1. Summary
We propose that neocortical columns learn maps of objects, similar to how entorhinal cortex and hippocampus learn maps of environments. Each column associates sensory inputs with the sense organ's location and orientation relative to the sensed feature. Columns can then arrange these features into novel objects.

2. Cortical columns model objects as entorhinal cortex models environments.
Entorhinal cortex represents locations and orientations relative to specific environments via grid and head direction cells. Each column associates sensory inputs with the sense organ's location and orientation relative to specific objects.

3a. L4 and L6a learn the sensory input at each sensor location.

3b. Through movement and sensation, ambiguity is eliminated.

4. L5thick and L6b arrange previously learned features into novel objects.
The cortical column builds a complete model of an object by recognizing each of its component features.

Predictions of the theory
- While an animal attends to an object, cells in the deep layers of sensory cortex have grid cell and head direction cell firing properties. Instead of anchoring to the environment, these cells anchor to the object.
- L5thick builds stable representations of objects that are invariant to pose. L6b builds representations that are invariant to the sense organ's location and orientation relative to the object.
- L5thick quickly builds these stable representations for novel objects if the novel objects consist of familiar features.
- L6a cell activity oscillates between being driven by movement cues and by sensory cues.

Current research
- How are objects and locations passed up / down / sideways in the hierarchy? How does this theory change the way we should think about the hierarchy?
- How do neurons do more advanced "transforms" between features and objects? (Not just translation, but rotation.)
- L5thick represents objects as arrays of transforms. L6b cell activity represents objects as arrays of transforms. L6b cell activity represents objects as arrays of transforms via autoassociation in L5 and via the L3->L5 projection. What's the detailed mechanism?

References