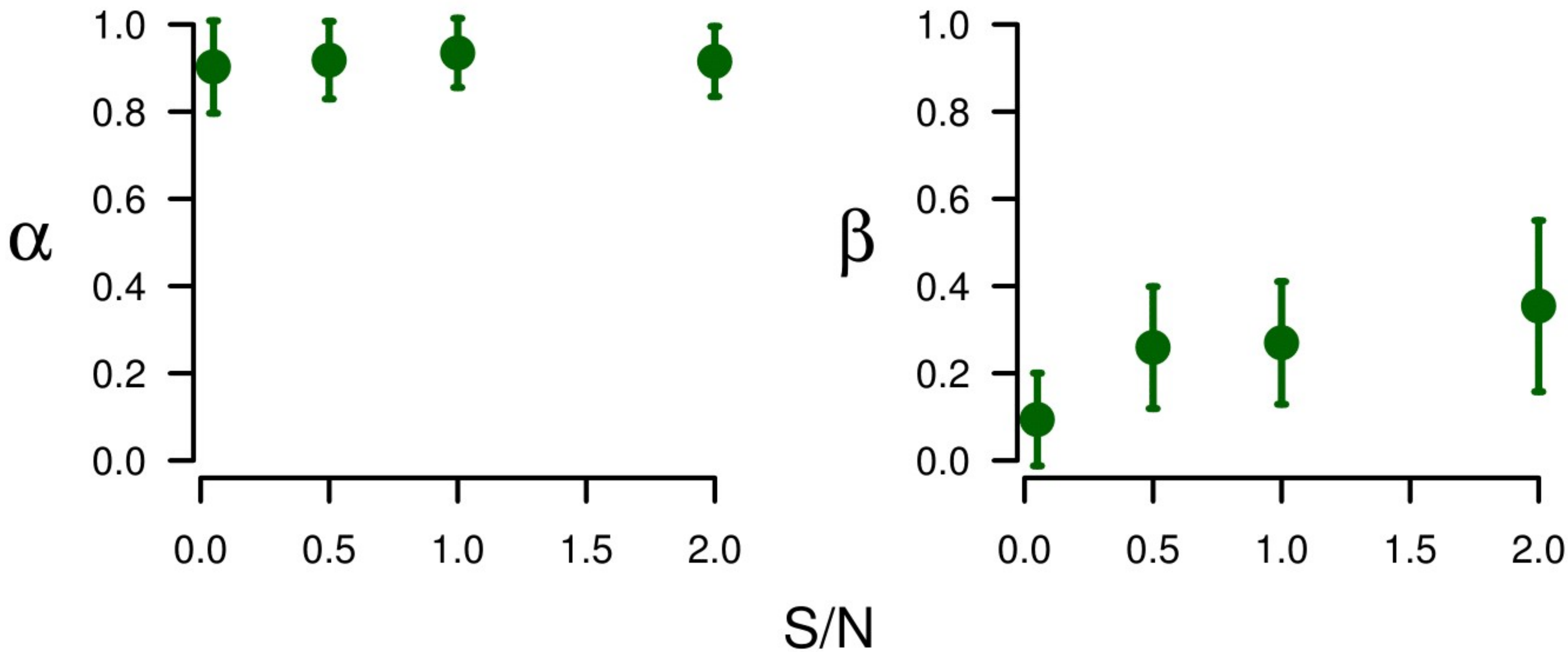

S/N = 0.05
S/N = 0.5
S/N = 1
S/N = 2
σ = 0.54
σ = 0.24
σ = 0.21
σ = 0.19
Proportion
0.25
0.20
0.15
0.10
0.05
0.00
-2
-1
0
1
2
Error (generative mean - response)

Learning rates.

- MCMC sampling was used to approximate posterior distributions of parameters.

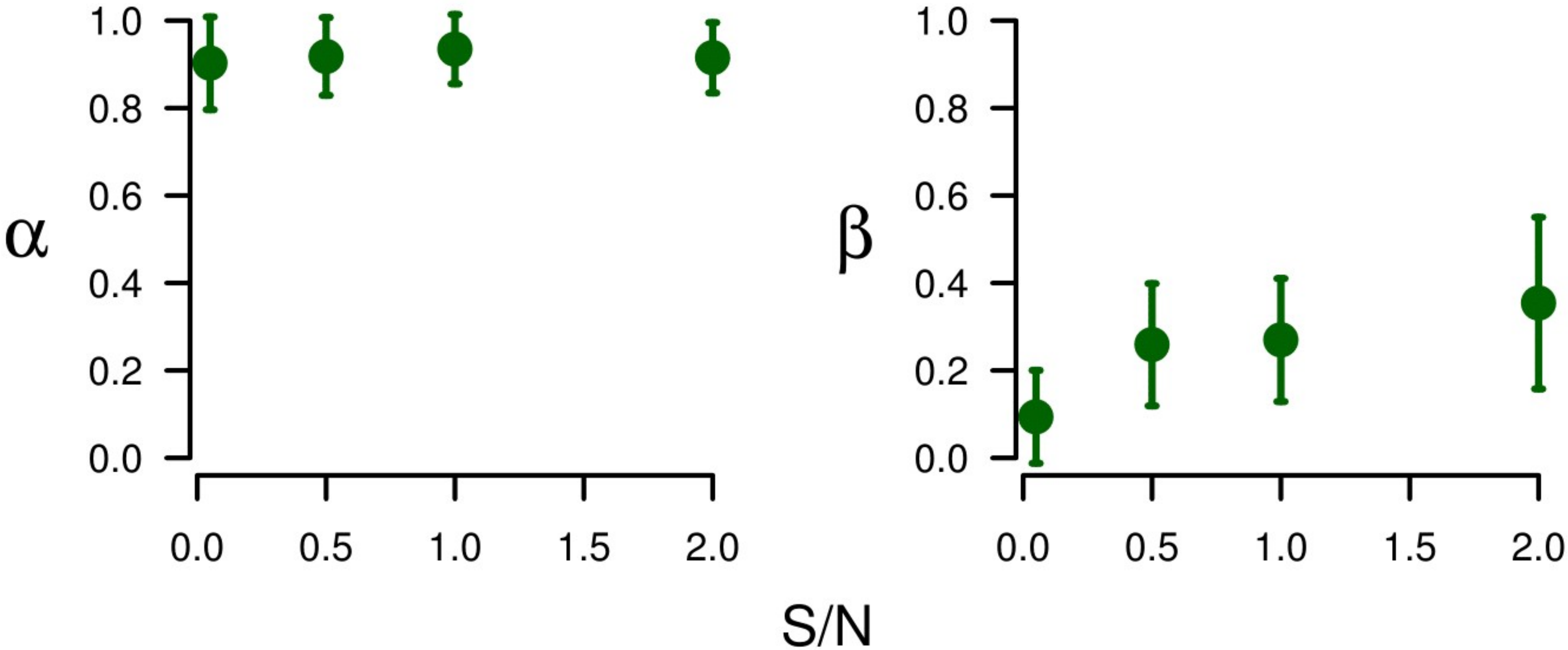
Learning rates.

- Maximum posteriors for position(**α**) and velocity (**β**) as a function of S/N.



Learning rates.

- Subjects took a value close to their last observation, and added the estimation for velocity, which was updated faster for less noisy conditions.

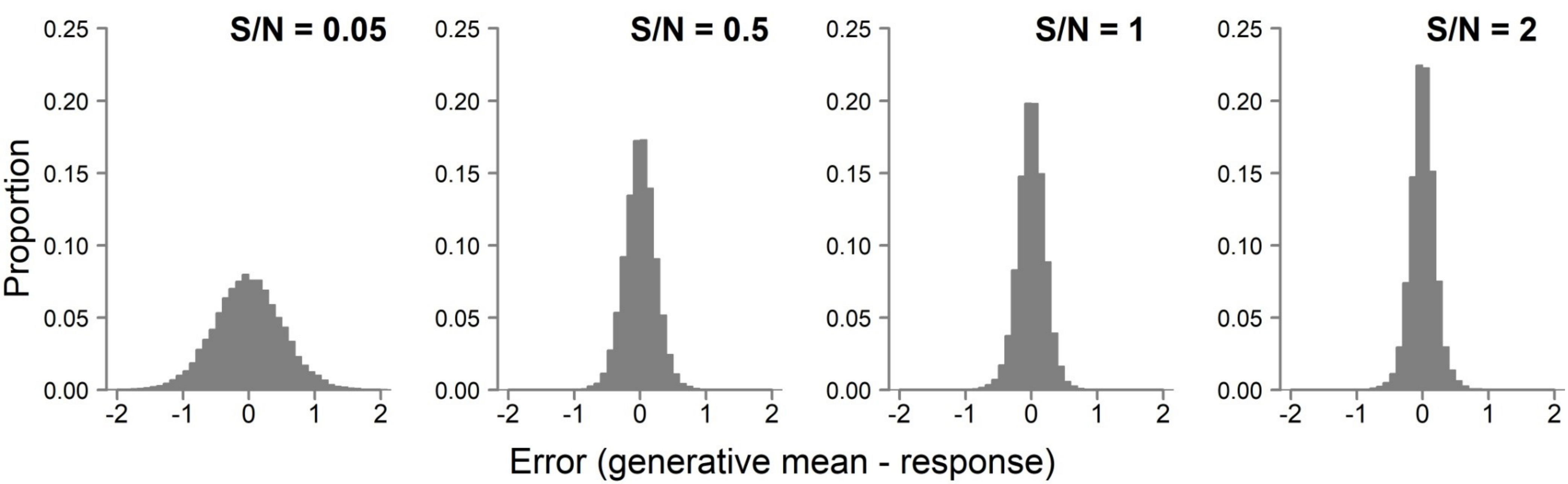


Descriptive accuracy.

- Simulations with samples from the joint posterior distribution and observations of subjects.

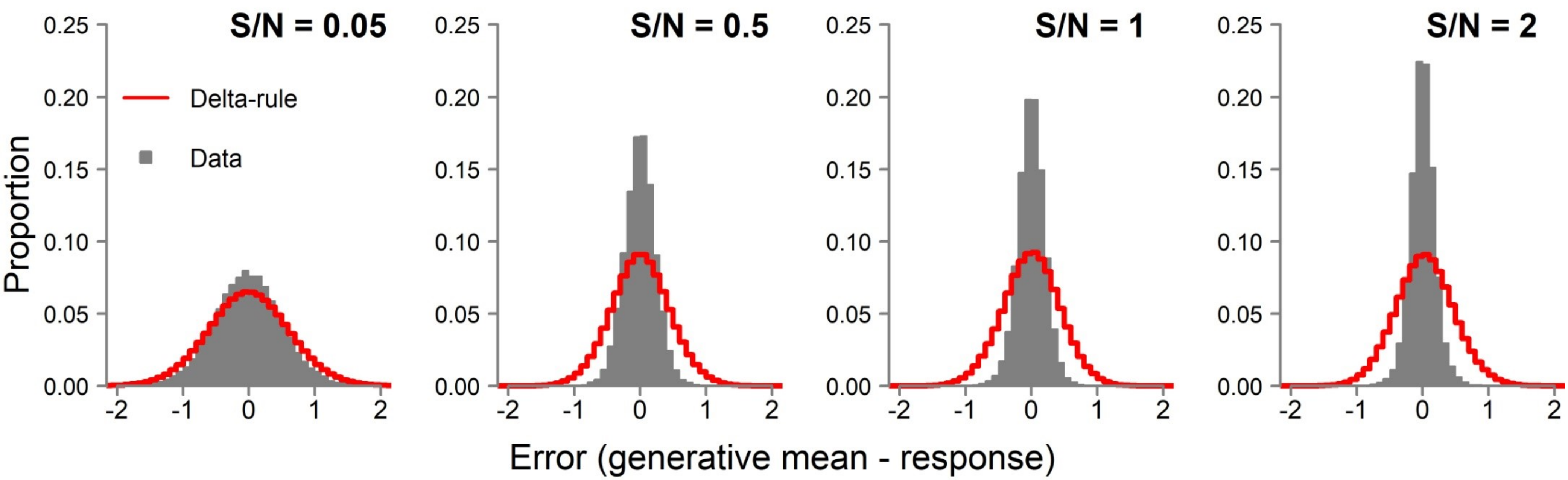
Descriptive accuracy.

- Simulations with samples from the joint posterior distribution and observations of subjects.



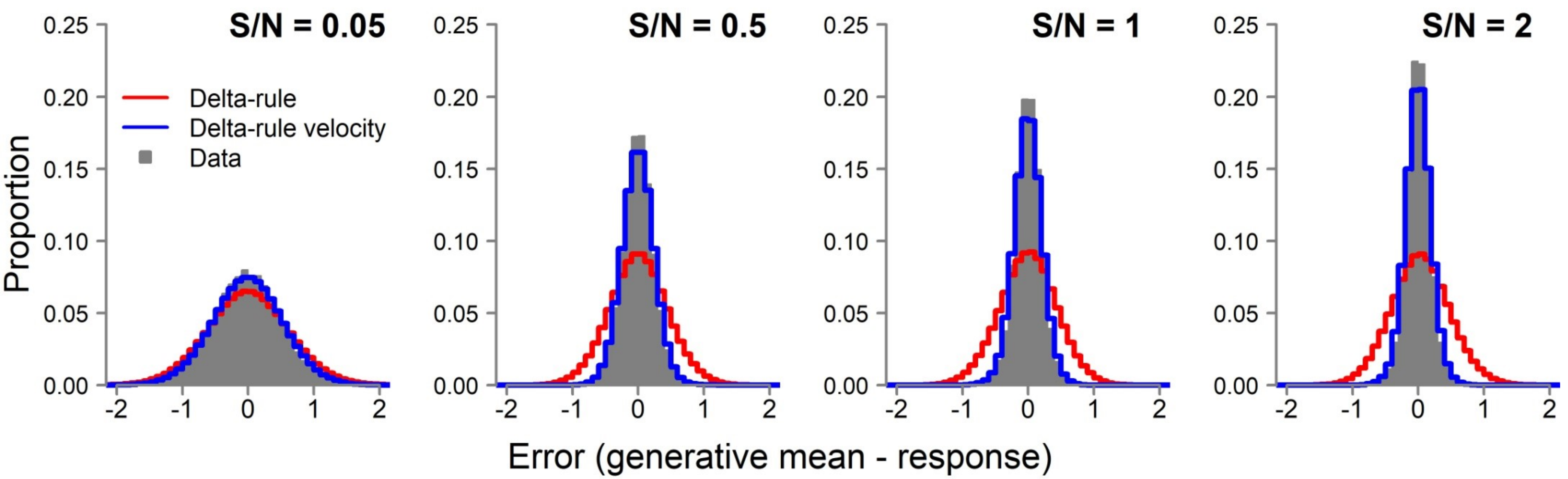
Descriptive accuracy.

- Standard delta-rule model.



Descriptive accuracy.

- Standard delta-rule model.
- Delta-rule with a velocity term.



Summary.

- A reinforcement learning model incorporating the estimation of velocity describes data better than a model that ignores this variable.

- Learning rates for position remain high in all conditions.

- Learning rates for velocity increase gradually with the signal-to-noise ratio.

Bayesian latent-mixture model.

- Proportion of participants favoring:

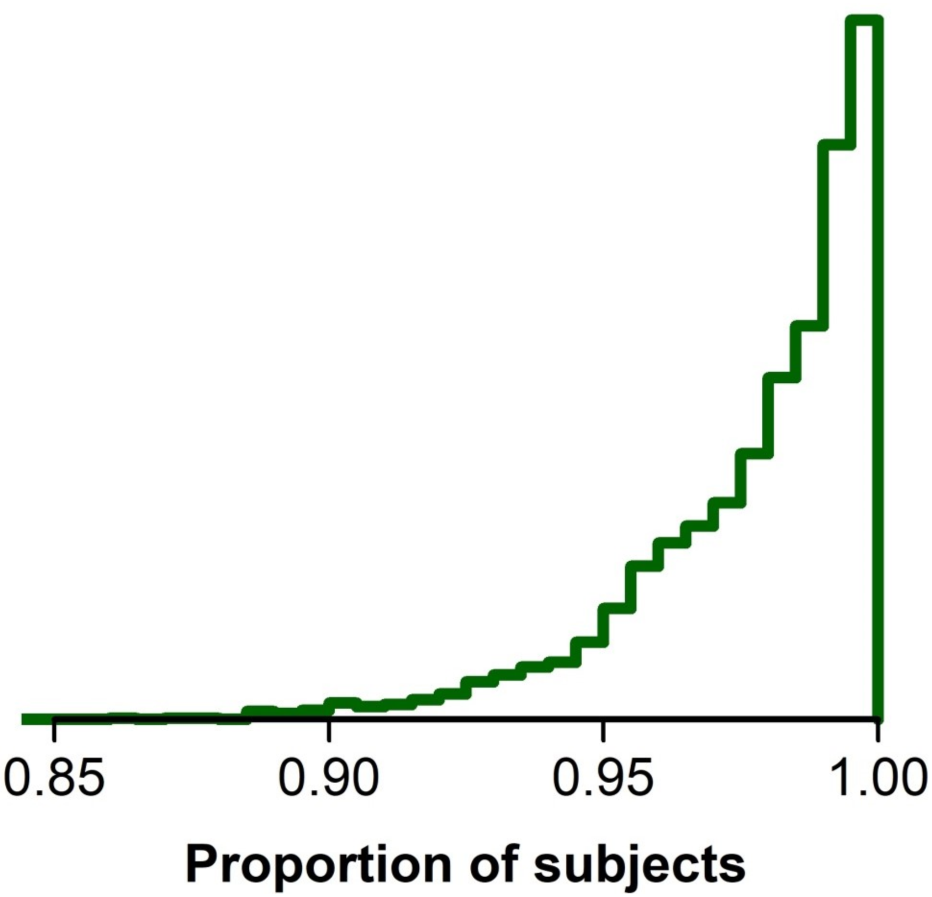
1. A reinforcement learning model incorporating a velocity term in comparison to the standard reinforcement model.

2. Gradual increase (hyperbolic) of learning rates for position as a function of S/N in comparison to a model that assumes no change.

3. Gradual increase (hyperbolic) of learning rates for velocity as a function of S/N in comparison to a model that assumes no change.

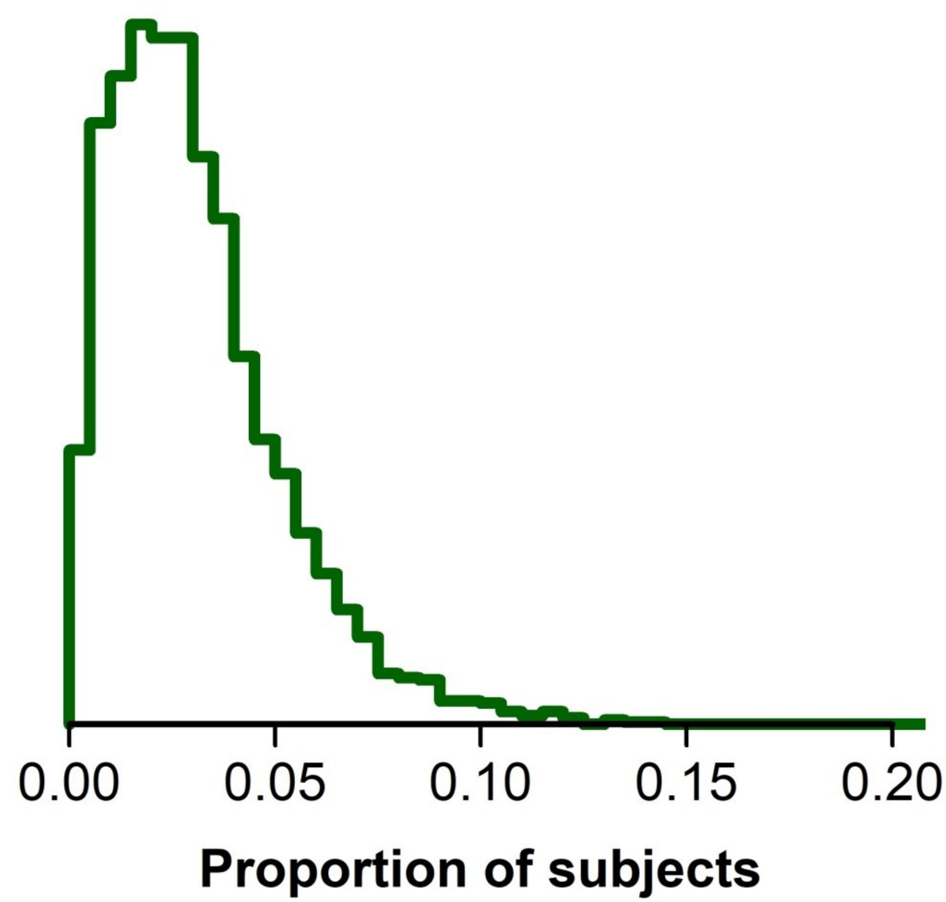
Bayesian latent-mixture model.

1. Estimated proportion of subjects favoring a model with a velocity term.



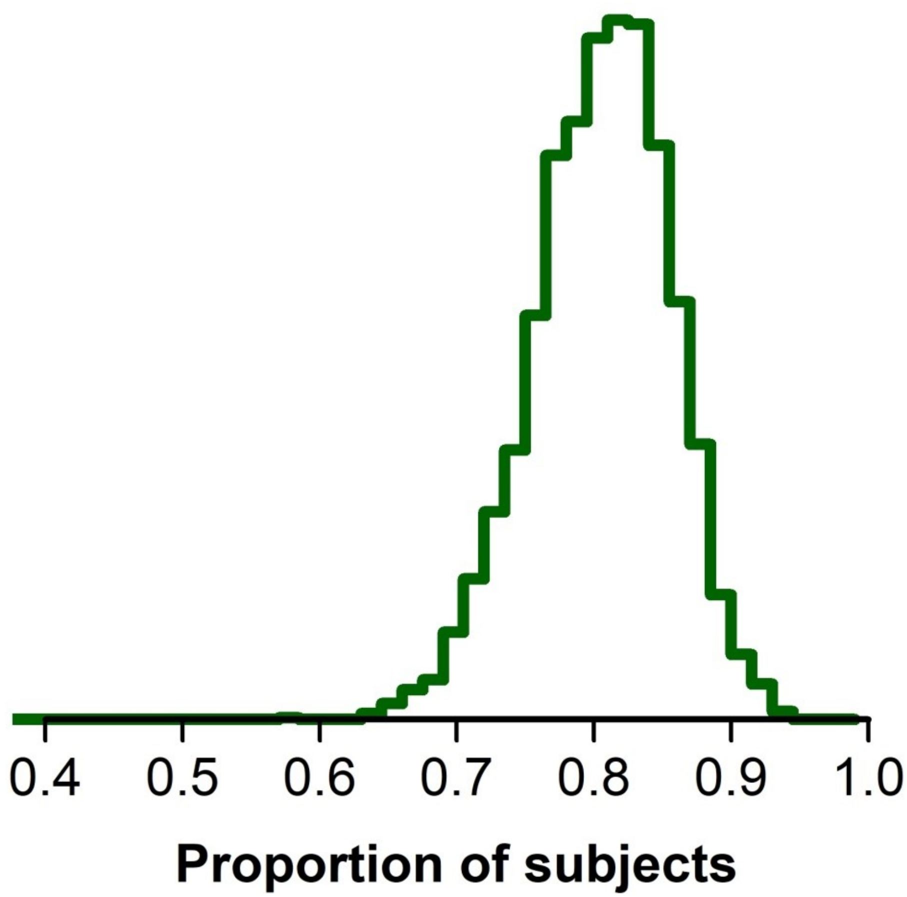
Bayesian latent-mixture model.

2. Estimated proportion of subjects favoring gradual increase of learning rates for position.



Bayesian latent-mixture model.

3. Estimated proportion of subjects favoring gradual increase of learning rates for velocity.



Conclusions.

- Incorporating the estimation of velocity to the standard reinforcement learning model allows to describe accurately participants’ predictions in an environment that changes at a variable velocity.

Conclusions.

- Most participants (≈80%) favor a gradual increment of learning rates for velocity as a function of S/N, which allows faster updates for this variable in less noisy conditions.

Conclusions.

- For most participants (≈98%) learning rates for position are high and do not change as a function of S/N.

- In all conditions participants added the estimation of velocity to a value close to the last observation they had.

Acknowledgements.

Research thesis:
bouzaslab25.com

Contact:
carlos.unamLab25@gmail.com