

Action, Detection, and Perception: A Computational Model of the Relation Between Movement and Orientation Selectivity in the Cerebral Cortex

Anthony F. Morse (anthony.morse@his.se)

Tom Ziemke (tom.ziemke@his.se)

COIN Lab, Informatics Research Center, University of Skövde, Sweden

Abstract

A fundamental tenet of enactive theories of cognition states that action is a necessary prerequisite to perception. In this paper we review the basis for this assumption and, with the help of a computational model of the famous Held and Hein kitten experiments, challenge the necessity of movement in subsequent detection. In normal development action does play an important role in setting up detection, but we aim here to widen our conceptions and consider the effect of correlations between non-motoric events.

Keywords: Action; detection; perception; enaction; actionism; embodiment; computational modeling; cortical hierarchy.

Introduction

The unit of analysis that many cognitive scientists consider relevant to our understanding of cognition has shifted from the view of cognition as purely internal computation that, at least to some degree, can be reduced to mapping sensory input to motor output, to the view of cognition as situated and embodied action that spans brain, body and environment (Clark, 1997; Clark & Chalmers, 1998; Hutchins, 1995; Suchman, 1987; Varela, Thompson, & Rosch, 1991). Thus embodiment calls our attention to the relationship between an agent or organism and its environment. One prominent example of this is the focus on sensorimotor knowledge in the enactive approach to perception (Noë, 2004; O'Regan & Noë, 2001; Thompson, 2007; Varela et al., 1991) in which perception is thought to be both dependent upon, and constituted by our possession of sensorimotor knowledge, i.e. "practical knowledge of the ways movement gives rise to changes in stimulation." (Noë, 2004).

Sensorimotor knowledge is not simply factual knowledge about a domain but is intimately about the relationship between an agent, its environment, and objects therein. Regularities in this relationship, such as the sensory consequences of actions in context, provide a grounded path to the discovery of affordances and through them to knowledge of the world around us (Gibson, 1979; Morse & Ziemke, 2007). Gallese and Lakoff (2005), though not explicitly enactive, propose a similar theory of conceptual knowledge underpinned by extensive neuroscientific evidence of the association between, and reactivation of, sensorimotor areas during cognitive tasks. In a similar vein, O'Regan and colleagues continue to produce experimental evidence of the role of sensorimotor dependencies in understanding space and colour (Philipona & O'Regan, 2006; Philipona, O'Regan, & Nadal, 2003).

Clearly then there is a wealth of evidence both theoretical and experimental originating from philosophy, psychology, and neuroscience, all highlighting the immensely important role that action plays in perception. Such a relationship is quite rightly positioned at the fore in enactive attempts to understand the normal development of cognition, conceptual knowledge, and perception. One problem, as we see it, is that action has come to dominate our understanding of how perception is brought fourth to the exclusion of other possible routes. By modelling the role of action in detection and highlighting possible mechanisms we challenge the claim that action and movement is necessary for, and by implication the only way in which, perception can arise. While such a view is not universally held, and we hope that to most our position seems reasonable, stronger claims of the necessity of action have been made, Noë for example, claims that "what is ruled out is the possibility of someone who lacked all sensorimotor comprehension having experiences with spatial content (or for that matter, any content)." (2004, p. 91).

That action is under normal development the dominant means of exploring and discovering a distal world is not surprising. For many the paradigmatic human sense is vision, which by its very nature is inherently spatial and therefore requires motor exploration of the environment in-order to discover its spatial content. According to Noë, "the claim is that by sampling the way appearances change as you move through this appearance space, we encounter the invariants." (2004, p. 86). As Noë argues, vision is misleading as we are unaware of just how active a process it is. Instead he suggests that a more appropriate candidate, at least for helping us to understand the mechanisms underlying perception, is touch, a more obviously motoric modality than vision. If, as many have argued, the role of action is in exposing correlations or 'encountering the invariants' then surely any other means to encounter the invariants could equally lead to perception and should be investigated as such. We further suggest that such investigations have the potential to lead to stronger accounts of non-spatial concepts and aspects of cognition.

In the remainder of this paper we first briefly review a body of work with animals (kittens) demonstrating the calibrating role of proprioception. We then introduce a robotic model of those experiments, the results of which, while fully supporting the role of action in the development of detection capacities, equally highlight and demonstrate that a similar role can be played by correlations existing between other non-motoric modalities such as passive vision and touch. We conclude with a discussion of the

dominance of spatial aspects of thought and cognition and the potential implications of acknowledging another route to perception.

Background: “Blind” Kittens

A body of work often referred to in the enactive cognitive science literature originates with the experiments of Held and Hein (1967) in which kittens raised in the dark, and unable to control their own movements during exposure to light, exhibit severe deficiencies in visually guided behavior. The experiments involved two groups of kittens both raised in the dark. However, before exposure to light, kittens from group A were placed (individually) in a gondola and held there by a neck yoke and body clamp. Kittens from group B were also placed in a neck yoke and body clamp but not a gondola and so were able to control their own movements. The two kittens were connected via the body clamps to a mechanism and pivot such that the gross movements of the kitten able to control its movement were transferred to the other kitten (see Figure 1 below).

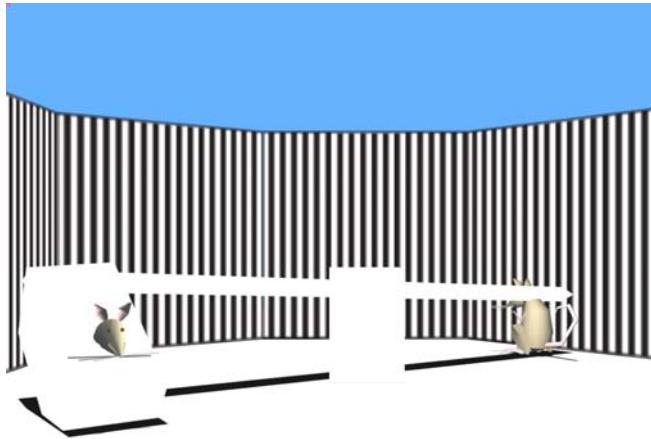


Figure 1: Replication of the Held and Hein kitten experiments using the ICEAsim rat like robots. While one kitten or robot is able to move freely, the other kitten or robot is restrained and the gross movements of the first kitten or robot are transferred to the second.

Following such limited exposure to light (and removed from the apparatus) kittens that were able to control their movements (group B) were able to make visually guided paw placements and could also avoid a visual cliff. The kittens from the gondola (group A) however were unable to perform either task and behaved as if blind. In subsequent unrestrained exposure to light, all the kittens developed normally. For Held and Hein this confirmed their thesis that “self-produced movement with its concurrent visual feedback is necessary for the development of visually-guided behavior.” (1967, p. 875). They further consider and dismiss the possibility that these results were due to either anatomical / physiological deterioration (as both groups were free to behave normally in the dark) or behavioral

inhibition (as no signs of shock, excitement, or fright were observed).

The important aspect of this experiment is to note that purely afferent visual exposure is insufficient for the constitution of visually guided behavior (Gapenne, In Press). This result should hardly be surprising as it requires the coordination of visual and behavioral modalities that the kitten has never experienced in any coordinated way before. However this result has been further explored in studies by Buisseret and Imbert (1976). In this work kittens were similarly raised in the dark for 6 weeks before a 6 hour exposure to light. Extracellular recordings were taken from the visual cortex (visual area 17) and analyzed according to cellular selectivity for the orientation of a visual stimulus. Their results showed that the selectivity of cells to the orientation of a visual stimulus was absent both in kittens that have never been exposed to light and in kittens deprived of movement (including ocular movements) during exposure to light. In contrast, kittens able to freely move during exposure to light did develop orientation sensitivity. Finally, if movement is limited to one plane, then sensitivity develops to features orthogonal to that plane (for a detailed review of this an other related work see (Gapenne, In Press).

Experimental Setup and Model Design

In replicating the experiment of Buisseret and Imbert (1976) we use the Webots based ICEAsim, a simulated rat-like robot in a 3D simulation environment (developed in the ICEA project, www.iceaproject.eu). The robot provides visual input from two cameras and tactile information from a set of six movable whiskers and has a further 12 degrees of freedom. The robot is placed in a simple simulated environment consisting of a single round room with a repeating pattern on all walls. This pattern consisted of parallel black and white stripes which could be rotated to any angle (see Figure 2 below).

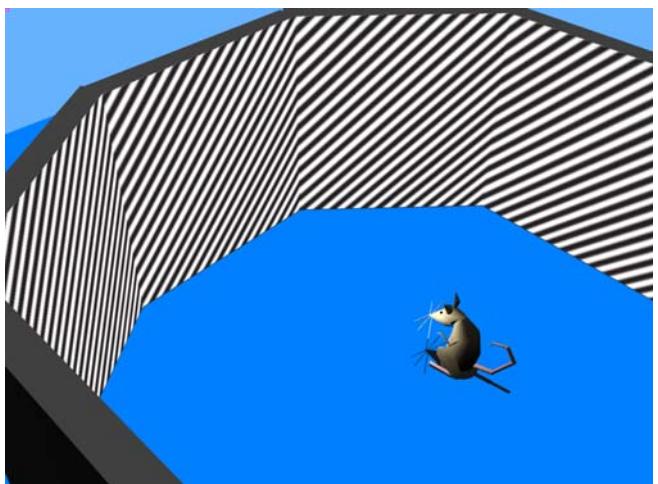


Figure 2: Screenshot of the ICEAsim rat in a simulated environment with 45 degree striped black and white walls.

As the computational model of the cortex underlying these experiments and its neurophysiological basis is non-trivial we delay their discussion to a later section of this paper. Of relevance to this section however is that, deflection information from the robots whiskers provided input to a somatosensory / barrel cortex analogue, and pixel information from the cameras provided input to a visual cortex analogue. Furthermore motor signals controlling the robot's movements could optionally provide input to a motor cortex analogue (see Figure 3 for a connection diagram of the model).

A Computational Model of the Cortex

For many neuroscientists the basic unit of the cortex is the cortical micro-column, a structure of between 10 and 100,000 cells with inhibitory lateral connectivity between local micro-columns, and excitatory connections between micro-columns in different regions of cortex (Mountcastle, 1978; Swanson, 2003). Cortical micro-columns have complex and varying structure and internal connectivity, which we shall not attempt to model in any detail here. However, they are observed to be non-chaotic; they do not display stable attractor dynamics (activity decays on cessation of input); the size of input is small relative to the size of the micro-column; and the state space achieved by an active 'firing' micro-column is large and sensitive to its input (Gupta et al., 2002; Markram, Wang, & Tsodyks, 1998). All of these properties are also displayed by Echo State Networks (ESN) (Jaeger, 2002), which while not modeling all of the internal details of cortical micro-columns, do serve as abstract neurocomputational models to the extent that they capture these features. The ESN reservoir is a large and fixed recurrent neural network acting as a high dimensional excitable medium containing information about current and recent inputs in the trajectory of transient internal states. The 100 neuron ESN we use is derived from a random weights matrix populated with 30% connectivity and adjusted so as to have a spectral radius < 1 , i.e. $|\lambda_{\max}| < 1$, where λ_{\max} is the eigenvalue of \mathbf{w} which has the largest absolute value, thus the ESN is uniquely controlled by the input and the effect of initial states disappears. By observation this would also seem to be true of cortical micro-columns.

The ESN reservoir is cycled according to standard DTRNN equations: $\mathbf{a}_i = \sum_j \mathbf{y}_j \mathbf{w}_{ij} + i_i$ where neuron output is computed by: $\mathbf{y}_i = \tanh(\mathbf{a}_i)$ and input to the reservoir is provided via weights generated by the same method as the ESN weights.

Most excitatory connections between cortical micro-columns target the same regions and thus form major pathways through the cortex, connecting first unimodal regions and then polymodal regions and following a similar path independently of the particular modality (including the motor cortex) (Jones & Powell, 1970; Swanson, 2003). Herein we construct a simple model of the connectivity and regionalization of the rat cortex based on cortical maps from Brown and Aggleton (2001) (see Figure 3).

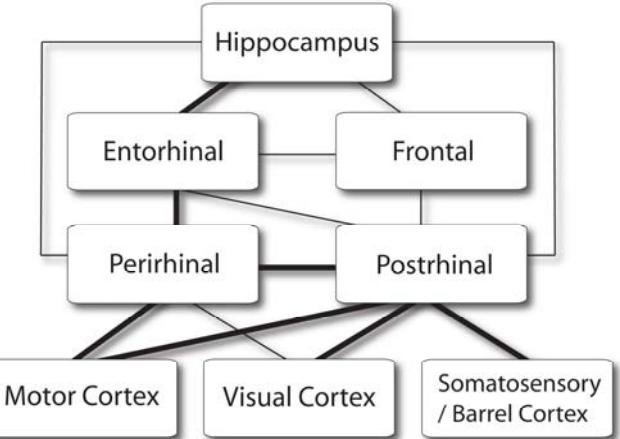


Figure 3: Connection diagram showing the major pathways by which sensory information reaches specific regions of the rat cortex. The thickness of the connecting lines indicates the size of the projection.

According to a theory of cortical processing proposed by Hawkins and Blakeslee (2004), each micro-column detects and classifies features in its input, passing these feature classifications onto the next region. While classifications of detected features flow up this hierarchy, top down connections project back along these pathways such that partial patterns are completed 'top down' providing anticipatory input based on the presence of other sensory features. Major pathways from different sensory regions converge in polymodal regions which are able not only to detect multimodal features, but also predict features in one modality based on information from another.

As we wish to avoid specifying which things are to be classified, and given that we know that information about current and recent inputs is present in the transient internal states of the reservoir, our ESN-based computational model passes on information about these states to connected columns following the map shown in Figure 1. Rather than pass on the full internal state of the reservoir (which would contradict the biological observation that the size of input is small relative to the size of the micro-column) we autonomously classify its state, preserving topology, using a Self-Organizing Map (SOM) (Kohonen, 1998). The SOM provides an ongoing approximation of the principle components of the state space, thus using a 2 dimensional map (in map space) we can extract the position of the winning node and provide information that co-varies with the main principle components of the state space of the reservoir. This provides a low dimensional output that maximally varies with the state of the reservoir. The SOM then also provides a normal input back into the ESN reservoir as shown in Figure 4.

Finally single layer perceptrons reading the ESN reservoir of one cortical hierarchy unit are trained using a standard delta rule: $\Delta \mathbf{w}_i = \alpha(\mathbf{t}^p - \mathbf{a}^p)\mathbf{x}_i^p$ to match the current activity of SOM units in connected columns. By allowing these predictions to activate SOM units in connected hierarchy

units, which in turn provide input to the ESN reservoirs of those units, the model can provide anticipatory input to those units based on the presence of sensory or motor features. Such feedback has already been identified as a mechanism by which inattentional blindness can be modeled and accounted for in these models (Morse, In Press; Morse, Lowe, & Ziemke, In Press-a, In Press-b). By comparison to most connectionist or evolutionary models this may seem rather complex, however we identified here a set of principles for generating mid to large scale neural models capturing aspects of biological cortex and able to display a variety of both neurological and psychological phenomena (cf. Morse, 2006).

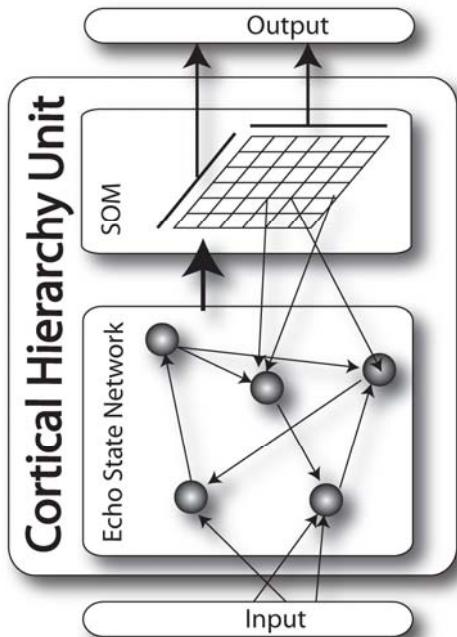


Figure 4: The basic unit of the cortical hierarchy. Input perturbs an ESN reservoir which is read by a SOM. The SOM also provides an input to the ESN and the location of the winning SOM unit in SOM space is provided as output.

Experiment 1: Unrestrained Movement

In experiment 1, the rat was driven by a simple wall following program in the environment shown in Figure 1 for 2000 time steps as the neighborhood size of each SOM reduced linearly to 0. During this period, in condition A, the motor activity provided input to the motor cortex, in condition B no input was provided to the motor cortex, and in condition C random input was provided to the motor cortex. Following this exposure period all learning was stopped and the model was tested on its ability to distinguish the angle of the stripes by linear regression (trained single layer perceptrons) of the activity of the ESN reservoir in the visual cortex only. During testing the robot continued to drive using the same behavioral program while

the angle of the stripes on the walls were periodically adjusted.

Experiment 2: Movement in one Plane

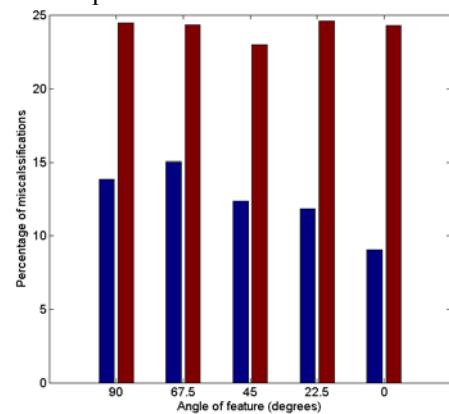
In experiment 2 we followed the same procedure as in experiment 1 but changed the motor program so that the robot turned its head from full left to full right and back again along a single horizontal plane of movement, thus replicating Buisseret and Imberts' experiments in which kittens' ocular movements were restricted to a single plane.

Experiment 3: Sensory-Sensory Corelations

In experiment 3 the robot remained motionless while identical objects (e-puck robots) moved towards it. 50% of these objects collided with the robot causing whisker movement and hence stimulation of the sensorimotor / barrel cortex. Experiment 3 followed a similar design to experiments 1 and 2, having three conditions. In condition A, whisker activity stimulated the sensorimotor / barrel cortex; in condition B, the sensorimotor / barrel cortex received no stimulation; and in condition C the sensorimotor / barrel cortex received random stimulation. During testing objects continued moving toward the robot and 50% of these collided with it causing whisker deflection.

Results and Analysis

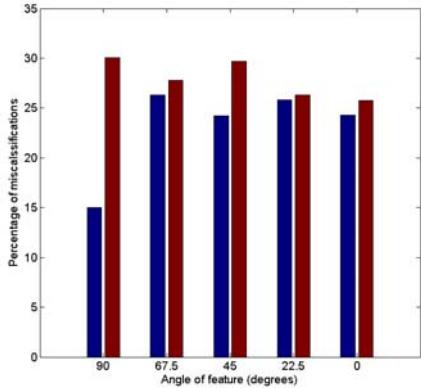
In all conditions of experiments 1 and 2, we recorded the full activity of the ESN reservoir of the visual cortex at every time step as well as the corresponding angle of the stripes on the walls of the environment. For each different angle of the stripes we performed a separate linear regression (on the ESN reservoir activity over time) to distinguish time steps with that particular angle of stripes from time steps in which other angles of stripes were used. Discrimination performance was generally quite high, but we noted particular performance differences between the conditions and experiments.



Graph 1: Showing the results from Experiment 1 in Condition A (motor input)(left bars), and Condition B (no motor input)(right bars). Results shown are the percentage of misclassifications of the linear regression detecting for different angles of the stripes.

Firstly, in both experiments no significant differences were found between discriminatory abilities for any angle of stripes between condition B and condition C, thus the difference between random input to the motor cortex and no input to the motor cortex has no significant effect on the discriminatory abilities of the visual cortex in our model. However, a significant effect ($p < 0.1$) was found in experiment 1 whereby an increase in discrimination performance was observed in all tested angles in condition A over conditions B and C (see Graph 1). This improvement was in the range of a 9% to 15% decrease in the percentage of misclassifications by the trained perceptrons. This shows increased sensitivity to the orientation of the stripe stimuli in the visual cortex of our model when the activity in its motor cortex relates to actual movements as opposed to being either random or absent.

In experiment 2 the movement of the robot was restricted to a single horizontal plane. While general performance at detection was again quite good in all conditions, and no significant differences were found between performance at detection of any angle between condition B and condition C, condition A displayed a 15% improvement at detecting the vertical stripes only, all other tested stripes showed no significant difference between conditions. This demonstrates an improvement in detection of stripe stimuli orthogonal to the plane of movement and thus accurately models the results of Buisseret and Imbert.



Graph 2: Showing the results from Experiment 2 in Condition A (motor input)(left bars), and Condition B (no motor input)(right bars). Results shown are the percentage of misclassifications of the linear regression at detecting different angles of the stripes.

In Experiment 3, as with experiments 1 and 2 we again recorded the full activity of the ESN reservoir of the visual cortex at every time step. We further recorded for each time step whether the approaching object eventually collides with the robot or not. In comparing the performance of trained perceptrons in conditions A, and B, we found a significant improvement in distinguishing whether an approaching object would collide with the robot or not, if the activity of the whiskers stimulated the sensorimotor / barrel cortex (average of 17.14 % misclassifications Vs an average of 38.40 % misclassifications). Given that correlations induce performance increases (Morse et al., In Press-a) which also

facilitate conditioning (Morse & Aktius, 2008), then experiment 3 demonstrates a non-motoric route to detection where events or features in one sensory modality facilitate detection of events or features in another.

Discussion

The cortical model we use is not a neuroscientific model in that it does not attempt to accurately model the internal circuitry of biological cortical micro-columns. As such no specific ‘detectors’ corresponding to neuroscientific findings are produced. However many aspects of cortex and regional interconnectivity are present in the model and our results are based on those aspects of the model. We suggest that increases in the separability of specific environmental features (measured by performance increases in disambiguation by linear regression) would in biological counterparts facilitate the creation of such detectors. We hypothesize that it is these top-down projections that lead to the development of detectors rather than improved separation per se, though the detector would be of those features exhibiting such improved separation.

To summarize, we have replicated, in experiment 1 and 2, Buisseret and Imbert’s findings that controlling your own movement is necessary to establish detection of spatial features and furthermore that if movement is restricted to a single plane then detectors are established only for spatial features orthogonal to that plane of movement. Furthermore, in experiment 3, we have shown that correlations between non-motoric modalities can similarly lead to improvements in performance that we associate with the development of detectors in biological cortex. These results support the biological relevance of our cortical model and further provide an account of the cognitive mechanisms responsible for these well known effects, specifically the top down projection of anticipatory signals.

Our model fully supports the role that action in an environment plays in directing the sensitivity of detection and we presume that in more complex environments this would extend to the discovery of affordances and object recognition. Our third experiment however demonstrates that correlations between non-motoric sensory information can play the same role as action in leading to the discovery of invariants. While the normal human mental schema is dominated by spatial information we believe that it is important to remember that other routes to detection, and by implication perception, exist. What we propose is that sensorimotor knowledge is partnered by sensory-sensory knowledge, the application of which can equally lead to perception. While to many this may seem obvious, such a route to perception has been overlooked by some enactive cognitive theorists. The benefits of considering this alternative route are apparent in the directness of the accounts that can be given of certain perceptions. For example; a sensorimotor account of the perception of an impeding collision can be given in terms of simulating the sensory consequences of performing various behaviors; by contrast sensory-sensory knowledge triggered by the

presence (or simulated presence) of certain stimuli provides a direct route to such perceptions. Such parsimony is not limited to collision detection either, we argue that once sensorimotor knowledge is possessed many perceptual attributes and object characteristics can be perceived in this way and we plan further experiments to demonstrate precisely this.

Acknowledgments

This work has been supported by a European Commission grant to the project “*Integrating Cognition, Emotion and Autonomy*” (ICEA, IST-027819, www.iceaproject.eu) as part of the European *Cognitive Systems* initiative.

References

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, 2, 51-61.

Buissert, P., & Imbert, M. (1976). Visual cortical cells: their developmental properties in normal and dark reared kittens. *The Journal of Physiology*, 255(2), 511-525.

Clark, A. (1997). *Being there - putting brain, body and world together again*. MA: MIT Press.

Clark, A., & Chalmers, D. J. (1998). The Extended Mind. *Analysis*(58), 10-23.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in reason and language. *Cognitive Neuropsychology*, 22, 455-479.

Gapenne, O. (In Press). Kinaesthesia and the construction of perceptual objects. In J. Stewart, O. Gapenne & E. Di Paolo (Eds.), *Enaction: A new paradigm for cognitive science*: MIT Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gupta, A., Silberber, G., Toledo-Rodriguez, M., Wu, C. Z., Wang, Y., & Markram, H. (2002). Organizing principles of neocortical microcircuits. *Cellular and Molecular Life Sciences*.

Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*: Times Books.

Held, R., & Hein, A. (1967). Dissociation of the Visual Placing Response into Elicited and Guided Components (Vol. 158, pp. 390-392).

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Jaeger, H. (2002). *Tutorial on Training Recurrent Neural Networks, Covering BPPT, RTRL, EKF and the "echo State Network" Approach*: GMD-Forschungszentrum Informationstechnik.

Jones, E. G., & Powell, T. P. S. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Journal of Anatomy*, 93, 793-820.

Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, 21(1-3), 1-6.

Markram, H., Wang, Y., & Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons (Vol. 95, pp. 5323-5328): National Acad Sciences.

Morse, A. F. (2006). *Cortical Cognition: Associative Learning in the Real World*.: DPhil Thesis, Department of Informatics, University of Sussex, UK.

Morse, A. F. (In Press). *Neural Models of Prediction and Sustained Inattentional Blindness*. Paper presented at the NCPW11 the Neural Computation and Psychology Workshop, Oxford.

Morse, A. F., & Aktius, M. (2008). Dynamic Liquid Association: Complex learning without implausible guidance. *Neural Networks*, doi : 10.1016/j.neunet.2008.10.008.

Morse, A. F., Lowe, R., & Ziemke, T. (In Press-a). Manipulating Space: Modelling the Role of Transient Dynamics in Inattentional Blindness. *Connection Science*.

Morse, A. F., Lowe, R., & Ziemke, T. (In Press-b). *A Neurocomputational Model of Anticipation, Priming and Sustained Inattentional Blindness*. Paper presented at ABiALS, Munich.

Morse, A. F., & Ziemke, T. (2007). *Cognitive Robotics, Enactive Perception, and Learning in the Real World*. Paper presented at the CogSci 2007 - The 29th Annual Conference of the Cognitive Science Society.

Mountcastle, V. B. (1978). An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System. In Edelman & Mountcastle (Eds.), *The Mindful Brain*: MIT Press.

Noë, A. (2004). *Action in Perception*. Cambridge, Mass: MIT Press.

O'Regan, K., & Noë, A. (2001). A sensorimotor account of visual perception and consciousness. *Behavioral and Brain Sciences*, 24, 939-1011.

Philipona, D. L., & O'Regan, J. K. (2006). Color naming, unique hues, and hue cancellation predicted from singularities in reflection properties. *Visual Neuroscience*, 23(3-4), 331-339.

Philipona, D. L., O'Regan, J. K., & Nadal, J. P. (2003). Is There Something Out There? Inferring Space from Sensorimotor Dependencies (Vol. 15, pp. 2029-2049): MIT Press.

Suchman, L. A. (1987). *Plans and Situated Action: The Problem of Human-Machine Communication*. New York: Cambridge University Press.

Swanson, L. W. (2003). *Brain Architecture: Understanding the Basic Plan*: Oxford University Press.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*: Harvard University Press.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.