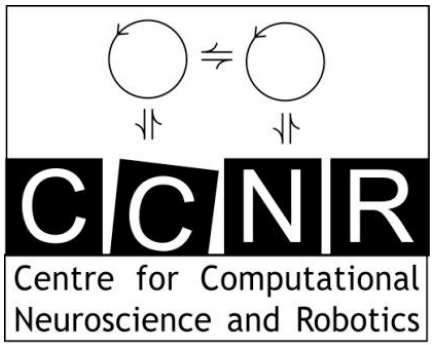
enaction begins in autonomy

Ezequiel A. Di Paolo
Centre for Computational Neuroscience and Robotics
Centre for Research in Cognitive Science
University of Sussex



what is enaction?

collaborators

Hanne De Jaegher

Marieke Rohde

Hiroyuki Iizuka

autonomy

evolutionary robotics

dynamical systems - Ashby

autonomy in the enactive approach

enactivism

- Five central ideas:
  - *autonomy*,
  - *emergence*,
  - *embodiment*,
  - *experience*,
  - *sense-making*

(Varela, Thompson, Rosch, 1991, Thompson, 2007, Di Paolo, Rohde, De Jaegher, 2007).

autonomy

- A cognitive agent is autonomous, it gives itself its own laws.
- How? Only by being able to affect its own constitution this is possible. Only a system able not just to modify itself, but to build itself as an entity.
- *A precarious, self-sustaining process of identity generation.* Classical example: autopoiesis, but others are possible.
- Implication: Forget about *internal drives* and *stimulus-driven cognition*. Mind has a proper, *irreducible* level, that of the autonomous cognitive identity.
- In practice: more attention to endogenous dynamics, value and motivation as dynamical principles.
- A good task for exploring autonomous behaviour: *The setting of own goals*

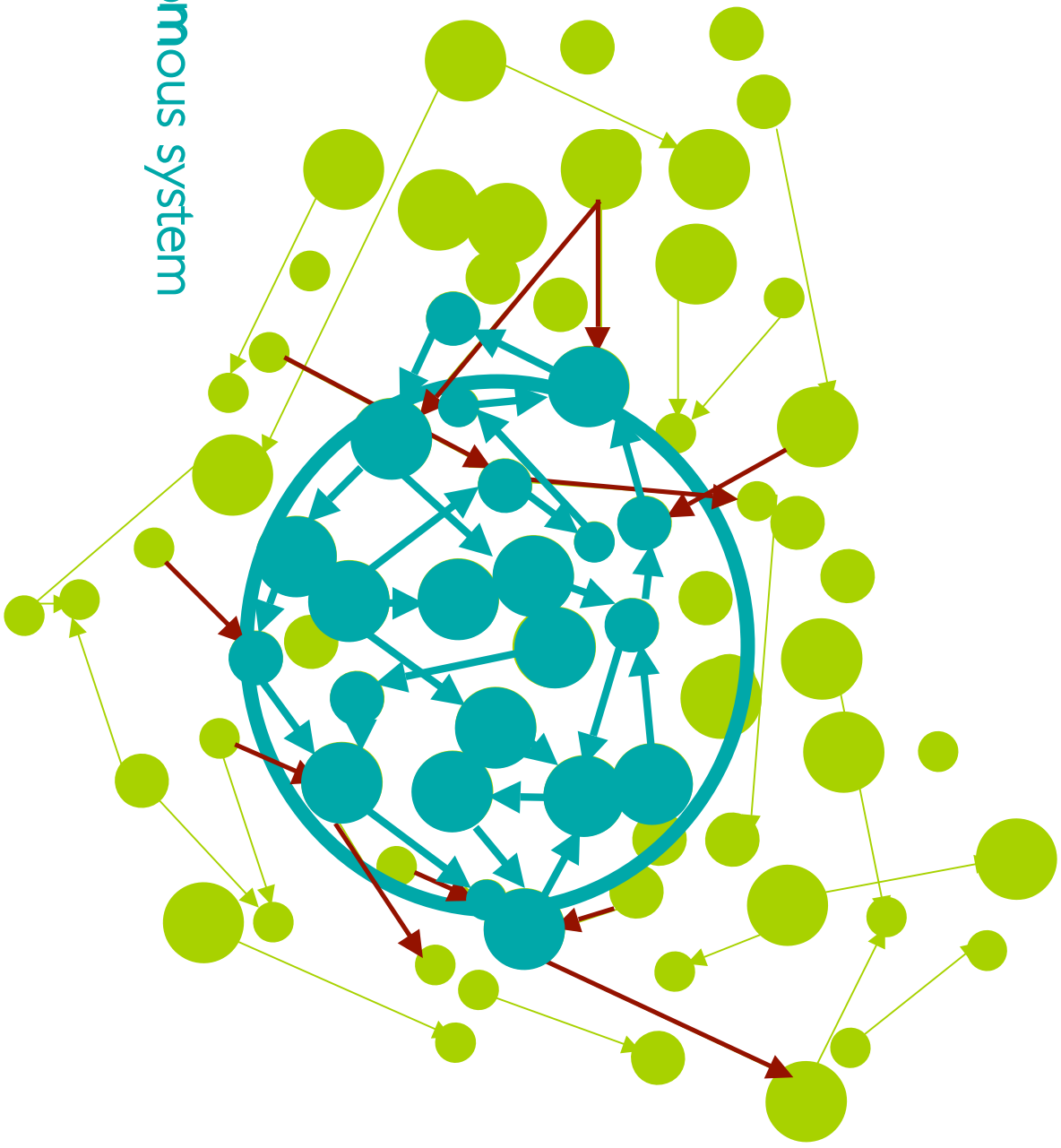
autonomy

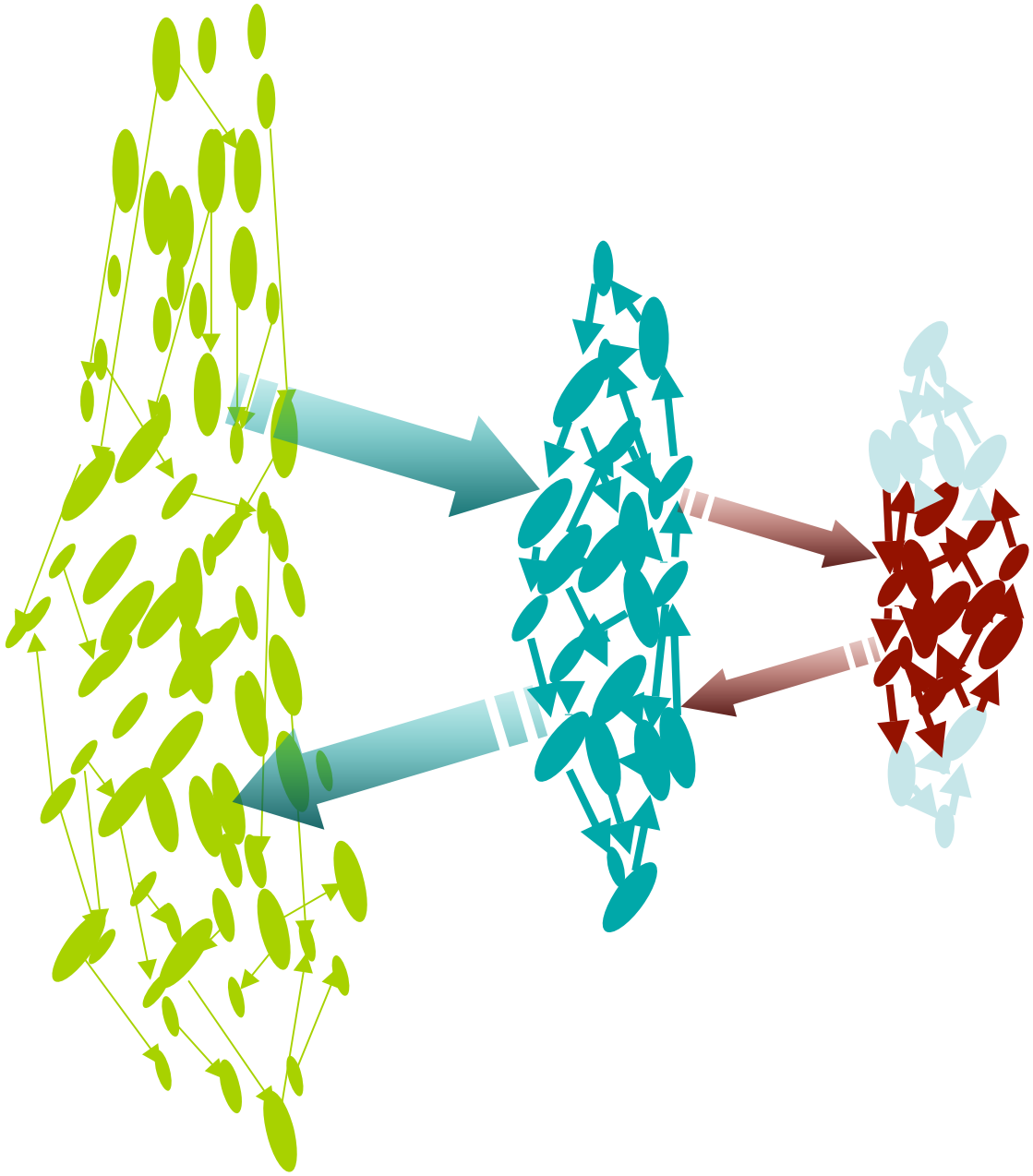
- Before the enactive approach, autonomy was mysterious.
- An agent was always obeying either internal or external demands.
- No proper sense of an identity. Agents defined by convention.
- Enaction establishes at the centre of all its theoretical developments the concept of a self-generating identity.
- The agent is now the proper centre of cognition, not an entity by decree or convention.

definition

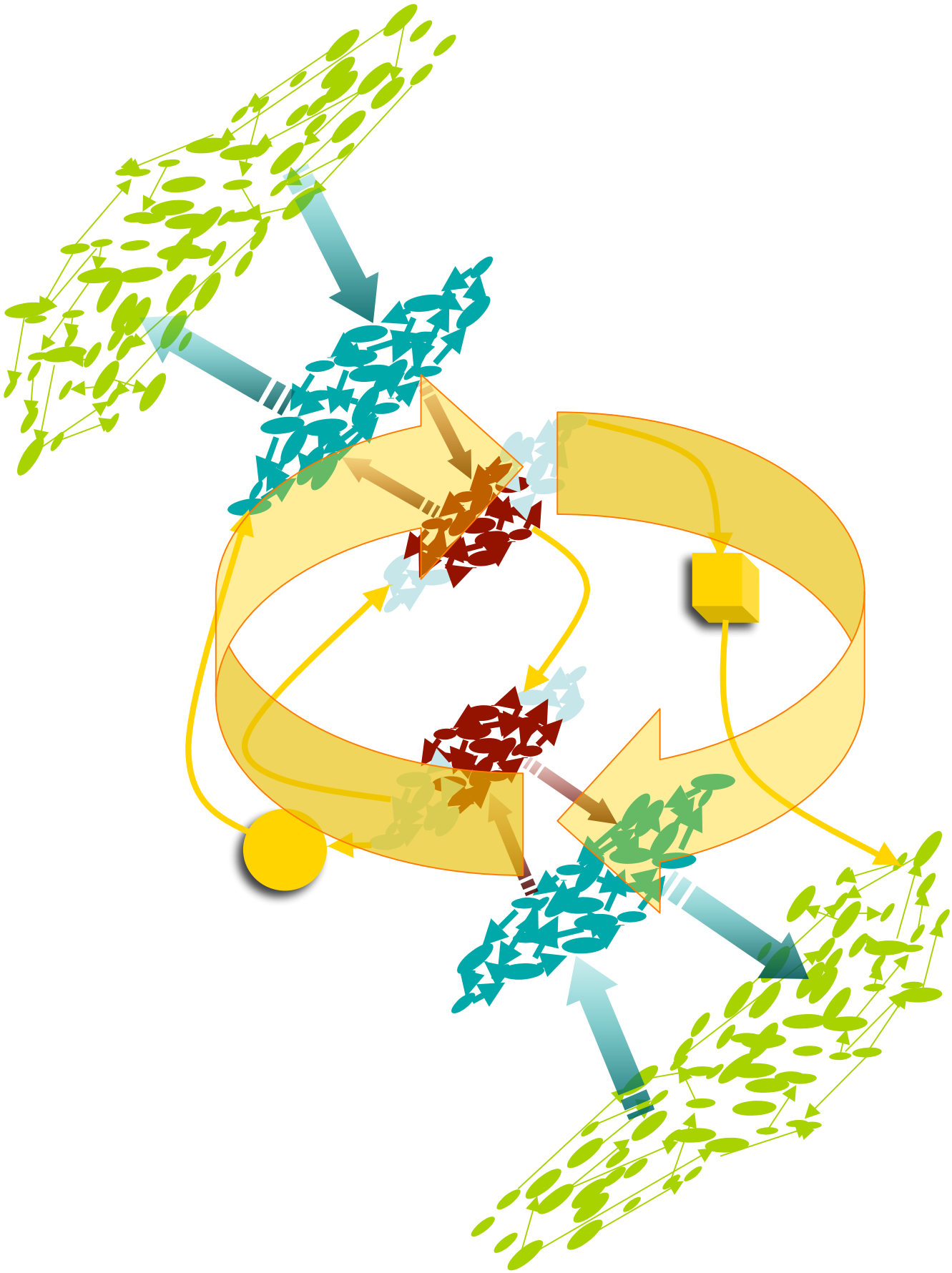
An autonomous system is defined as a system composed of several processes that actively generate and sustain an identity under precarious conditions. By identity we refer to the property of operational closure. Operational closure indicates the property that among the enabling conditions for any constituent process in the system one will always find one or more other processes in the system (i.e., there are no component processes that are not conditioned by other processes in the network, which does not mean, of course, that other conditions external to the system are not necessary as well for such processes to exist). An autonomous system is self-distinct, i.e., a process/component either belongs or not to such a network of enabling conditions (i.e., it is the relation of closure that defines whether a process/component belongs or not to the system), and more strongly, of actively affirming the identity of the system by its own operation. By precarious we mean the fact that in the absence of the organization of the system as a network of processes, under otherwise equal physical conditions, isolated component processes would tend to run down or extinguish.

An Autonomous system





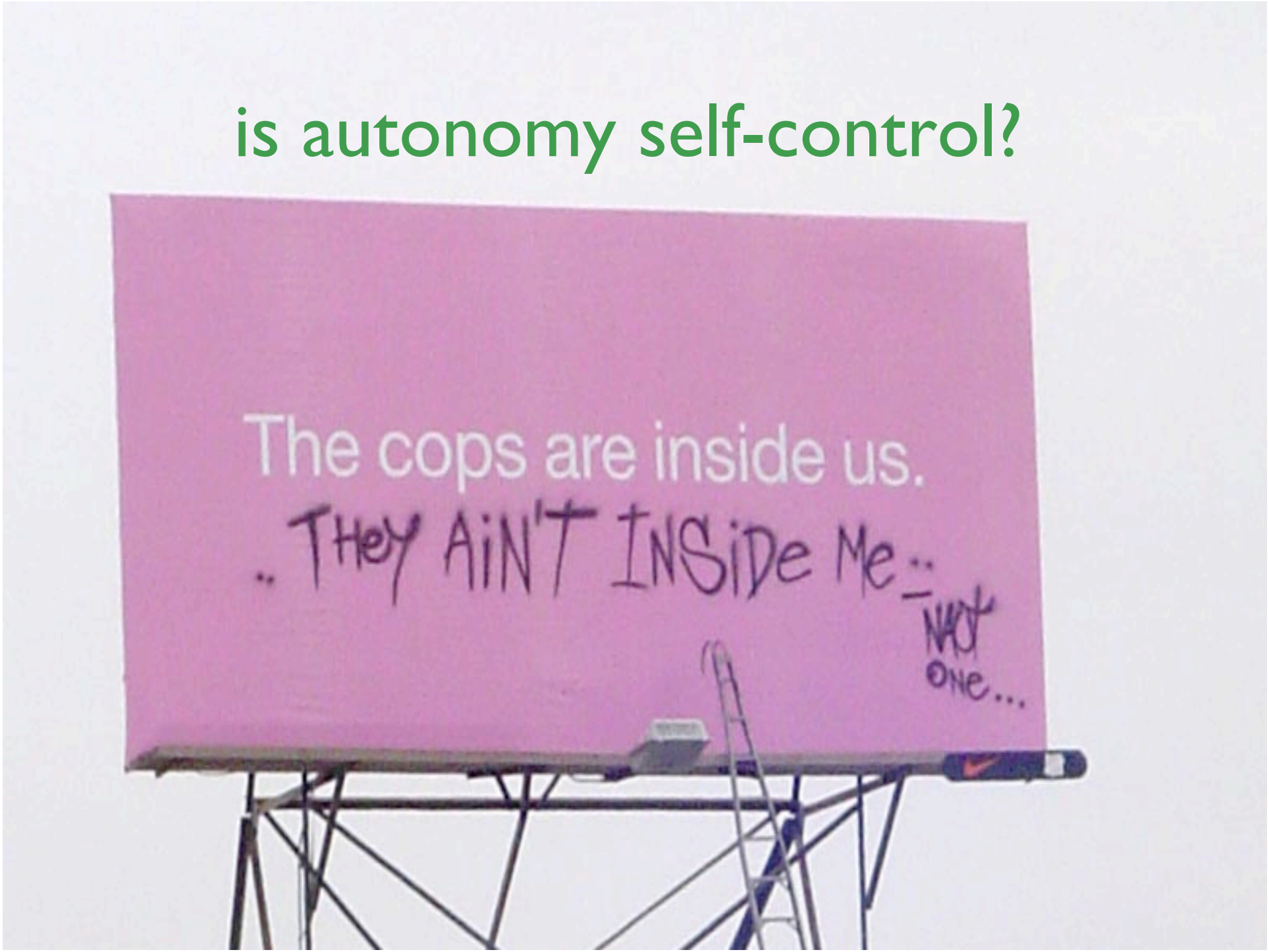
Ezequiel DI PAOLO - Ecole thématique Enaction 2007



social interaction

- A dynamical perspective on the interaction process reveals its transient autonomy.
- An interaction is therefore partially responsible for its own outcome (even, often, to the frustration of the interactors).
- Patterns of coordination and breakdown are the component processes of social interaction.
- But interaction stops being social if the autonomy of the interactors is destroyed.
- Work in collaboration with Hanne De Jaegher (De Jaegher & Di Paolo, *Phenomenology and the Cognitive Sciences*, in press).

is autonomy self-control?



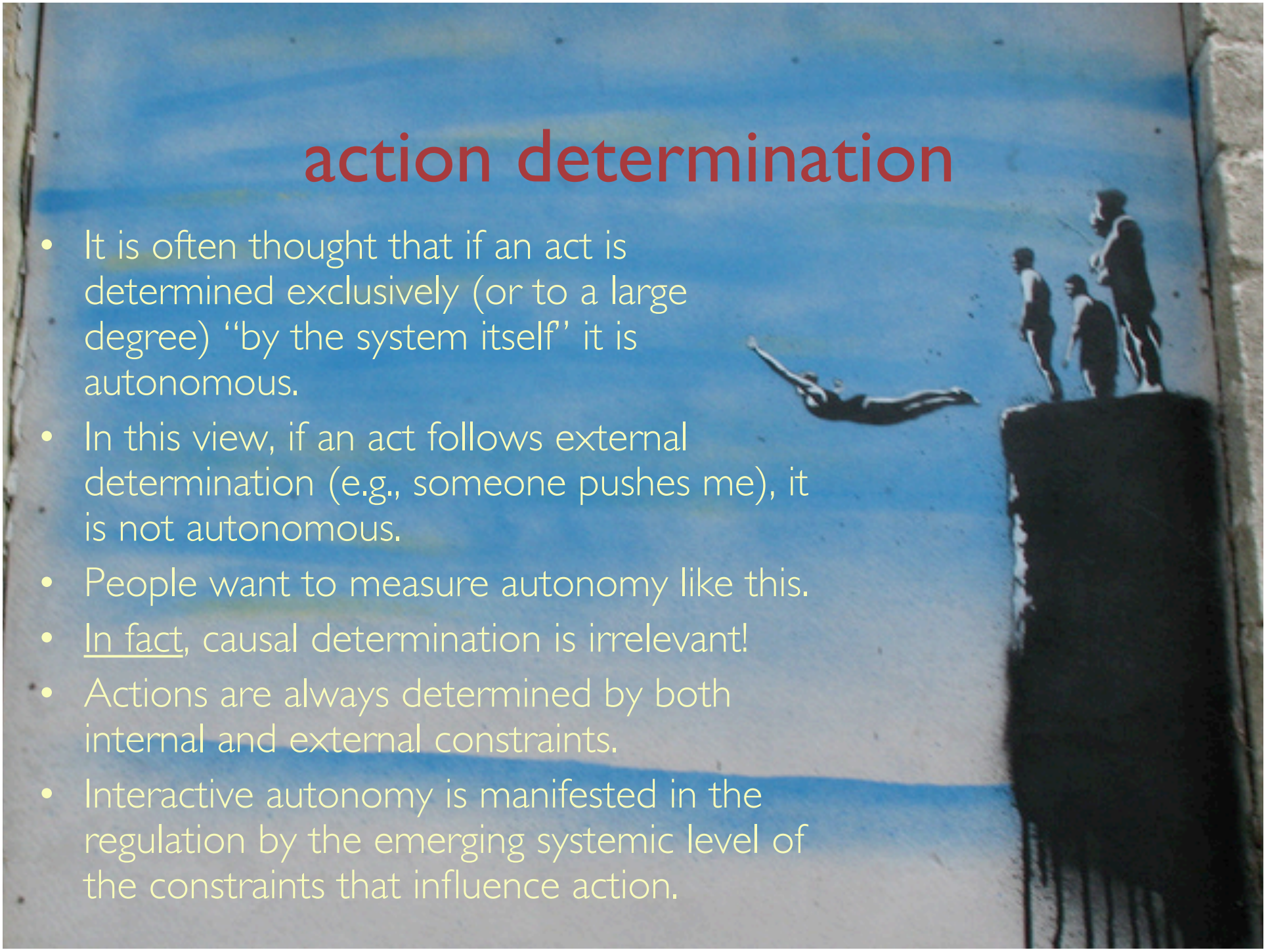
no

- Autonomy is self-law. The determination not of what an agent does, but of the laws that the agent follows.
- Strictly speaking, behaviour, in itself, is not autonomous, only a system can be. It is not a property of actions, but of systems as wholes (i.e., not a property of any component eihter).
- What we experience as *self-control* (e.g., going against the apparent bodily desires) is in fact the result of an interaction between overlapping identities (e.g., socio-linguistic and bodily selves). Possibly only humans can do it.


Submitting to external causation is not to lose one's autonomy

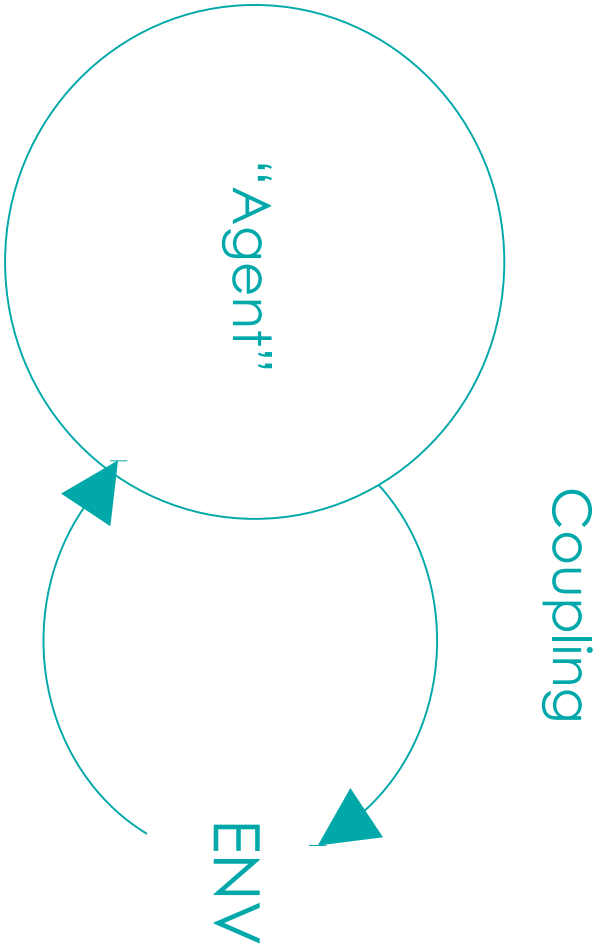
action determination

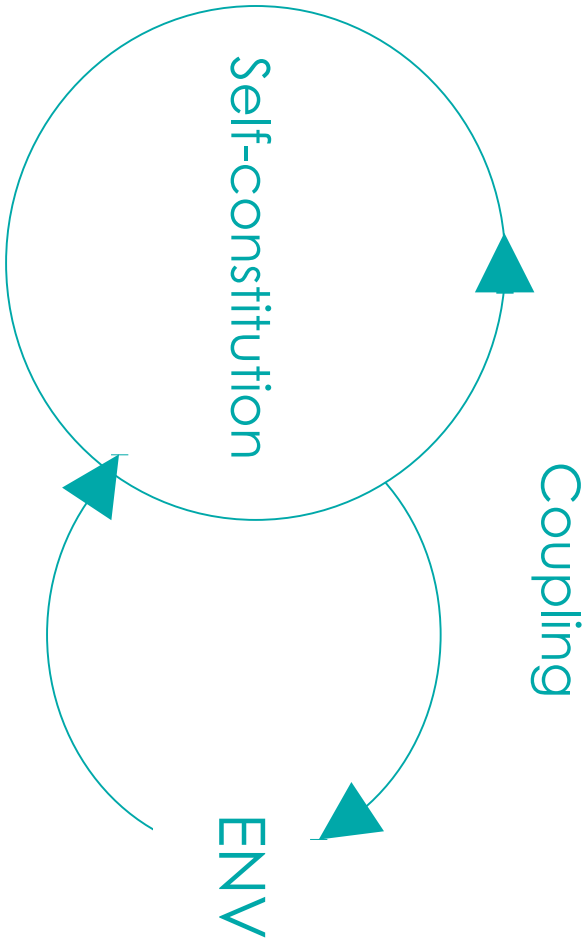
- It is often thought that if an act is determined exclusively (or to a large degree) "by the system itself" it is autonomous.
- In this view, if an act follows external determination (e.g., someone pushes me), it is not autonomous.
- People want to measure autonomy like this.
- In fact, causal determination is irrelevant!
- Actions are always determined by both internal and external constraints.
- Interactive autonomy is manifested in the regulation by the emerging systemic level of the constraints that influence action.

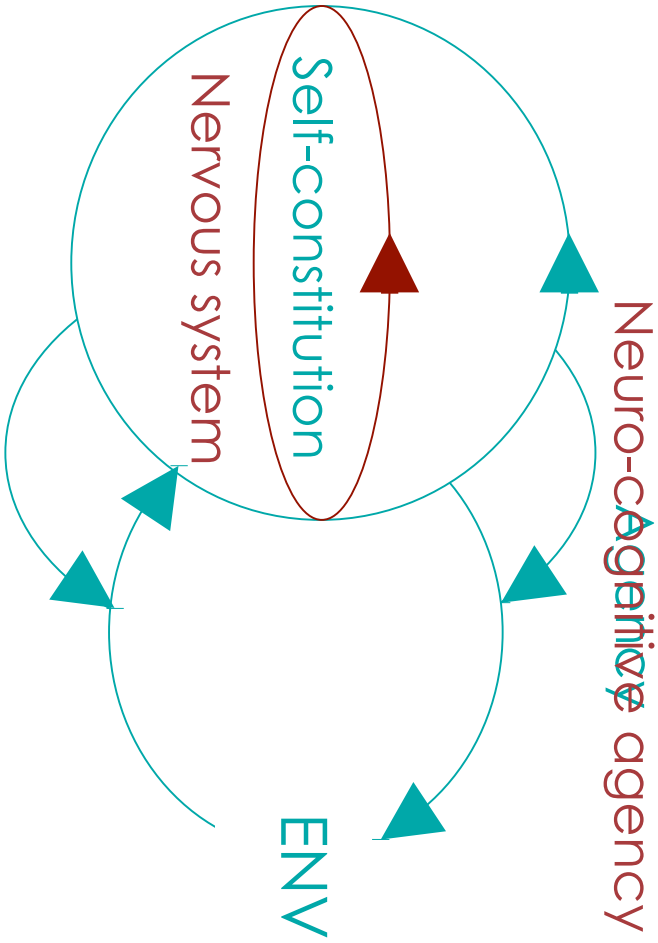


sense-making

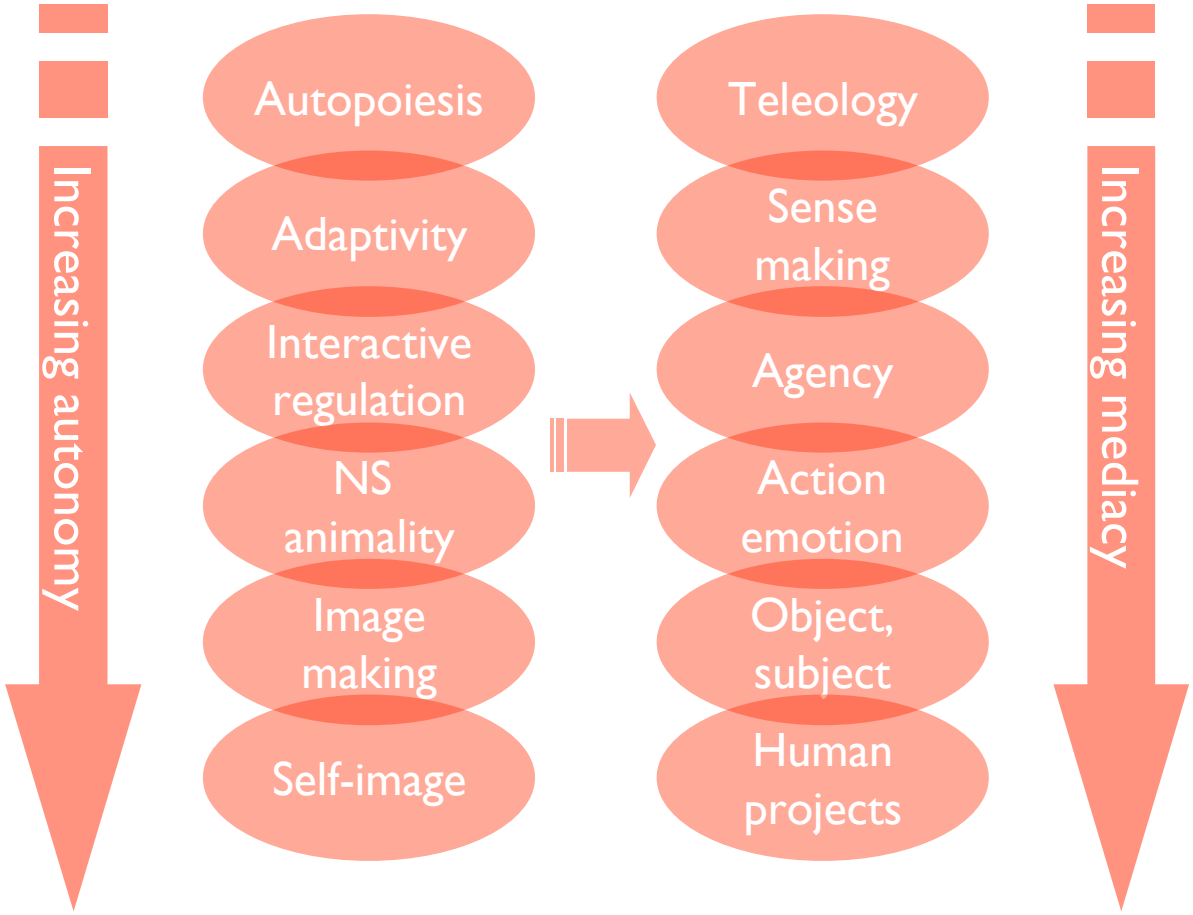
- A self-generated identity implies a normativity with respect to interactions with the world.
- If the mechanisms are present that allow regulation guided by this normativity, the system is now capable of *Sense-making, the active engagement with the world in terms of meaning and value.*
- This property underlies all cognition. Necessary and sufficient.
- *Agency:* sense-making in the interactive domain: when the system adaptively regulates its coupling with its world
- *Behaviour:* The control and selection of what physical exchanges to suffer. It has intentional structure, it may fail or succeed.


"Agent"
Coupling
ENV


Self-constitution
Coupling
ENV


Nervous system
Self-constitution
Neuro-cognitive agency
ENV

life/mind continuity



- The more mediacy, the more autonomy, the more freedom, … (positive freedom in the political sense)
- But also, the more likely the co-existence of …

  … transient, …

  … contemporaneous, …

  … overlapping, …

  … and conflicting, identities.

animality

is about

movement

animality

- spatial behaviour
- sensorimotor correlation building up to a spatial know-how.
- proprioception essential to build a knowledge of space
- intentional distance (tension/satisfaction) is regulated by temporal distance (how much is the tension sustained, how fast is satisfaction attained) and in animals they both acquire a spatial logic, a properly spatial distance so that now and there correlate to later and here.
- emotion comes into being with their correlates of action and perception. It's the unfolding of basic sense-making into a spatio-temporal bodily matrix.
- an animal has a lived body (*Leib*).

from vision to objectivity

- vision becomes a sophisticated spatial sense.
- it responds to a very high-dimensional manifold.
- distance becomes de-centralized. It matters not only how far or near that object is, but also what's the relative position to another object.
- This is one of the bases for what Trevarthen calls secondary inter-subjectivity (my understanding of a sharedness of my world and yours).
- In turn, it is the basis for more objective forms of know-how (think of the act of throwing a stone).

how was the first image ever created?

- Image-making serves as historical evidence of this transition from concrete to expanded situatedness. (Hans Jonas).

- Image-making is an embodied activity. But it's meaning becomes detached from the concrete material form due to its resemblance to the depicted object. But by being different from it, it gains its generality. A pictured horse represents many horses.

- Crucially: such transitions (as the transition to animal motility) can only be understood as the emergence of new value-generating modes.

image-making

world project

self project

Dasein

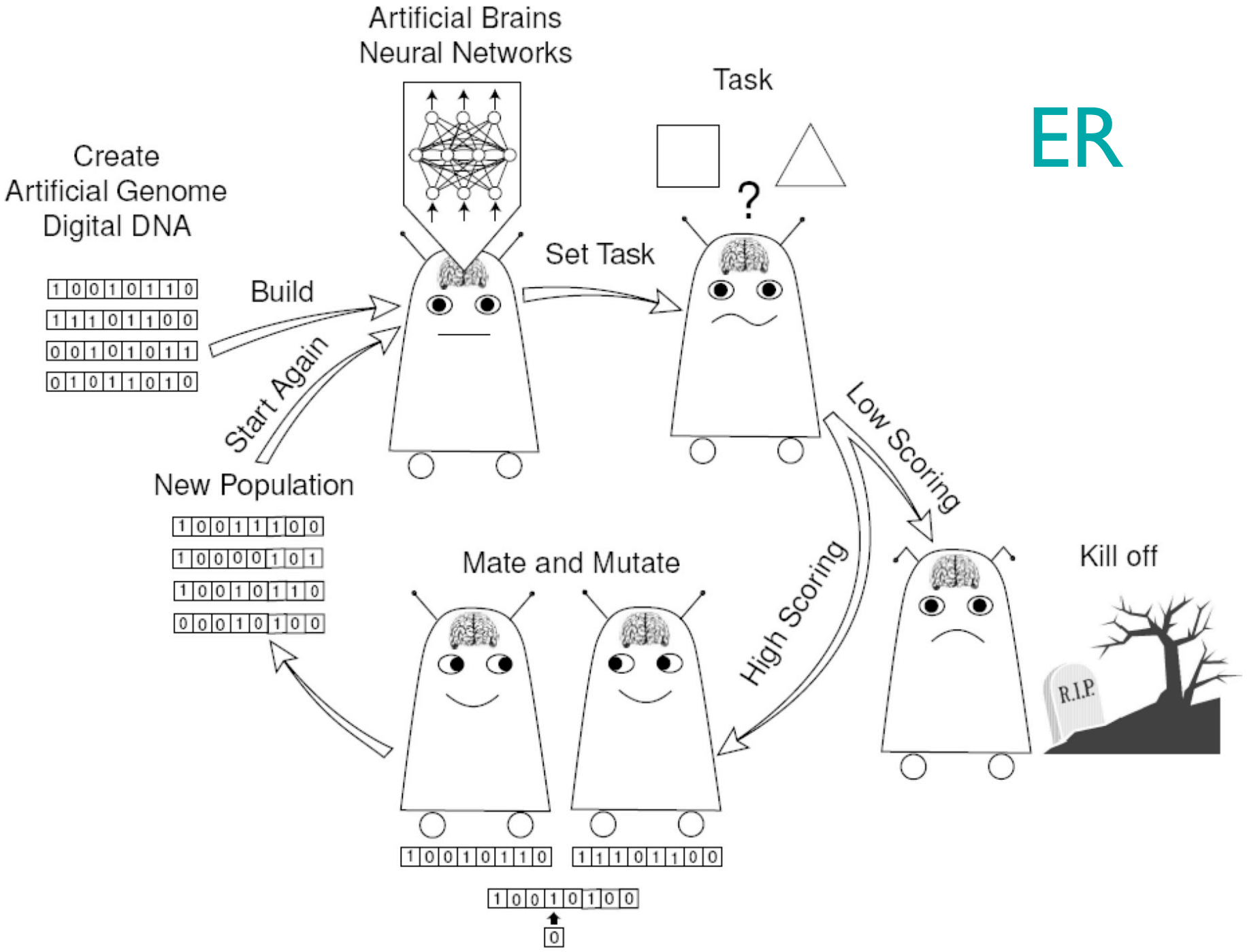
A being for whom its own being is an issue. A 'self-projecting project'. A person.

unfreedom

- At all levels of identity conflict is possible.
- As personal identity is social in nature, it is malleable and integrate social constraints.
- But as social processes themselves are autonomous conflict easily erupts.
- In today Western, liberal societies, the vocabulary of freedom has been high-jacked by powers that limit personal and social autonomy.
- We lack the language to describe our unfreedom (Zizek).


putting ideas to work

evolutionary robotics

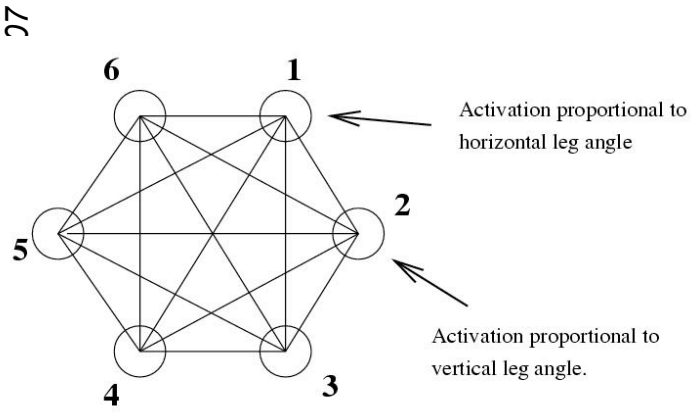


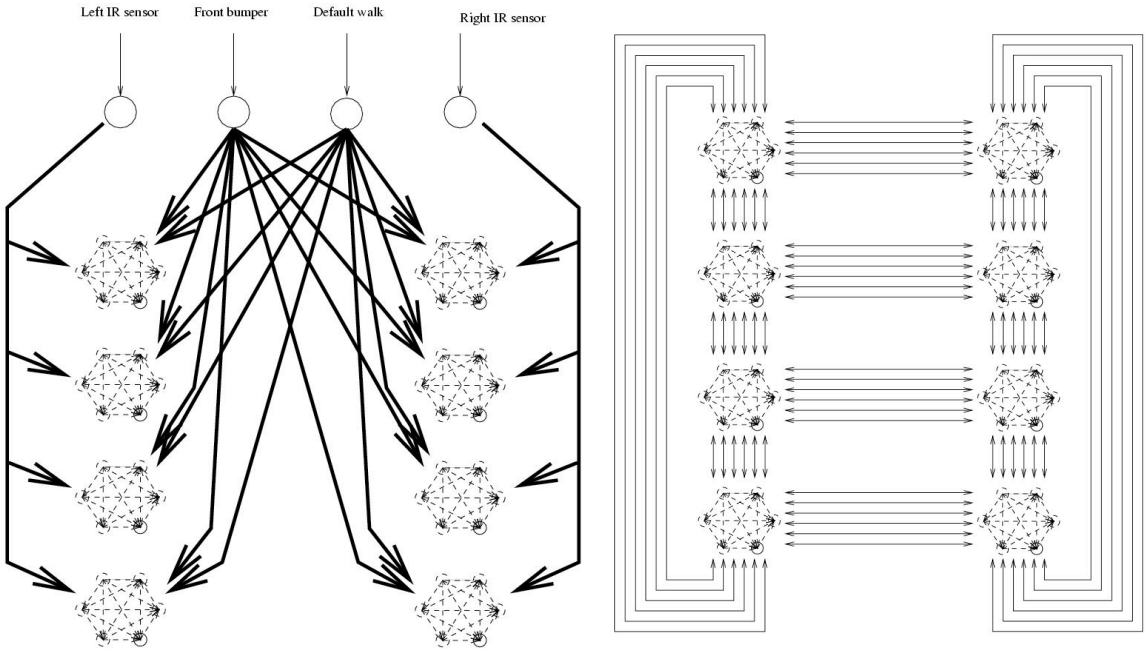
not just obstacle-avoidance

- visually-guided discrimination
- planning
- developmental dynamics
- social coordination
- the perception of social contingencies
- path integration in insect navigation
- arbitrary body morphologies
- multi-joint motor coordination with biological muscle properties and linear synergies
- the role of gaseous neurotransmitters
- neural homeostasis
- spike-timing dependent plasticity

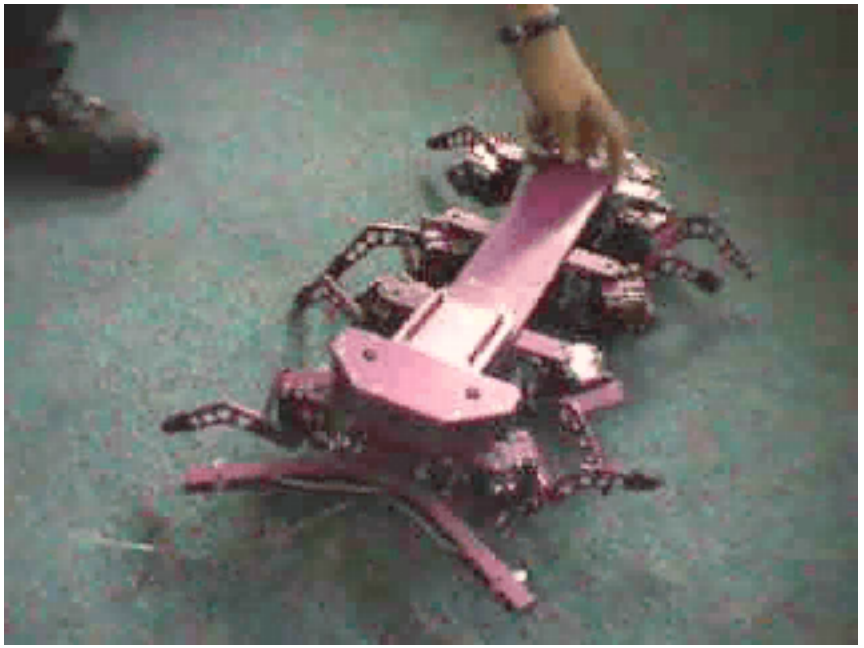
Octopod

- Walk straight as fast as possible. Turn left on the spot when obstacle appears on right-hand side and vice versa. Walk backwards if front bumper is hit.



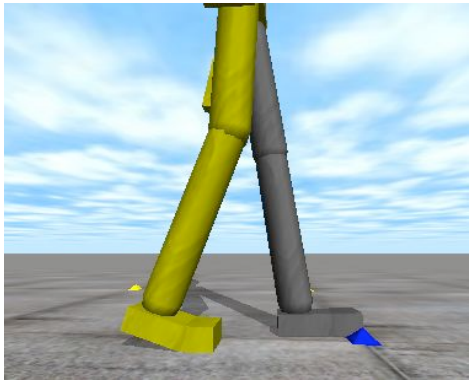


Octopod video



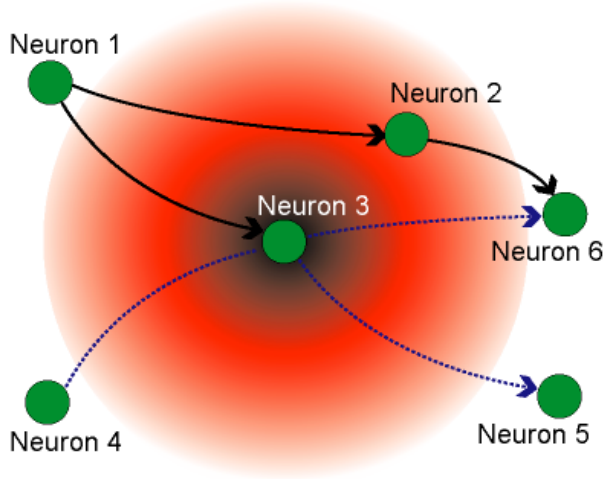
Evolving natural walking

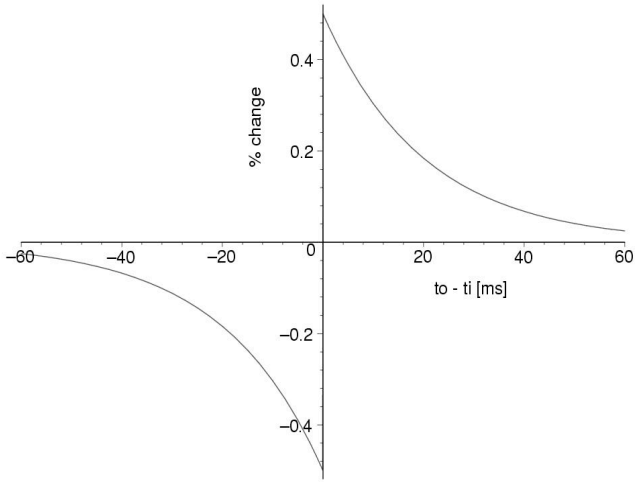
- Full body 32 degrees of freedom including spine
- Physical version currently under construction.
- Vaughan, Harvey, Di Paolo



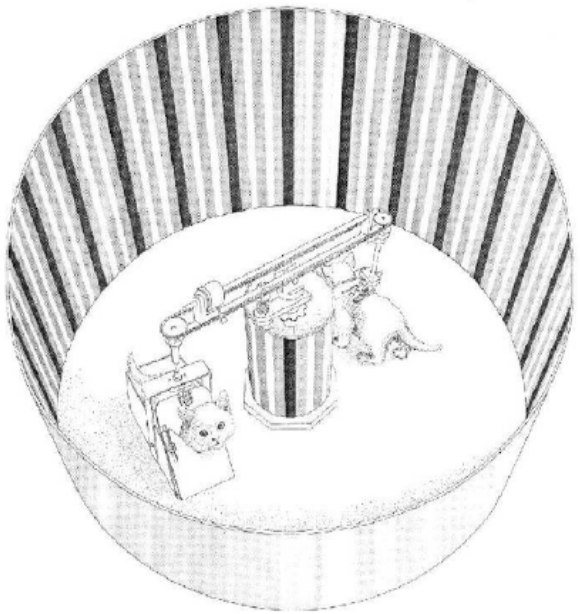
ER and neuroscience

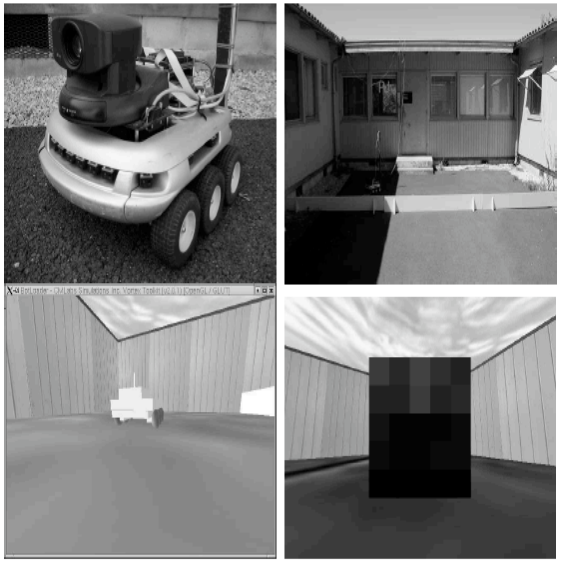
- GasNets. Neural controllers with diffusible gaseous neuromodulation, Husbands et al.
- Spike-time dependent plasticity, Di Paolo, 2004.

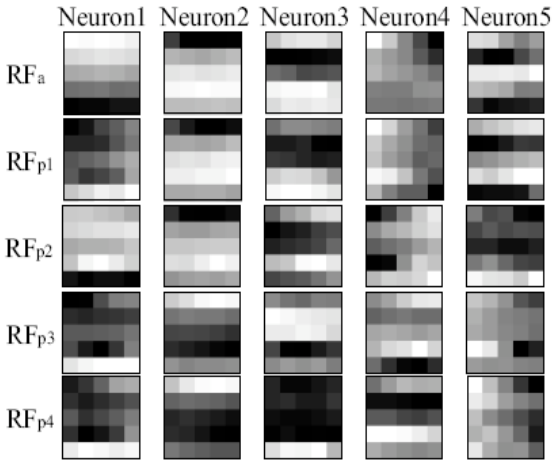




ER and development



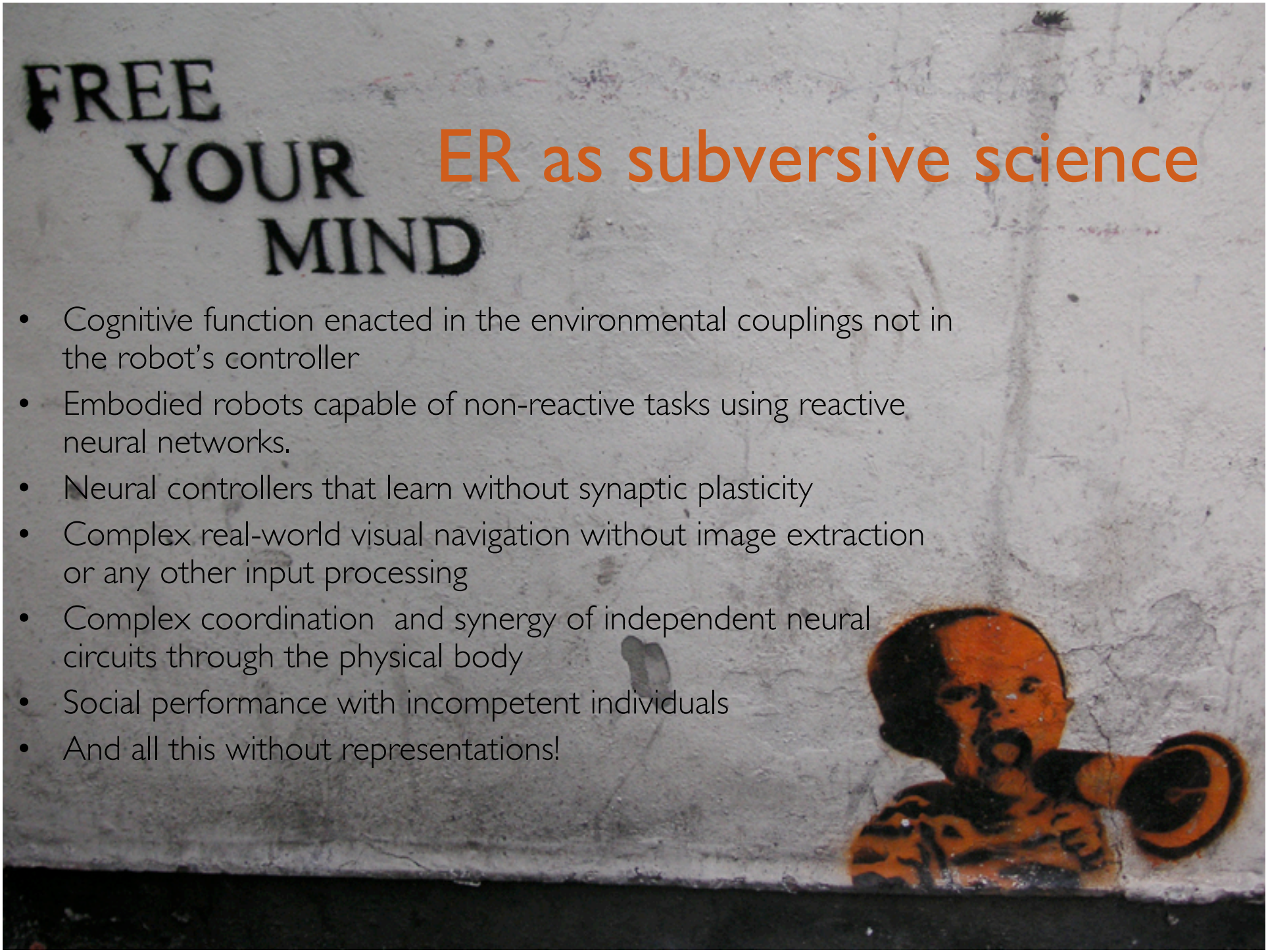




- Replication of Held and Hein experiment on visuomotor development. Suzuki, Floreano, Di Paolo, 2005

common to all of these

- Evolved solutions take advantage of the dynamics of the body in interaction with the environment. These solutions are literally impossible to design by traditional engineering methods.
- The realisation that cognition is never solely inside an agent but at the interface between body and world.
- That representational architectures hardly ever evolve. Only when we constrain the problem so much that there's no other option.
- That cognition is always-already embodied-embedded.



ER as subversive science

- Cognitive function enacted in the environmental couplings not in the robot's controller
- Embodied robots capable of non-reactive tasks using reactive neural networks.
- Neural controllers that learn without synaptic plasticity
- Complex real-world visual navigation without image extraction or any other input processing
- Complex coordination and synergy of independent neural circuits through the physical body
- Social performance with incompetent individuals
- And all this without representations!

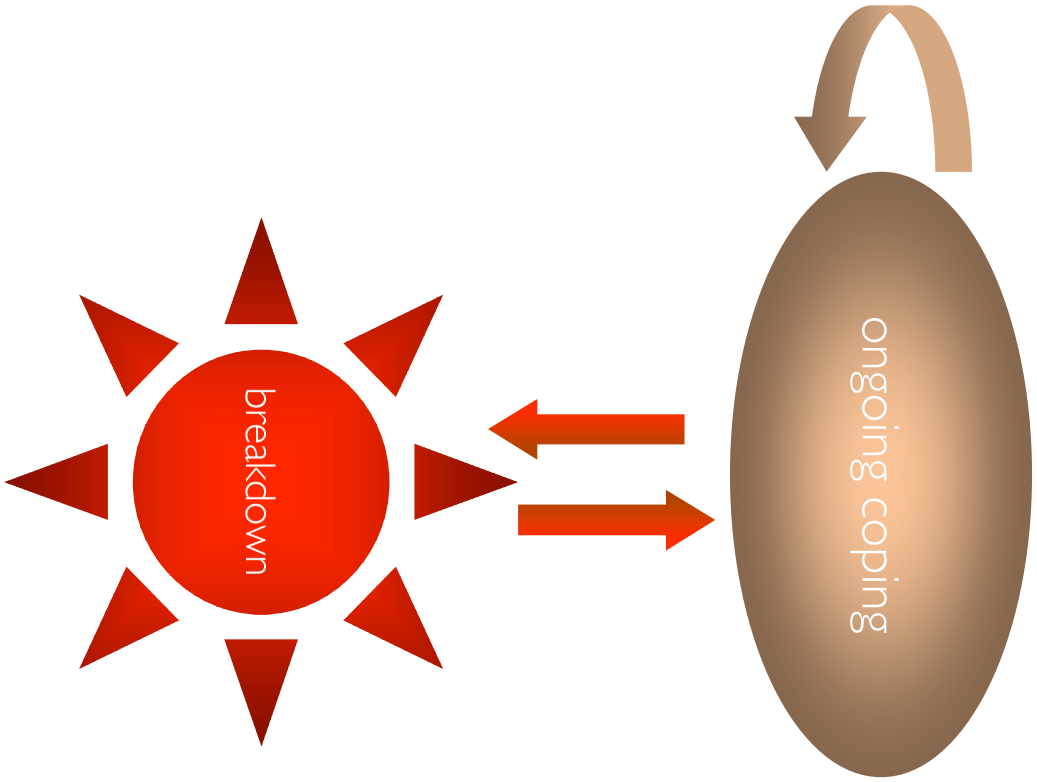
representation recedes

- Anti-computational emphasis on embodiment, situatedness and dynamics downplays *representation-hungry* cognition. (Clark & Toribio, 1994, Clark 1997). Cases of direct engagement with the situation are "cognitively marginal"

- Later Clark (2005) says his own argument is extreme. Offline reason is not de-contextualized even if it is disengaged.

- If by *disengaged* we mean that the target of performance is not present, it is simply not true that this is cognitively marginal. Most animal intelligence (including humans) falls into this category. Clark is wrong.

- But there is more to human cognition. Clark is right. But this is poorly described as disengaged (offline).


breakdown
ongoing coping

a challenge to enactivism

- Cognizing outside the current environmental dynamics.
- For representational thinking, this is not a major problem. Manipulation of representations to deal with the here-and-now isn't fundamentally different from manipulation of representation to deal with the there-and-then, or with non-temporal, non-spatial concepts (in fact it's harder!).
- Cognitivism is based on non-temporal, non-spatial, unsituated mechanisms!! So, no surprises there.
- But, the enactive approach emphasizes situatedness. In doing so, cognition appears as "glued" to the here-and-now.

robots are too safe! we need:

- A process of self-sustaining identity generation leading to
- A unitary, bounded organization that actively self-distinguishes from its environment and so it produces
- A norm of value in its interactions.
- Also: adaptive processes that allow the evaluation of the current state of viability in a graded manner leading to the capability of sense-making and therefore a world.
- And: the extension of adaptive regulation to the interactions with the world so that the system can be said to properly act and perceive (behave intentionally) and be called an agent.

an ashbyan framework for adaptive dynamics

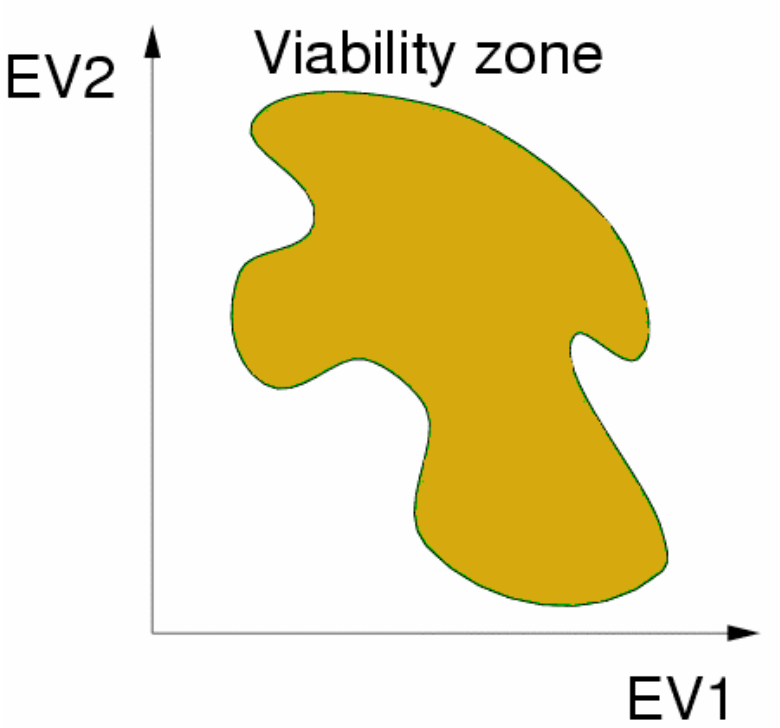
Ezequiel DI PAOLO - Ecole thématique Enaction 2007

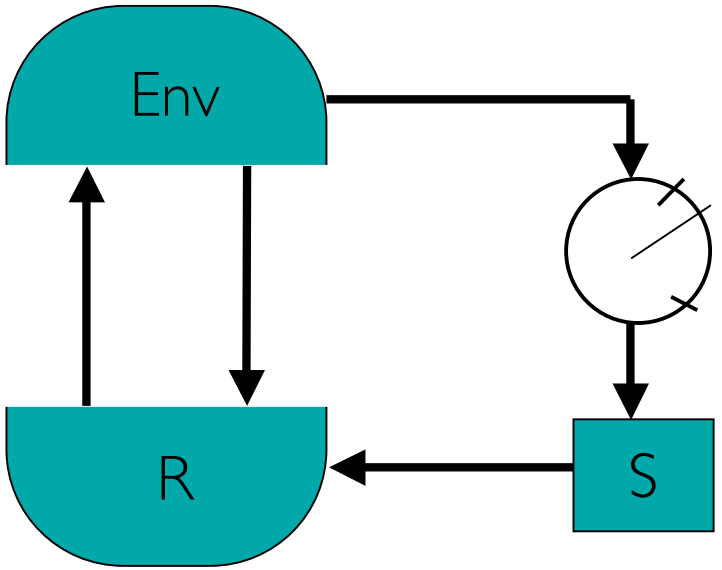
adaptation as stability

- Adaptive behaviour stems from the dynamical principle of ultrastability.
- Plastic systems conserve their identity by remaining within viability constraints.
- Crossing (or approaching) viability boundaries induce plastic changes in the system
- Plastic changes that result in the system remaining non-viable induce more plastic changes.
- Plastic changes that restore viability do not induce further changes.
- W. Ross Ashby. *Design for a brain: The origin of adaptive behaviour*, 1952/1960, London: Chapman and Hall.


W. Ross Ashby

ultrastability

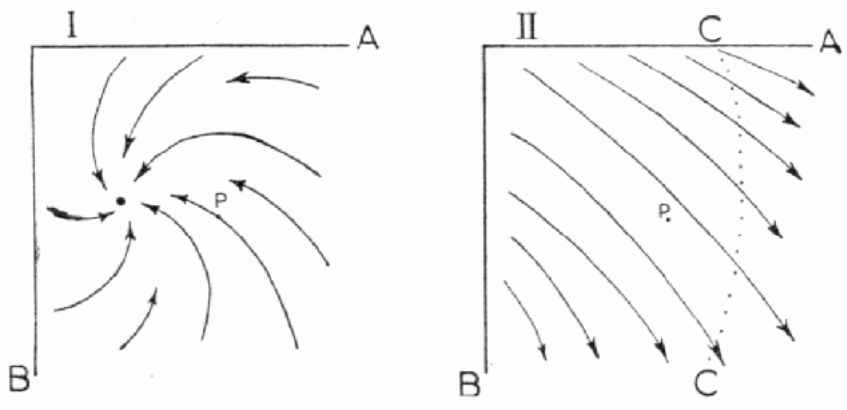




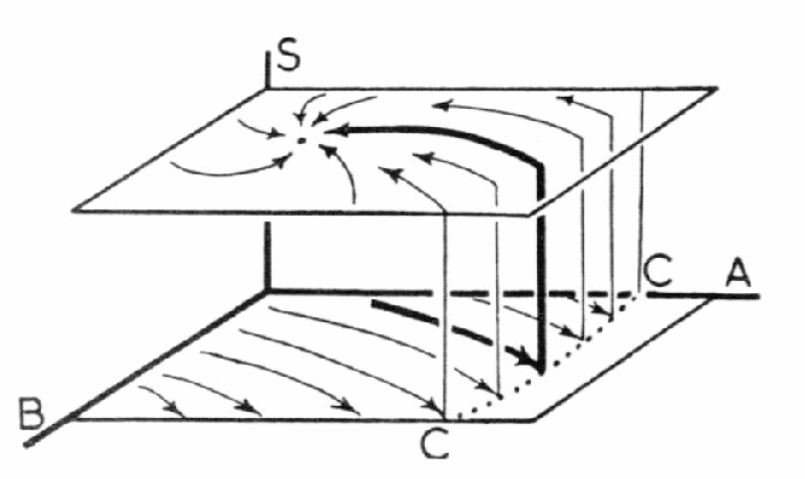
Ashby's framework for adaptive behaviour

ultrastability

In the unstable case, state trajectories will reach a critical condition (right). If parameters were different (left) the system could still be stable under the new environmental pressure.

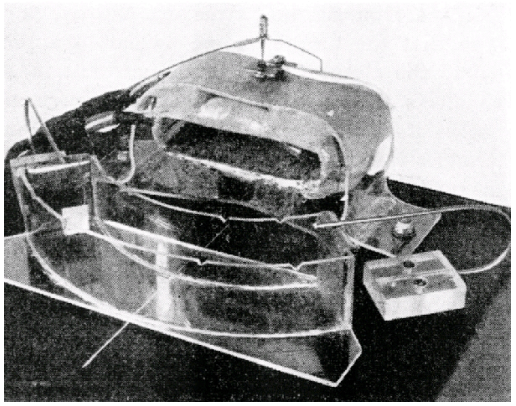


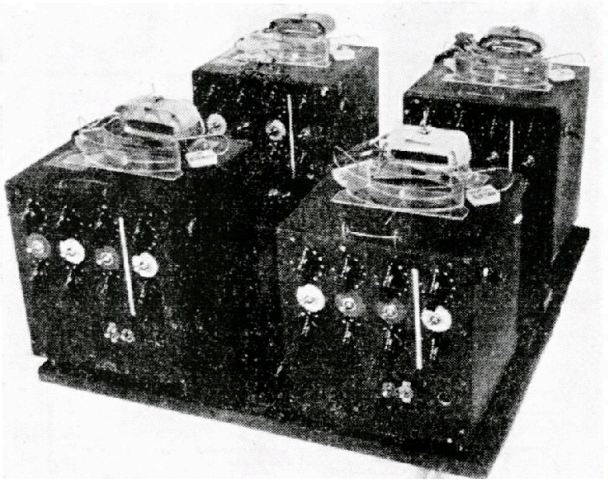
Steps functions acting through secondary feedback could take the dynamics from one field to the other.

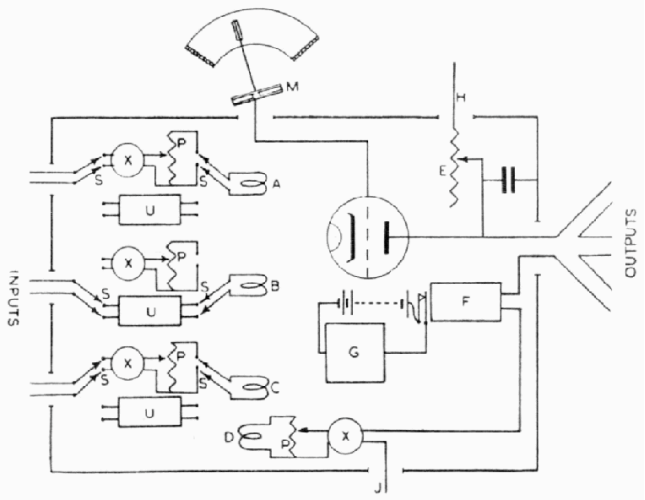


the Homeostat

- Electromagnetic device consisting of 4 ultrastable units that could be coupled in different ways
- Many experiments including habituation, reinforcement learning.

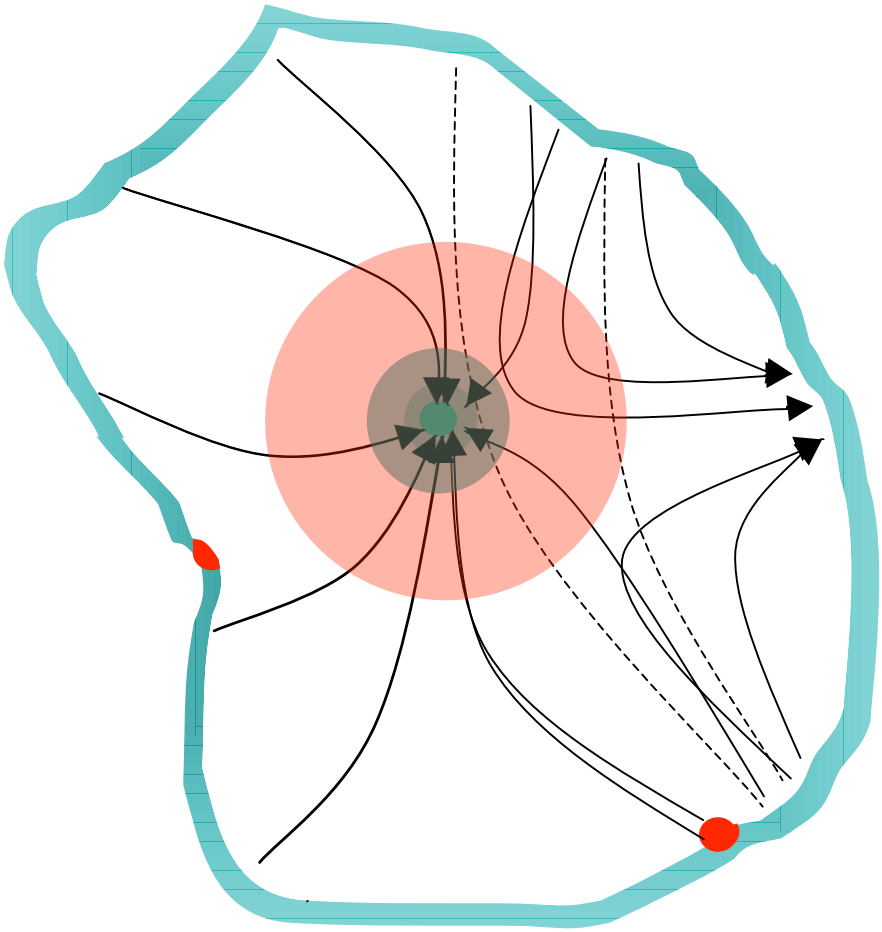




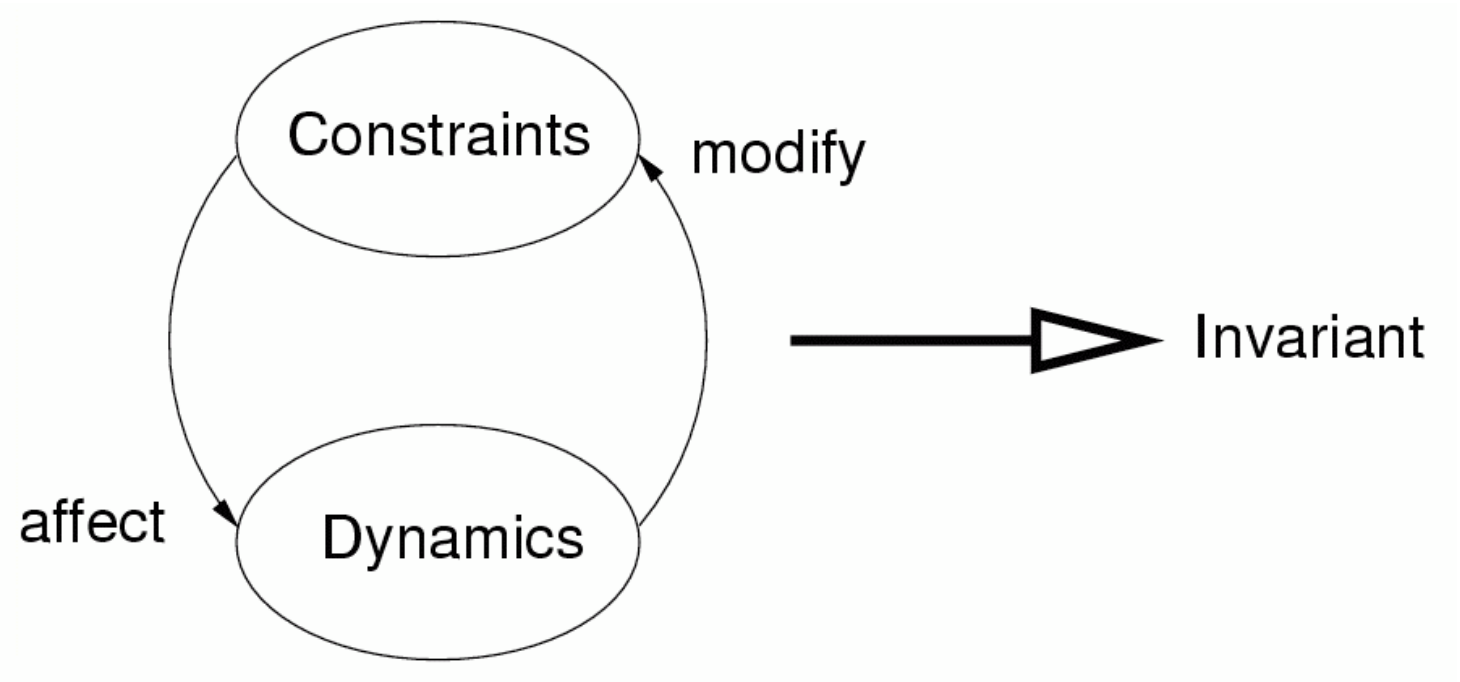


groovedigging: a corollary

- In an ultrastable system with a finite viability zone and a given distribution of fluctuations adaptation ensues as long as the viability zone is sufficiently bigger than the typical variance in essential variables.
- Over time the system may be knocked off equilibrium by infrequent bigger fluctuations.
- It may later return to the viability zone.
- By chance trajectories may move deeper into the zone making it harder for large fluctuations to disrupt the system.
- The process digs itself into a groove.
- [Related ideas: canalisation, mutational robustness of the wildtype]



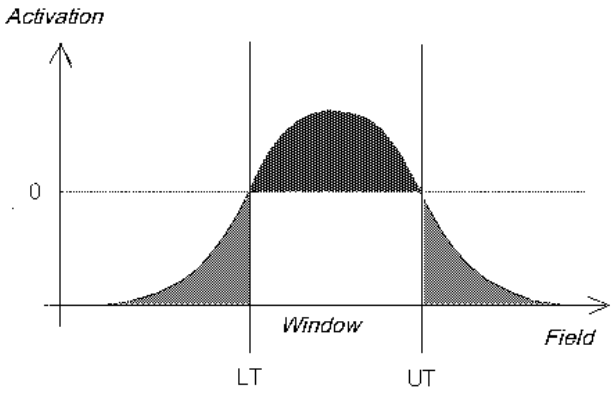
historical processes

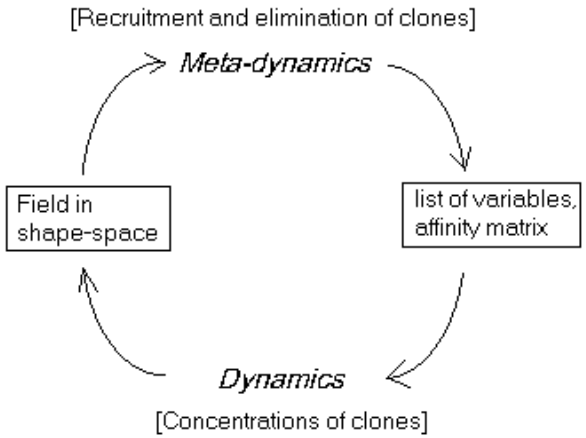


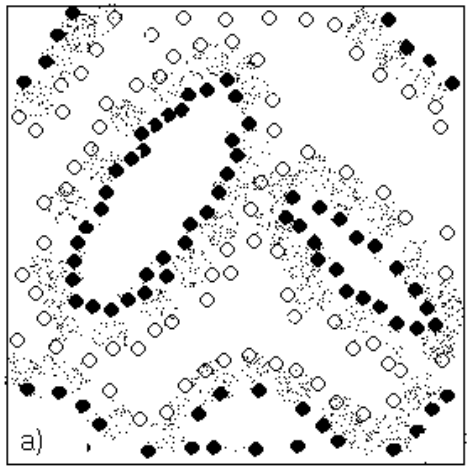

Walking
constrains
forms
Path

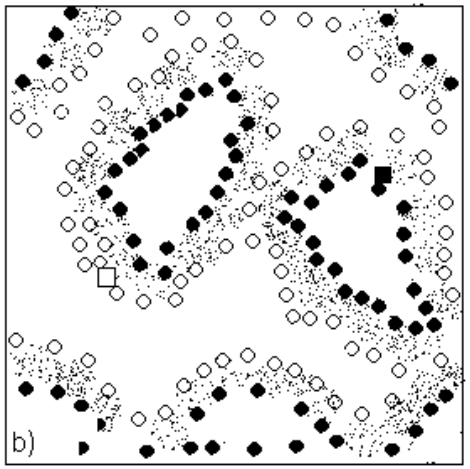
ontogeny in idiotypic networks

- Work by A Coutinho, F Varela, J Stewart et al.



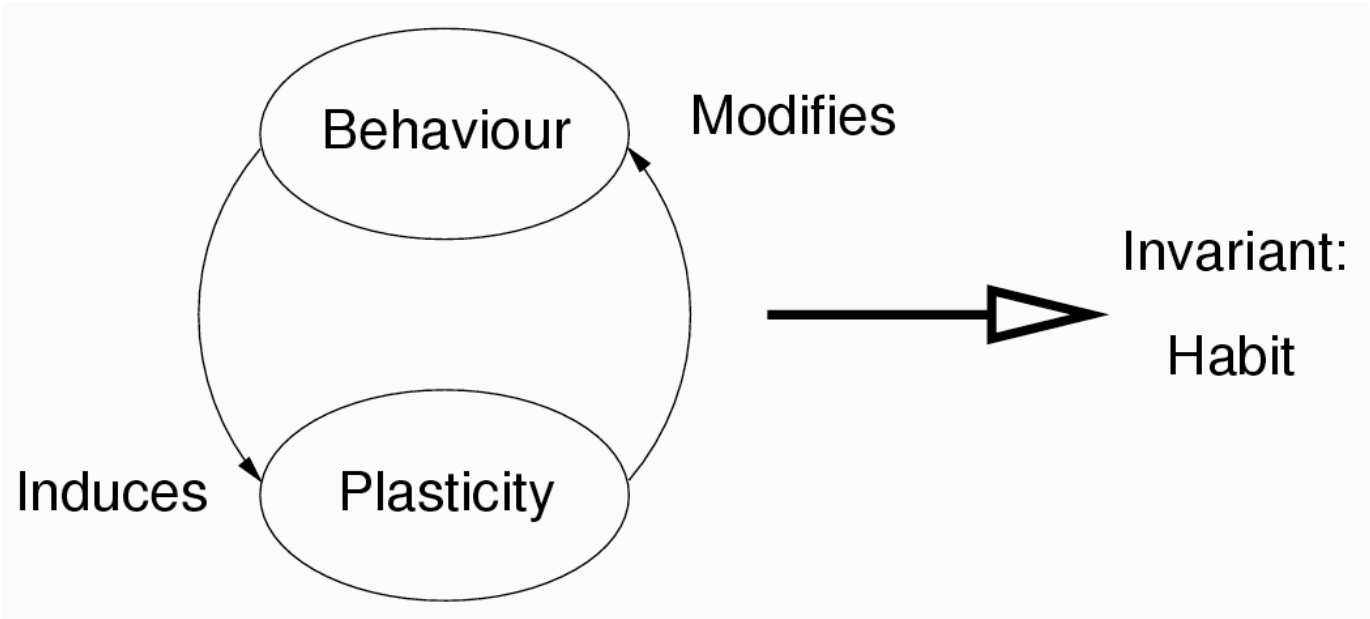






beyond the organism: habits

K. Goldstein, J. Dewey, W. James, M. Merleau-Ponty, P. Guillaume, N. Berstein, I. Kohler and others have used the term 'habit' to describe how the body, as an ecological entity, sets itself into preferred patters of action and perception.



habits as value generators

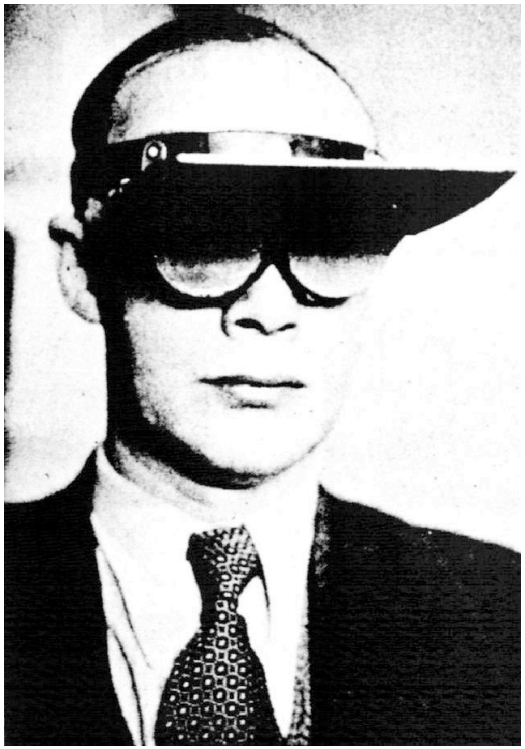
- "Preferred" does not sneak in teleology. Behaviour induces plastic changes which facilitate further behaviour, which induces plastic changes,..., and so on, until a behaviour pattern is found that induces little or no plastic change.
- However, by their process of formation and their self-conservation, habits also generate value. (Something is "good" if it serves the conservation of this autonomous ecological entity).
- Out of all the possible viable behaviours those produced by habit will be conserved (Bernstein's problem, perceptual constancies, etc.)

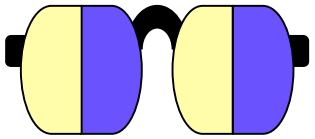
breakdown and recovery

- Habits have different degrees of organisation. From sensorimotor coherences to the concatenation of movement into whole acts, and further.

- Small disturbances can be compensated for, due to the stability of the invariant structure; but radical disruptions may either lead to the *total loss* of the habitual pattern or may allow for the conservation of some *invariant residue* of the original organisation. From this a modified organisation may "grow".

the world upside-down

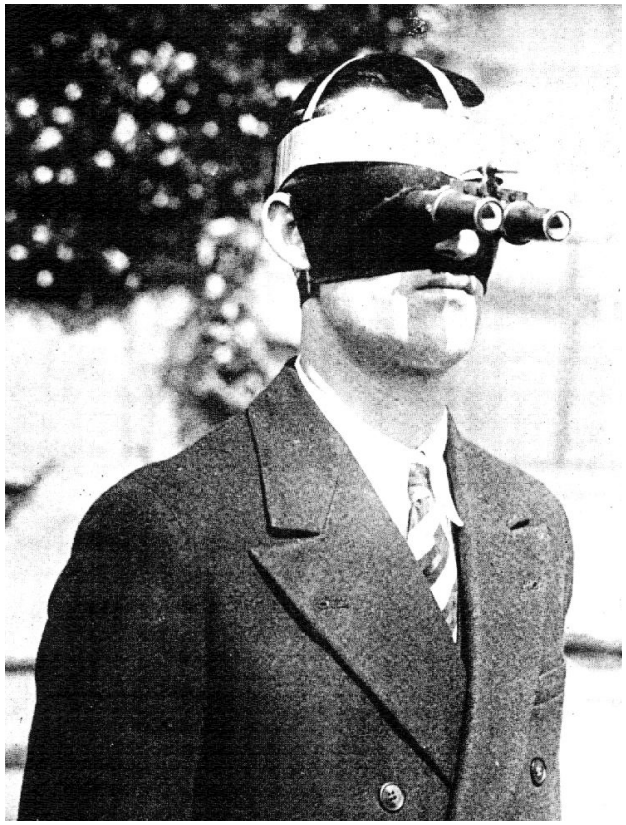




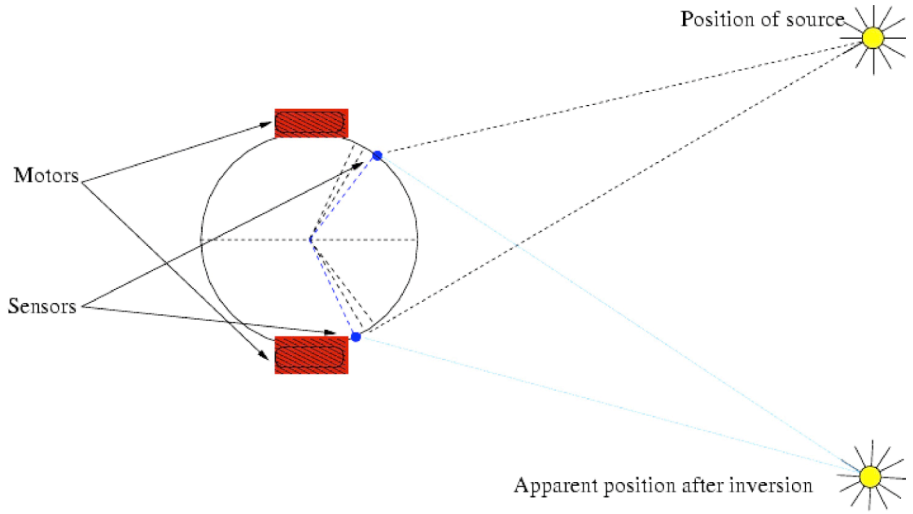
Experiments on the distortion of the visual field (Stratton, Ewert, Erismann & Kohler, Taylor, Dolezal)

adaptation to visual inversion

- Adaptation to left/right visual inversion in a phototactic robot using the individual activity of neurons as the essential variables, (Di Paolo, 2000 following ideas by J. G. Taylor, 1964).
- Neurons facilitate local plasticity when their activity is too high or too low.
- Robots evolved to perform only normal phototaxis and to be internally stable (minimize internal change).
- When sensors are inverted a robot becomes unstable and starts to change. Eventually phototaxis is regained.



adaptation to visual inversion



CTRNN

$$\tau_i \dot{y}_i = -y_i + \sum_j w_{ji} z_j + I_i$$

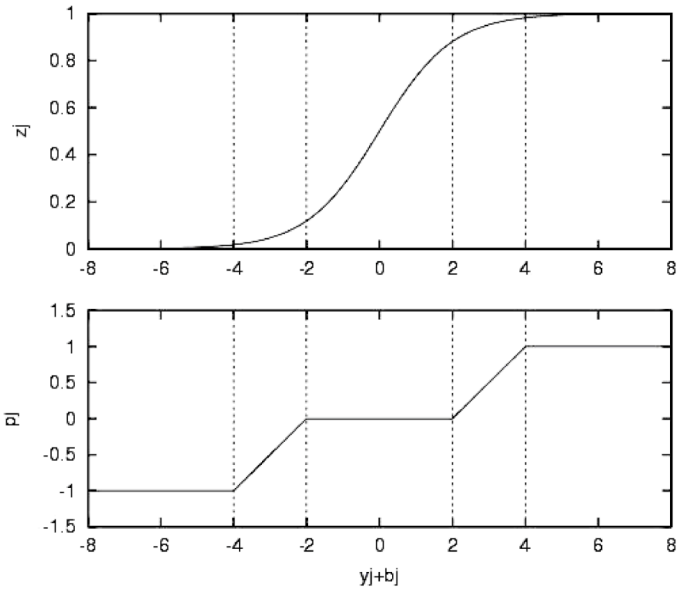
$$z_j = \frac{1}{1 + \exp[-(y_j + b_j)]}$$

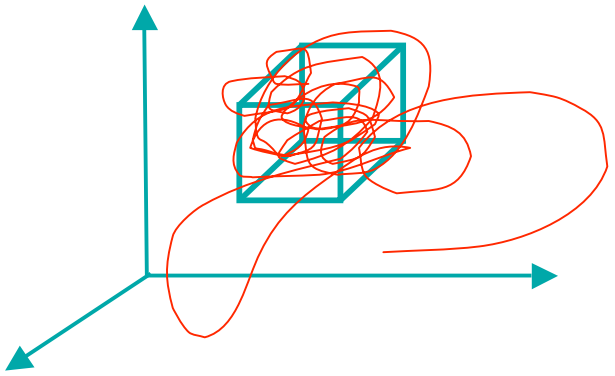
Evolved plastic rules

$$\begin{aligned} R0 : \Delta w_{ij} &= \delta\, \eta_{ij}\, p_j\, z_i z_j, \\ R1 : \Delta w_{ij} &= \delta\, \eta_{ij}\, p_j\, (z_i - z^o_{ij}) z_j, \\ R2 : \Delta w_{ij} &= \delta\, \eta_{ij}\, p_j\, z_i (z_j - z^o_{ij}), \\ R3 : \Delta w_{ij} &= 0, \end{aligned}$$

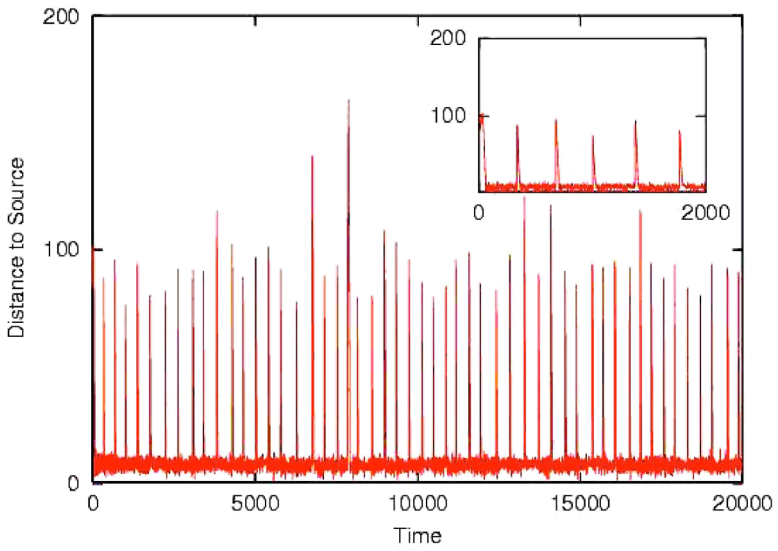
homeostatic adaptation

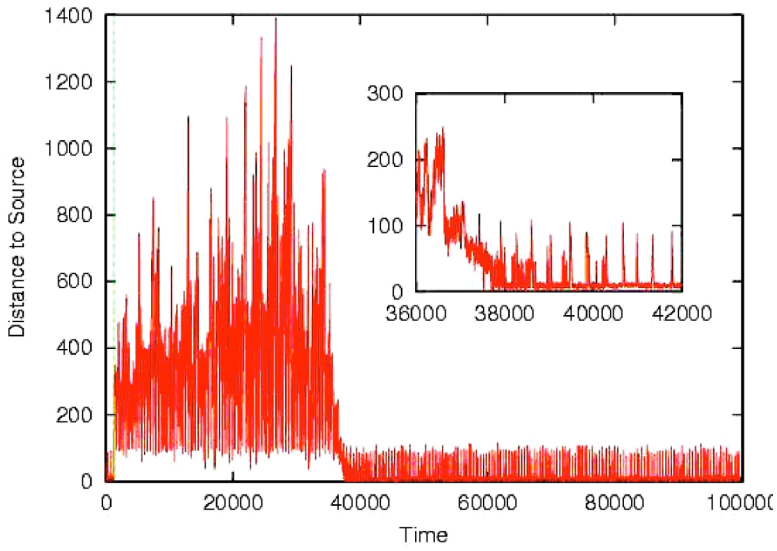
- Local plastic changes triggered by the activity of single neurons acting as essential variables, (Di Paolo, 2000).
- Select only for the desired behaviour and maximum internal stability. Test for adaptation to different perturbations.





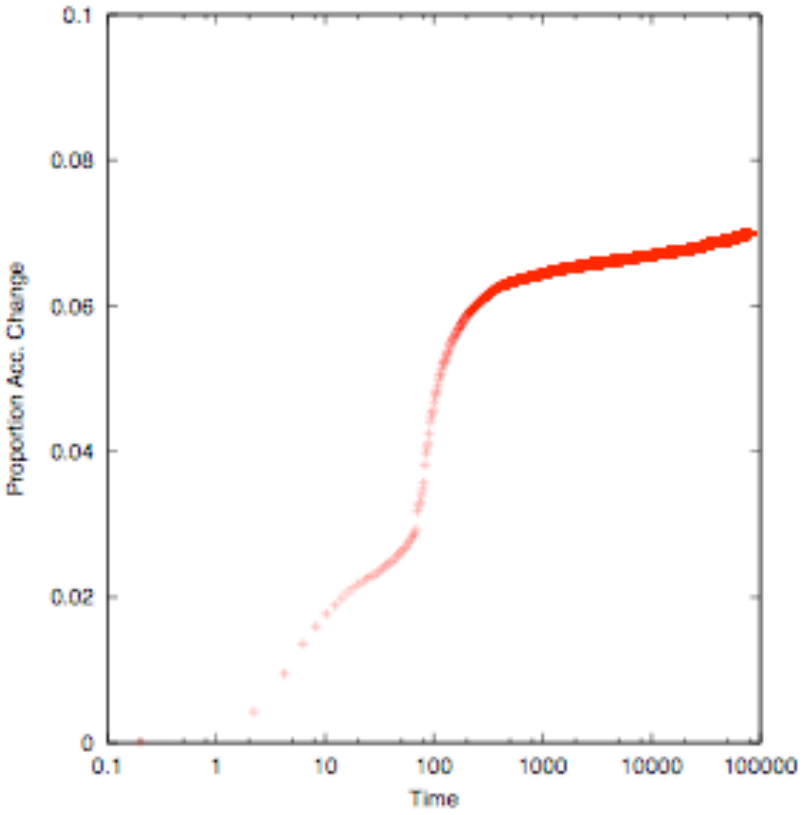
homeostatic adaptation



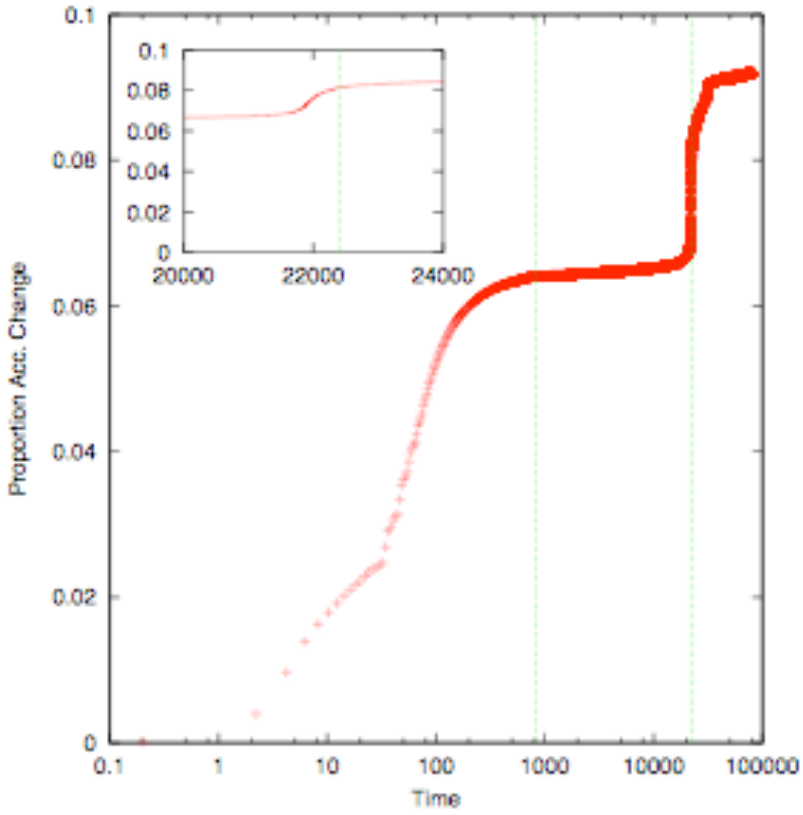


- Adaptation to left-right visual inversion
- Replicated on Khepera (J. Bird, Sussex) and applied to legged locomotion tasks (C. Linder, Bielefeld)

punctuated plastic change



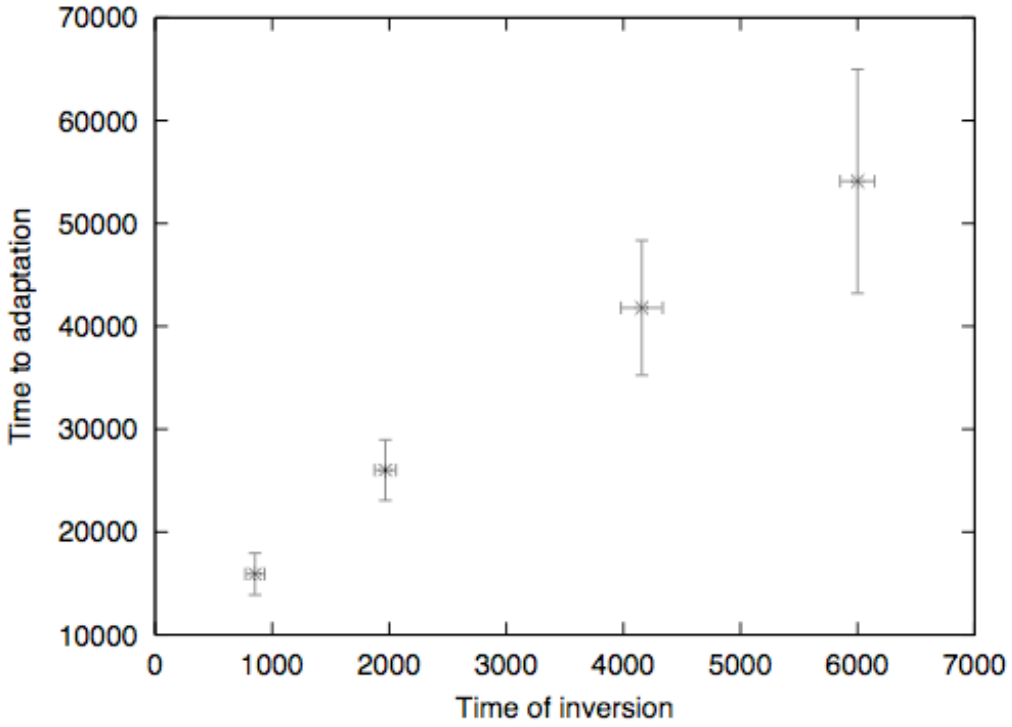
Normal

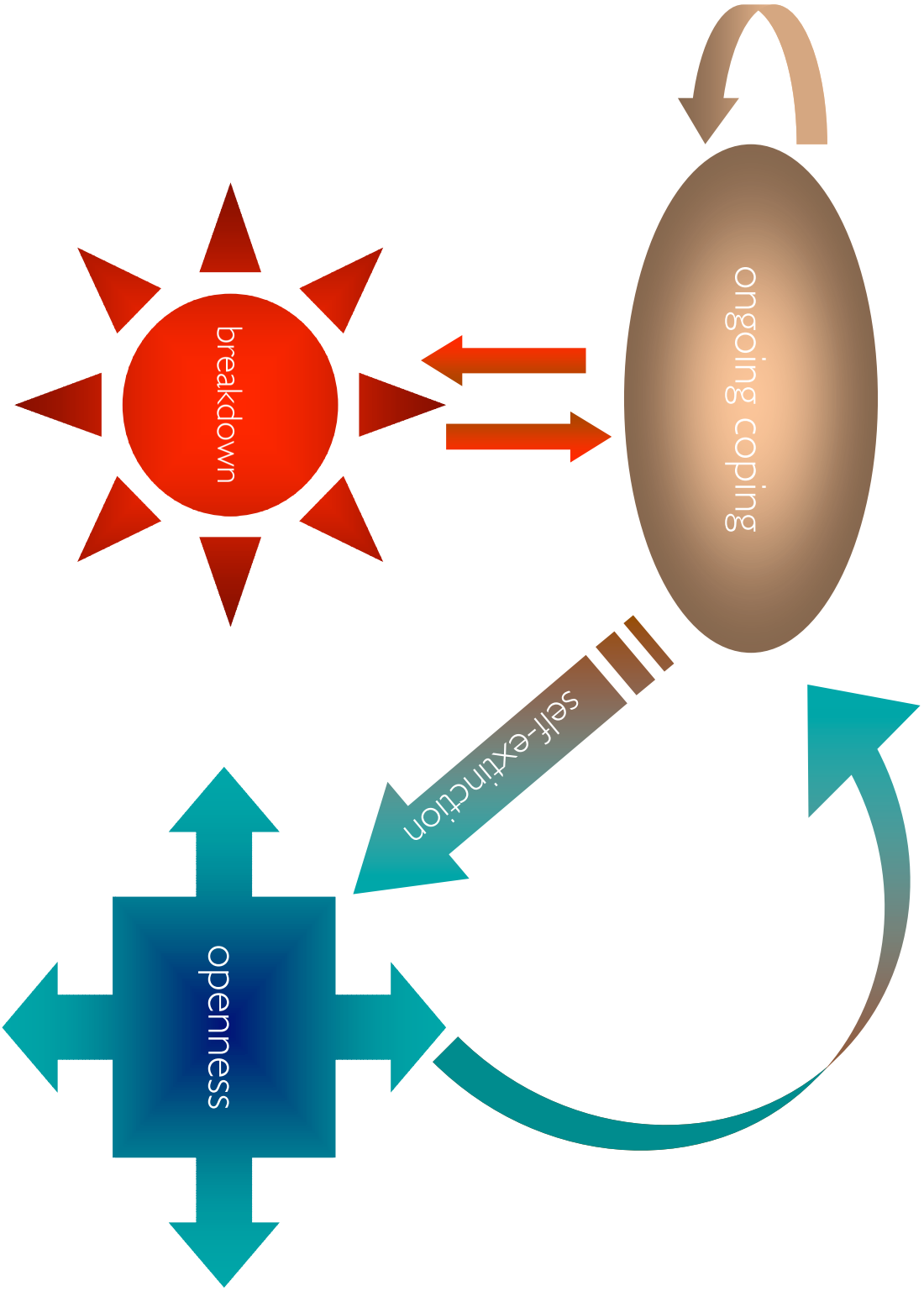


Inverted

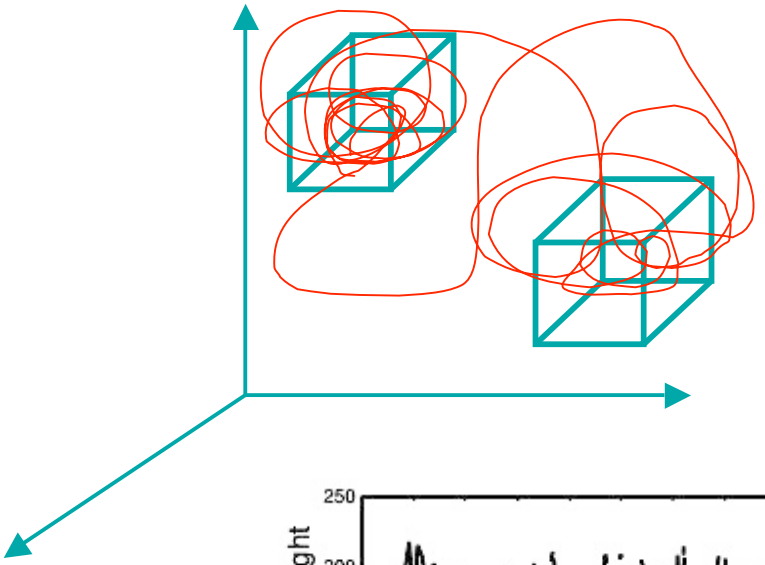
critical periods

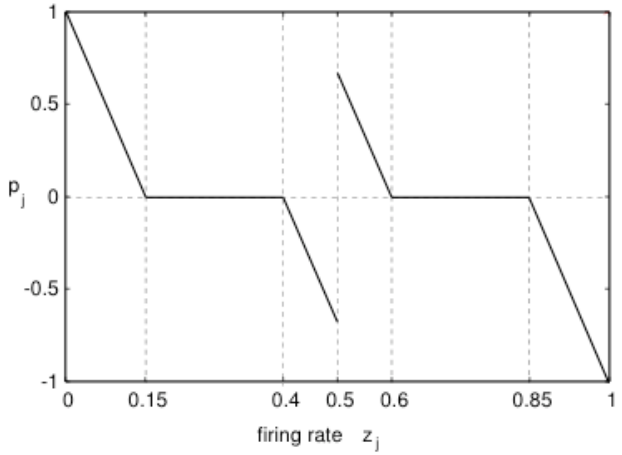
- Plasticity is found to decrease with age.
- This is not a feature that is selected for.

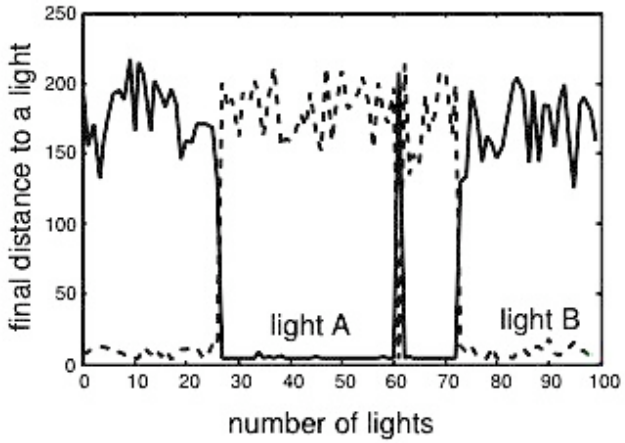


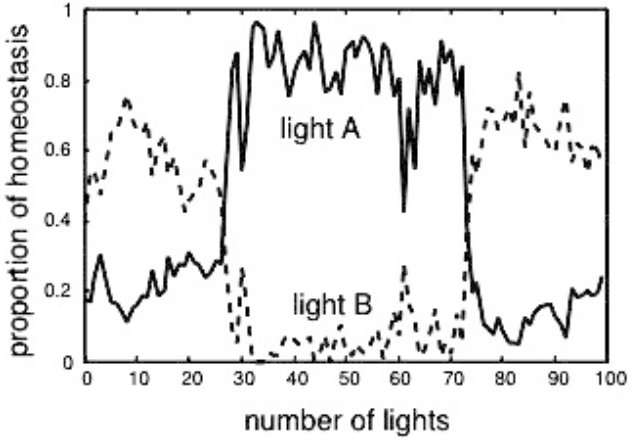

breakdown
ongoing coping
self-extinction
openness

preferential switching



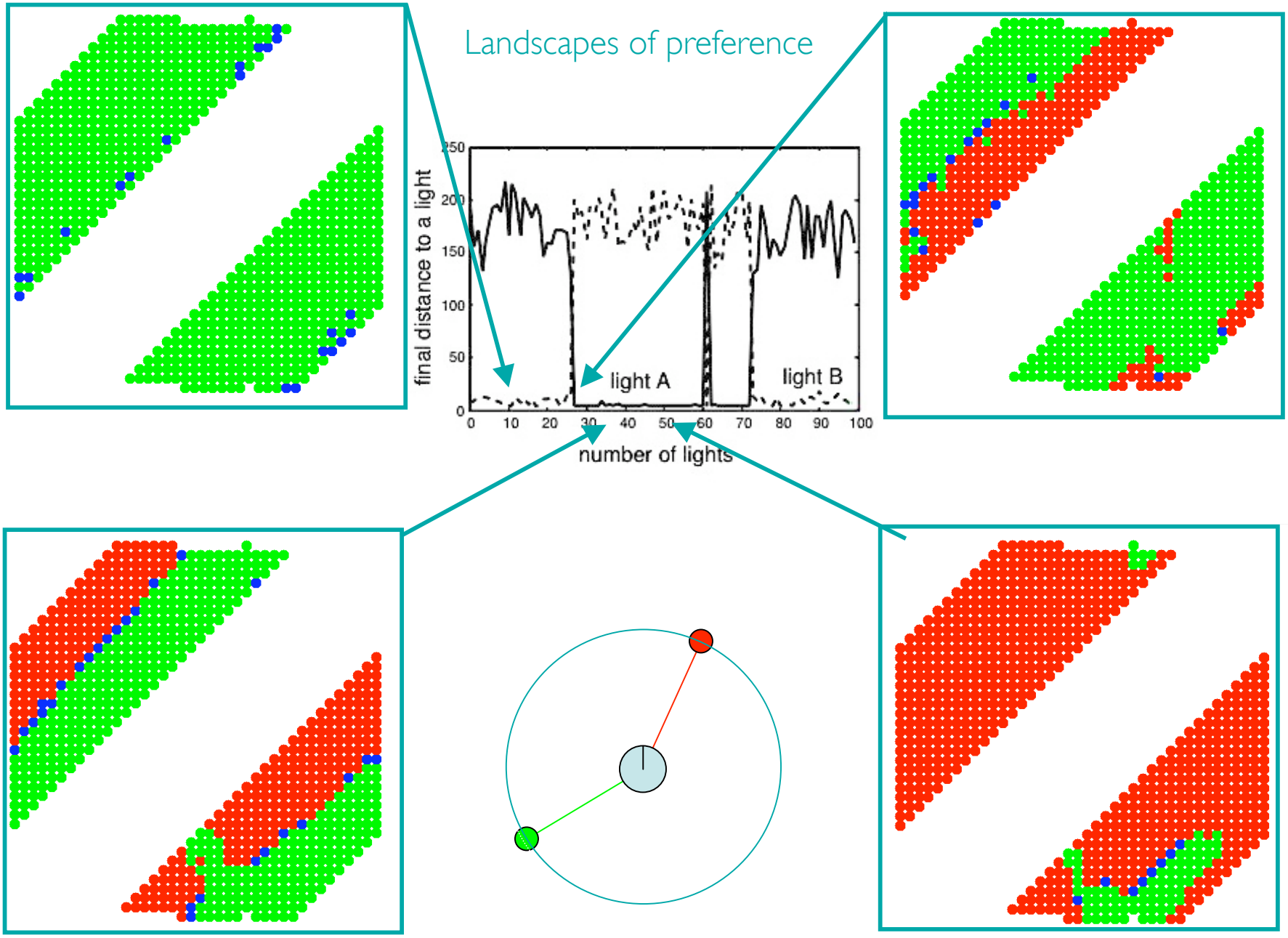




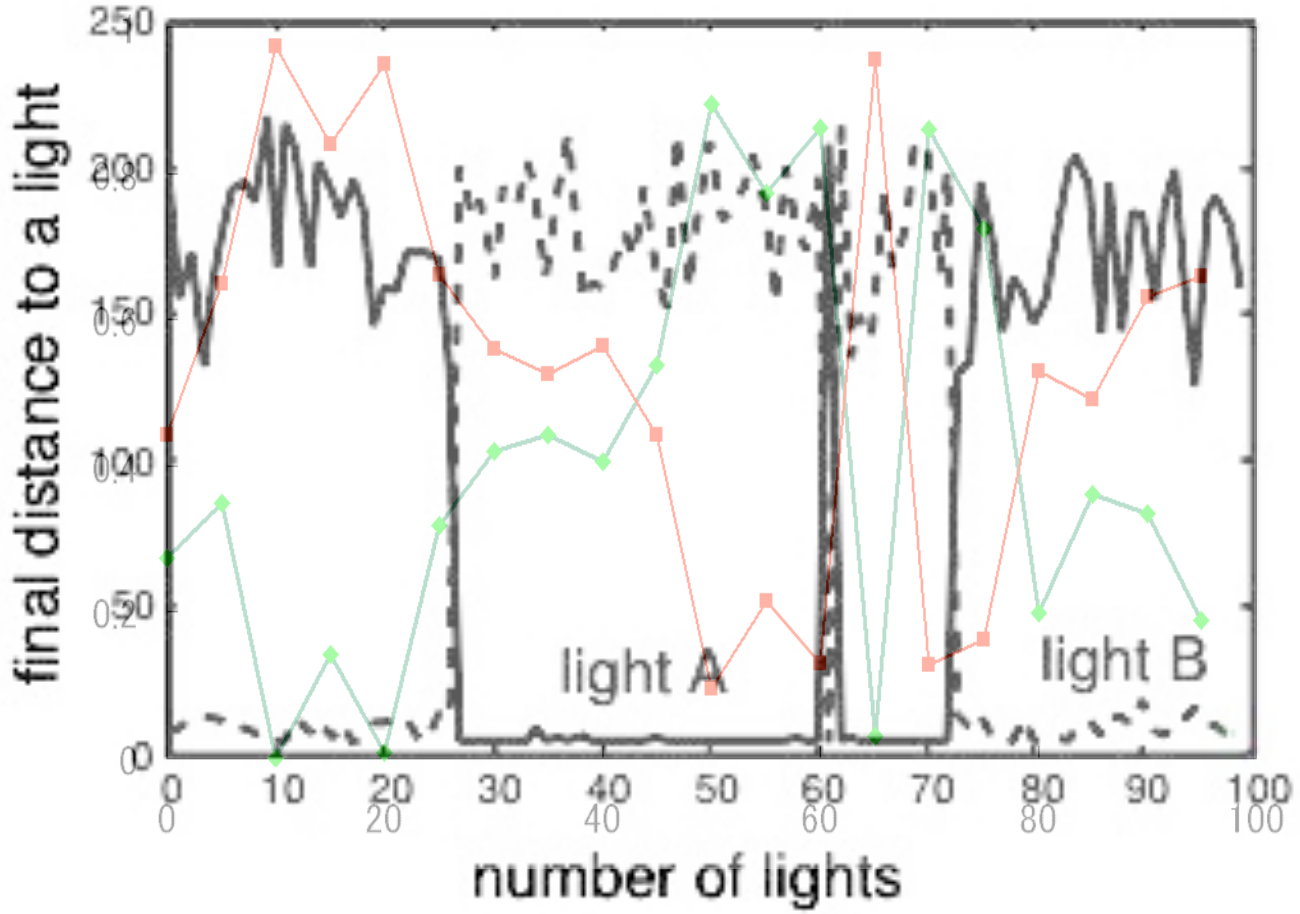


Work in collaboration with Hiro Iizuka

Landscapes of preference

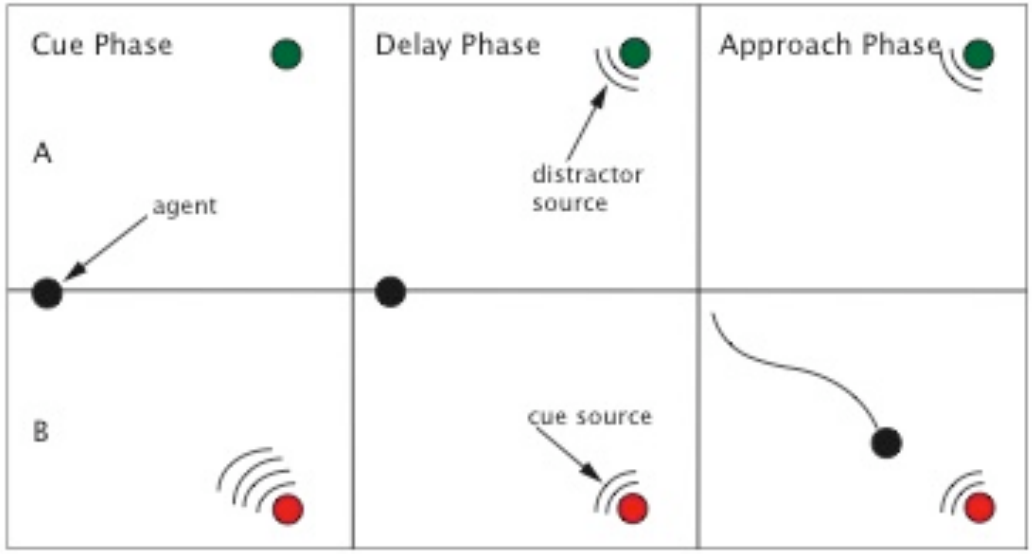


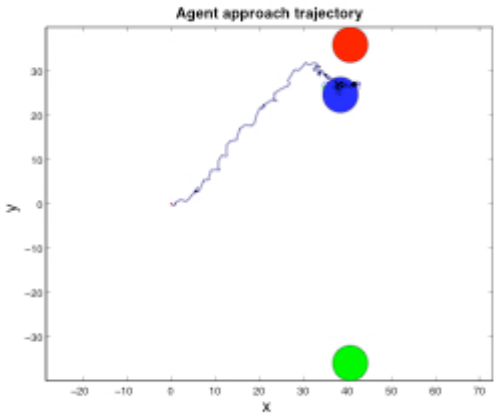
Ezequiel DI PAOLO - Ecole thématique Enaction 2007

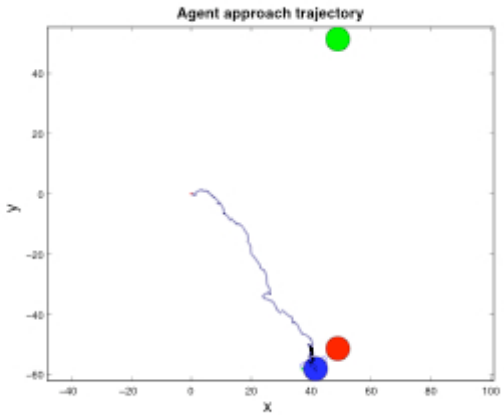

final distance to a light
number of lights
light A
light B
0
10
20
30
40
50
60
70
80
90
100
250
200
150
100
50
0

modelling a-not-b error

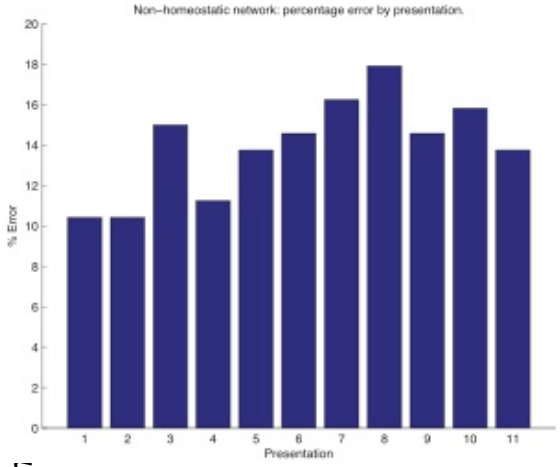
- Well studied experimental paradigm whereby infants of about 12 months mistakenly reach for a toy they have seen being hidden behind a different location (Piaget, Thelen & Smith).
- Work in collaboration with Rachel Wood (Wood & Di Paolo, 2007)

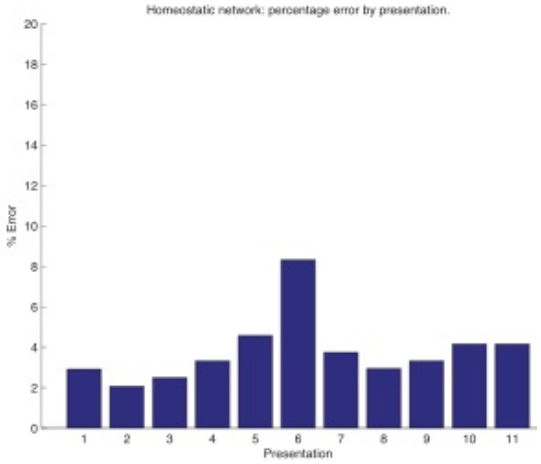


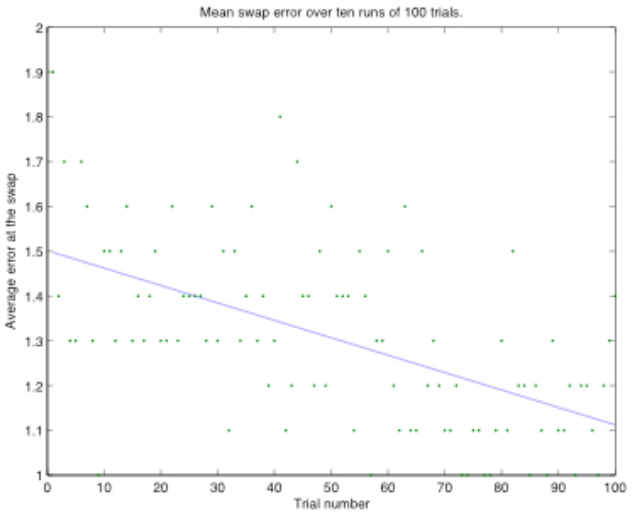




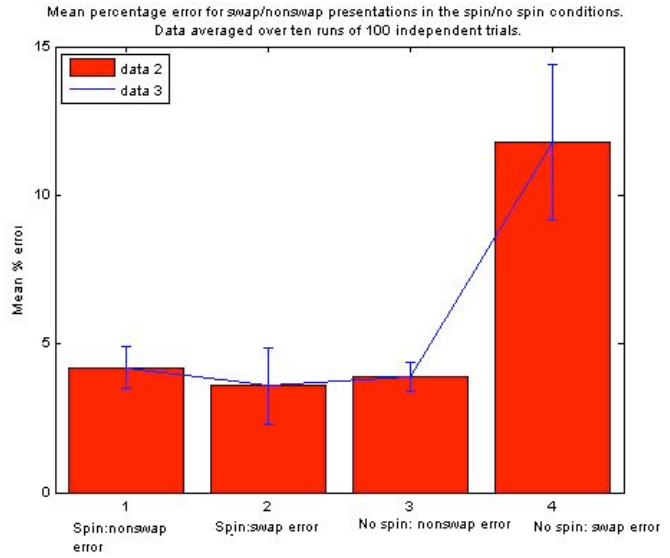
modelling a-not-b error







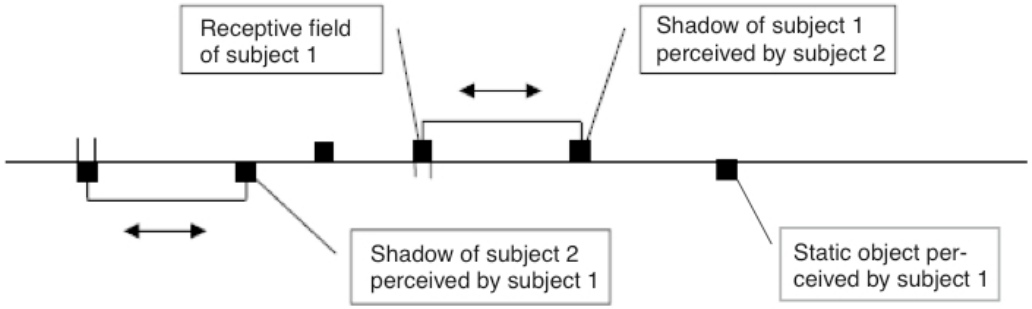
- Adding homeostatic regulation (again 2 boxes as in preference model) replicates pattern of error found in infants.
- "Shake" the agent before the swap and the swap error doesn't happen.



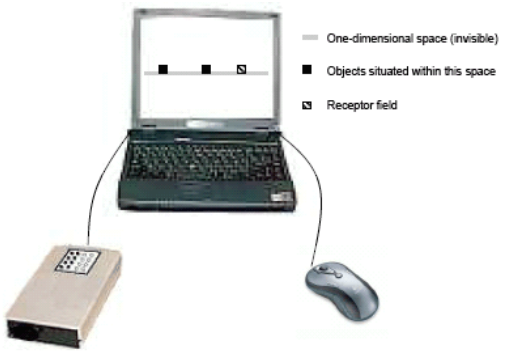

How is meaning created and transformed in interaction?
participatory sense-making

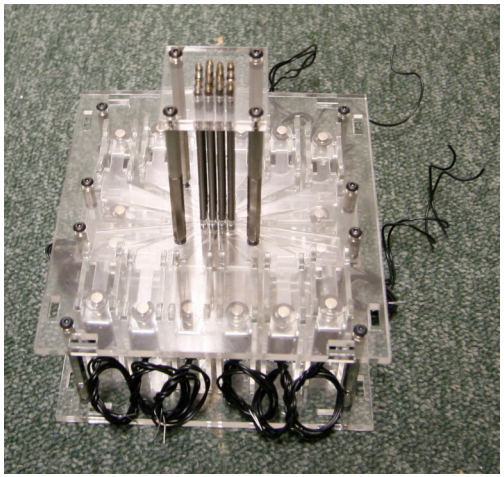
modelling perceptual crossing

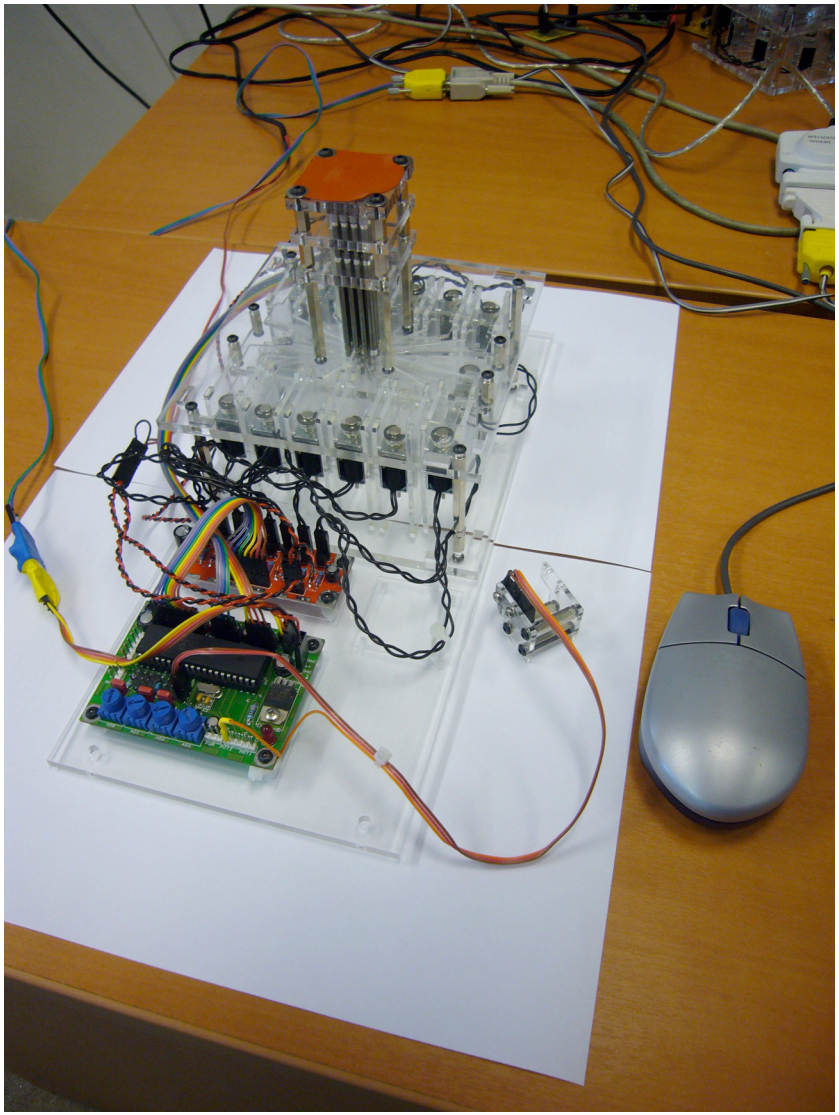
- Experiments using perceptual supplementation by Charles Lenay, John Stewart and Olivier Gapenne et al. at Compiegne

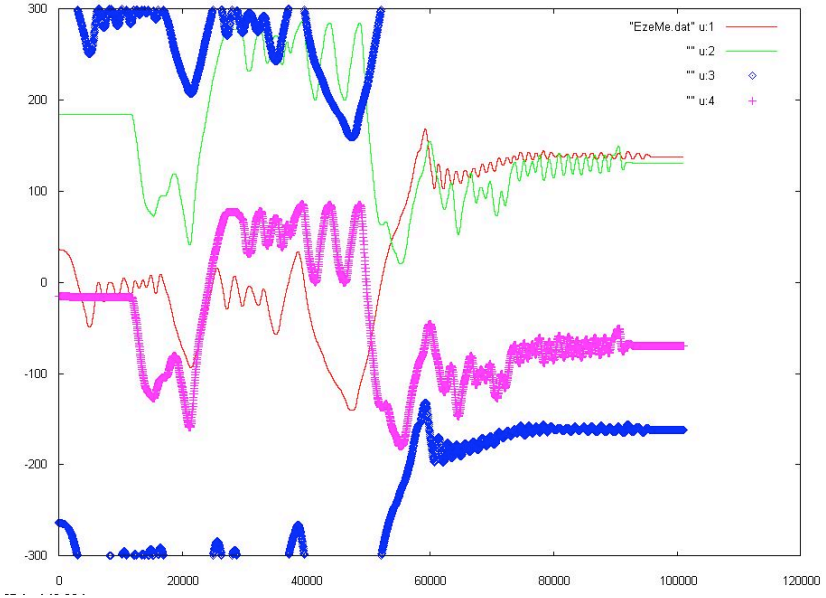


- Modelled at Sussex in collaboration with Marieke Rohde and Hiro Iizuka (*New Ideas in Psychology*, in press) (De Jaegher and Di Paolo, *Phenomenology and the Cognitive Sciences*, in press)

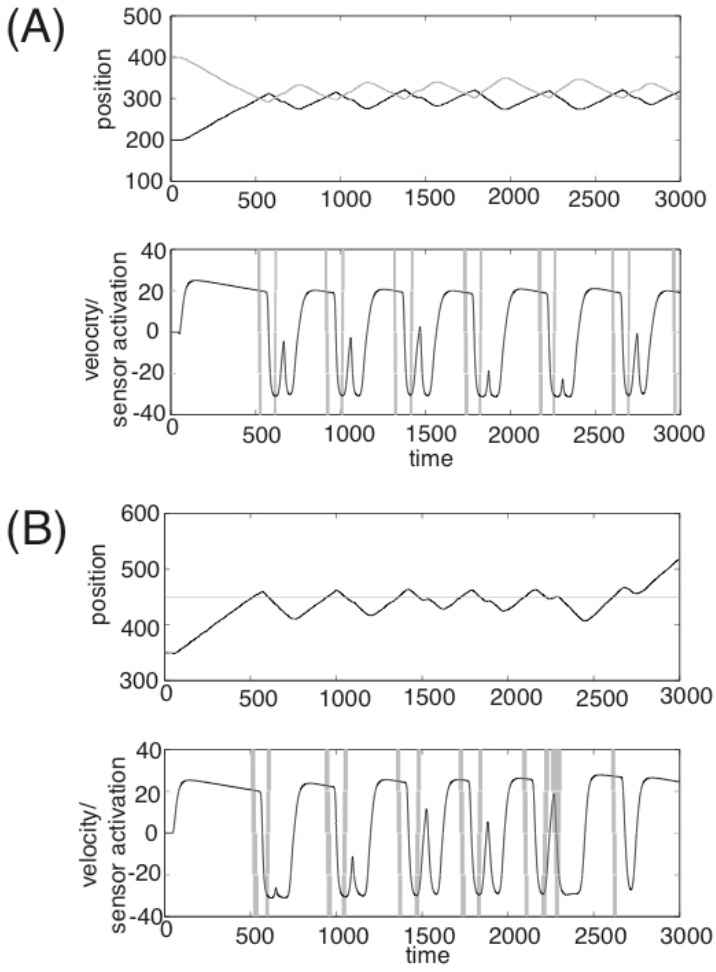


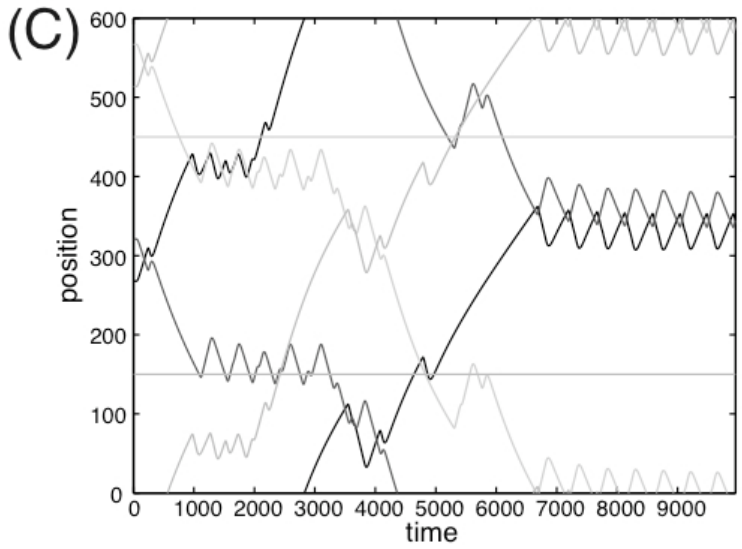


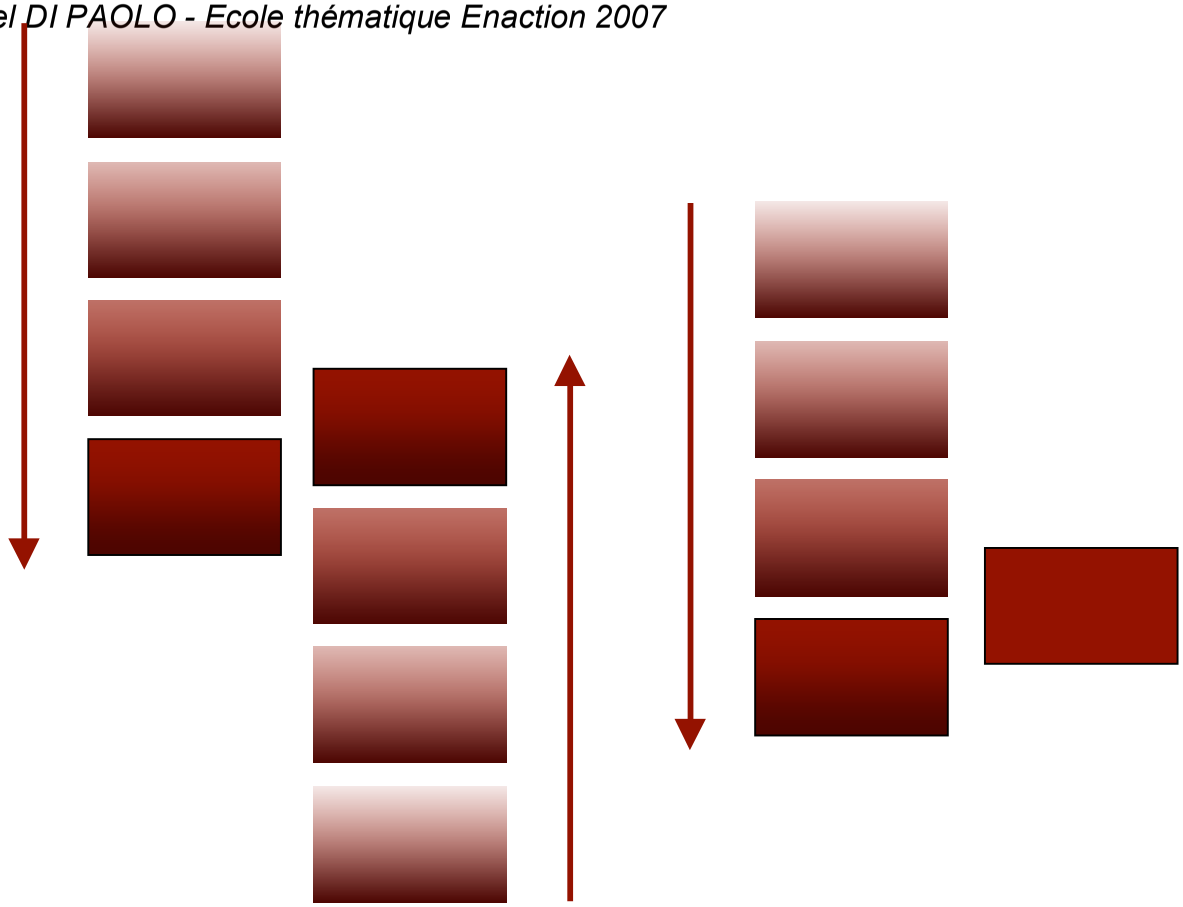


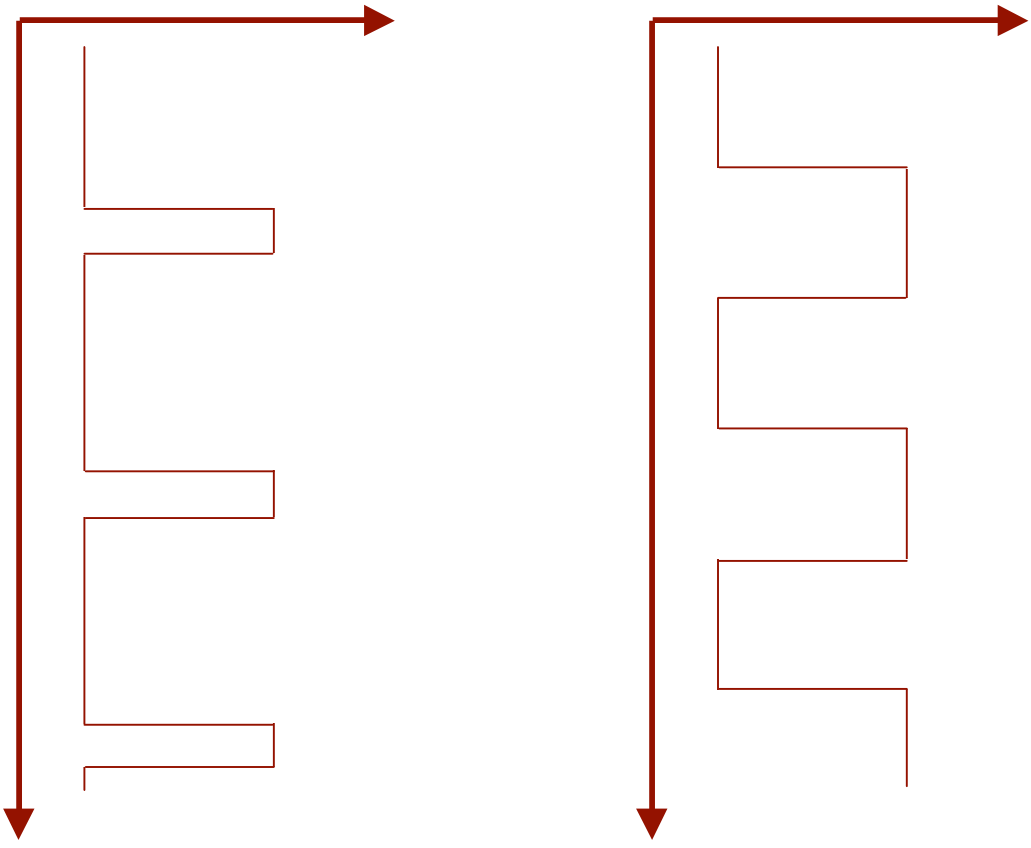

"EzeMe.dat" u:1
"" u:2
"" u:3
"" u:4
300
200
100
0
-100
-200
-300
0
20000
40000
60000
80000
100000
120000

modelling perceptual crossing









detecting social contingency

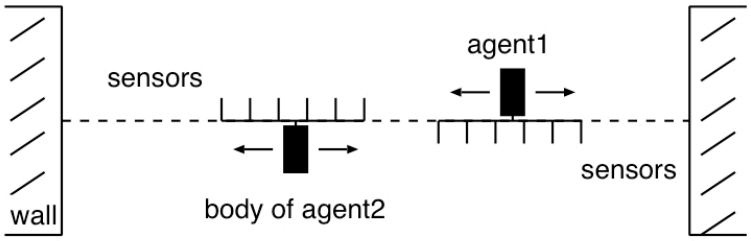
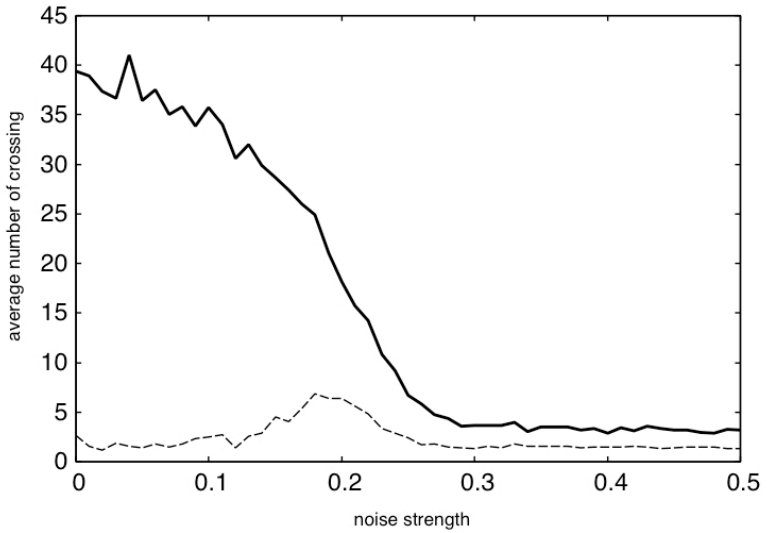
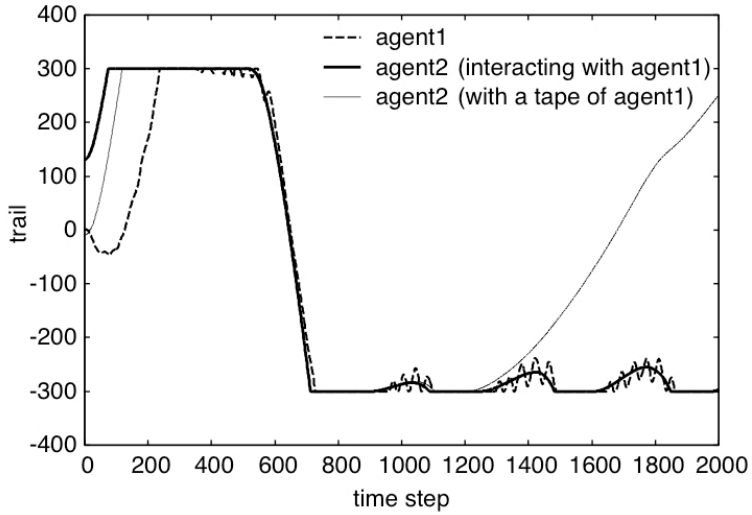
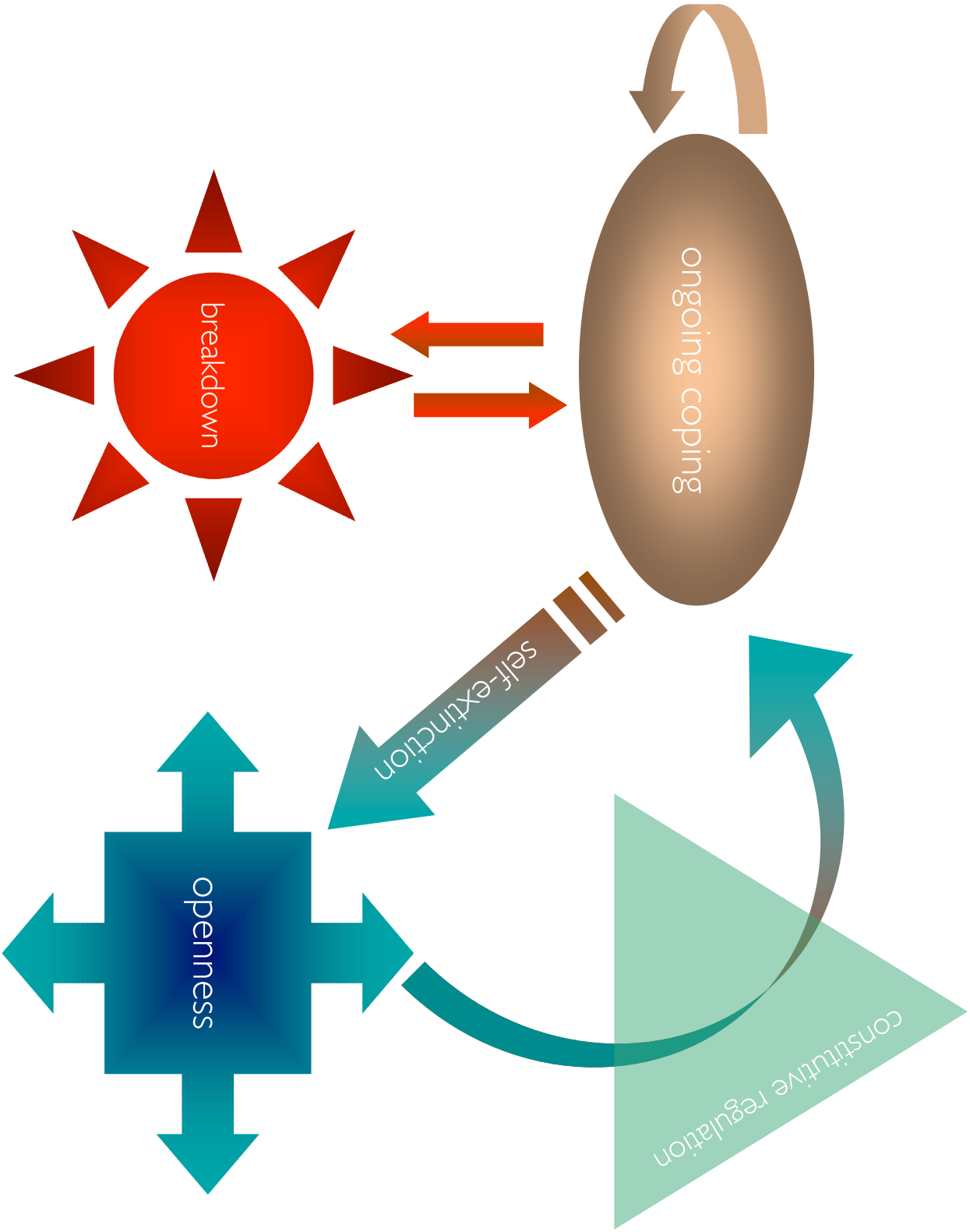


Figure 4.





- Social contingency detection is *not* done by the individuals but by the collective dynamics. Role of noise.


breakdown
ongoing coping
self-extinction
openness
constitutive regulation

conclusion

conclusion

clusion

ER; a tool for thinking

- a powerful tool to test pre-conceptions and elaborate ideas that are initially half-baked.
- It results in concrete systems that work and generate novel intuitions.
- Models of adaptation to unforeseen body perturbations, goal-generation, perseverative reaching and detection of social contingency.
- Interesting consequence of using ER is in the questions that are open by this practice: what is an agent? What is an intention in dynamical terms? The combination of enactive ideas and modelling tools can fuel a paradigm change in cognitive science and AI.

autonomy

- makes enaction a truly novel paradigm
- can be formalized
- comes in degrees, levels, and jumps
- relates to identity generation
- is not a function or a component
- is not about internal or external causation
- can be approached by modelling (e.g., preference, social contingency)

thank you

- De Jaegher, H, and Di Paolo, E. A., (in press) Participatory sense-making: an enactive approach to social cognition. *Phenomenology and the Cognitive Sciences.*
- Di Paolo, E. A. (2005) Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4, 429 - 452.
- Di Paolo, E. A. and Iizuka, H. (in press) How (not) to model autonomous behaviour. *BioSystems*.
- Di Paolo, E. A., Rohde, M. and De Jaegher, H. (2007) Horizons for the enactive mind, in: Gapenne, O. Stewart, J. and Di Paolo E. A. (Eds) *Enaction: Toward a new paradigm for cognitive science*. MIT Press.
- Di Paolo, E. A., Rohde, M. and Iizuka, H. (in press). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology*.
- Iizuka, H. and Di Paolo, E. A. (in press) Toward Spinozist robotics: Exploring the dynamics of minimal behavioural preference. *Adaptive Behavior.*
- Rohde, M. and Di Paolo, E. A. (in preparation) *The problem with semantic drift: A close inspection of dedicated learning mechanisms in evolutionary robotics.*
- Wood, R. and Di Paolo, E. A. (2007) New models for old questions: Evolutionary robotics and the 'A not B' error. *ECAL 2007*, Lisbon, Portugal.

http://www.informatics.sussex.ac.uk/users/ezequiel/