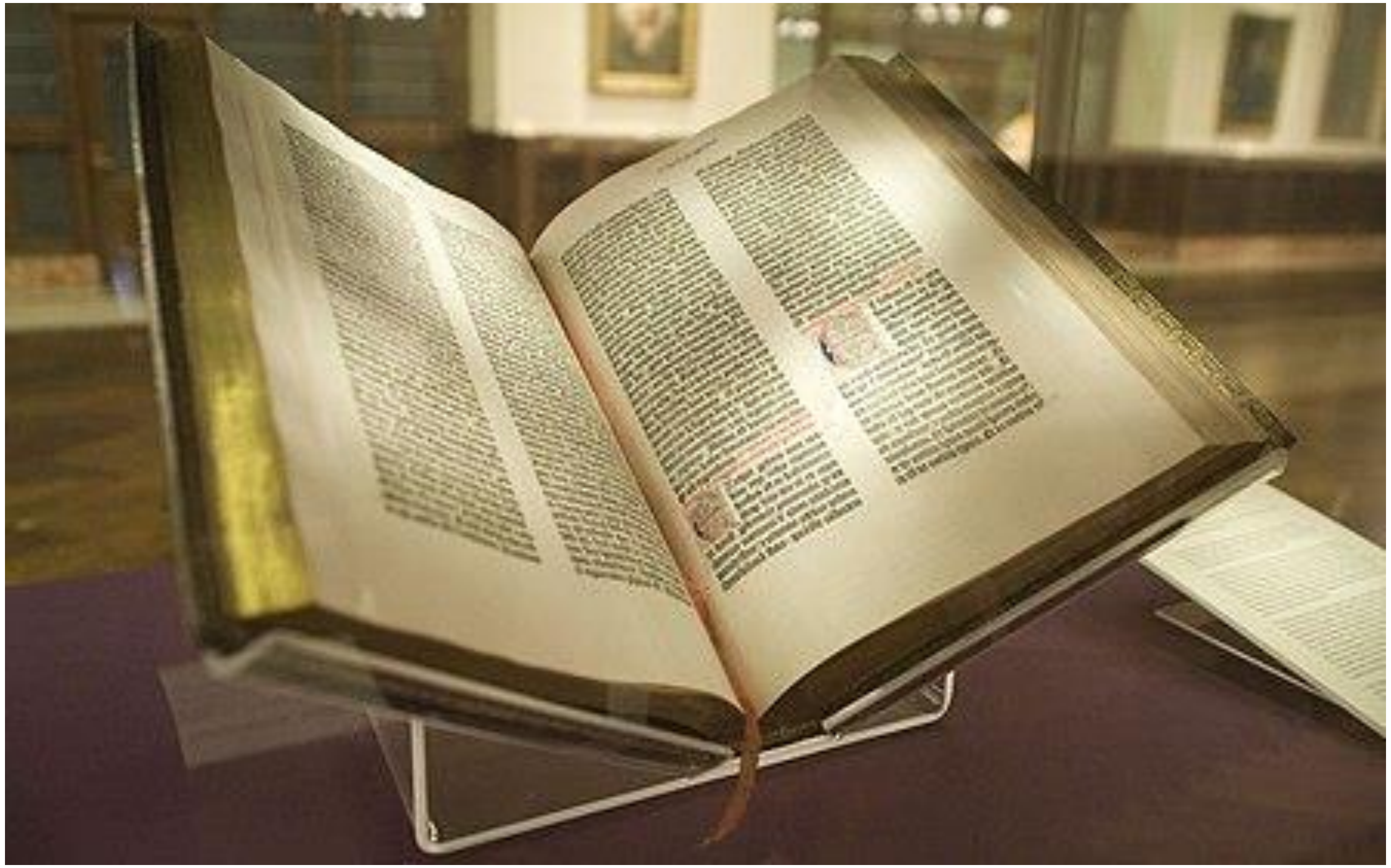
# Dynamics for working memory and time encoding

Zeyuan Ye

Christopher J. Cueva, et al. PNAS (2020)



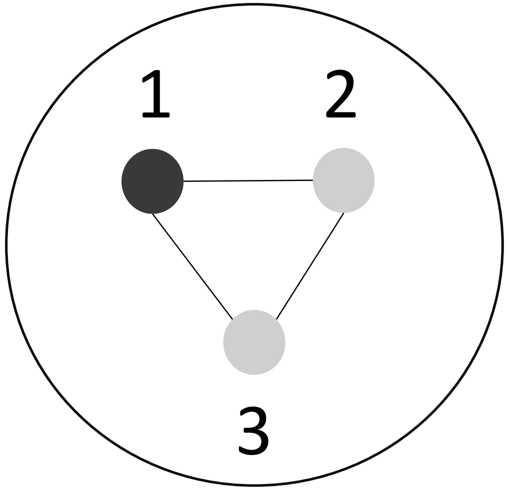
The mechanism of working memory?

1. Concepts of dynamics

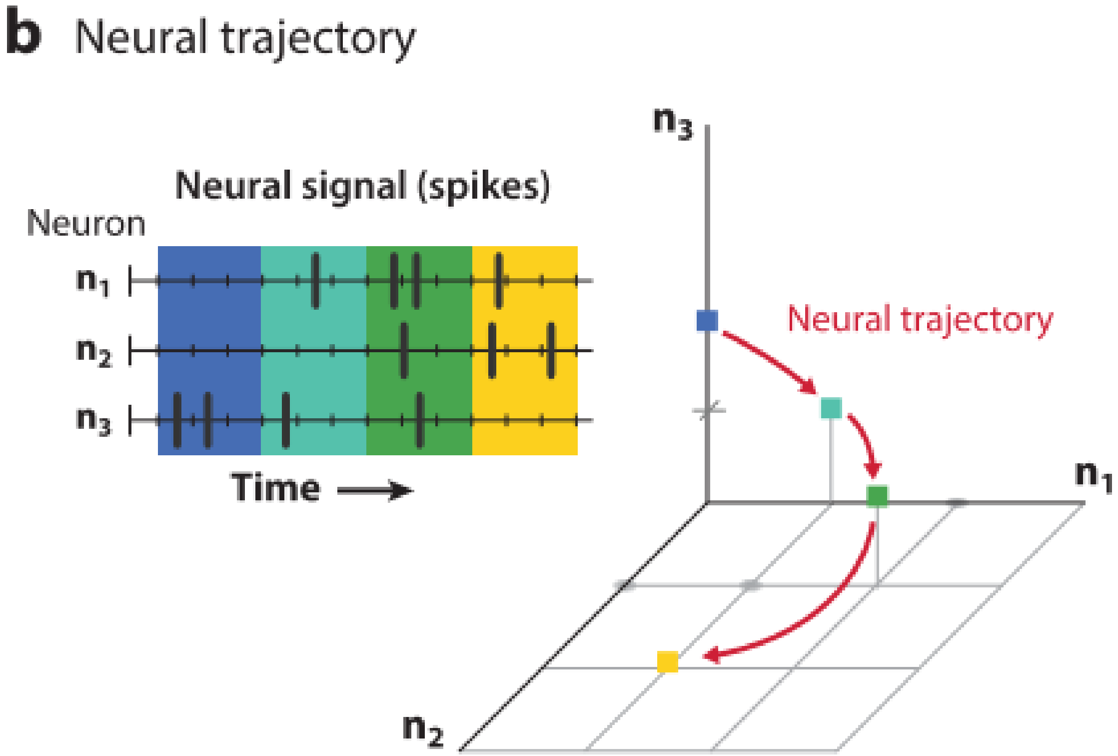
2. Three hypotheses of the working memory

3. Experiments

# 1. Concepts of dynamics




(1, 0, 0)




$x(t)$: N dimensional vector, N = # of neurons

# 1. Concepts of dynamics

$$\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij}\phi(x_j)$$

$$\frac{\mathbf{dx}}{\mathbf{dt}} = f\left(\mathbf{x}\left(t\right), \mathbf{u}\left(t\right)\right),$$

Vyas, S., Golub et al. (2020), Annual Review Neuro.

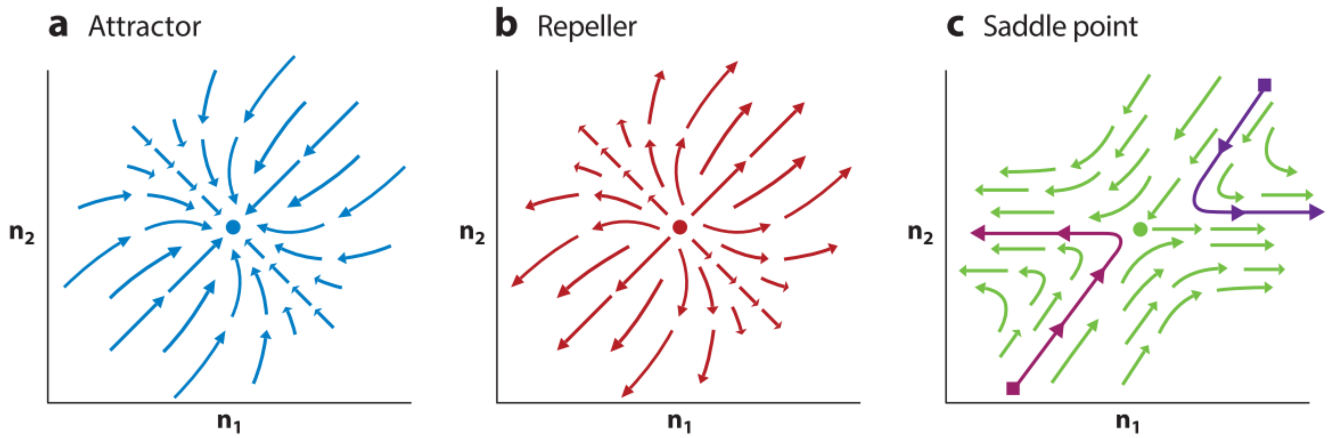
# 1. Concepts of dynamics

$$\frac{\mathbf{dx}}{\mathbf{dt}} = f\left(\mathbf{x}\left(t\right),\mathbf{u}\left(t\right)\right) = 0$$

Taylor Expansion

$$\frac{\mathbf{d}\left(\boldsymbol{\delta}\mathbf{x}\right)}{\mathbf{dt}} = \mathbf{A}\left(\mathbf{x}^{*}\right)\boldsymbol{\delta}\mathbf{x}\left(t\right),$$

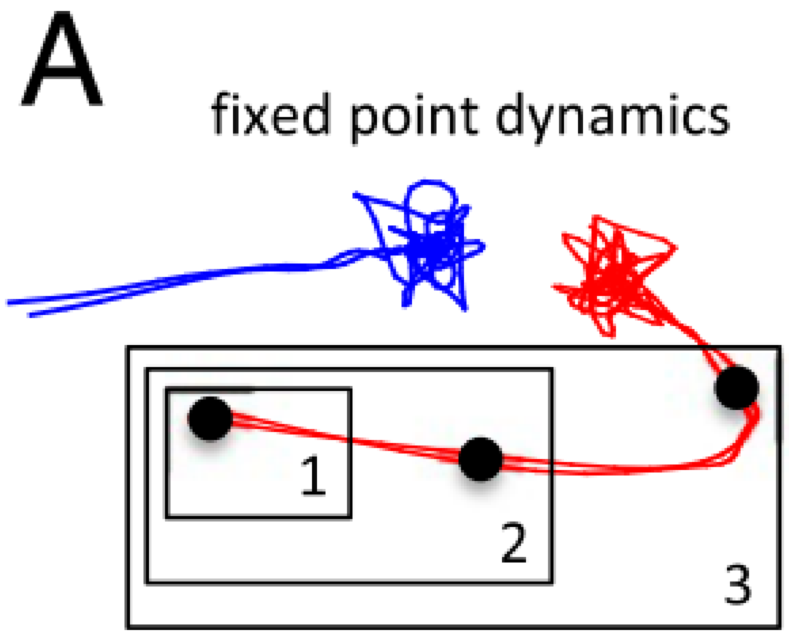
## Fix Point




Vyas, S., Golub et al. (2020), Annual Review Neuro.

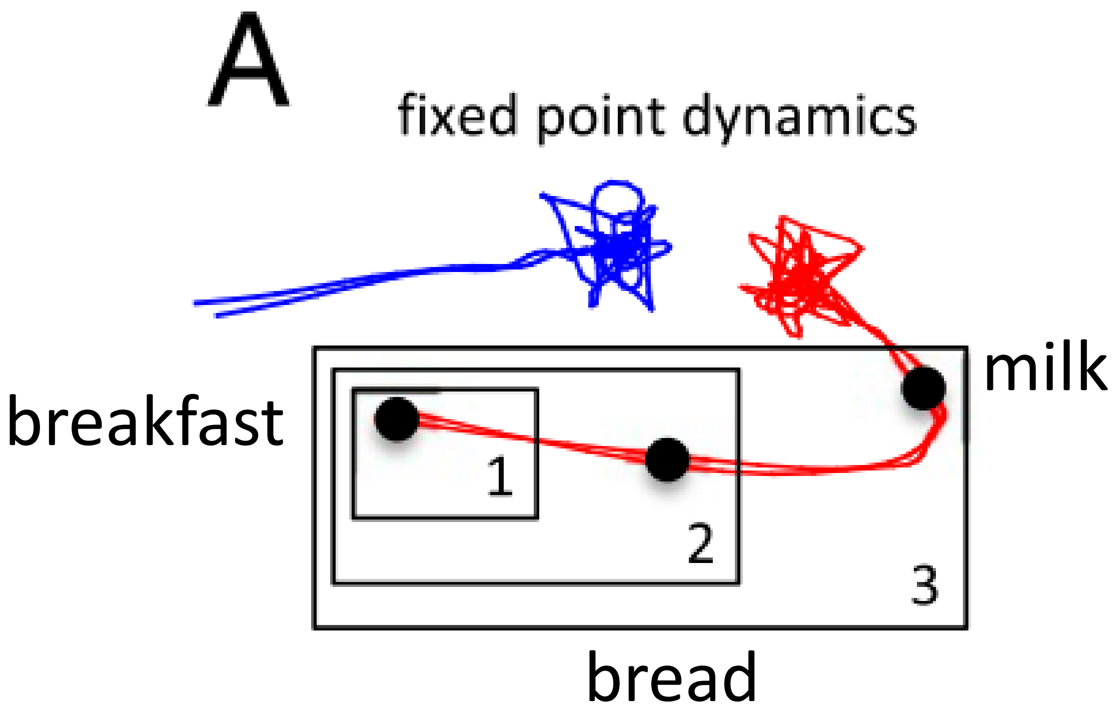
1. Concepts of dynamics
2. Three hypotheses of the working memory
3. Experiments

# 2. Attractor system




Christopher J. Cueva, et al. PNAS (2020)

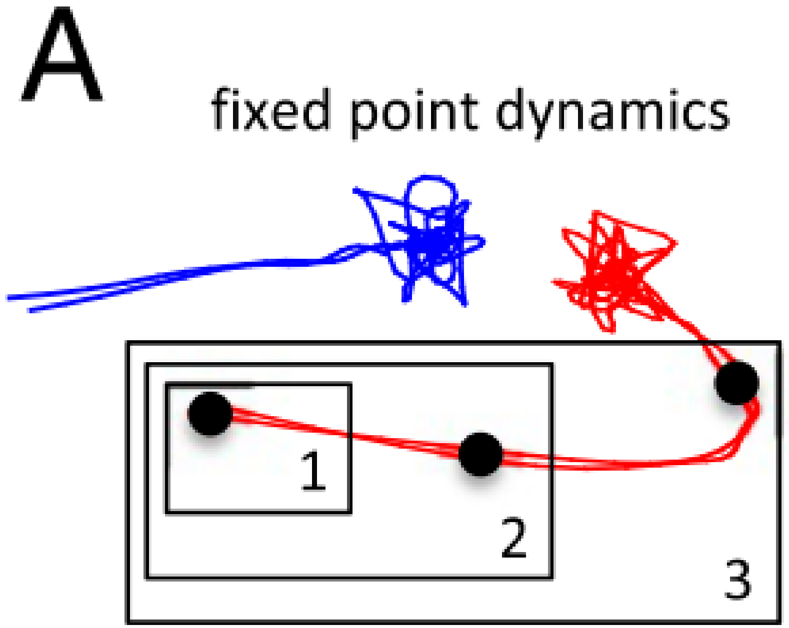
# 2. Attractor system




Associated Memory

Christopher J. Cueva, et al. PNAS (2020)

# 2. Attractor system




Trial Average

Christopher J. Cueva, et al. PNAS (2020)

# 2. Attractor system

How the brain realize the attractor system?

# 2. Attractor system

## Hopfield networks

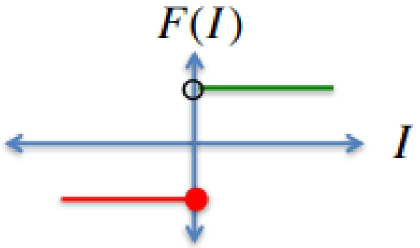
- We started with this dynamical equation

$$\tau_n \frac{d\vec{v}}{dt} = -\vec{v} + F\left[\vec{h} + M\,\vec{v}\right]$$

- We are going to simplify this as follows:

$$\vec{v}(t+1) = F\left[M\,\vec{v}(t)\right] \qquad v_i(t+1) = F\left[\sum_{j=1}^{N} M_{ij} v_j(t)\right]$$

where the neuronal activation function is




binary threshold neuron

$$F(x) = \mathrm{sgn}(x) = \begin{cases} 1 \; if \;\; x > 0 \\ -1 \; if \;\; x \le 0 \end{cases}$$



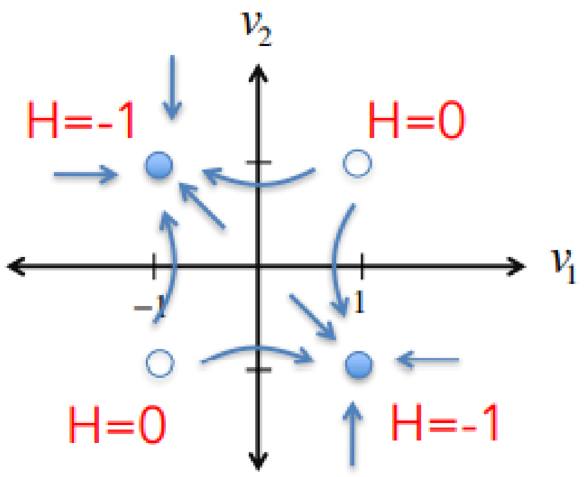
[Lecture 20: Hopfield Networks - 9.40 Introduction to Neural Compuatation (mit.edu)]

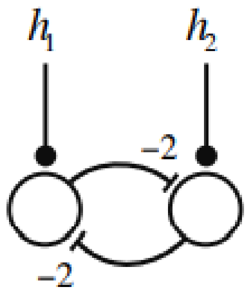
# 2. Attractor system

## The Energy Function

- Each possible state of the network has an energy given by:

$$H = -\frac{1}{2}\,\vec{v}^T M \vec{v} \qquad M = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}$$



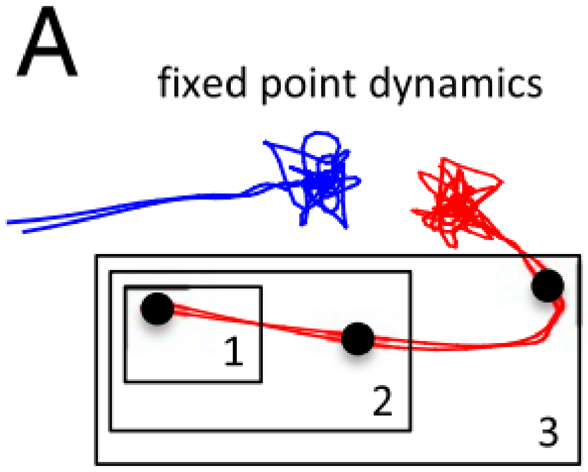




[Lecture 20: Hopfield Networks - 9.40 Introduction to Neural Compuatation (mit.edu)]

# 2. Attractor system

Conclusion:

1. Properties of attractor system




2. One realization of attractor system

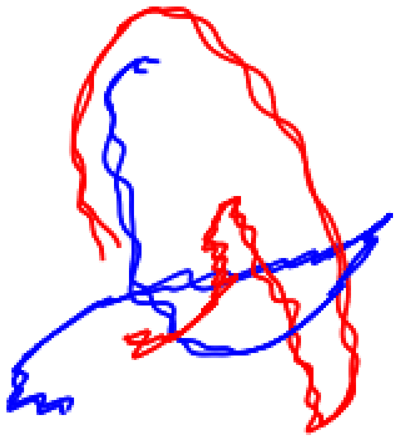
Hopfield networks

$$\vec{v}(t+1) = F\left[M\,\vec{v}(t)\right]$$

[Lecture 20: Hopfield Networks - 9.40 Introduction to Neural Compuatation (mit.edu)]

# 2. Trajectory

Neural trajectory
stabilized random RNN



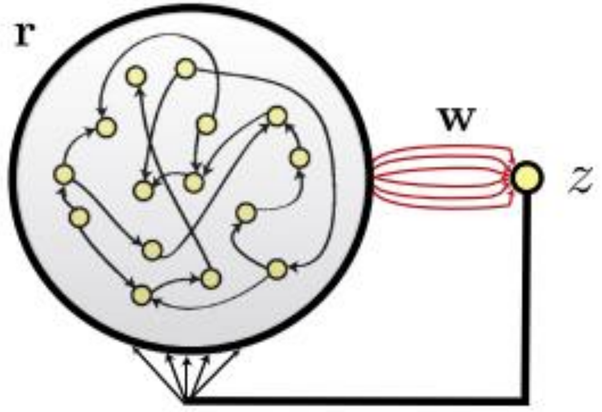
Time encoding

Christopher J. Cueva, et al. PNAS (2020)

# 2. Trajectory

How the brain realize trajectory hypothesis?

# 2. Trajectory

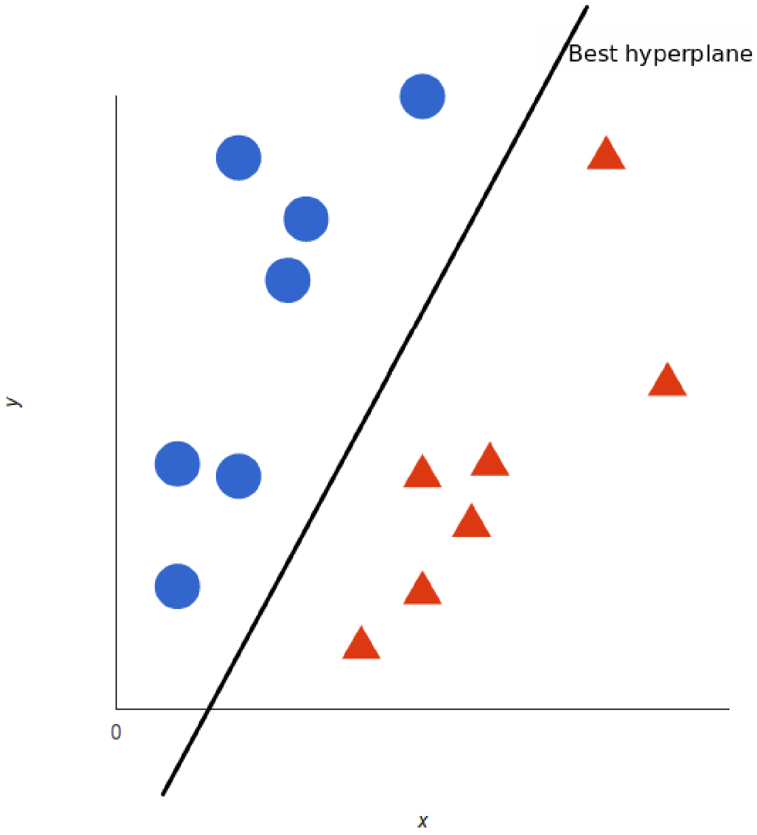



Echo state network

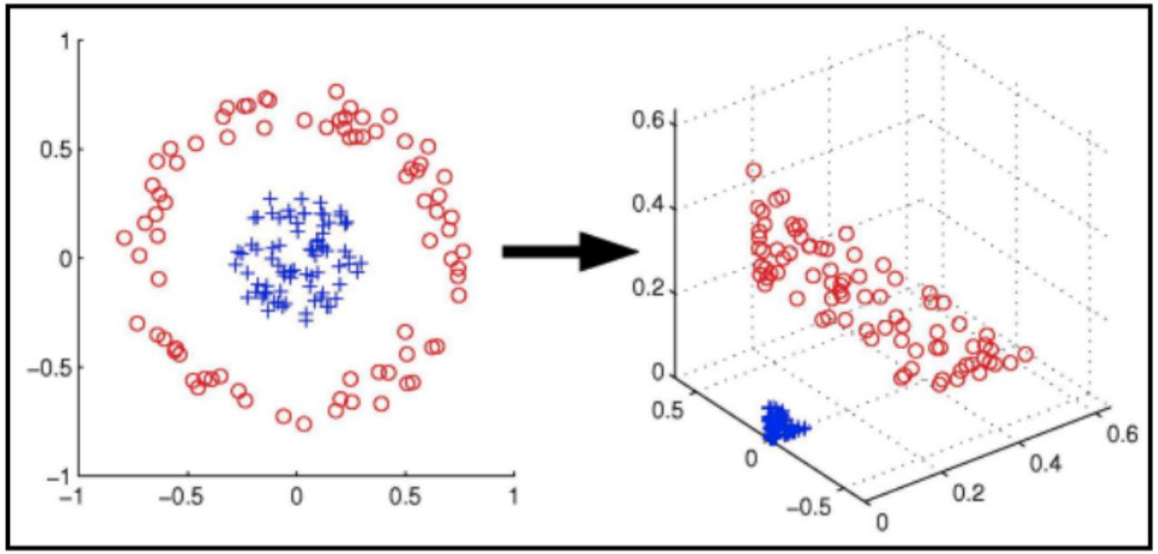
How this network works?

# 2. Trajectory

SVM linear kernel

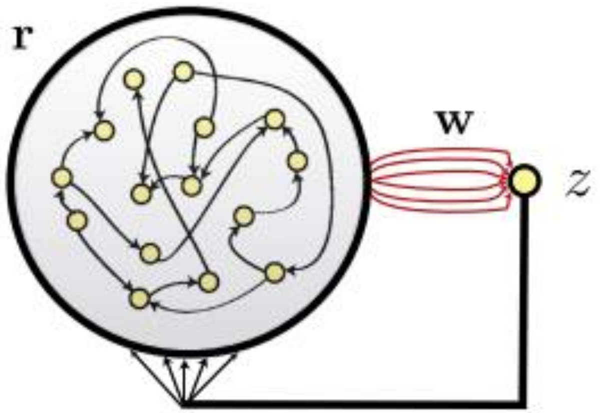


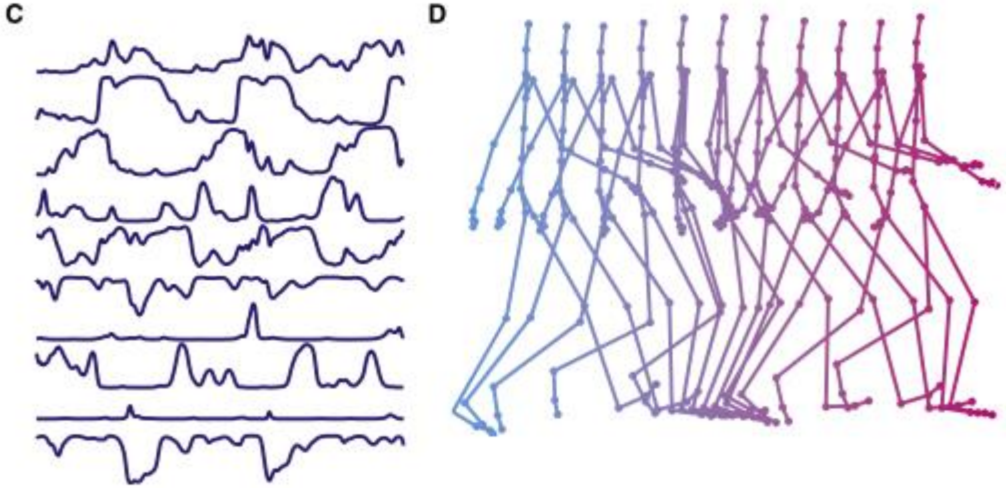

u = ax + by




u = ax + by + cz

# 2. Trajectory




High dimensional mapping

Sussillo, D., & Abbott, L. F. (2009). *Neuron*, *63*(4), 544–557.

# 2. Trajectory

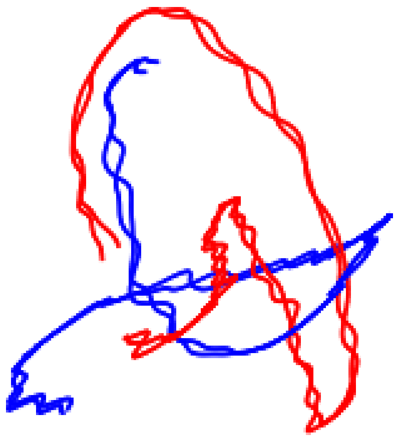
Conclusion:

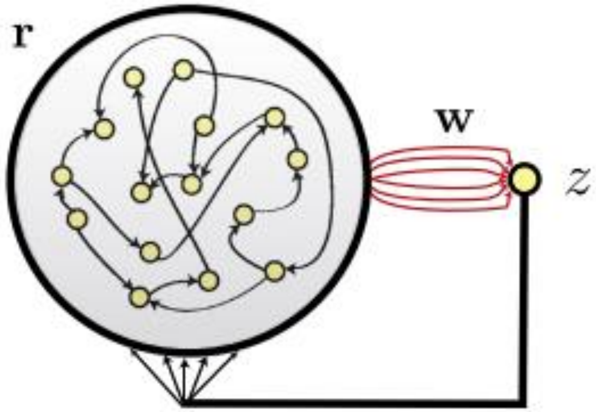
1. Neural Trajectory is easy for time encoding

2. One realization of trajectory hypothesis

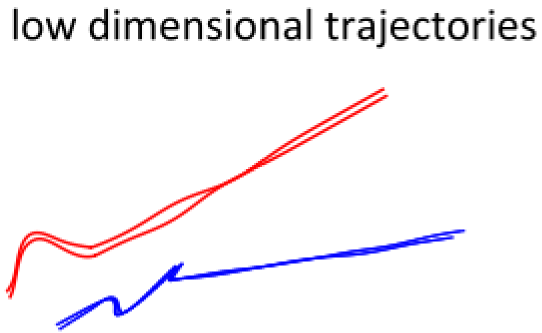
Neural trajectory

stabilized random RNN

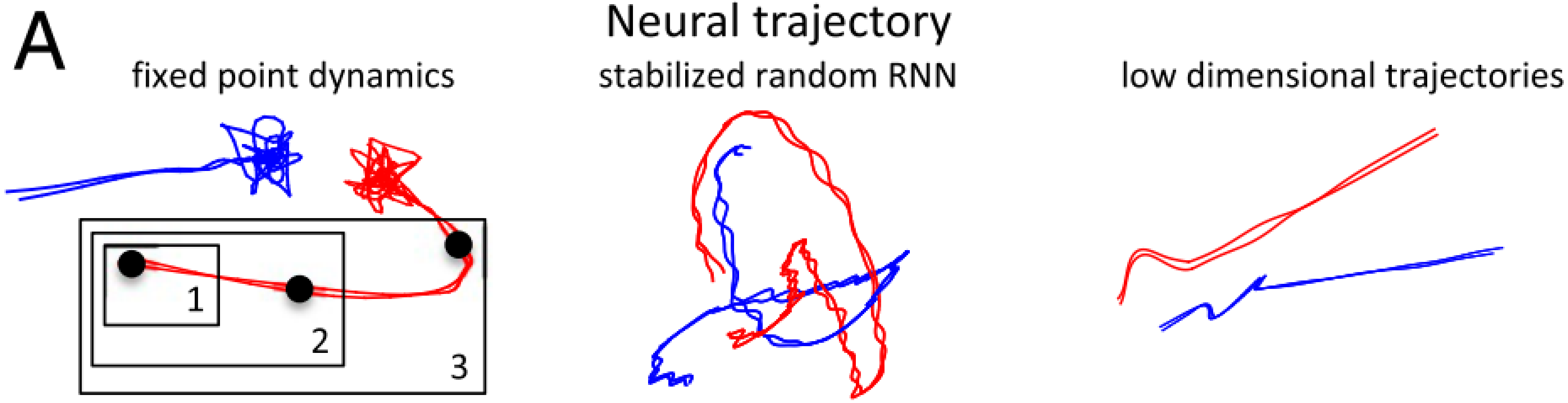




# 3. Low dimensional trajectory



# Neural trajectory



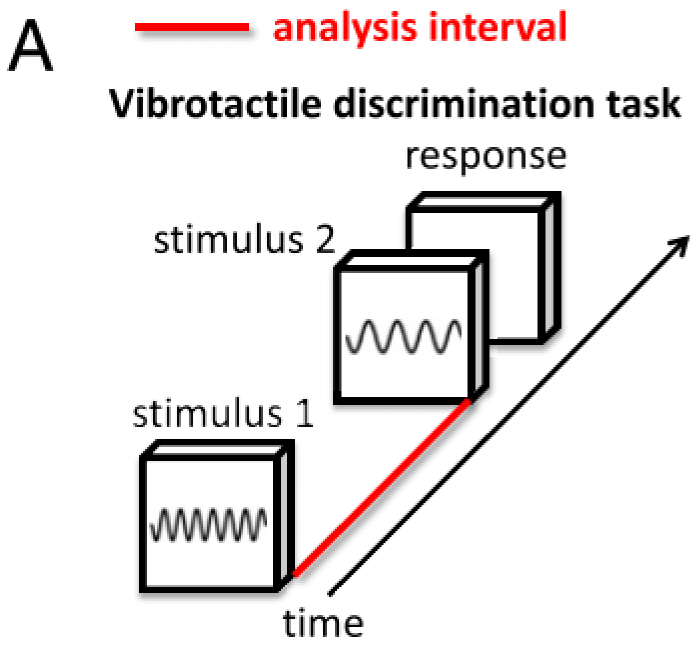

Experiment?

1. Concepts of dynamics
2. Two hypotheses of the working memory
3. Experiments

# 3. Experiment

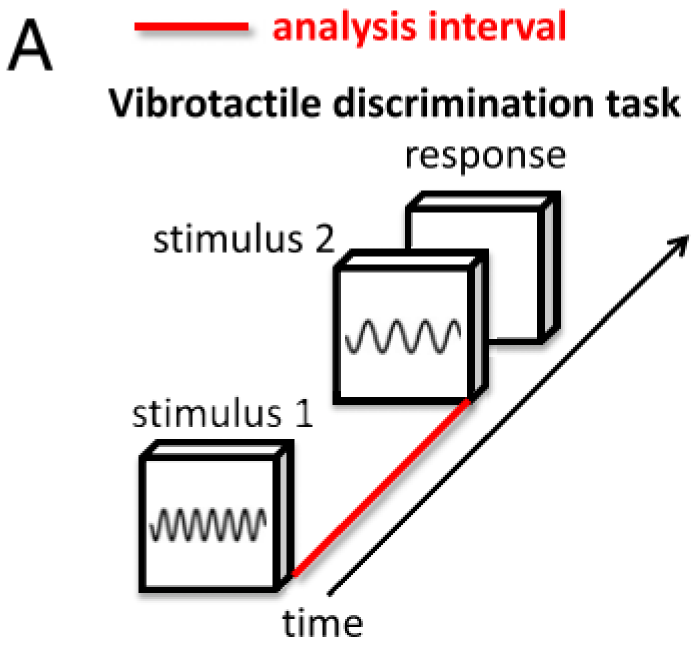
## Low-dimensional dynamics for working memory and time encoding

Christopher J. Cueva[a,b,c,1], Alex Saez[a], Encarni Marcos[d,e], Aldo Genovesio[e], Mehrdad Jazayeri[f,g], Ranulfo Romo[h,i,1], C. Daniel Salzman[a,c,j,k,l], Michael N. Shadlen[a,c,j,m], Stefano Fusi[a,b,c,j,1]




Christopher J. Cueva, et al. PNAS (2020)

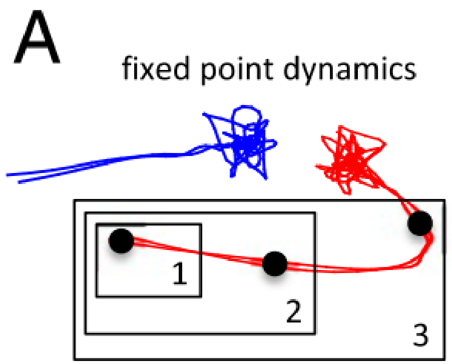
# 3. Experiment



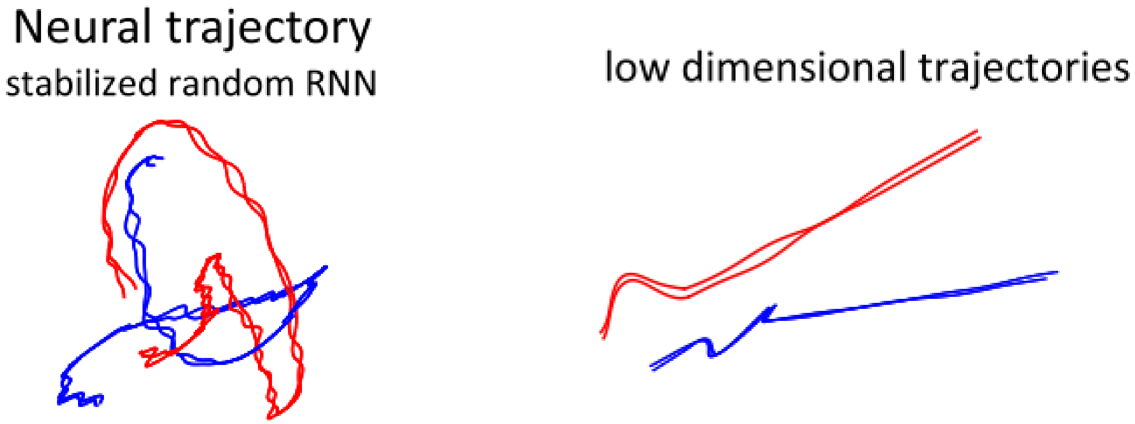

Christopher J. Cueva, et al. PNAS (2020)

## Prediction from two hypothesis

1. Neural state will not changed

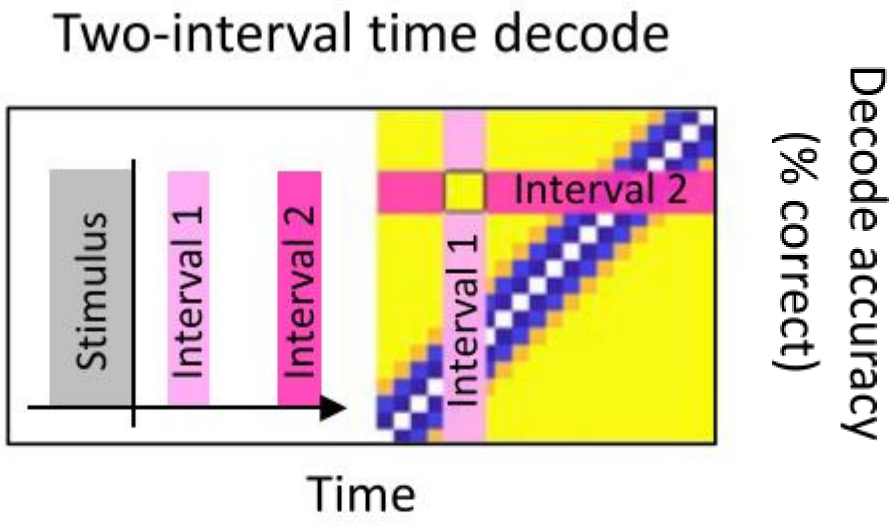


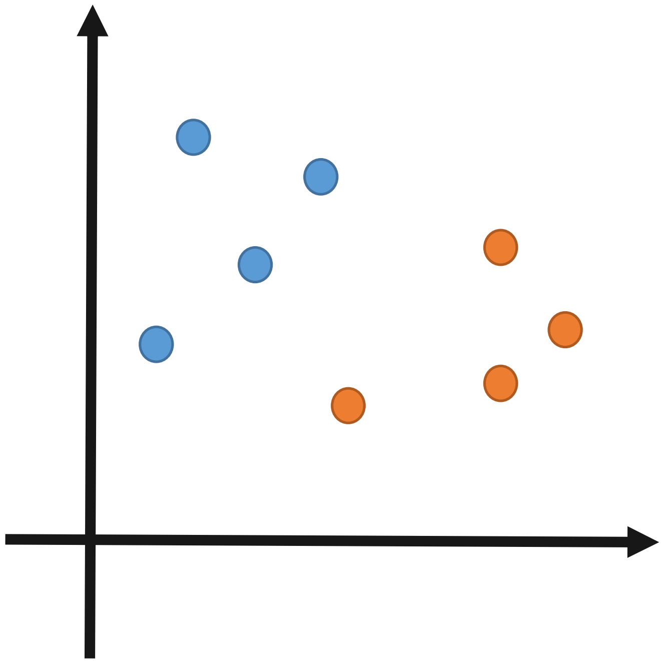

2. Neural state will change



# 3. Experiment

How to tell if the neural state change over time or not?

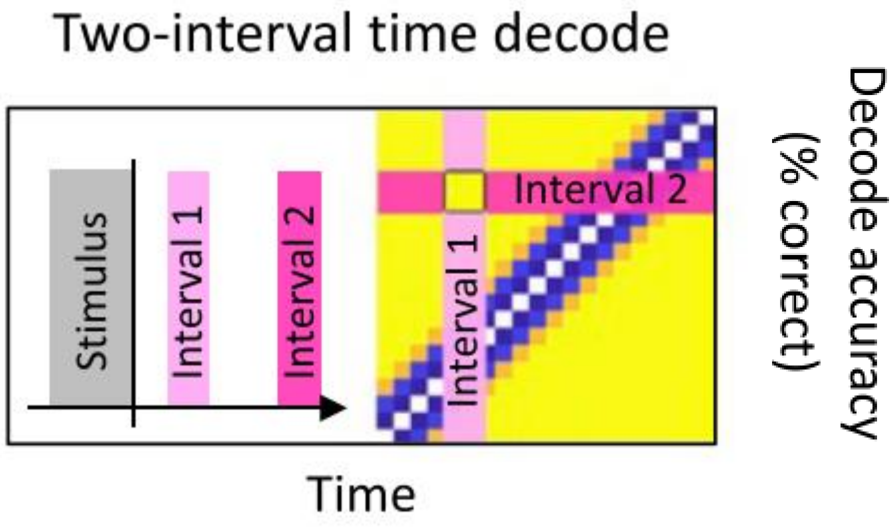


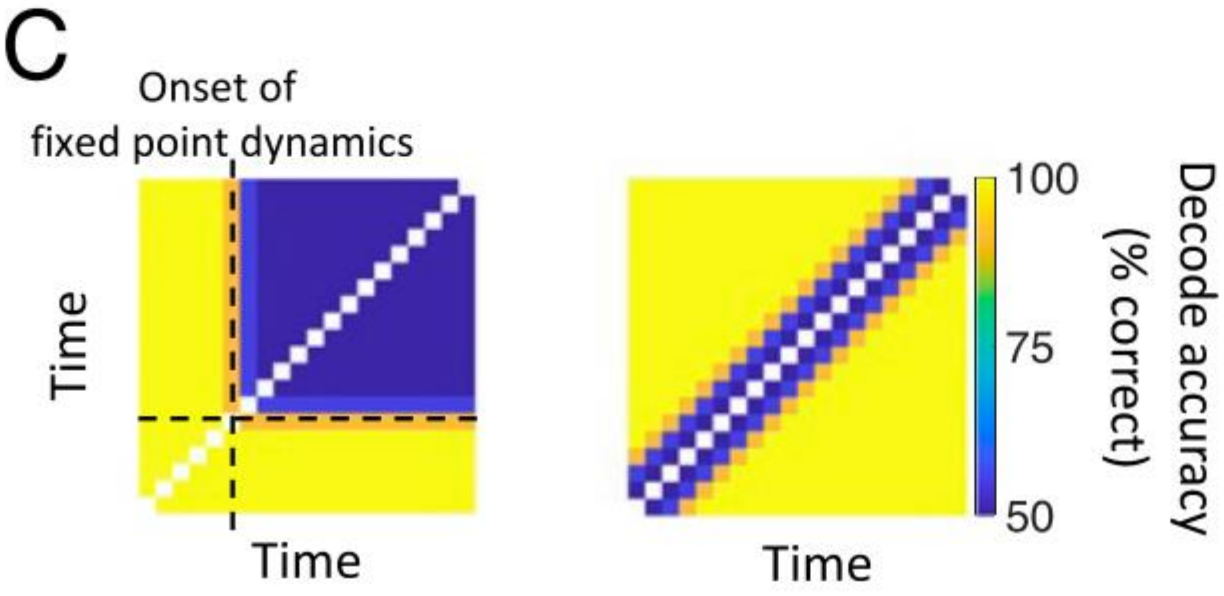



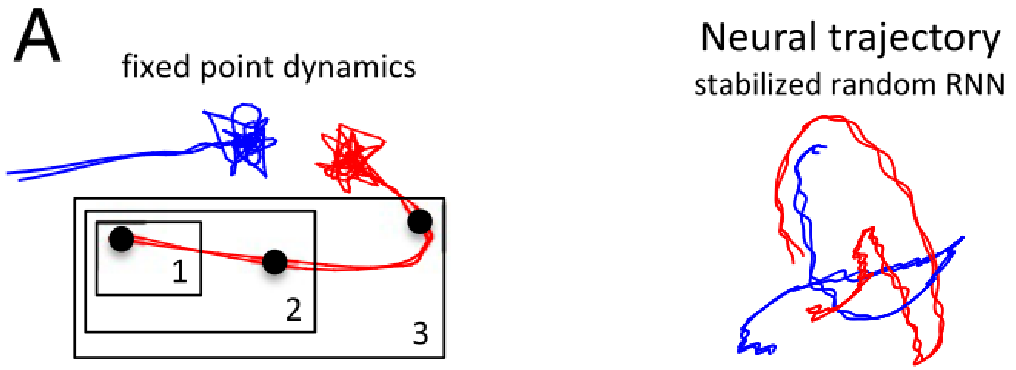
Christopher J. Cueva, et al. PNAS (2020)

# 3. Experiment

How to tell if the neural state change over time or not?

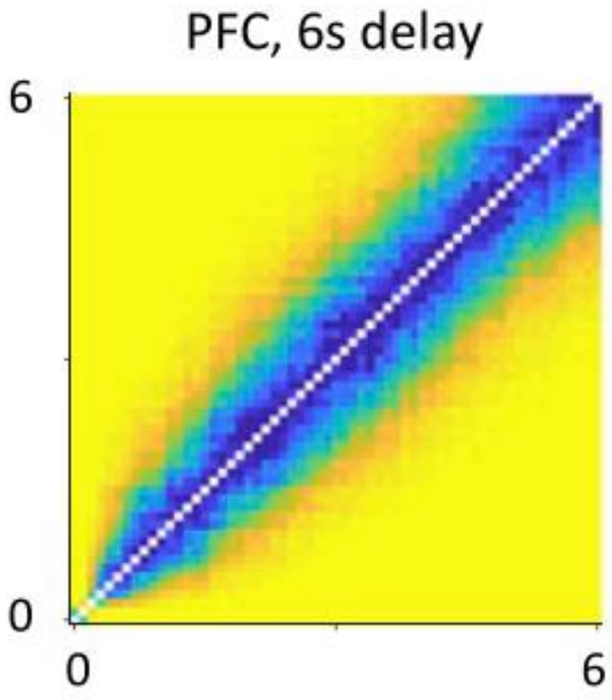


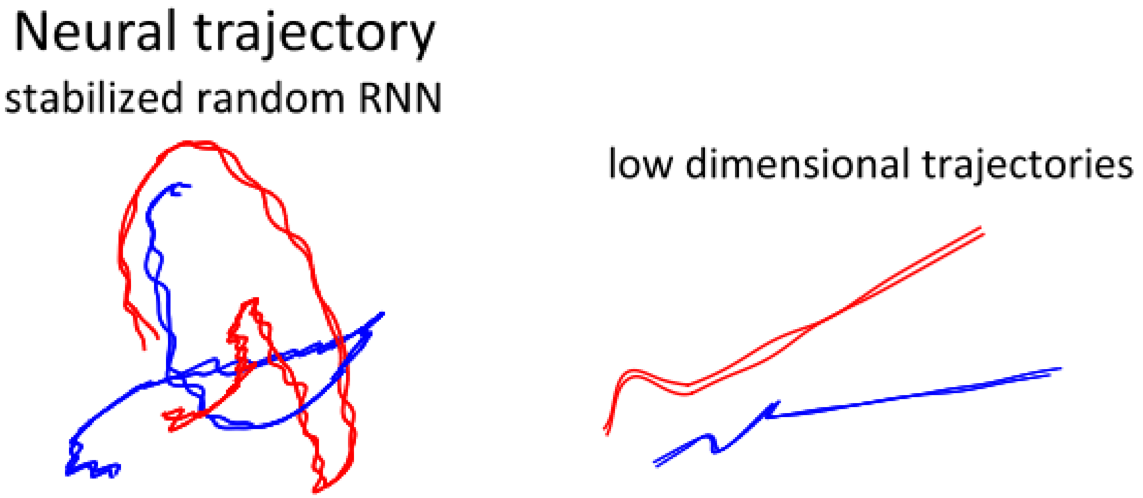




Christopher J. Cueva, et al. PNAS (2020)

# 3. Experiment

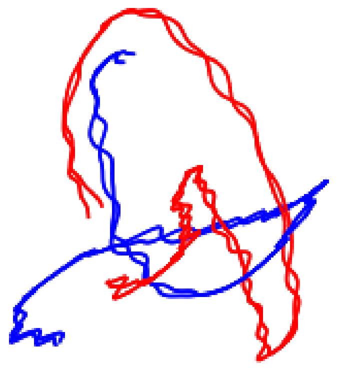


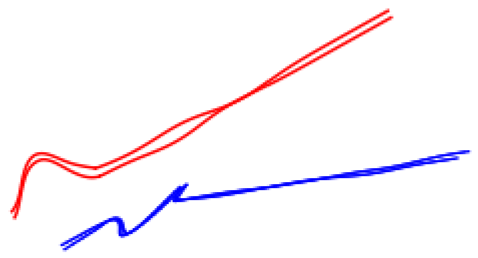

Christopher J. Cueva, et al. PNAS (2020)

# Neural trajectory

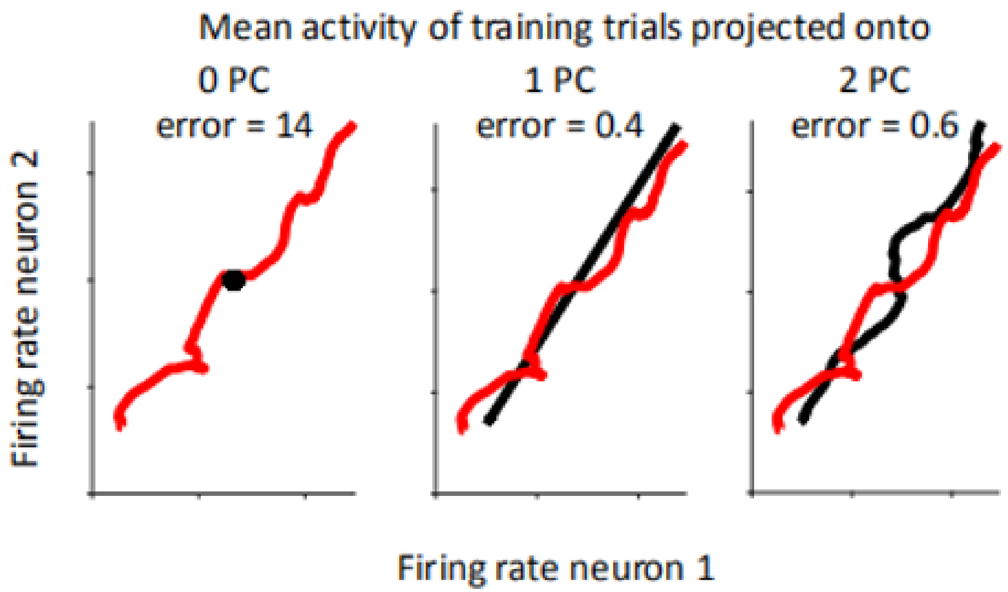
stabilized random RNN



low dimensional trajectories

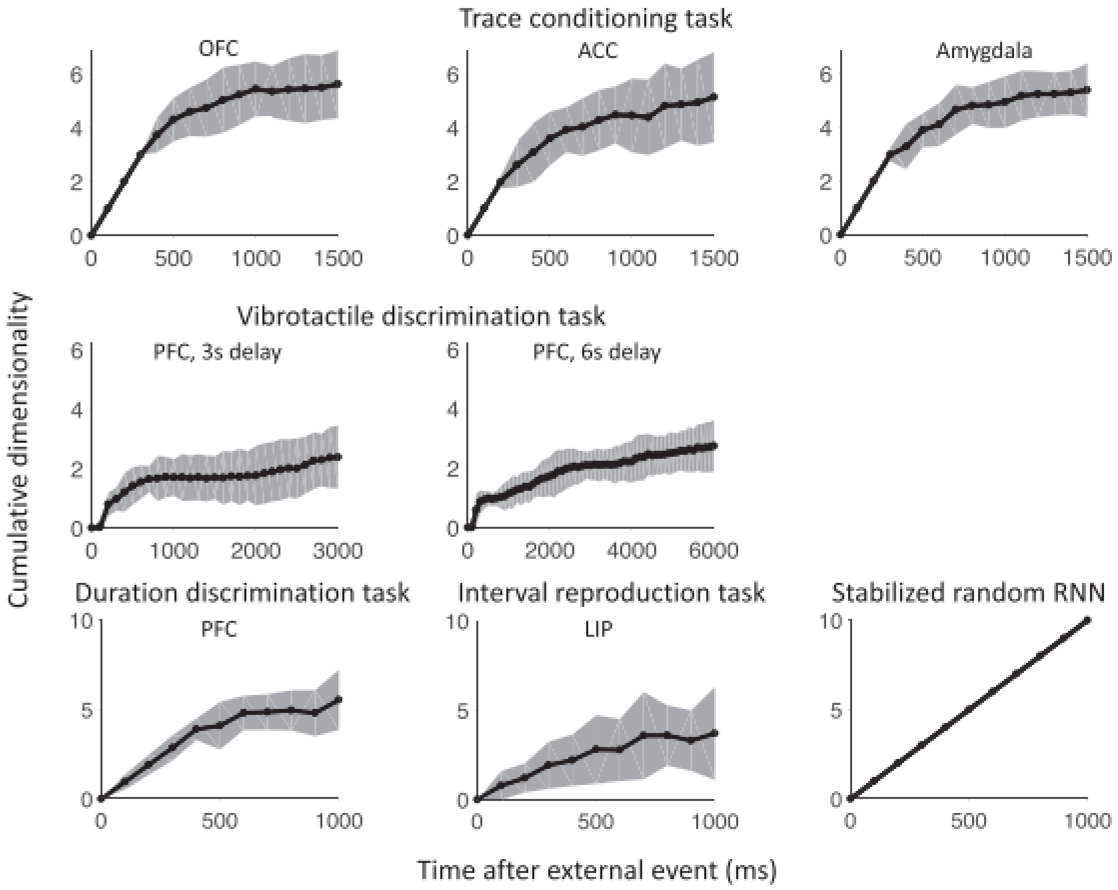


# 3. Experiment



# 3. Experiment

## Experiment




## RNN

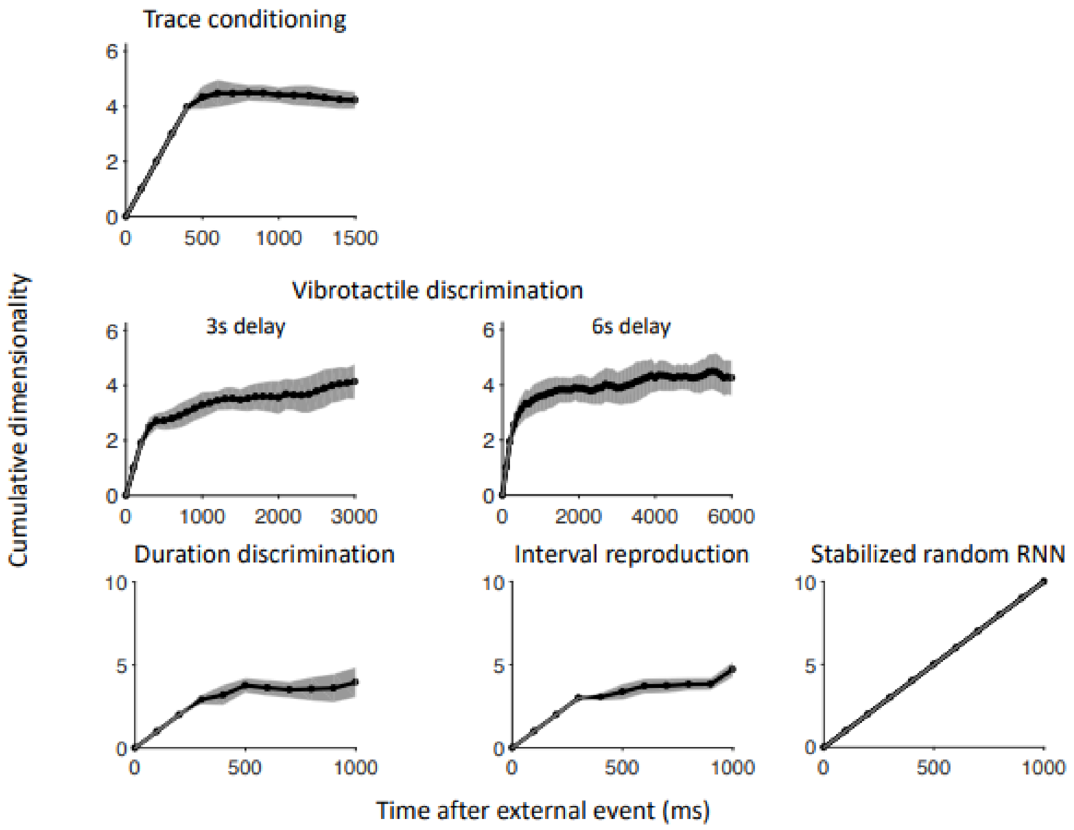
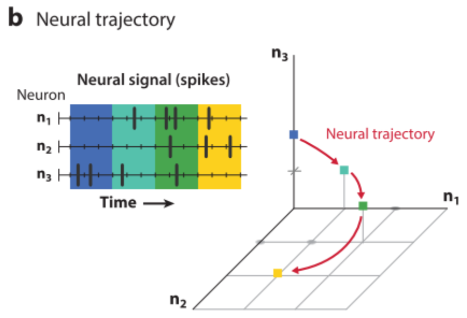


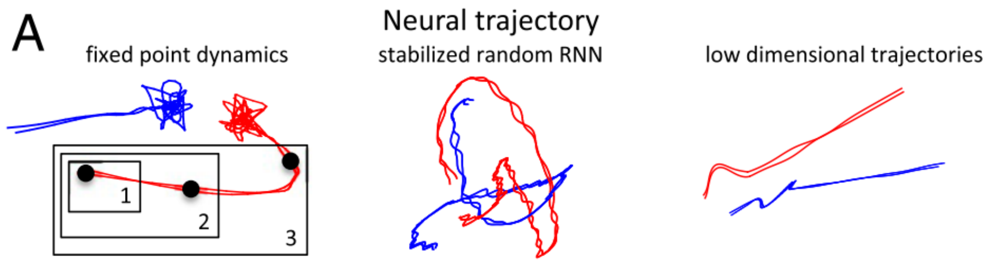

**Fig. S14.** Cumulative dimensionality for the RNN models. Error bars show one standard deviation.

# 4. Conclusion

1. Concepts of dynamics




2. Three hypotheses of the working memory




3. Experiments