

Persistent versus fading memory in serial recall: An analysis through recurrent neural networks

Lucas Stordeur

Pierre Sacré, Steve Majerus, Guillaume Drion



Recurrent neural networks use a different representation from other models

Most models

- Separate representation: **Item vs Order**
- Support from human data

Recurrent Neural Networks

- Unified representation
- Support from more recent monkey data

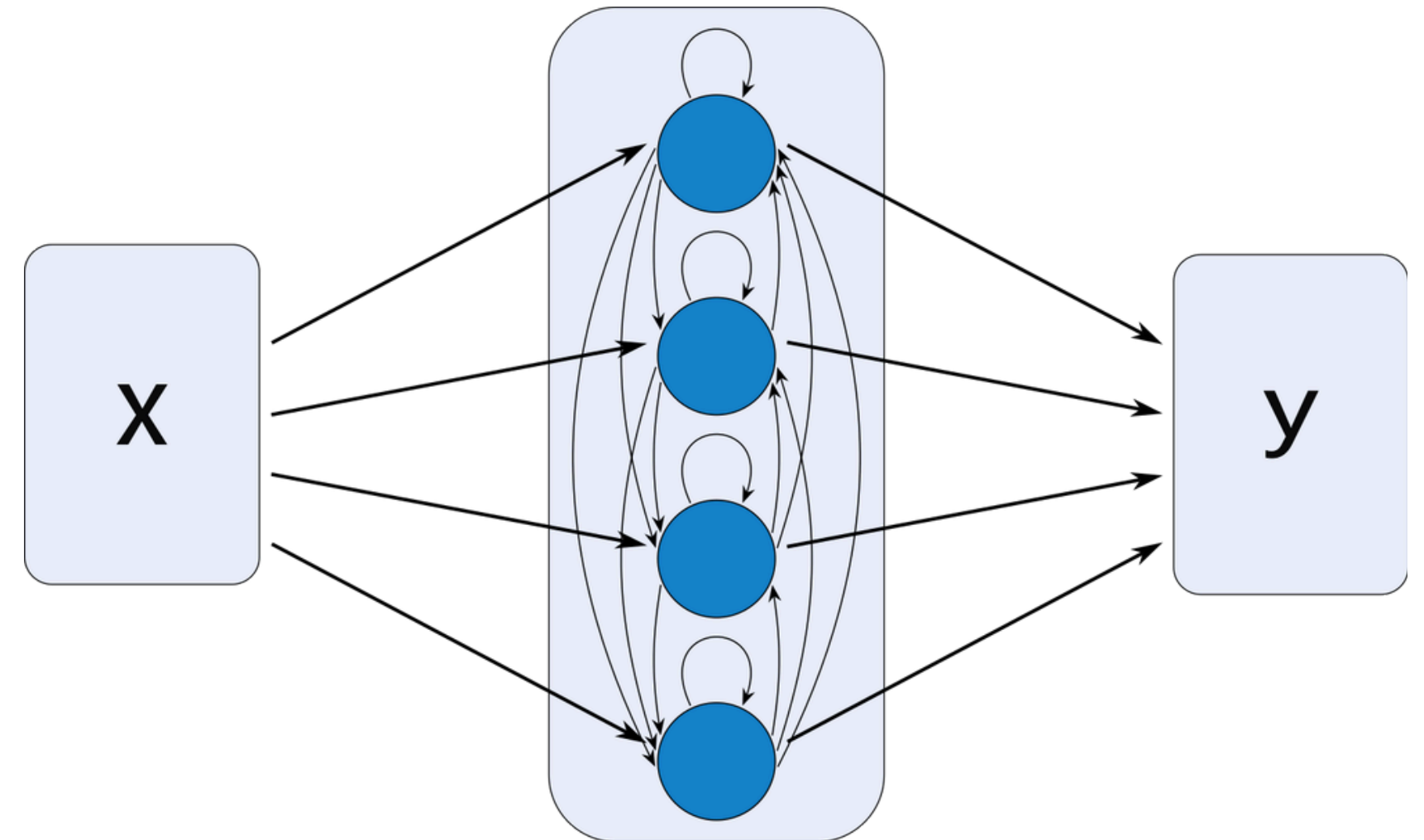
The new cell dynamics encode the information differently

$$\mathbf{h}_t = \phi(\mathbf{x}_t, \mathbf{h}_{t-1}, \theta_t)$$

$$\mathbf{y}_t = \psi(\mathbf{h}_t, \theta_t)$$

There exists an old version of the model

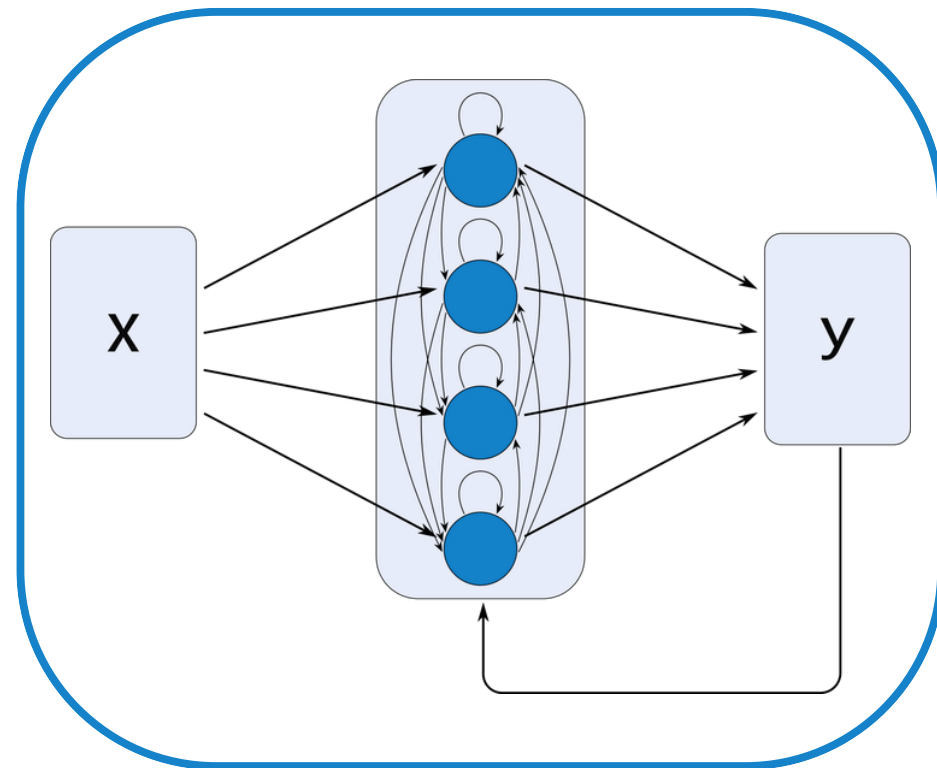
There are new cell dynamics



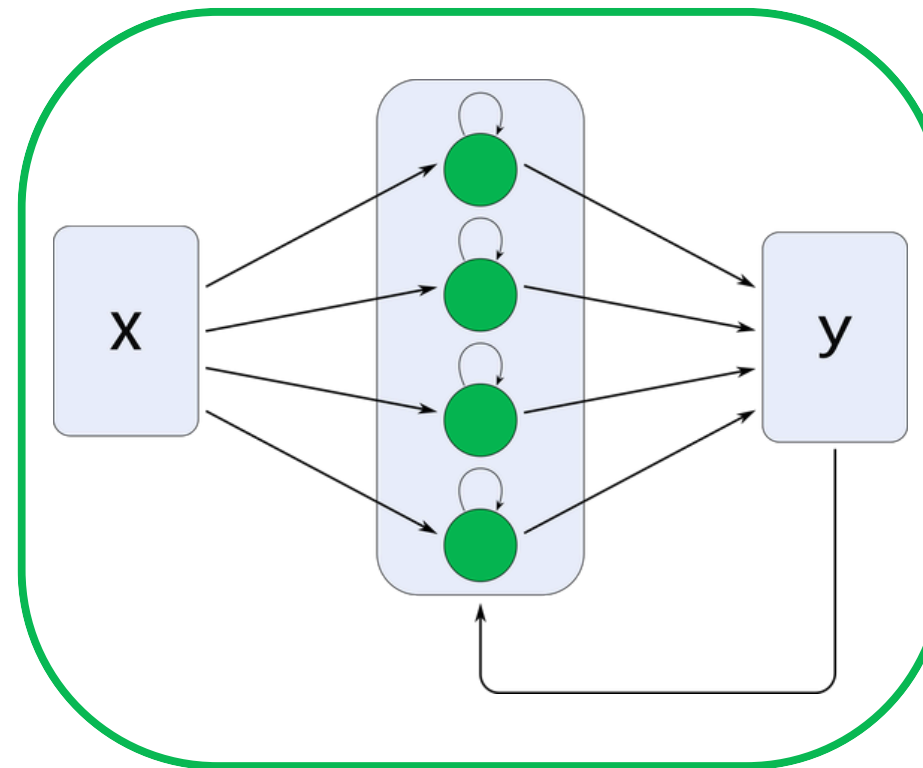
**No one study the impact of these dynamics
on the model**

Cell dynamics shape which serial recall effects a RNN can reproduce

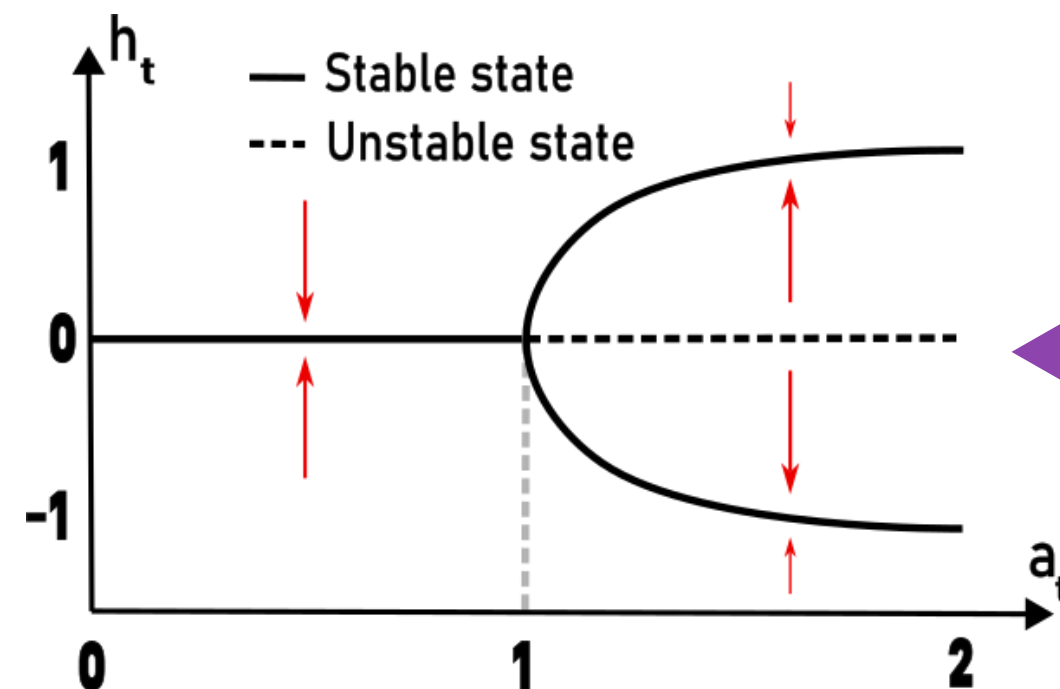
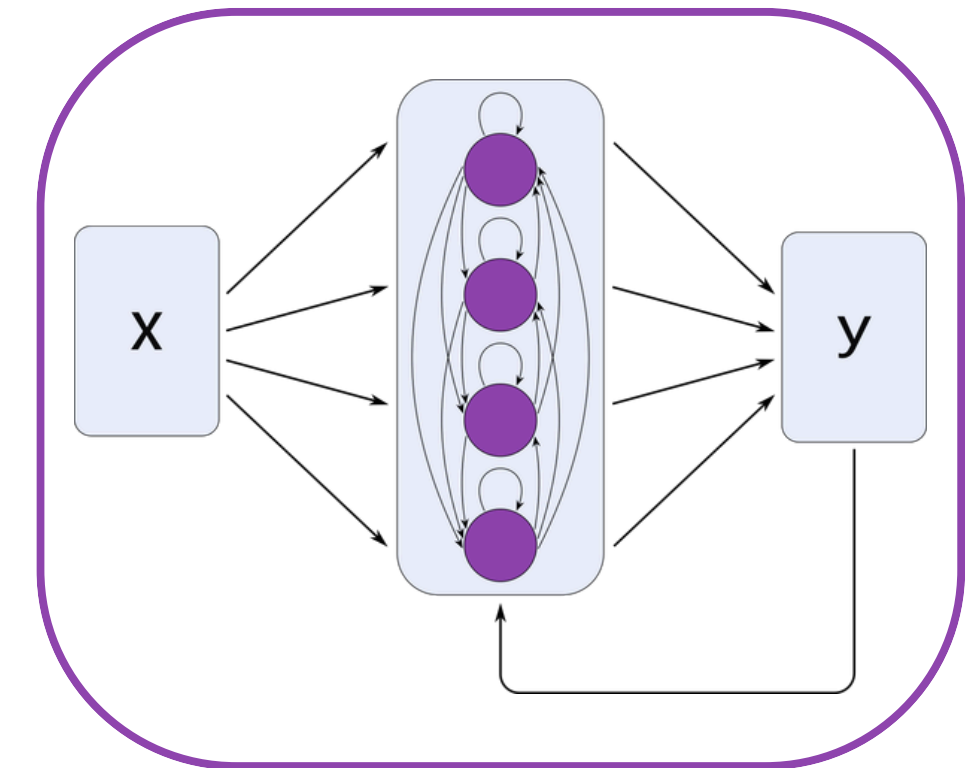
Gated recurrent unit (GRU)



Bistable recurrent cell (BRC)



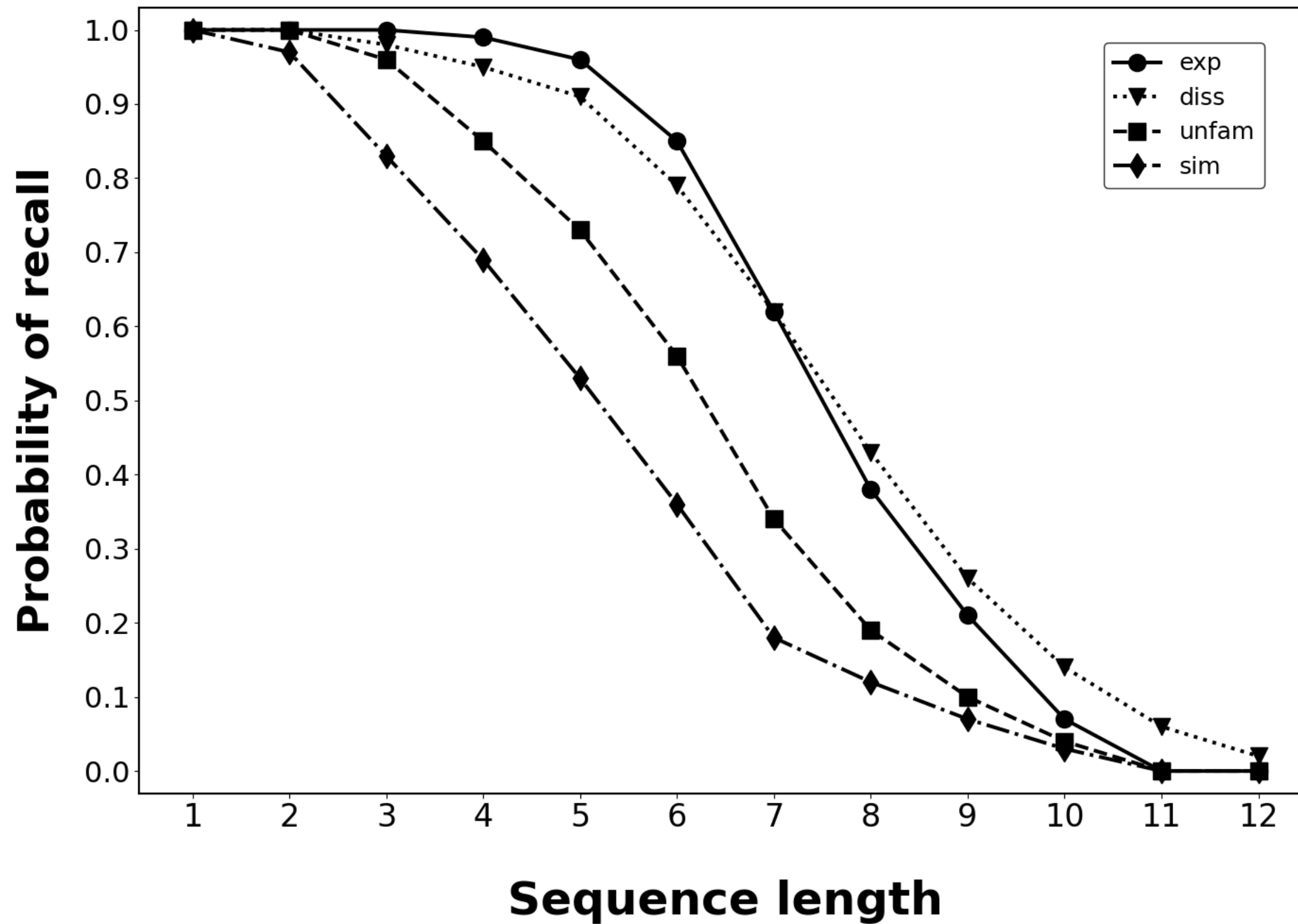
Neuromodulated BRC (nBRC)



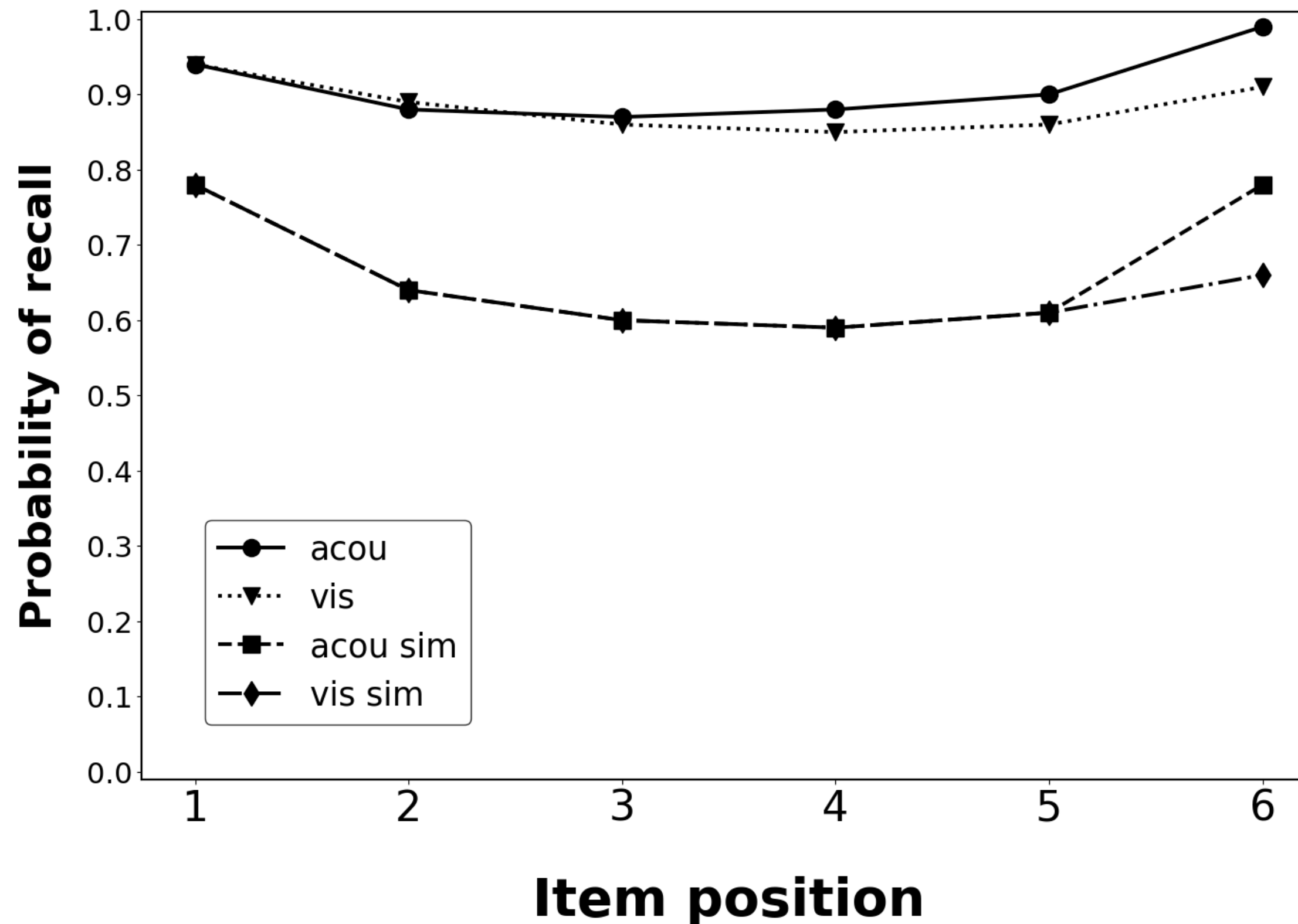
Outline

1. Serial recall and a RNN model of the task
2. Cell dynamics reproduce different effects
3. Combining the dynamics captures the full range of effects

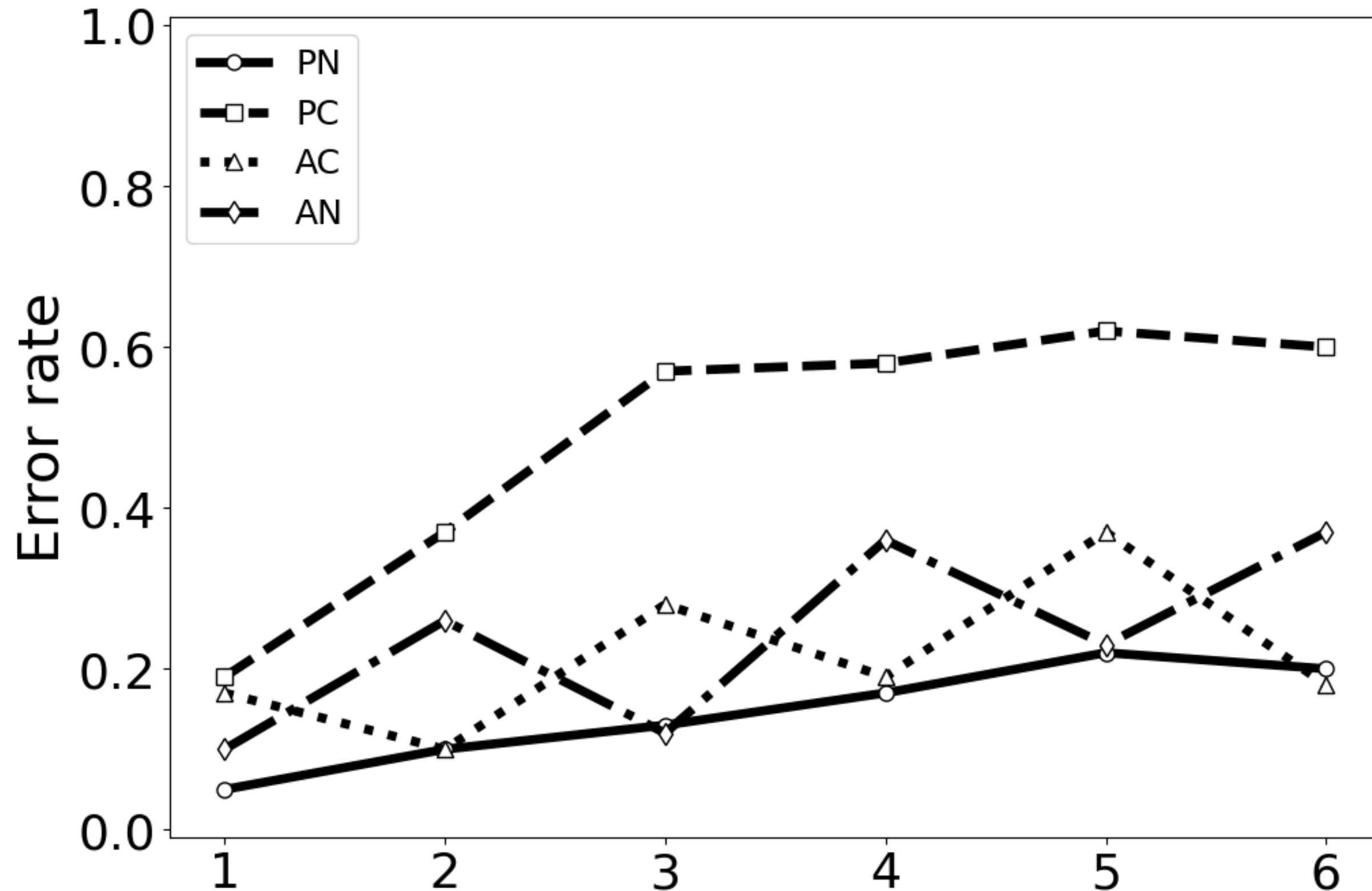
Recall probability drops sharply beyond 7 items



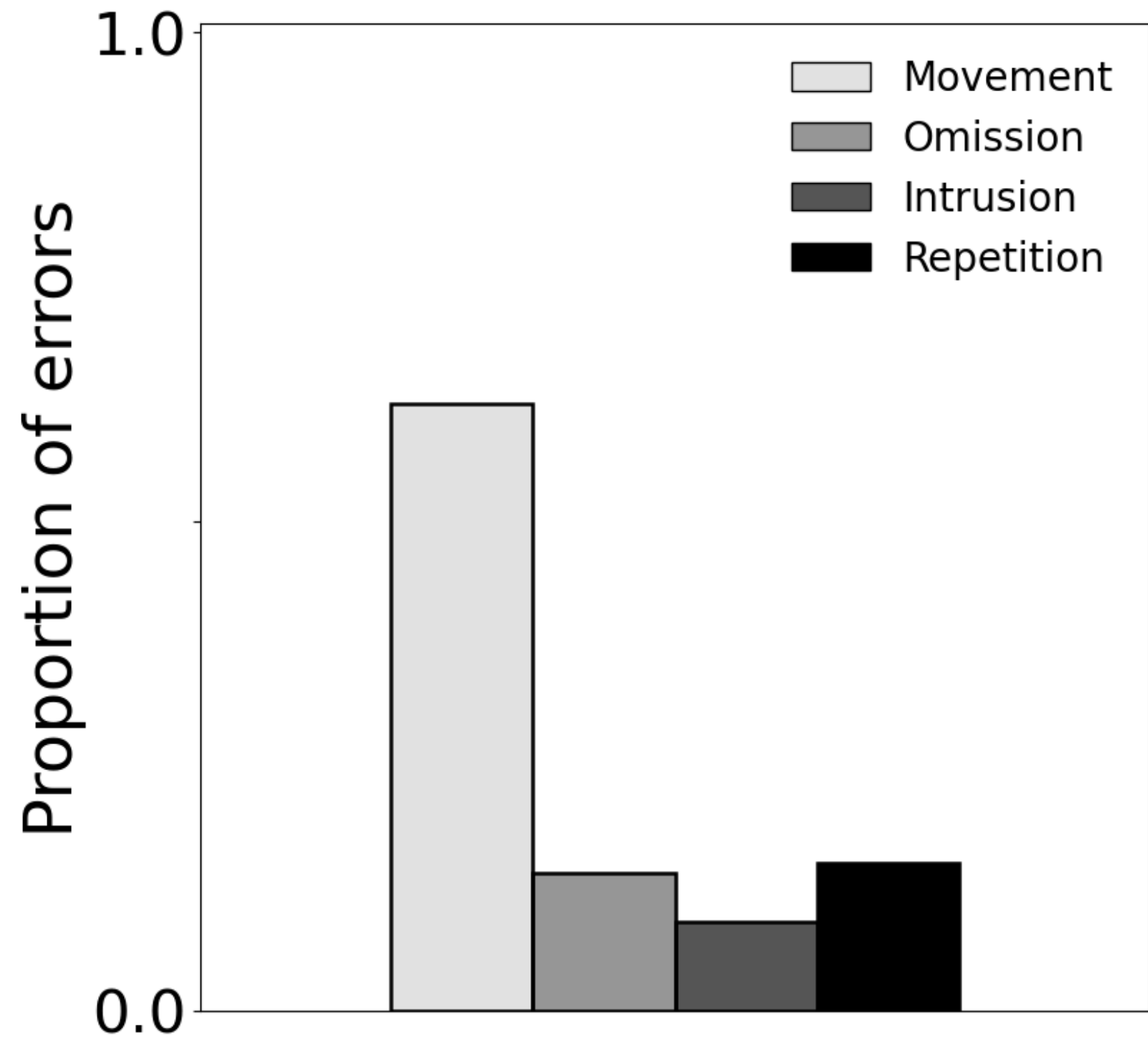
Items at the beginning and end of the sequence are better recalled



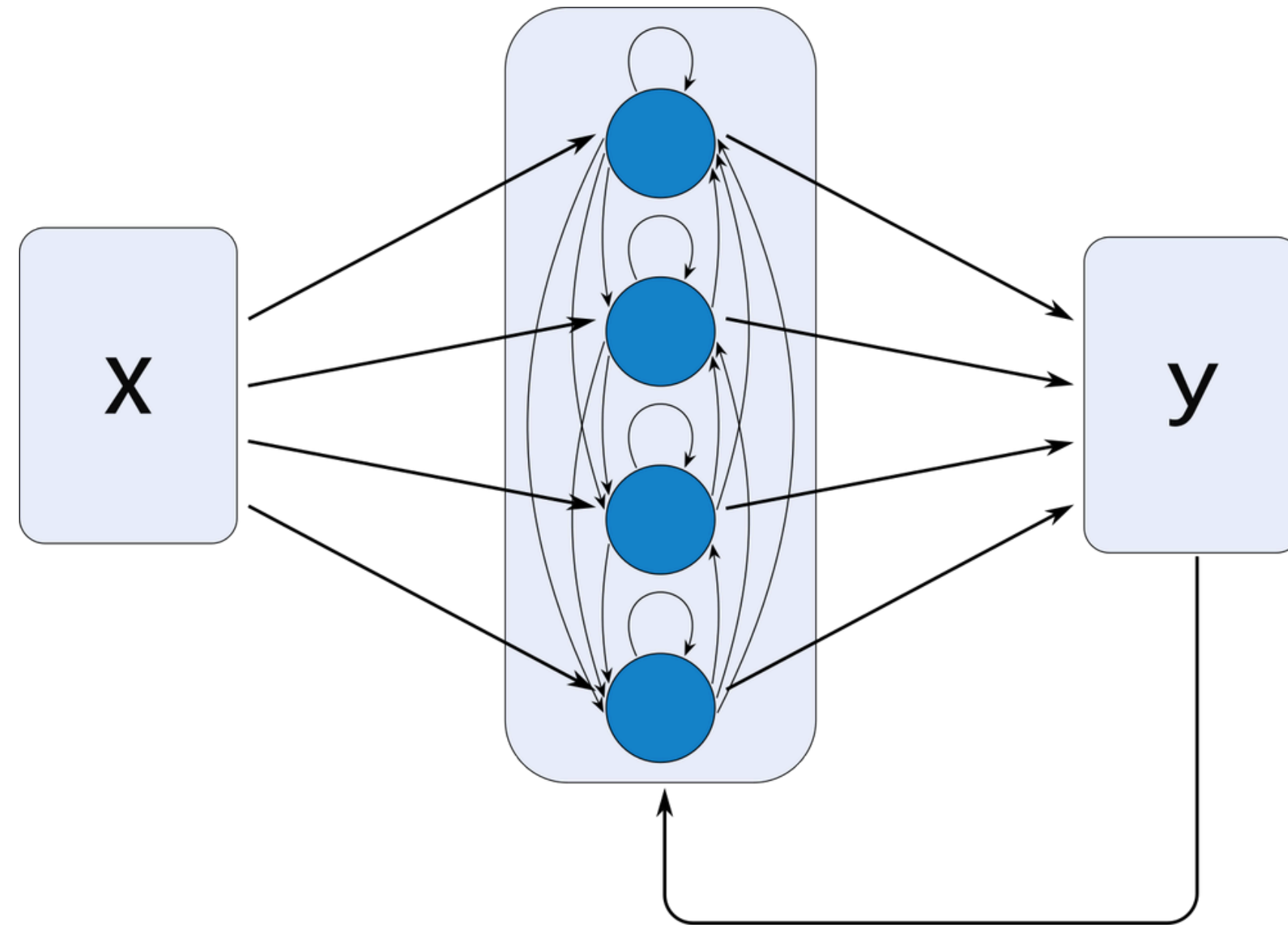
Recall of non-confusable items are not affected by the confusable items



Various types of errors can be made

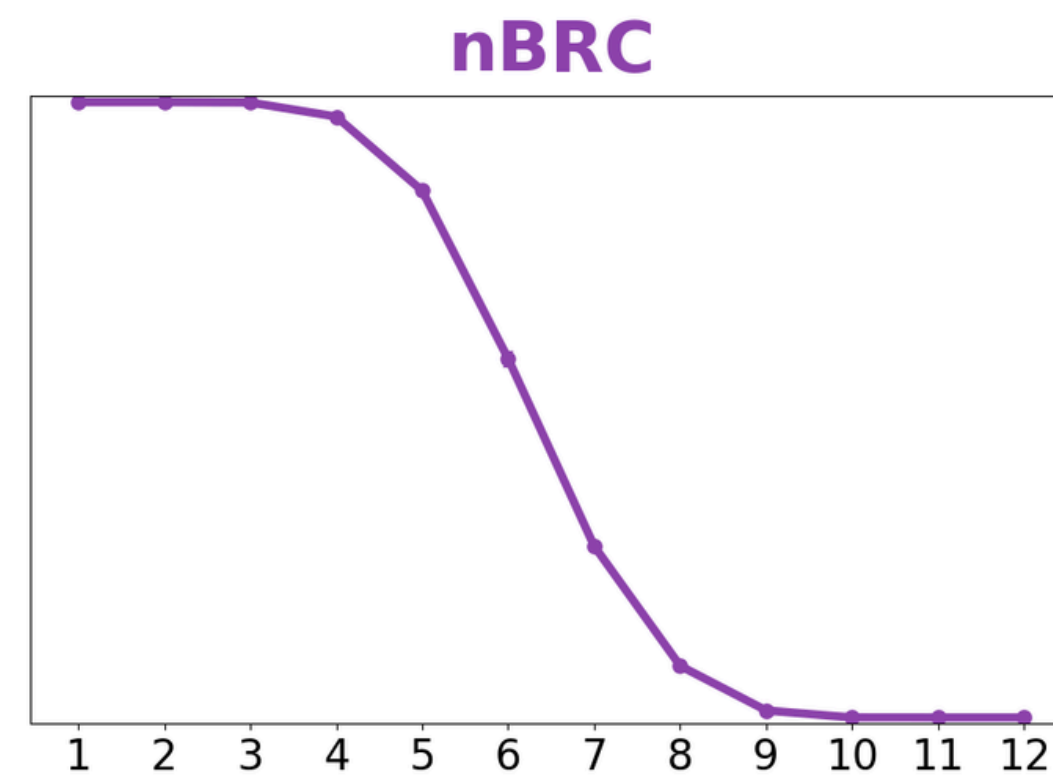
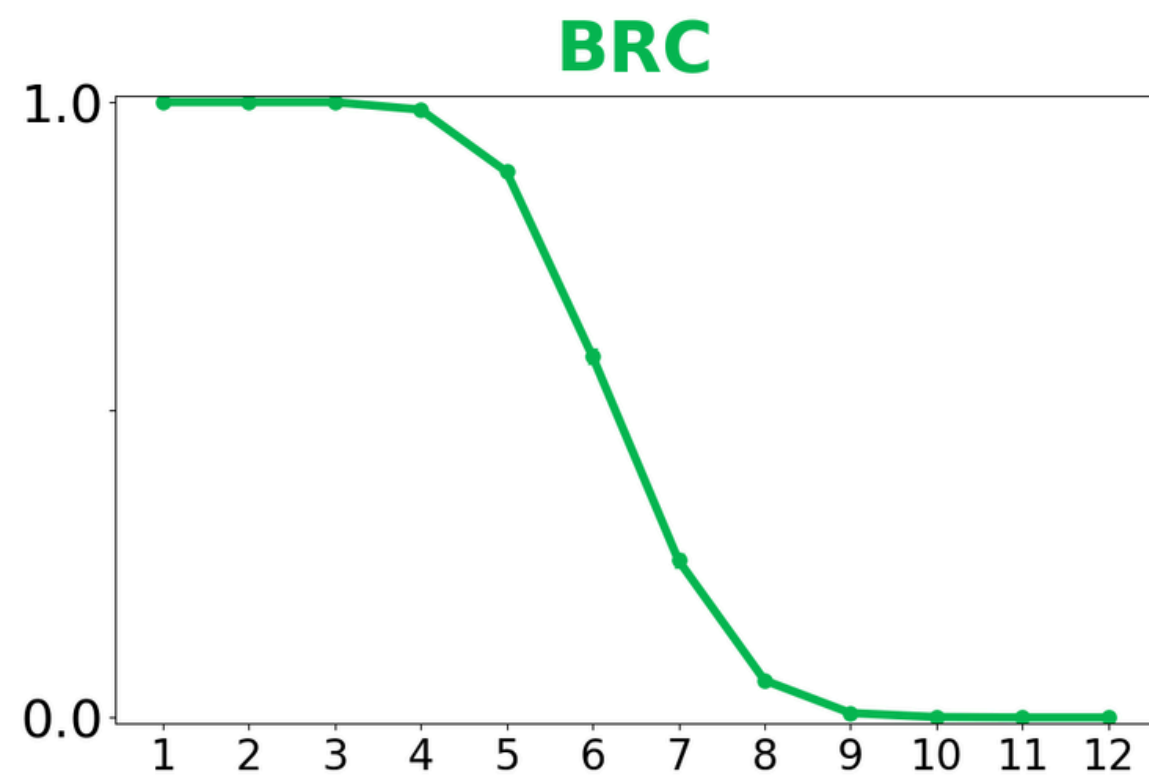
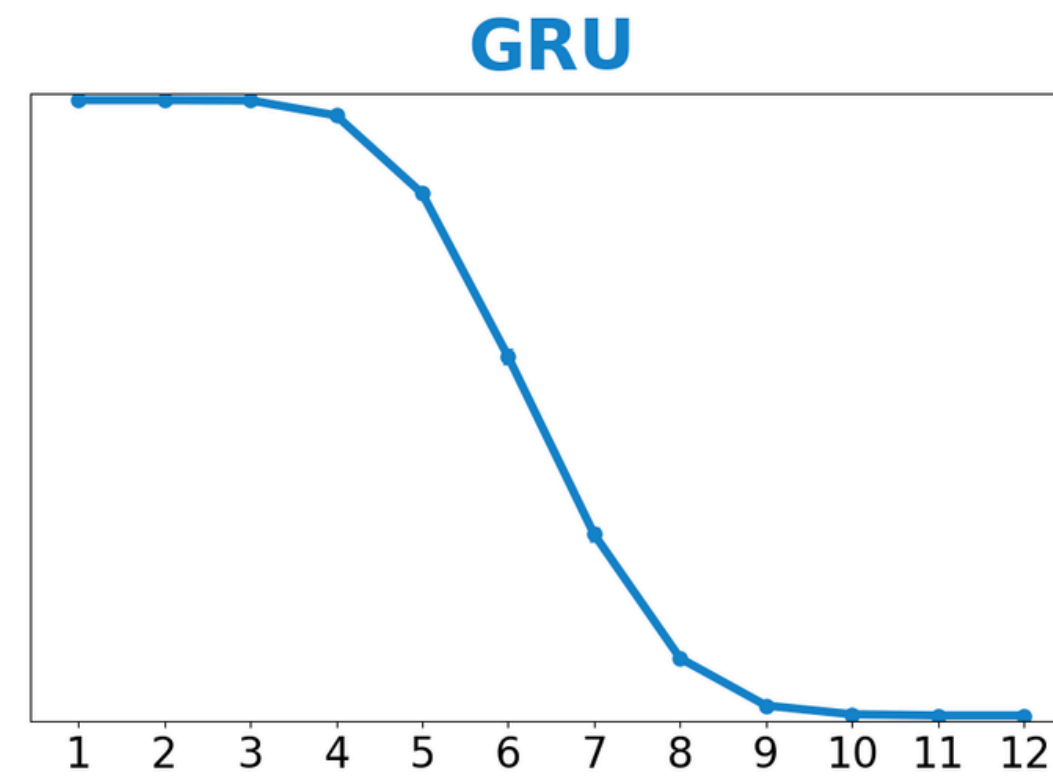
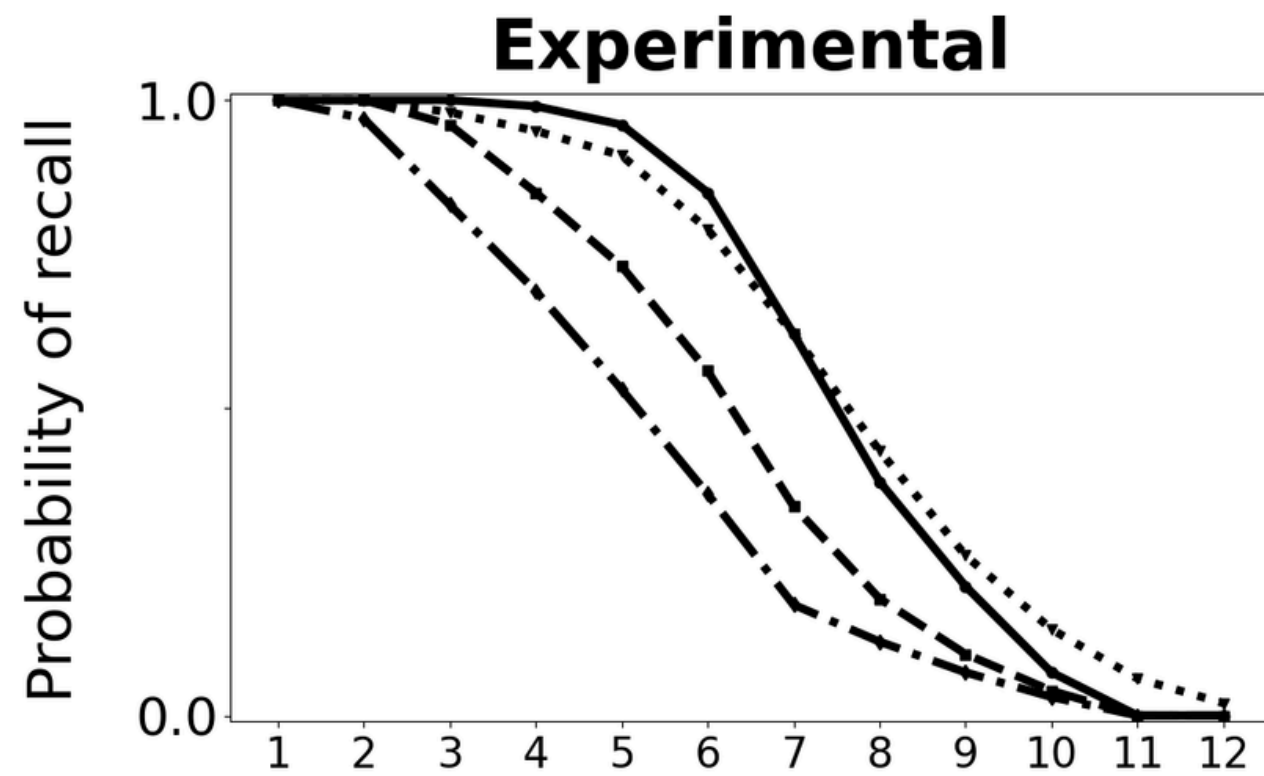


**The goal of this model is not to solve the task
as best as possible**

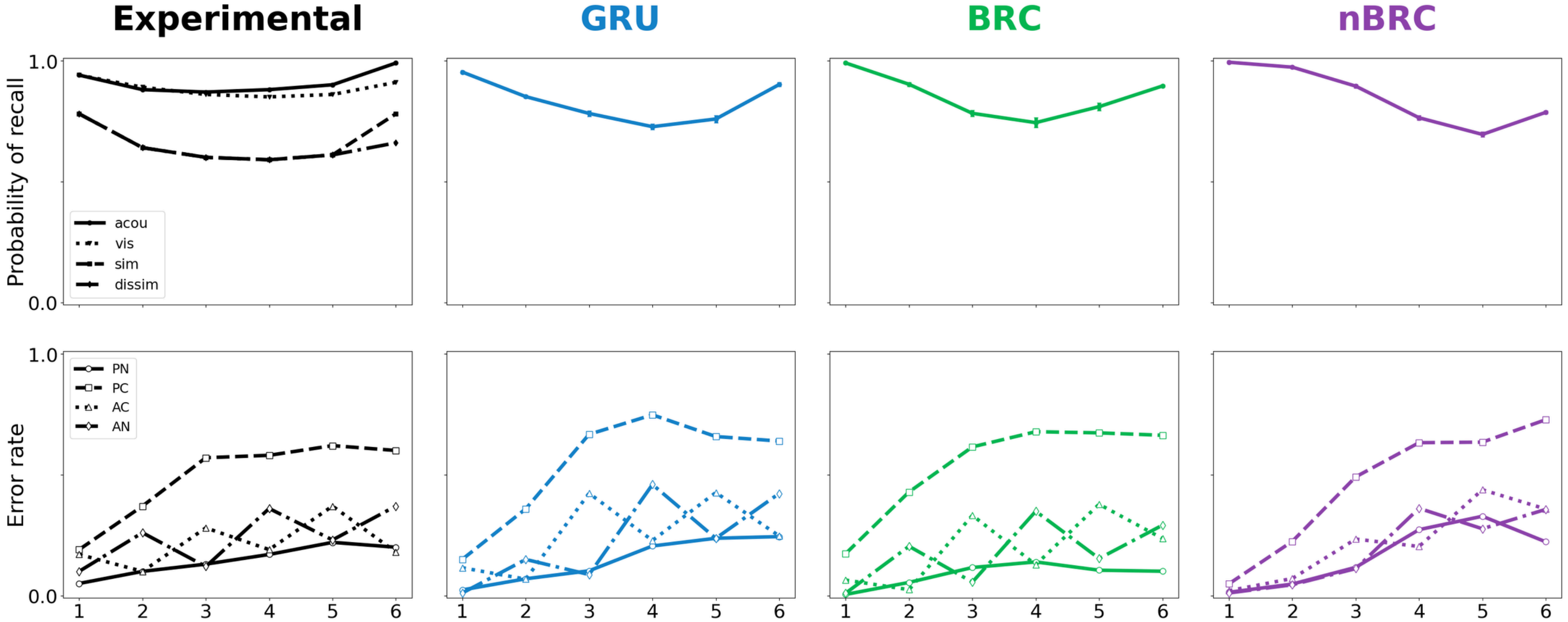


$$\mathbf{h}_t = \phi(\mathbf{x}_t, \mathbf{h}_{t-1}, \mathbf{y}_{t-1}, \theta_t)$$

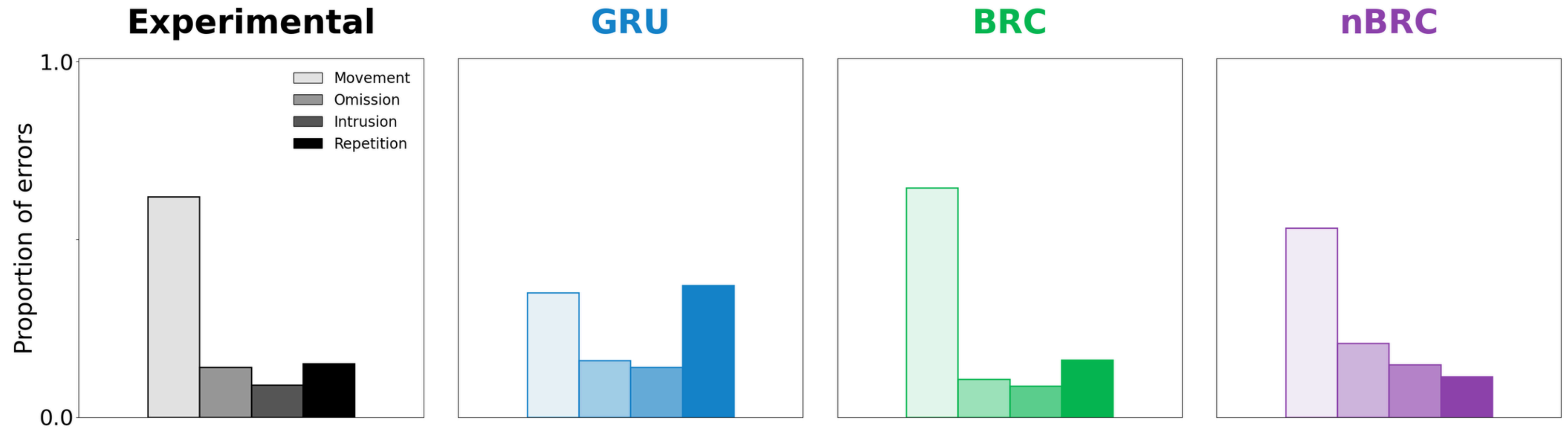
All cells seem to have similar performance on the different sequence lengths



BRC network cannot reproduce confusability



GRU network fails to reproduce the proportion of errors



None cell dynamic can represent all experimental results

	GRU	BRC	nBRC
<i>Decrease of performance with sequence length</i>			
<i>Serial position recall curve</i>			
<i>Transposition gradients</i>			
<i>Confusability</i>			
<i>Proportion of errors</i>			

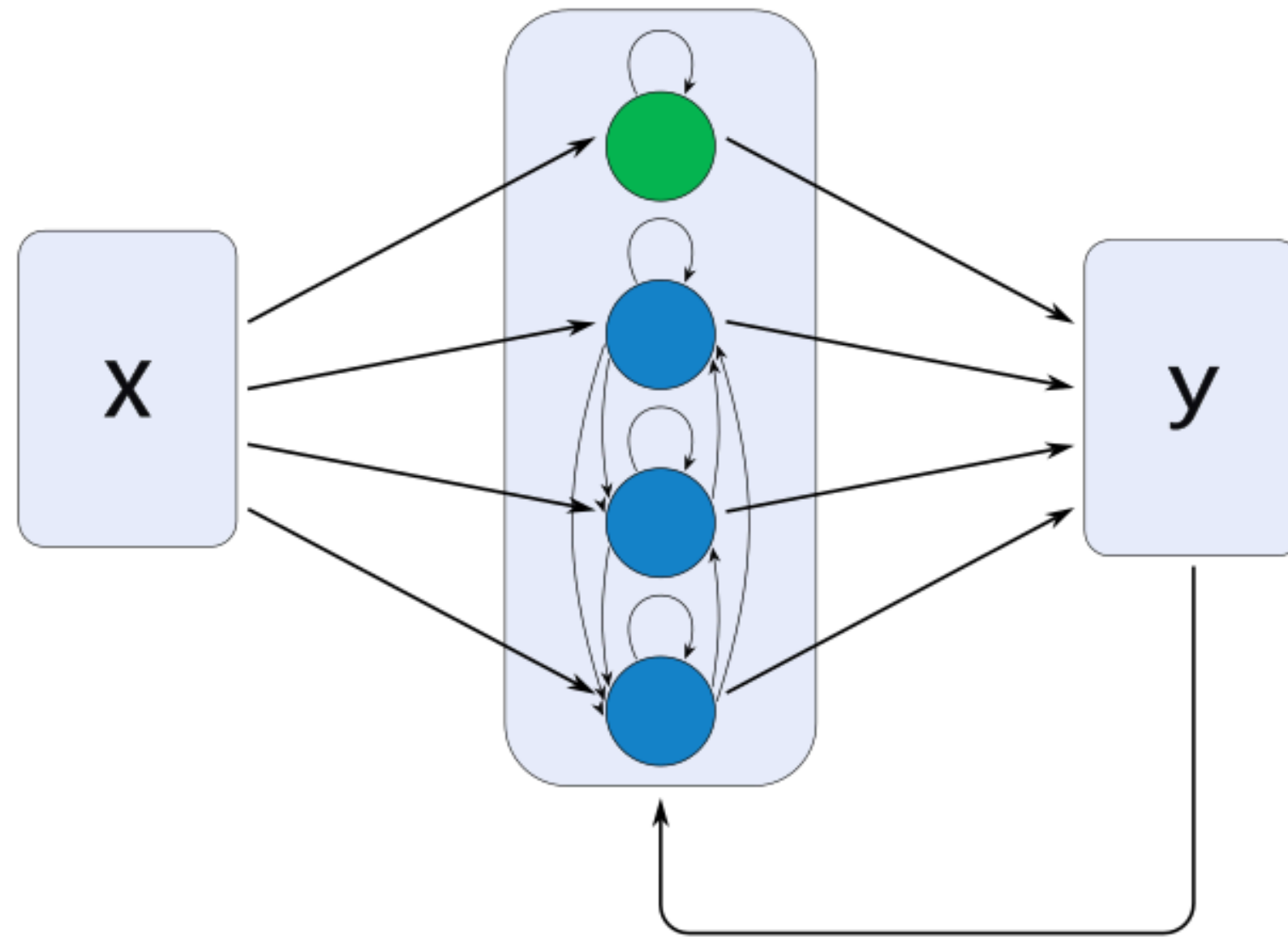
The GRU and BRC encode the information differently

GRU cells: Encode the dynamic of the task

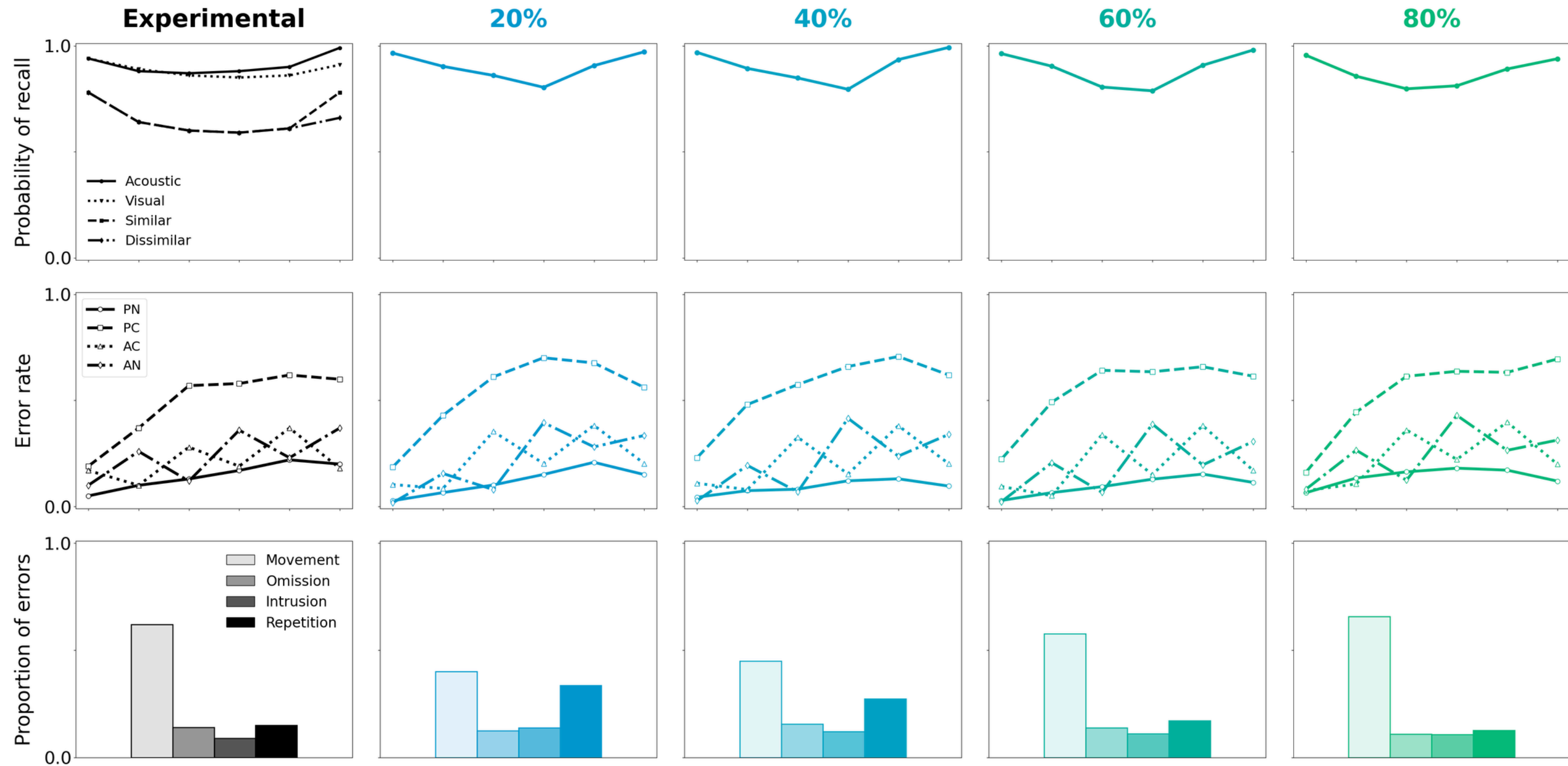
BRC cells: Encode the items

Each encoding method better capture certain effects of the task

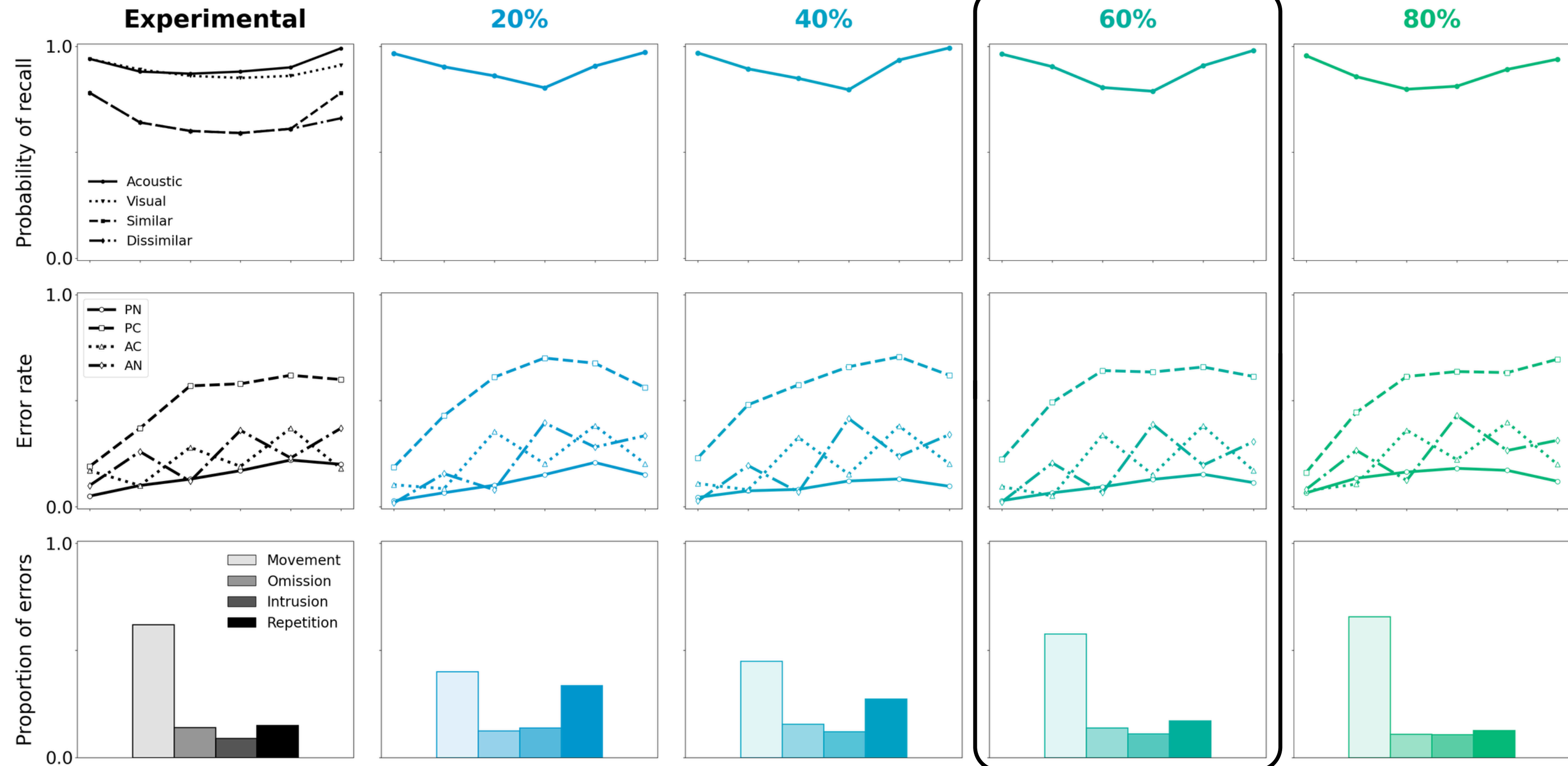
BRC and GRU can be combined to take the advantages of both cells



BRC and GRU can be combined to take the advantages of both cells



BRC and GRU can be combined to take the advantages of both cells



Conclusion

Recurrent neural networks are a valid model to represent serial recall

New cell dynamics can represent different effects

Combining both types of encoding allows to represent the different effects

**Now we know a bit more about how you just did it...
or at least how a RNN would do it...**

