From Egocentric to Object-Centered Reference Frames: Grounding Visuo-Spatial Cognition in the World

https://escholarship.org/uc/item/7cd6v67q


1069-7977

Simmering, Vanessa R.
Spencer, John P.
Schoner, Gregor

2007

Peer reviewed
From Egocentric to Object-Centered Reference Frames: 
Grounding Visuo-Spatial Cognition in the World 

Vanessa R. Simmering (vanessa-simmering@uiowa.edu) 
John P. Spencer (john-spencer@uiowa.edu) 
Department of Psychology, University of Iowa 
Iowa City, IA 52242 USA 

Gregor Schöner (gregor.schoener@neuroinformatik.ruhr-uni-bochum.de) 
Institut für Neuroinformatik, Ruhr-University 
44780 Bochum, Germany 

Keywords: spatial cognition; visual reference frames; dynamic field model 

Research in spatial cognition has established that people anchor memories of object locations in a local task space to an object-centered reference frame. Although a number of models of spatial memory capture systematic biases near such reference frames, no previous models have captured the processes that transform spatial information from an initially egocentric (e.g., retinal) frame into an object-centered frame. This paper focuses on one model of spatial cognition that has shown promise in this regard, the Dynamic Field Theory (DFT; see Spencer, Simmering, Schutte, & Schöner, 2007, for review). Our previous work offered an initial account of how reference frames are established and maintained. Here we provide a more complete account that is fully grounded in neural principles. 

The DFT is a process-based theory of spatial cognition implemented in a neural network model (for a complete description, see Spencer et al., 2007). The full model consists of seven layers of spatially-tuned neurons. Here we focus on interactions of the top four layers: a retinal or egocentric perceptual field (PF_ego), a field that translates activation to an object-centered frame of reference (Shift), an object-centered perceptual field (PF_obj), and a long-term memory field associated with PF_obj (LTM_PF_obj). 

To understand how the model aligns reference frames, consider a typical spatial recall task (e.g., Spencer & Hund, 2002). The participant sits at a large, homogeneous tabletop, and a target is presented at the to-be-remembered location. Adults’ performance suggests that the edges and midline symmetry axis of the table serve as reference frames for remembering target locations. In the model, we present these inputs as Gaussian activation profiles corresponding to the relevant spatial locations. These inputs are presented to PF_ego within a retinal reference frame; therefore, movements of the head and eyes will change where these inputs are in the field. To use these inputs to remember a location relative to the table, information must be transformed from this retinal frame into an object-centered reference frame. 

This transformation occurs in the Shift field, where activation in PF_ego is effectively compared to activation in PF_obj to find the shift value with the highest overlap in activation (i.e., greatest spatial correlation). For example, in the recall task described above, the edges and symmetry axis of the table provide constant inputs at three locations. These inputs can be initially mapped into the object-centered frame at arbitrary values. When the input to the retina shifts (e.g., due to an eye movement), however, the input to the object-centered frame must be shifted accordingly so the three peaks at new locations in PF_ego are correctly mapped to the three peaks at the original locations in PF_obj. To achieve this transformation, we use a two-dimensional Shift field that maps input from PF_ego onto all possible shifts of PF_obj (which we achieve by pre-structuring the weight matrix in the 2D Shift field). The best match in the Shift field always occurs along a particular diagonal axis. Moreover, the amount the selected diagonal axis is shifted relative to the “no shift” axis (i.e., the diagonal axis that maps input one-to-one from the retinal to the object-centered frame) specifies the shift amount. 

The model can also re-align reference frames after longer delays, by using LTM_PF_obj. In this case, the peaks in PF_obj are built from activation sent from LTM_PF_obj, rather than being established online in the task. In this case, the alignment process functions exactly the same; the only difference is the source of the initial activation in PF_obj. This feature of the model allows for previous experience with a given object-centered frame of reference to influence behavior over multiple encounters, rather than only within a given session. 

Acknowledgments 
This work was funded by NIMH RO1 MH62480, NSF BCS 00-91757, and NSF HSD 0527698, awarded to JPS. 

References 