Modeling the Evolution of Decision Rules in the Human Brain

Daniel S. Levine
Department of Psychology
University of Texas at Arlington
Arlington, TX 76019-0528
levine@uta.edu
www.uta.edu/psychology/faculty/levine

(Most of this work appears in Levine, D. S., Angels, devils, and censors in the brain, *ComPlexus*, in press.)

Selfishness vs. Cooperation

Eisler and Levine (*Brain and Mind*, 2002):

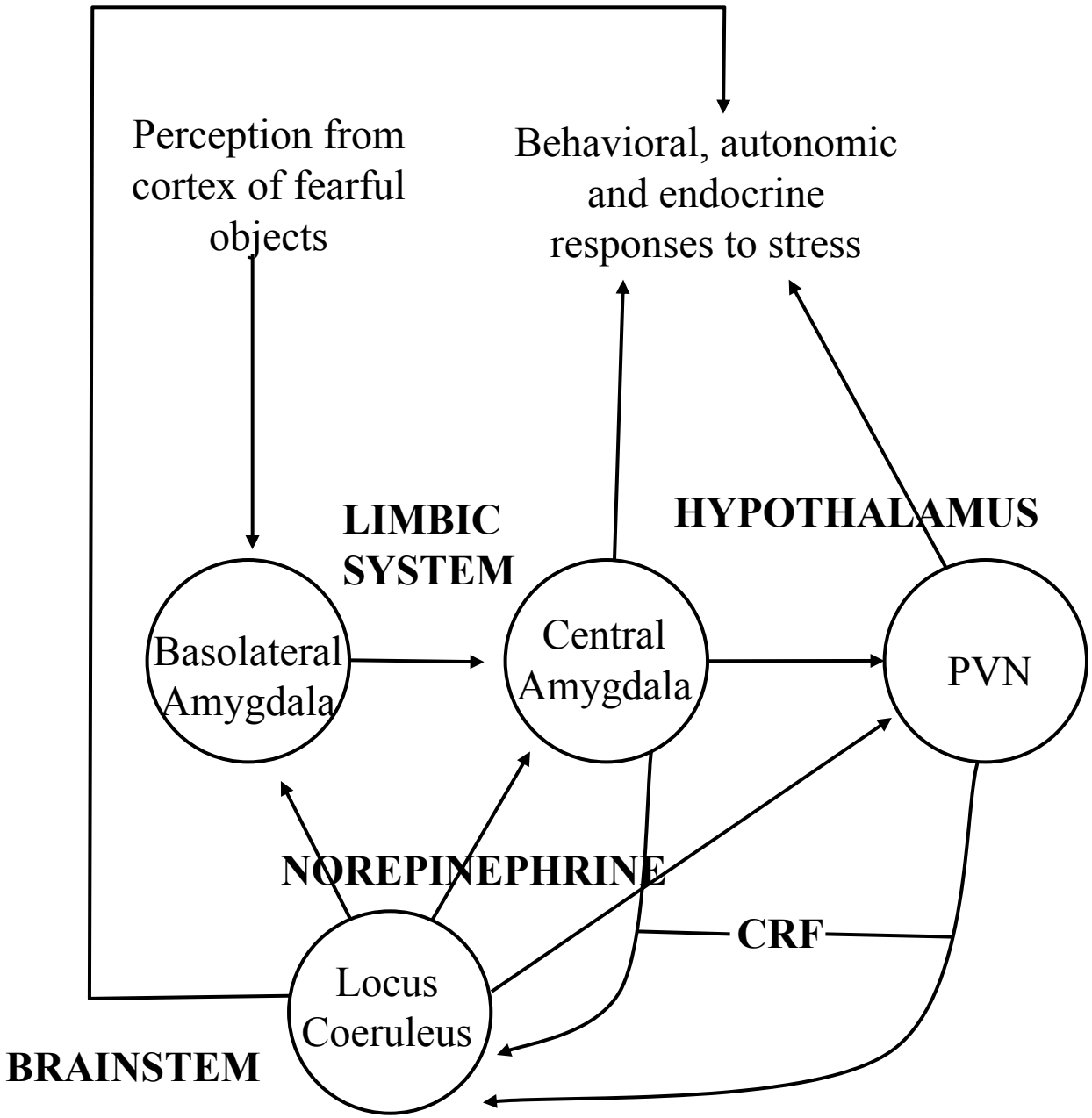
Cortical-subcortical neural pathways for behavioral patterns of

- Fight-or-flight
- Dissociation
- Bonding (tend-and-befriend)

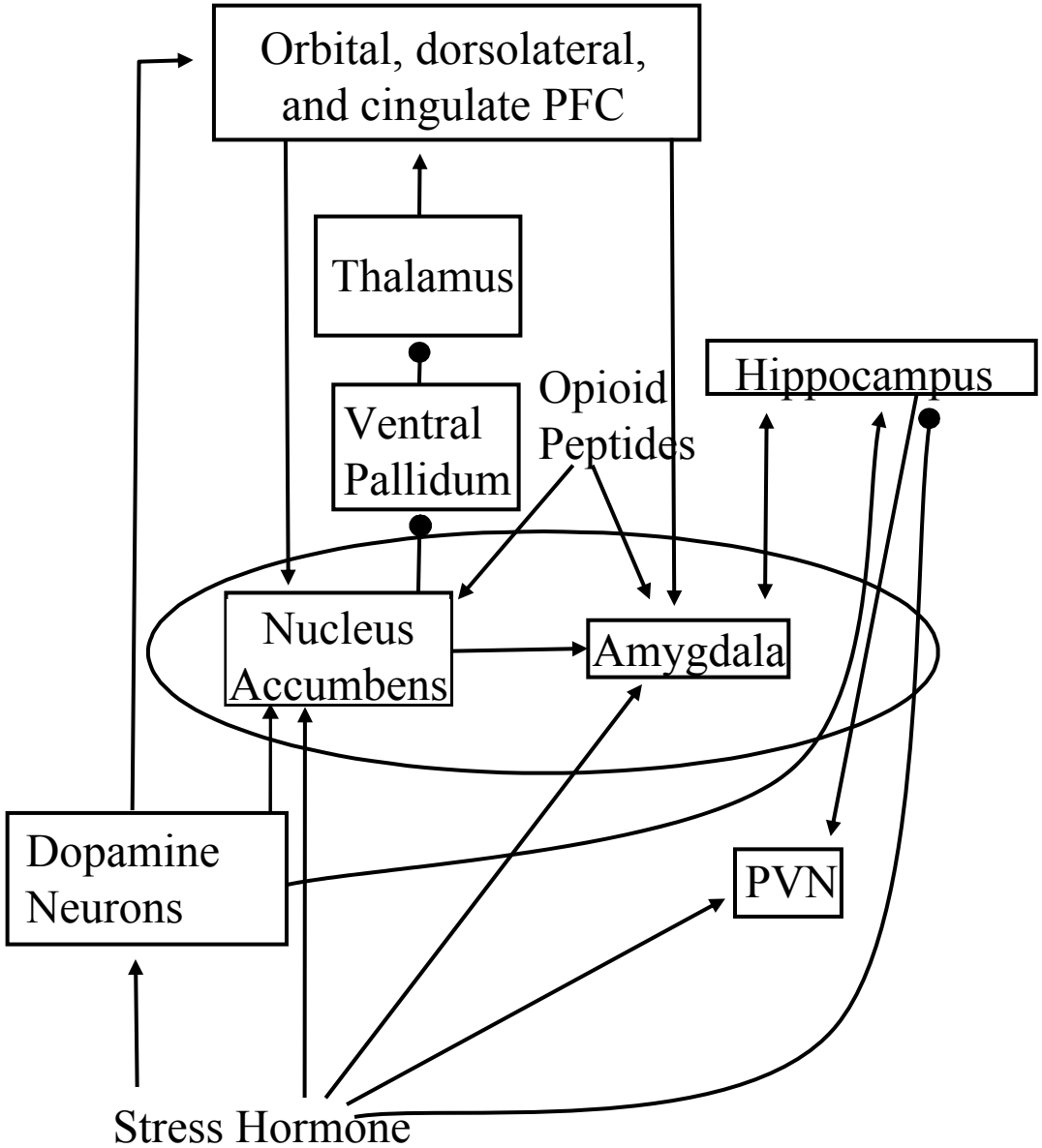
Orbital prefrontal cortex is main area for deciding between these patterns based on context.

NATURE *AND* NURTURE!

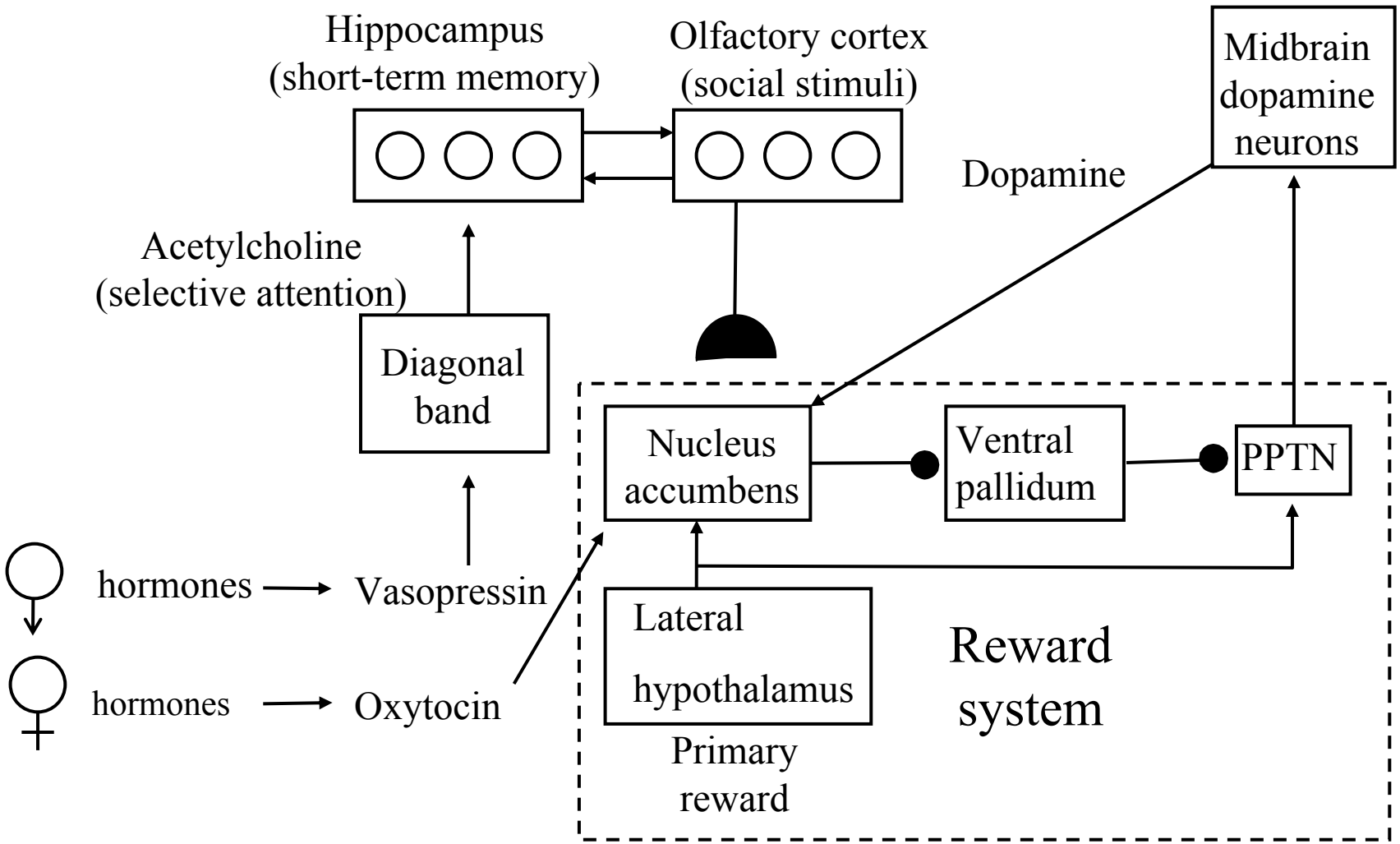
Possible fight-or-flight network



Possible dissociation network



Possible tend-and-befriend network



The Orbitomedial Prefrontal Cortex and Choice

19th century patient Phineas Gage lost the ability to make plans and appropriate social responses after being injured in the orbitofrontal cortex by a railroad accident in which an iron rod went through his cheek and out the top of his head.

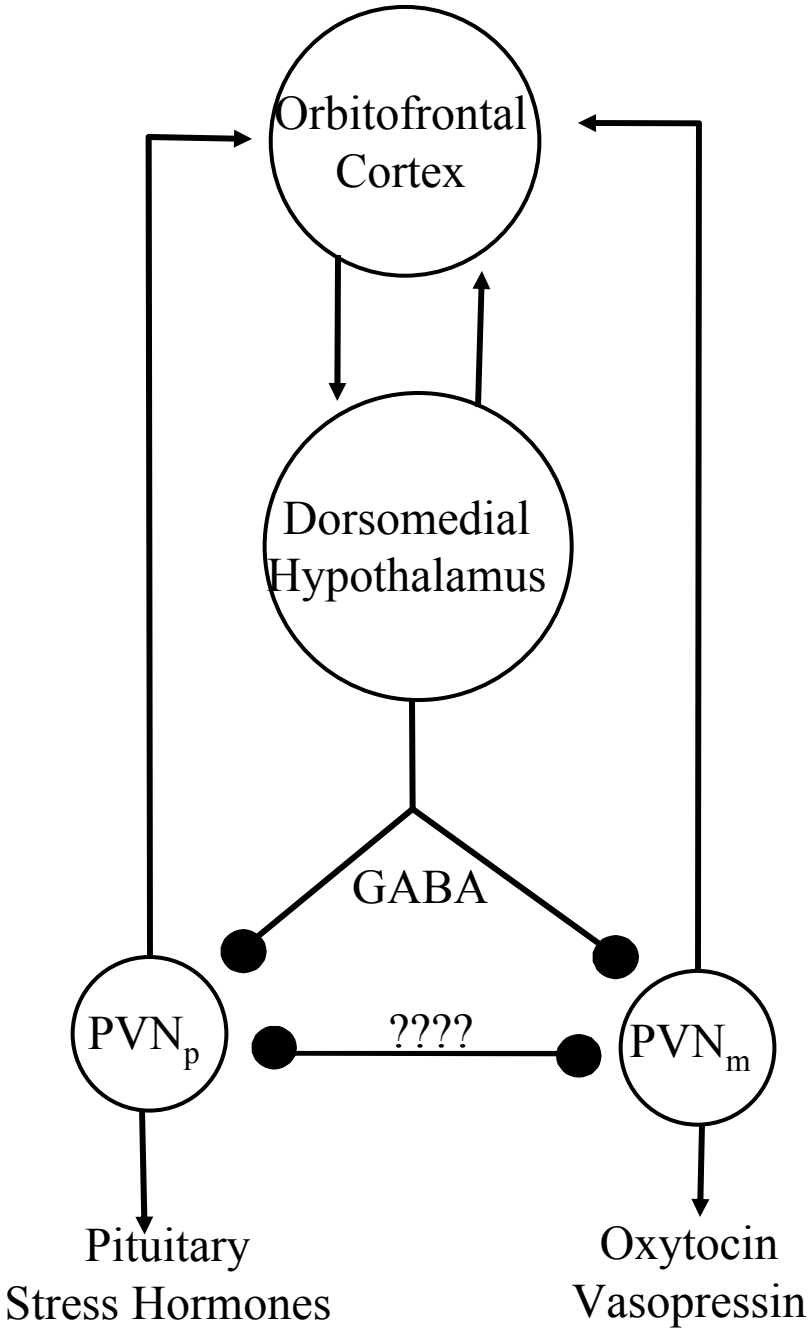
From Gage's case and other patient studies (Damasio, 1994) and animal lesion studies, neuroscientists believe orbitofrontal cortex forms and sustains mental linkages between specific sensory events in the environment (e.g., people or social structures) — and positive or negative affective states.

This region creates such linkages via connections between neural activity patterns in the sensory cortex that reflect past sensory events, and other neural activity patterns in subcortical regions that reflect emotional states

How might OFC mediate activation of large classes of responses?

Orbitofrontal connects reciprocally with a part of hypothalamus called the *paraventricular nucleus* (PVN).

Different parts of PVN contain various hormones including oxytocin, vasopressin, and CRF, the precursor of the stress hormone cortisol. Orbitofrontal synapses onto an area called the *dorsomedial hypothalamus* that sends inhibitory neurons to PVN that are mediated by the inhibitory transmitter *GABA* (*gamma-amino butyric acid*), This influences selective activation of one or another PVN hormone-producing subregion (picture on next slide).

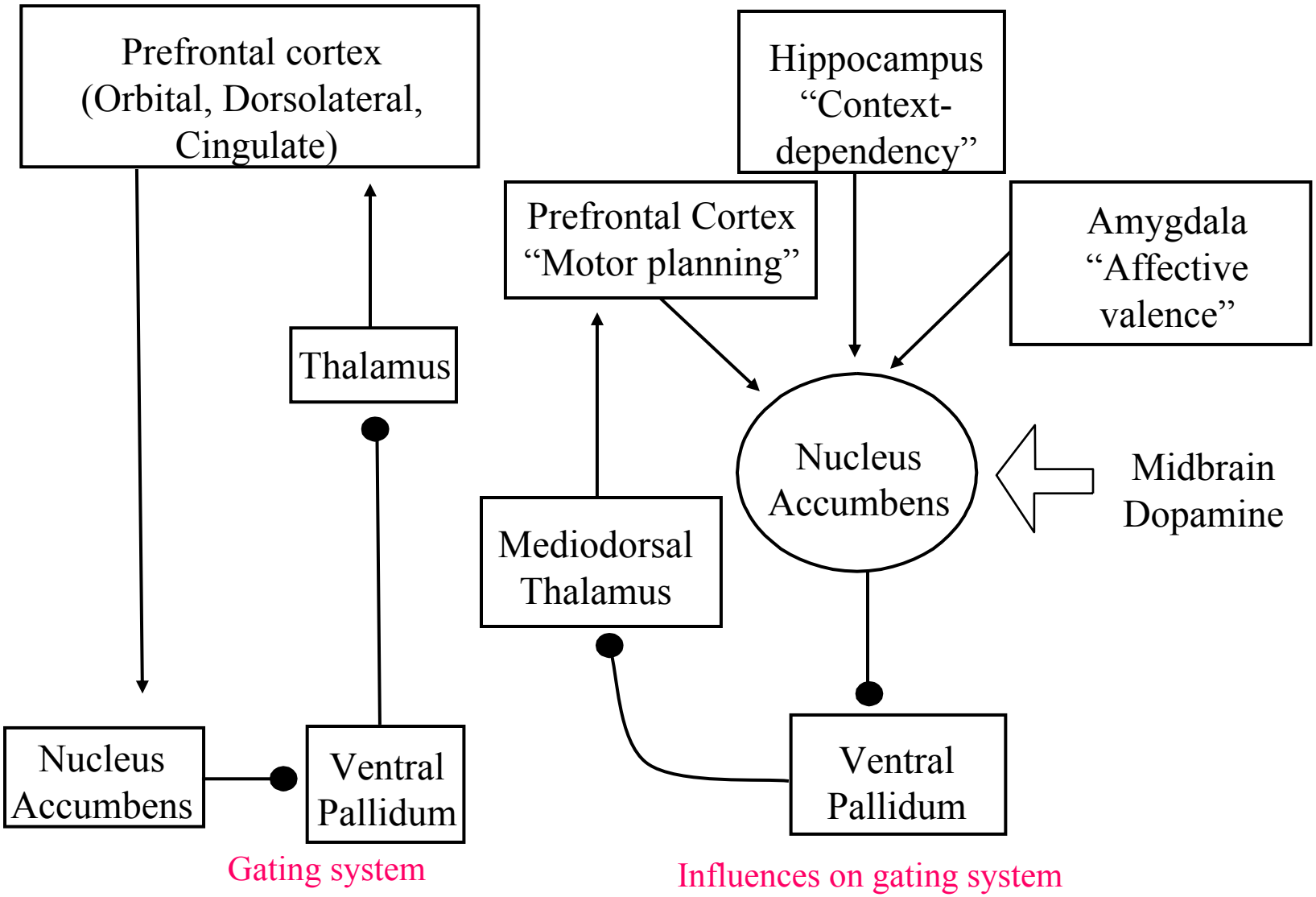

Orbitofrontal Cortex
Dorsomedial Hypothalamus
GABA
PVNp
????
PVNm
Pituitary Stress Hormones
Oxytocin Vasopressin

But how do context and personality affect these choices?

A mechanism is still needed to translate positive and negative emotional linkages into action tendencies or avoidances (the “angels” and “devils” of my article).

Gating system in pathways between the prefrontal cortex, basal ganglia, and thalamus (Frank, Loughry, & O’Reilly, 2002). Link from basal ganglia to thalamus disinhibits (based on contextual signals) performance of actions whose representations are usually suppressed.

Gating system



Personality as a Dynamical System

Cloninger (1999):

Components of

CHARACTER (largely developed)

and

TEMPERAMENT (largely inherited)

Character:

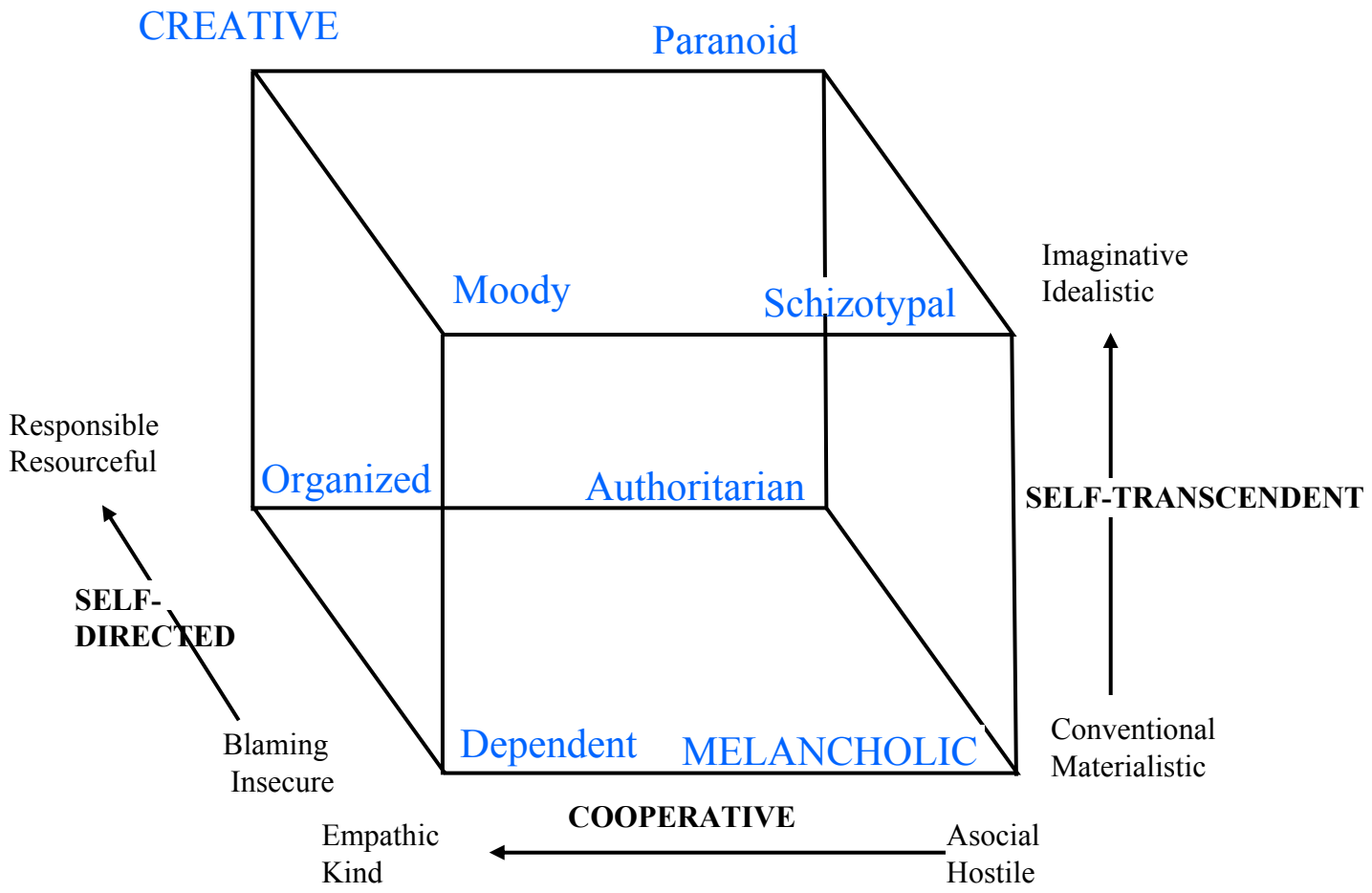
- Self-directedness (acceptance of the self)
- Cooperativeness (acceptance of other people)
- Self-transcendence (acceptance of nature)

Temperament:

- Novelty-seeking
- Harm-avoidance
- Reward-dependence
- Persistence

Character Cube

Reproduced with permission from Center for Psychobiology of Personality, Washington U., St. Louis



Dynamical system description

Each corner of the cube is an ATTRACTOR for the dynamical system of personality.

Cloninger describes the attractors as points with 0 and 1 values for his three character dimensions (creativity is (1, 1, 1), moodiness is (0, 1, 1), melancholia is (0, 0, 0), et cetera). Yet each attractor is really a different **state of a high-dimensional system representing connection strengths at many brain loci.**

What is the Goal of Psychotherapy?

To move the individual from other attractors toward the creative attractor.

Switches from less to more optimal states have been described in neural networks by SIMULATED ANNEALING (Kirkpatrick, Gelatt, & Vecchi, 1983; Hinton & Sejnowski, 1986; Levine, 1994).

Levine's (1994) Network Theory of Self-actualization

$$\frac{dx_i}{dt} = a_i(x_i)[b_i(x_i) - \sum_{k=1}^{n} c_{ik} d_k(x_k)]$$

Cohen-Grossberg equations for a competitive neural network: Each x_i excites itself, inhibits the others.

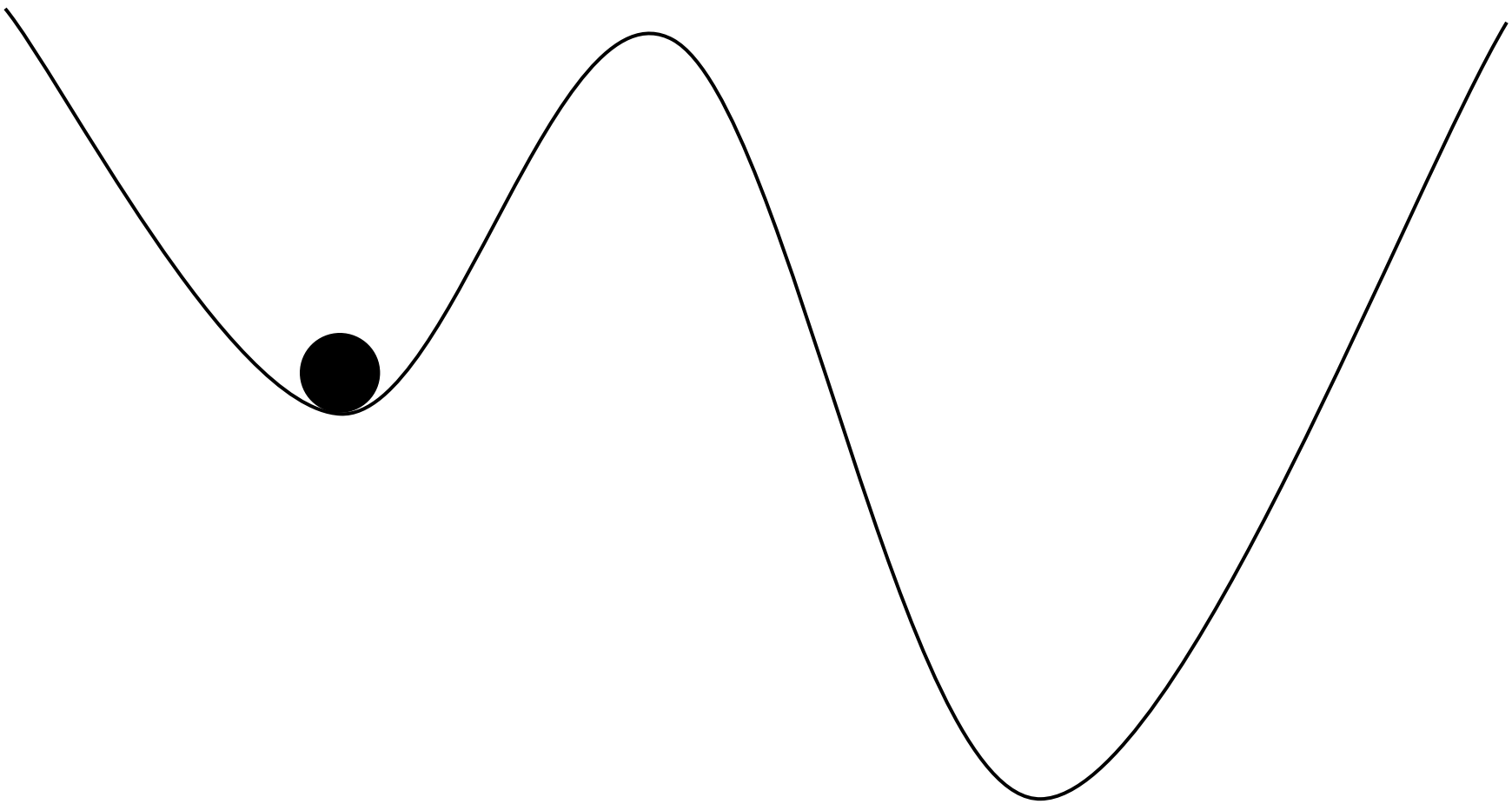
As time increases, the system always goes to a steady state (point attractor) because there is a system energy function or Lyapunov function, called V, that decreases along trajectories.

Now what does that theorem mean for decision making?

The system reaches a LOCAL minimum for V, but it may not be the GLOBAL minimum.

Kirkpatrick et al. (1983) and Hinton and Sejnowski (1986) interpreted GLOBAL minimum as OPTIMAL state.

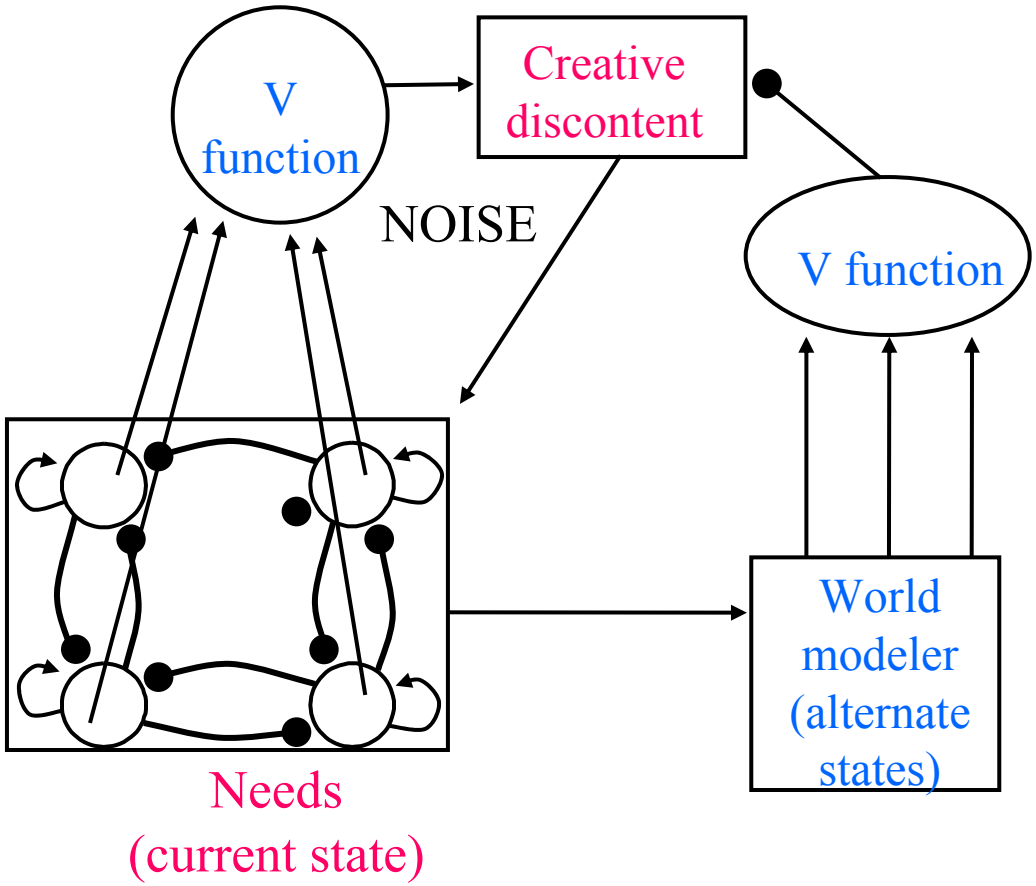
“Ball-bearing” analogy: systems (or people) can get trapped in local minima



Simulated Annealing (Noise)

Noise is added to the system to “shake” the “ball bearing” loose from the local minimum and get it to go toward the global minimum — that is, toward the Creative state!

Network: the "Needs" Module Satisfies Cohen-Grossberg



How would simulated annealing work in a continuous system?

Work in progress (Levine, Hardy, & Long):

Denote the right hand side of the Cohen-Grossberg equation,

$$a_i(x_i)[b_i(x_i) - \sum_{k=1}^{n} c_{ik} d_k(x_k)],$$

by $F_i(t)$. Let $\mathbf{x}_0$ be the optimal state and $\mathbf{x}$ be the current state. Let V be the Lyapunov function.

Then the "annealed" Cohen-Grossberg equations are

$$\frac{dx_i}{dt} = F_i(t) + \frac{\nu}{1 - \exp\left(\frac{-\max\left(\sum_{k=1}^{n} \frac{\partial V}{\partial x_k} F_k, 0\right)}{T} \right)}$$

ν is white noise (normally distributed with mean 0 and standard deviation 1);
the temperature is T = $(V(\mathbf{x}) - V(\mathbf{x}_0))\, N(t)$,
where $N(t)$, roughly labeled "initiative," can vary with mood or interpersonal context.

Can we combine all these network fragments?

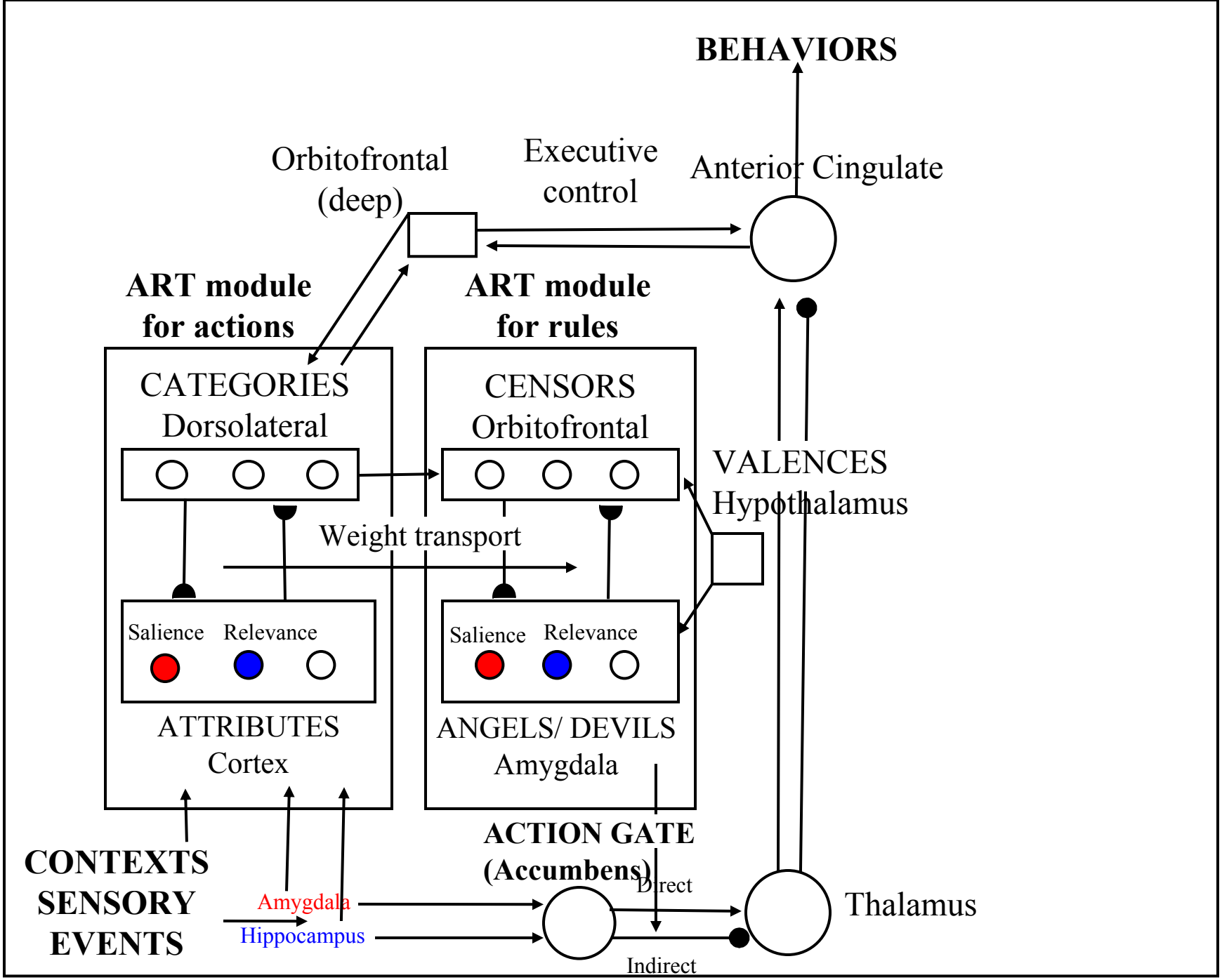
- "Angel behaviors" go through, and "devil behaviors" are actively barred from, nucleus accumbens gates.
- Hippocampus activates representation of current context, which in turn activates angel and devil representations relevant to that context.
- Longer-term storage of affective valences is likely to be at connections from **orbitofrontal cortex to amygdala** (Levine, Mills, & Estrada, IJCNN2005). Changes that affect behavior ("do" and "don't" instructions, approach toward or avoidance of an object) are likely to be at connections from **amygdala to medial prefrontal cortex (incentive motivation)** and from **orbitofrontal to nucleus accumbens (habit).**

Levels of complexity of decision rules

- In human development (Cloninger), neural representations associated with positive or negative valence become gradually more complex. These representations are at **all areas of prefrontal cortex**.
- Dehaene and Changeux (1991): dorsolateral prefrontal is *generator of diversity*, that is, creator of different possible decision rules. Orbitofrontal affective circuits "censor" possible rules based on rewards and punishments received from following these rules (Nauta, 1971; Damasio, 1994). EACH CLONINGER CORNER IS A DIFFERENT CENSOR!
- But developmental changes toward more complex angels and devils are not always total or permanent. They may be reversed under stress, or may depend on a specific mood or context for their manifestation.

The Big Picture (not all of it, I'm sure!)

- What is the relationship between the neural representations of these censors and the neural representations of the specific angels and devils the censors comprise?
- What are the neural mechanisms by which stress leads to reversal of the simulated annealing process? That is, how does stress move the system away from the creative corner of Cloninger's cube and toward less adaptive attractors on other corners?


BEHAVIORS
Orbitofrontal
(deep)
Executive
control
Anterior Cingulate
ART module
for actions
ART module
for rules
CATEGORIES
Dorsolateral
CENSORS
Orbitofrontal
VALENCES
Hypothalamus
Weight transport
Salience
Relevance
Salience
Relevance
ATTRIBUTES
Cortex
ANGELS/ DEVILS
Amygdala
ACTION GATE
(Accumbens)
CONTEXTS
SENSORY
EVENTS
Amygdala
Hippocampus
Direct
Indirect
Thalamus