20 YEARS OF NEURAL NETWORKS: A PROMISING START, A BRILLIANT FUTURE

Stephen Grossberg
Department of Cognitive and Neural Systems
Center for Adaptive Systems
Center of Excellence for Learning in Education, Science, and Technology (CELEST)
Boston University

IJCNN’07
August 13, 2007

steve@bu.edu
http://www.cns.bu.edu/Profiles/Grossberg

WE HAVE A LOT TO CELEBRATE!

Our field includes some of the most exciting and revolutionary science and technology in the world today!

Biological neural networks

show how brains give rise to minds: "last great frontier"

solve the age-old mind/body problem

clarify mental disorders

new algorithms for biologically-inspired technology

Artificial neural networks

solutions to engineering and
technology problems that require increasingly
autonomous and adaptive control
and better man-machine interfaces

KNOW THE PAST
TO UNDERSTAND THE FUTURE

To better know where we may be going, we need to know how we got here, historically and scientifically

THIS TALK: some history
some new computational paradigms
design principles
mechanisms
that can stimulate a lot of future research

TWO ANNIVERSARIES TO CELEBRATE!

20 YEARS OF THIS NEURAL NETWORK CONFERENCE

1987 IEEE International Conference on Neural Networks, San Diego
1987 Neural Networks journal
1987 INNS
1988 INNS Annual Conference, Boston
1989 IJCNN: INNS + IEEE, Washington, DC

BUILT ON 6 YEARS OF CONFERENCES BEFORE THEM

1980 - 1983 - 1985 - 1986 - 1986 - 1986

For this history, see IJCNN'07 web site and posters

20 Year Celebration
http://www.ijcnn2007.org/anniversary.htm

Reflections on the founding of INNS and IJCNN
http://www.ijcnn2007.org/anniversary.htm

I987 AND 1988 WERE EXCITING YEARS!

I announced the formation of INNS at the end of my plenary talk at the 1987 IEEE ICNN meeting

During the 14 months that I was INNS President,
INNS grew by 200 members each month without saturation

When I gave my plenary talk at the 1988 INNS meeting, there were

3071 INNS members
38 countries
49 states of the USA
20% in life sciences
19% in information and computer sciences
27% in engineering sciences
2% in business
7% in other fields

This membership reflected the INNS goal to be an
interdisciplinary forum
for linking psychological, neurobiological, mathematical,
computational, engineering, and technological research goals

IEEE FIRST ANNUAL INTERNATIONAL CONFERENCE ON NEURAL NETWORKS San Diego, California/June 21-24, 1987

The San Diego IEEE Section welcomes neural network enthusiasts in industry, academia, and government world-wide to participate in the inaugural annual ICNN conference in San Diego.

Plenary	*Speakers*
	Stephen Grossberg, **John Hopfield**, **Teuvo Kohonen**, **Carver Mead**, **Bernard Widrow**

Sessions	*Chairmen*
Network Architectures	**Micheal Cohen**, Boston University / **Shun-ichi Amari**, University of Tokyo
Learning Algorithms	**Bernard Widrow**, Stanford University / **James Anderson**, Brown University
Self-Organization	**Teuvo Kohonen**, Technical University Helsinki **Stephen Grossberg**, Boston University
Adaptive Resonance	**Gail Carpenter**, Northeastern University
Cooperative & Competitive Network Dynamics	**Morris Hirsch**, University of California, Berkeley
Neurobiological Connections	**George Sperling**, New York University
Cognitive Science Connections	**David Rumelhart**, University of California, San Diego **David Zipser**, University of California, San Diego
Electrical Neurocomputers	**Robert Hecht-Nielsen**, Hecht-Nielsen Neurocomputer Corporation **Andrew Penz**, Texas Instruments
Optical Neurocomputers	**Joseph Goodman**, Stanford University **Clark Guest**, University of California, San Diego
Knowledge Processing	**Bart Kosko**, VERAC Corporation
Vision	**Kunihiko Fukushima**, NHK Labs / **Ennio Mingolla**, Boston University
Speech Recognition & Synthesis	**Jeffrey Ellman**, University of California, San Diego **David Stork**, Clark University
Robotics	**Allen Stubberud**, University of California, Irvine **Behnam Bavarian**, University of California, Irvine
Combinatorial Optimization	**Harold Szu**, Naval Research Laboratory
Novel Applications	**Lee Giles**, Air Force Office of Scientific Research **Charles Kellum**, Department of Defense

NOTE TO THOSE CONTRIBUTING PAPERS (Deadline February, 1987)

Extended abstracts should be submitted by February 1, 1987 for conference presentation. Please submit abstract plus 4 clean copies. Abstracts must be neatly typed, single spaced, three to four pages. Abstracts will be carefully refereed. If your abstract is accepted, it will be printed and distributed at the Conference.

Final papers for publication in the book of proceedings are due June 1, 1987.

FINAL RELEASE OF ABSTRACTS AND PAPERS WITH RESPECT TO PROPRIETARY RIGHTS AND CLASSIFICATION MUST BE OBTAINED BEFORE SUBMITTAL.

Address all correspondence referring to abstracts and papers to:

Maureen Caudill
ICNN 10615G Tierrasanta Blvd.
Suite 346
San Diego, CA 92124
Tel. (619) 485-1809

INNS First Annual Meeting

Symposium and Plenary Speakers

Plenary

Stephen Grossberg
Carver Mead
Terrence Sejnowski
Nobuo Suga
Bernard Widrow

Vision and Pattern Recognition

Gail Carpenter
Max Cynader
John Daugman
Kunihiko Fukushima
Teuvo Kohonen
Ennio Mingolla
Eric Schwartz
George Sperling
Steven Zucker

Motor Control and Robotics

Jacob Barhen
Daniel Bullock
James Houk
Scott Kelso
Lance Optican

Cognitive and Neural Systems

James Anderson
Walter Freeman
Guenter Gross
Gary Lynch
Christoph von der Malsburg
David Rumelhart
Allen Selverston

Combinatorial Optimization and Content Addressable Memory

Daniel Amit
Stuart Geman
Geoffrey Hinton
Bart Kosko

Applications and Implementations

Dana Anderson
Michael Buffa
Lee Giles
Robert Hecht-Nielsen
Demetri Psaltis
Thomas Ryan
Bernard Soffer
Harold Szu
Wilfrid Veldkamp

The 1988 First Annual Conference of the International Neural Network Society (INNS) will bring together over 2,000 academic scientists, engineers, students, government administrators, industrial commercializers, and financiers in an open forum for the advancement of the full spectrum of significant neural network research and development, from biology through technology.

Formed in 1987 in response to the extraordinary international interest in neural network research, INNS includes among its founders many of the most distinguished leaders of the field. By Winter, 1988, INNS membership had grown to 1,600 of the field's most active researchers, from 33 countries and 45 states. These are the people who will determine the future of this strategic technology.

The INNS invites all those interested in the exciting and rapidly expanding field of neural networks to attend its 1988 Annual Meeting. The meeting includes plenary lectures, symposia, contributed oral and poster presentations, tutorials, commercial and publishing exhibits, government agency presentations, and social events.

Join us in Boston in September!

INNS
First
Annual
Meeting

…and
Birthday!

WHY WAS THIS GROWTH SO EXPLOSIVE?

A MAJOR PARADIGM SHIFT began in the late 1800's
when great scientists such as Helmholtz, Maxwell, and Mach
worked in both psychology and physics

This shift accelerated in the 1960's - 1980's
For reasons, see Grossberg (1988, Neural Networks, 1, 17)

What is this paradigm shift?

Understanding how an individual
adapts
on its own
in real time
to a complex and changing world

AUTONOMOUS adaptation to a NON-STATIONARY world

On-line adaptation to UNEXPECTED EVENTS

WHY WAS THIS GROWTH SO EXPLOSIVE?

Much previous science and technology discussed
EXTERNAL CONTROL of a STATIONARY world
e.g., optimal control theory, quantum theory

New INTUITIVE CONCEPTS and new
MATHEMATICAL EQUATIONS and METHODS
were needed to make this breakthrough
These are the most confusing kinds of scientific revolutions

Understanding AUTONOMY is taking a long time…decades!

Along the way, many BRAIN METAPHORS
attracted huge interest for awhile
faded when they failed to solve the big problems

telegraph circuit	catastrophy
hydraulic system	spin glasse
linear control system	back propagation network
hologram	Bayesian network

FROM PAST TO FUTURE: A STILL-EVOLVING THEME

NN: A MAJOR STEP FORWARD IN THE THEORY OF MEASUREMENT

Newton, celestial mechanics, absolute space and time

Einstein, relativity theory, measurement relative to each reference frame

Heisenberg, quantum mechanics, measurement alters the measured

Brain: a universal measurement device that continually and rapidly changes (develops, learns) as it interacts with the world

The Problem of Self-Organization

The Problem of AUTONOMY in a NON-STATIONARY World

All concepts not consistent with full autonomy or that depend on stationary hypotheses are classical or neo-classical

They do not fully capture the revolutionary potential of our field

TRENDS IN SCIENCE AND TECHNOLOGY THAT LOOK TO NEURAL NETWORK RESEARCH

WORLD

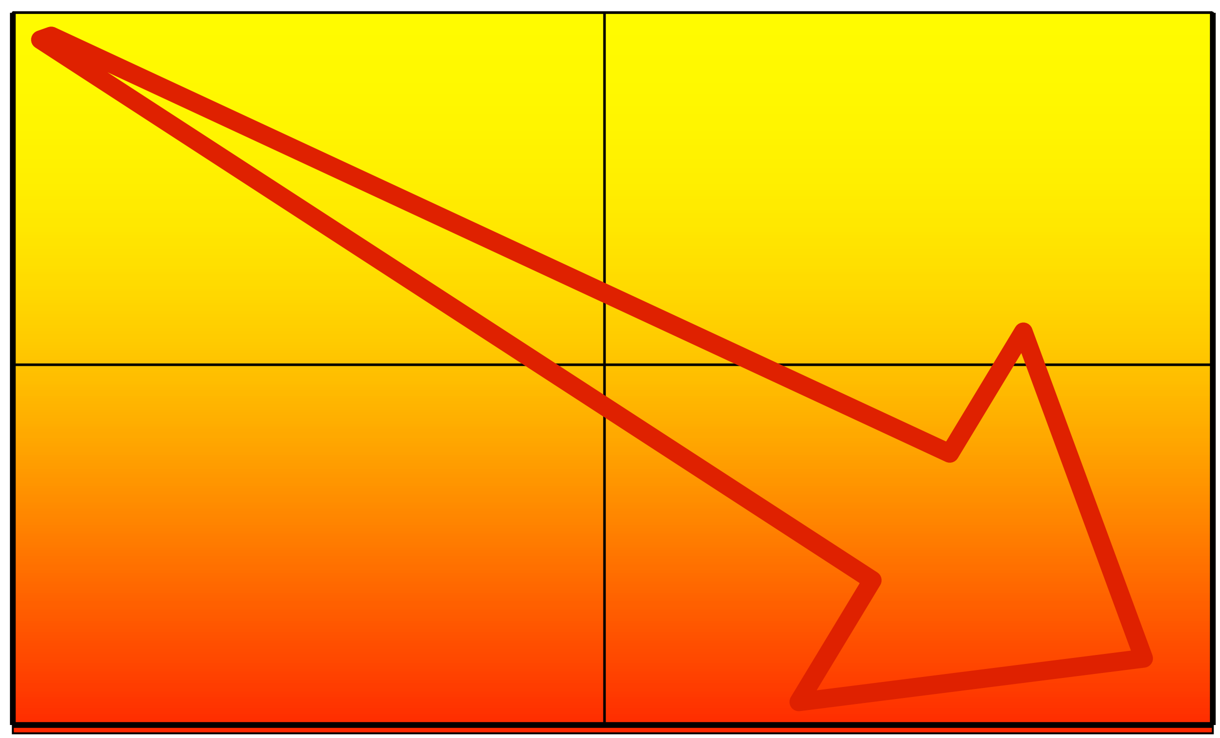
CONTROL

EXTERNAL (SUPERVISED)

AUTONOMOUS (UNSUPERVISED)

STATIONARY

NON-STATIONARY



TWO ANNIVERSARIES TO CELEBRATE!

To have big conferences, you need to have a lot to talk about!

50 YEARS OF NEURAL NETWORKS LINKING BRAIN TO BEHAVIOR

30+ years of neural network research preceded the 1987 IEEE meeting

1957 I introduced a new PARADIGM and a METHOD to theoretically link MIND to BRAIN

http://www.cns.bu.edu/Profiles/Grossberg/GrossbergInterests.pdf

This method accepts that

BRAIN evolution is driven by BEHAVIORAL success

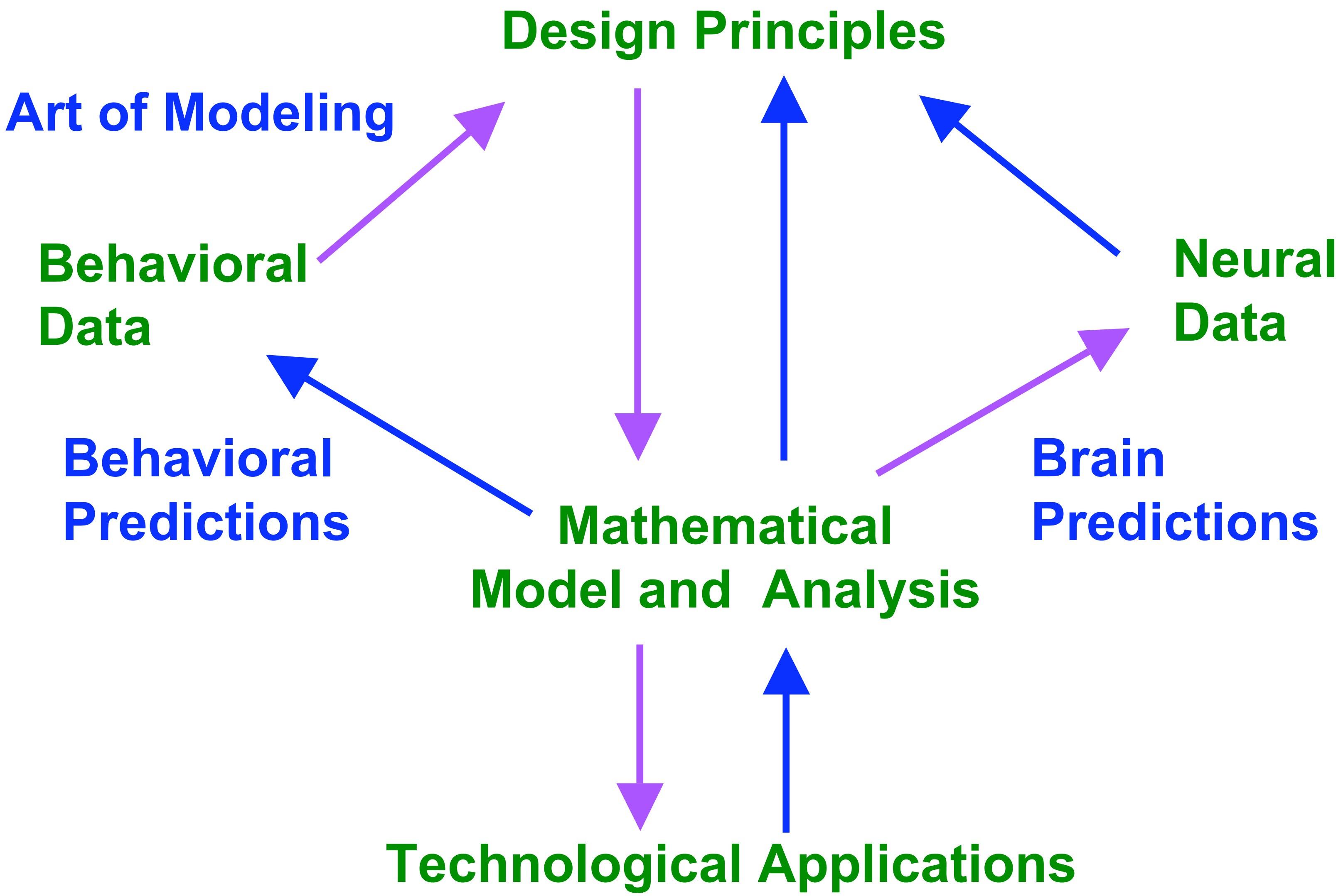
Discover the computational level that computes behavioral success

Fifty years of modeling show it is the NETWORK and SYSTEM levels

To understand brain design, start with BEHAVIORAL DATA & analyse how

individual adapts on its own in real time to a changing world

MODELING METHOD AND CYCLE



Embedding Principle: Repeat this cycle, leading to increasing model realism and explanatory power

TWO KEY CONCLUSIONS

1. Advanced brains look like they do to enable

REAL-TIME AUTONOMOUS LEARNING

Lesson: The Architecture is the Algorithm

2. Recent models show how the brain's ability to DEVELOP and LEARN greatly constrain the laws of

ADULT INFORMATION PROCESSING

Lesson: Learning and information processing need to be studied together as part of the same general problem

TWO ANNIVERSARIES TO CELEBRATE!

Back to 1957 Using this method, I derived
nonlinear neural networks as the
natural language to explain
autonomous adaptation to a non-stationary world

A single theoretical language to unify
Mind and Brain

Begin to solve the classical mind/body problem

Heady stuff for a 17 year old!

However!!!

Using brain models to explain psychological data was very controversial

NONLINEAR neural networks to explain ANYTHING was controversial!

25 years of rapid theory development before multiple factors led
to widespread interest starting in the 1980's

TWO ANNIVERSARIES TO CELEBRATE!

1957 I derived the ADDITIVE and SHUNTING network models:

interactions of STM (activation) and LTM (learning)

gated steepest descent learning used in SOM and ART

ADDITIVE MODEL

$$\frac{dx_i}{dt} = -Ax_i + \sum_k f(x_k)B_{ki}w_{ki} + I_i$$

$$\frac{dw_{ki}}{dt} = g(x_k)\left(-w_{ki} + h(x_i)\right)$$

This was very controversial!

How these equations were discovered is an unusual story:
I was a Freshman at Dartmouth College taking Psychology 1
http://www.cns.bu.edu/Profiles/Grossberg/GrossbergInterests.pdf

For those who complain about delayed reviews:
It took 10 years to get these results published!

Experimental support for gated steepest descent learning took 20 years longer:
Levy, 1985; Levy, Brassel, and Moore, 1983; Levy and Desmond, 1985;
Rauchecker and Singer, 1979; Singer, 1983

CONTINUOUS AND NONLINEAR NN DYNAMICS

Grossberg, PNAS, 1967, 1968

Short-term memory (STM) trace → $X_i(t)$

Activation

$f_i(X_i(t))\, B_{ij}$

Signal

$Z_{ij}(t)$ ← Long-term memory (LTM) trace

Adaptive weight

$X_j(t)$

STM EQUATION

ADDITIVE MODEL

$$\frac{d}{dt} X_i = -A_i X_i + \sum_{j=1}^{n} f_j(X_j) B_{ji} Z_{ji}^{(+)} - \sum_{j=1}^{n} g_j(X_j) C_{ji} Z_{ji}^{(-)} + I_i$$

PASSIVE DECAY — POSITIVE FEEDBACK — NEGATIVE FEEDBACK — INPUT

Special case:

$$\frac{d}{dt} X_i = -A_i X_i + \sum_j f_j(X_j) Z_{ji} + I_i$$

Cf. Hopfield (1984)

1980'S: A PERIOD OF MARKETING KNOWN MODELS TO ENTHUSIASTIC NEW AUDIENCES

Known models tended to become popular in the order in which they were historically discovered; e.g.,

Autoassociators:
Grossberg (1967-1972)
Amari (1974)
Cohen & Grossberg (1982, 1983)
Hopfield (1982, 1984)

Competitive learning and self-organizing maps:
Grossberg (1972-1976)
Von der Malsburg (1973-1978)
Kohonen (1982, 1984)

Back propagation:
Werbos (1974)
Parker (1982)
Rumelhart, Hinton & Williams (1986)

Our field needs a process to incrementally codify its history

WHY IS THIS NOT JUST OLD NEWS?

Two big reasons:

1. NEW PARADIGMS have been introduced over the past 50 years that have hardly begun to realize their revolutionary potential

2. Scores of experimental PREDICTIONS took 5 - 30 years to get supported, and many more have yet to be tested

Why such a big lag in a world with "instant global communication"?

A new way of THINKING that has not yet been fully assimilated

A huge problem of interdisciplinary literacy

More interdisciplinary infrastructure is needed to transmit high-level theory in depth to students and researchers

IJCNN can make a difference!

RECONCILING PROBABILITY AND DETERMINISM

cf., current popularity of Bayesian models

COMPUTING WITH PATTERNS

Individual pixels are meaningless

Patterns embody CONTEXT as well as FEATURES

1965 -1976: A series of theorems show

The Unit of STM is a SPATIAL PATTERN

The Unit of LTM is a SPATIAL PATTERN

The network tries to learn

SYNCHRONOUS

activities of spatially distributed patterns

RECONCILING PROBABILITY AND DETERMINISM

A link between cells, patterns, normalization, and synchrony

Synchrony is still a hot topic today: Wolf Singer's plenary talk

These connections are still not often discussed together

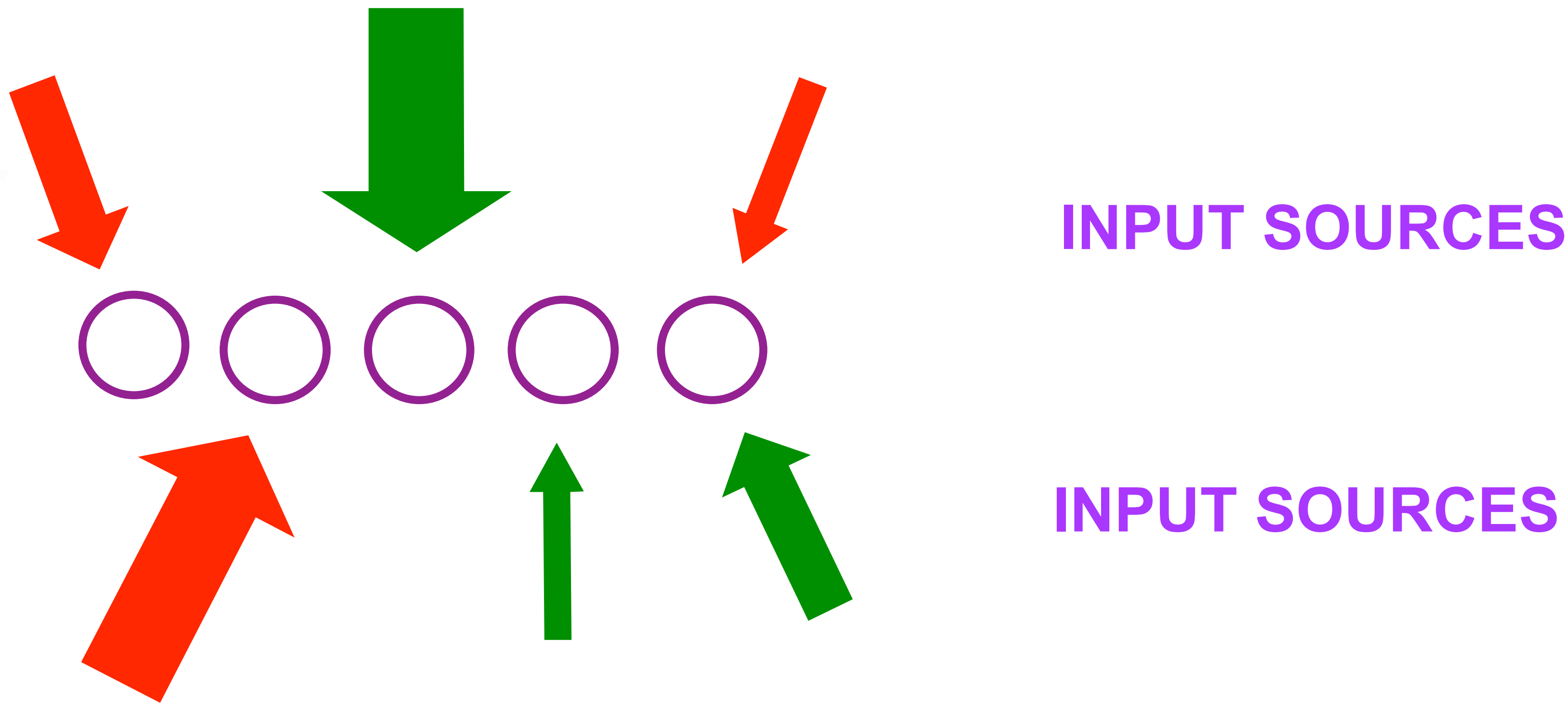
WHAT IS A CELL?

It contains a finite number of active and inactive sites
Infinity does not exist in biology!

NOISE-SATURATION DILEMMA (1968-1973)

How are feature patterns processed
in noisy cells with finitely many sites
without being contaminated by either noise or saturation?

PATTERN PROCESSING BY CELL NETWORKS

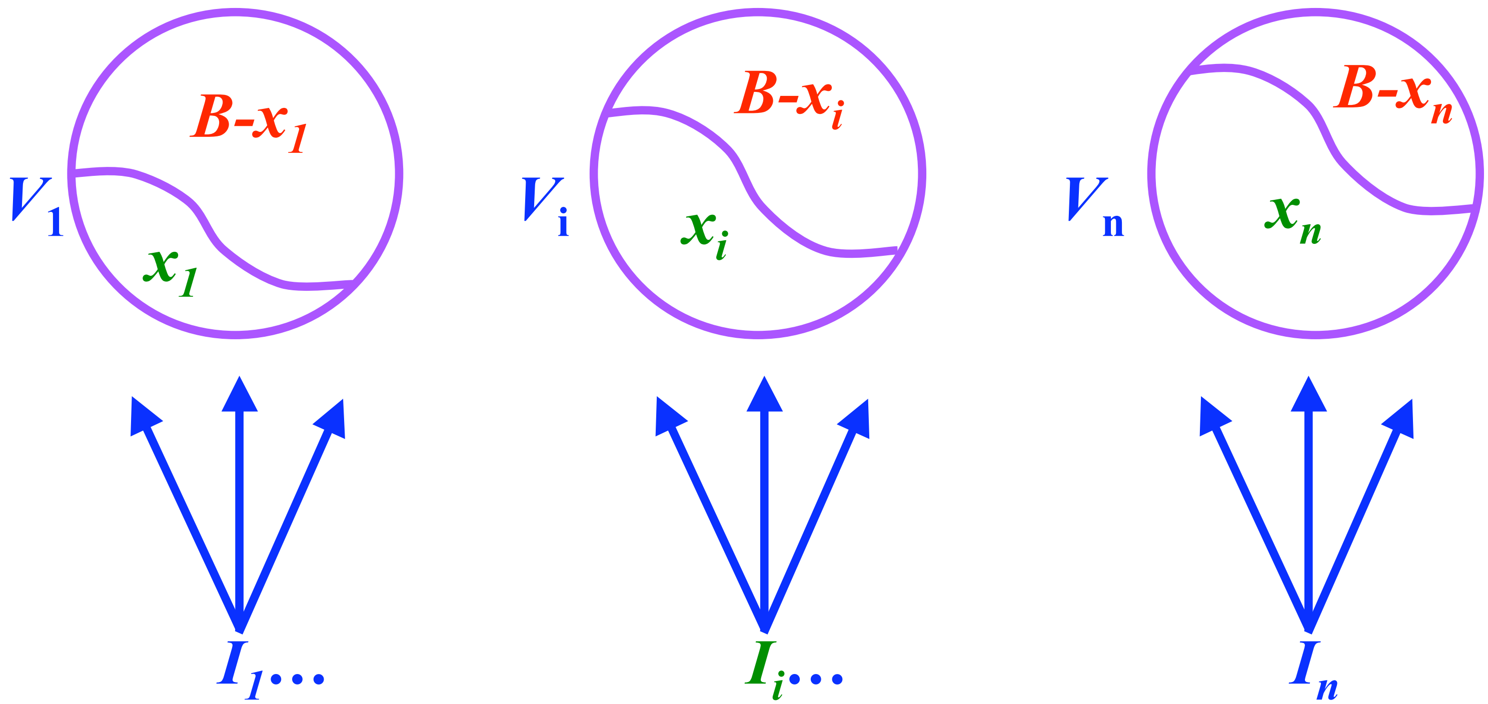


Total NUMBER and SIZE of inputs to each cell can vary wildly through time.

How do cells maintain their SENSITIVITY to input PATTERNS whose overall SIZE changes wildly through time?

COMPUTING IN A BOUNDED ACTIVITY DOMAIN

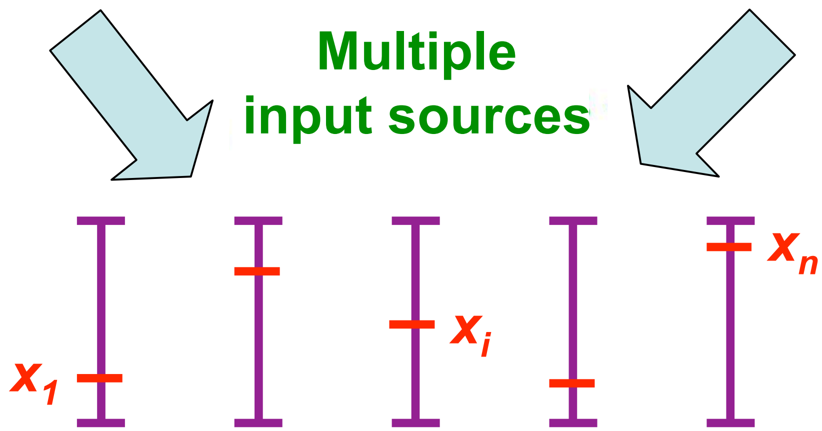
Thought experiment



B	excitable sites
$x_i(t)$	excited sites (activity, potential)
$B\text{-}x_i(t)$	unexcited sites

NOISE-SATURATION DILEMMA

Grossberg, 1968-1973



BOUNDED ACTIVITIES

If activities x_i are sensitive to SMALL inputs, then why don't they SATURATE to large inputs?

If x_i are sensitive to LARGE inputs, then why don't small inputs get lost in system NOISE?

SOLUTION: SHUNT + COMPETITION

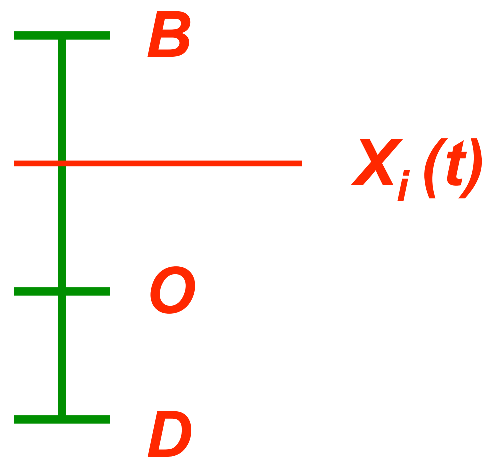
Join SHUNTING or MEMBRANE EQUATION dynamics to ON-CENTER OFF-SURROUND anatomy

SHUNTING MODEL

MASS ACTION, MEMBRANE EQUATIONS

Bounded activations

Automatic gain control



$$\frac{d}{dt}X_i = -A_iX_i + (B - X_i)\Big[\sum_{j=1}^{n} f_j(X_j)C_{ji}Z_{ji}^{(+)} + I_i\Big]$$

$$- (X_i + D)\Big[\sum_{j=1}^{n} g_j(X_j)E_{ji}Z_{ji}^{(-)} + J_i\Big]$$

INCLUDES THE ADDITIVE MODEL

RECONCILING PROBABILITY AND DETERMINISM

Cells, patterns, normalization, and synchrony

A shunting on-center off-surround network
SELF-NORMALIZES
its activities in response to a spatial pattern

It tracks input RATIOS

It computes "real-time probability distribution"

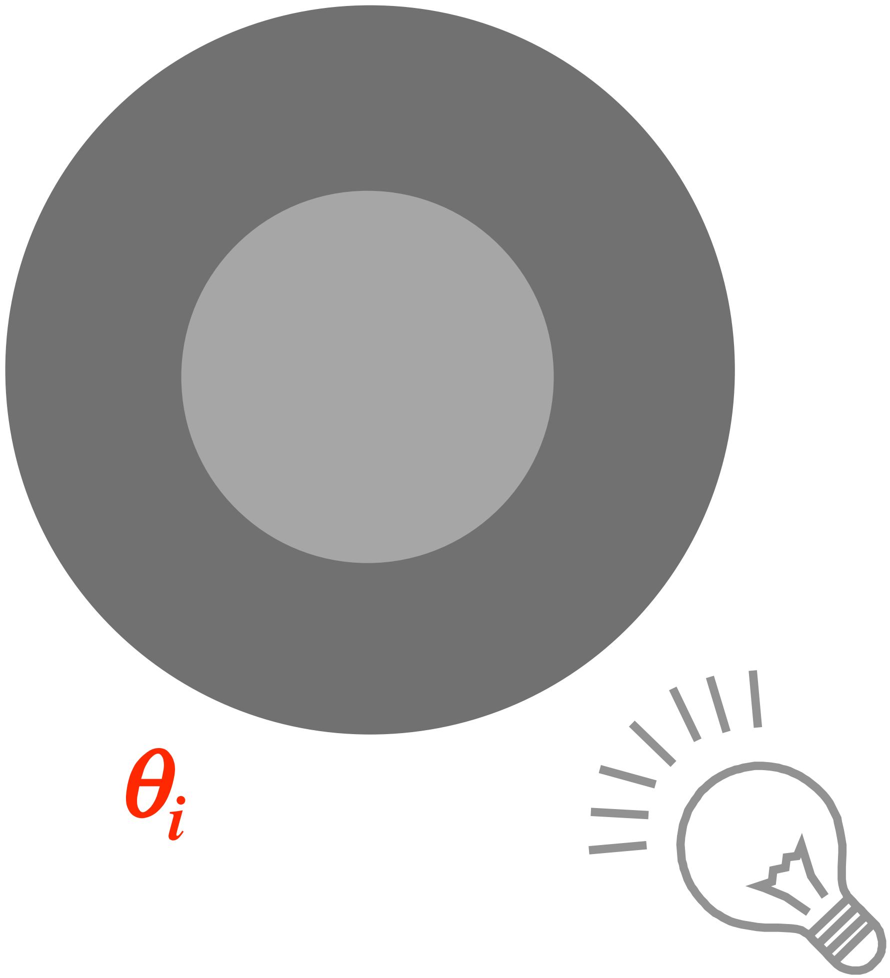
It processes the SYNCHRONOUS part of the distributed pattern

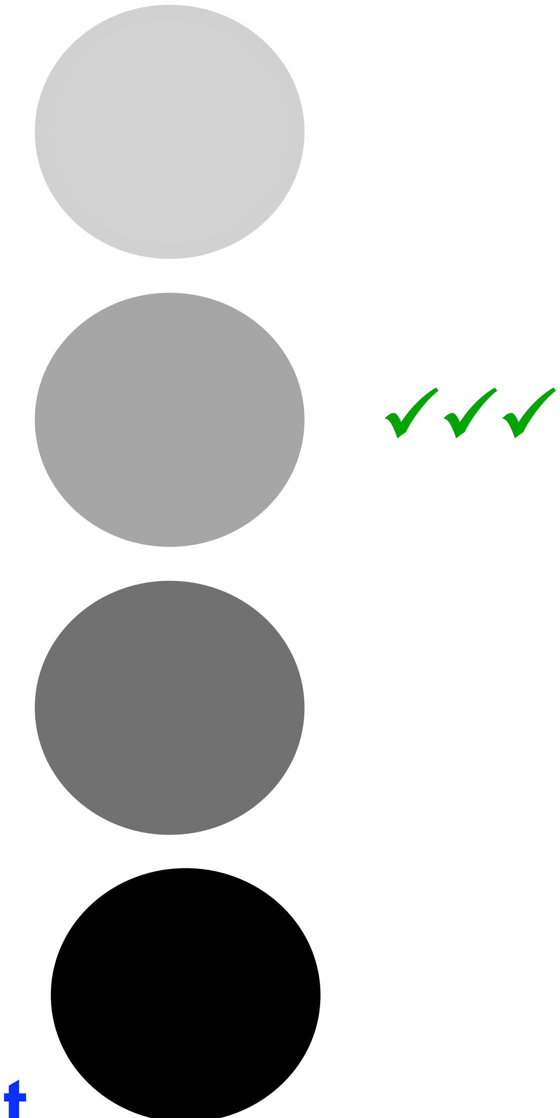
The brain carries out a kind of real-time probability theory and hypothesis testing that leads to SELF-ORGANIZATION in NON-STATIONARY environments

This competence goes beyond classical probabilistic concepts

VISION: BRIGHTNESS CONSTANCY

Contrast Normalization





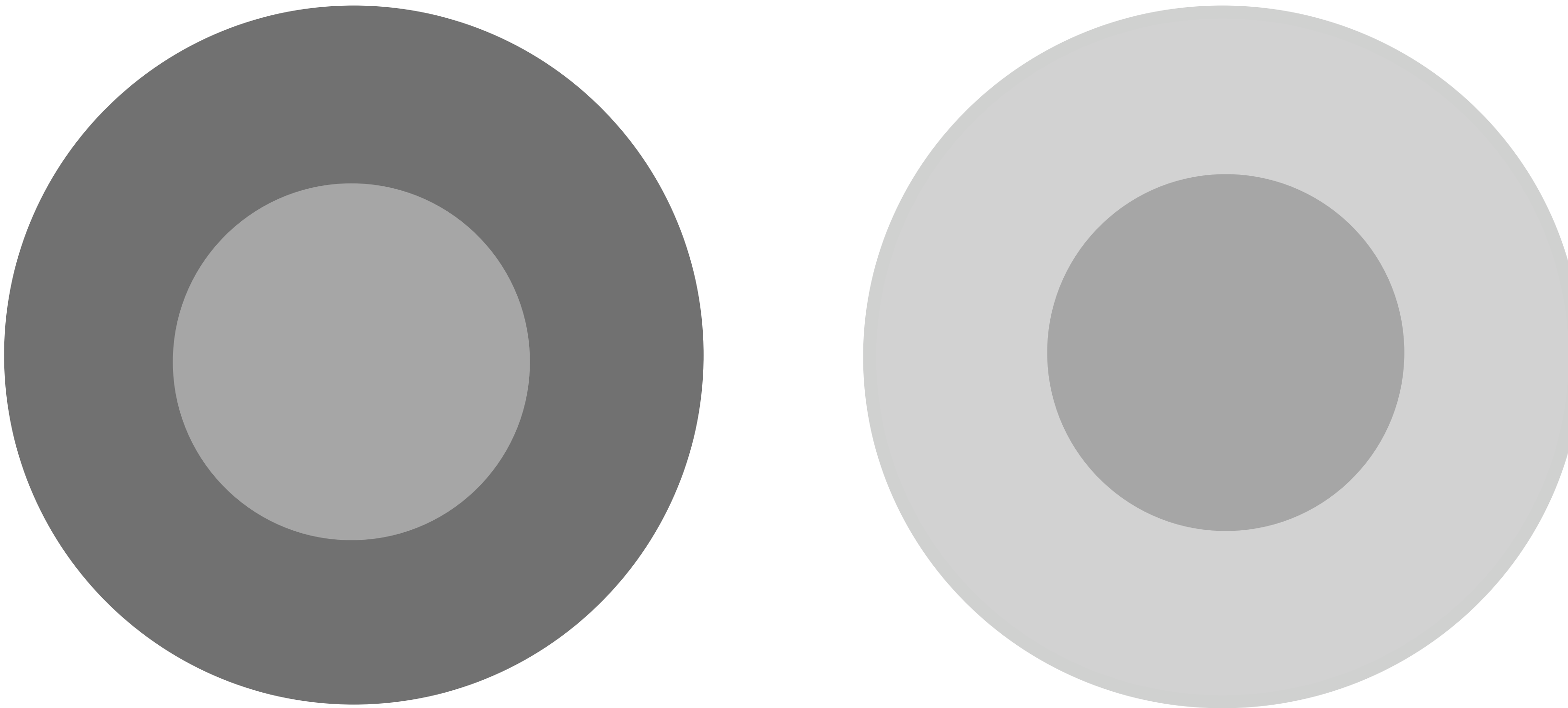
Compute RATIOS of reflected light

Reflectance processing

VISION: BRIGHTNESS CONTRAST

CONSERVE A TOTAL QUANTITY

Total Activity Normalization



LUCE Ratio scales in choice behavior

ZEILER Adaptation level theory

MULTIPLE TIME SCALES

FAST: Activation, or short-term memory

SLOW: Learning, or long-term memory

MULTIPLE TIME SCALES

MEDIUM: Habituation, or medium-term memory

Grossberg, PNAS 1967+

Chemical transmitters control UNBIASED transduction between cells as they habituate to sustained inputs

Enables INTRACELLULAR ratio processing and adaptation
antagonistic REBOUNDS for reset & error correction
INVERTED U properties to tune network sensitivity

Vision, speech, cognition, emotion, mental disorders,…

Recently called DEPRESSING SYNAPSES…a hot topic again!

Visual Cortex: Abbott et al. (1997)

Somatosensory Cortex: Markram & Tsodyks (1997)

Another 30 year delay…shows power of the modeling method

DO THESE EQUATIONS JUST GO ON AND ON?

Is the brain just a BAG OF TRICKS?

V.S. Ramachandran

NO!

TRUE THEORIES ARE EMERGING

A small number of EQUATIONS

e.g., shunting activation dynamics (STM)
habituative transmitter gates (MTM)
activity-gated learning (LTM)

A larger number of MODULES

e.g., on-center off-surround nets
resonant matching nets
opponent processing nets
spectral timing nets
boundary completion nets
filling-in nets…

Specialized combinations of modules, using a few basic equations, are assembled in architectures that solve modal problems

A still larger number of MODAL ARCHITECTURES

e.g. vision
audition
smell
touch
cognition
emotion…

WHAT PRINCIPLES DETERMINE HOW MODAL ARCHITECTURES ARE DESIGNED?

BREAKTHROUGHS IN BRAIN COMPUTING

Models that link detailed BRAIN CIRCUITS to the ADAPTIVE BEHAVIORS that they control

Mind/Body Problem

Describe NEW PARADIGMS for brain computing

INDEPENDENT MODULES
Computer Metaphor

COMPLEMENTARY COMPUTING
What is the nature of brain specialization?

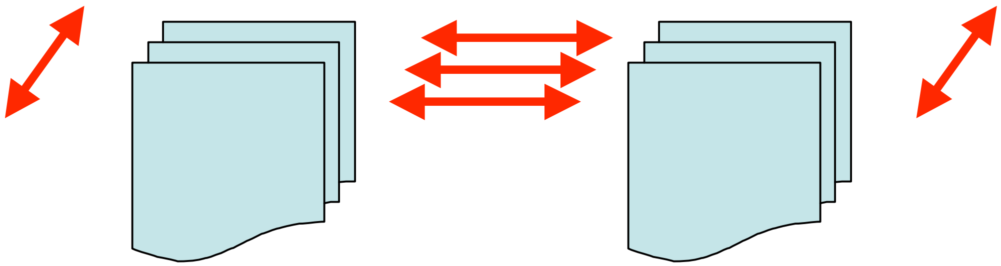
LAMINAR COMPUTING
Why are all neocortical circuits laminar?
How do laminar circuits give rise to biological intelligence?

COMPLEMENTARY COMPUTING

New principles of

UNCERTAINTY and COMPLEMENTARITY

Multiple Parallel Processing Streams Exist in the Brain



HIERARCHICAL INTRASTREAM INTERACTIONS

UNCERTAINTY PRINCIPLES operate at individual levels

Hierarchical interactions resolve uncertainty

PARALLEL INTERSTREAM INTERACTIONS

Each stream computes COMPLEMENTARY properties

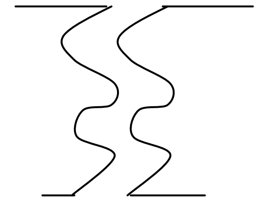
Parallel interactions overcome complementary weaknesses

ADAPTIVE BEHAVIOR = EMERGENT PROPERTIES

WHAT ARE COMPLEMENTARY PROPERTIES?

Analogies:

Lock and key, puzzles pieces fitting together



Computing one set of properties at a processing stage prevents that stage from computing a complementary set of properties

Complementary parallel processing streams are BALANCED against one another

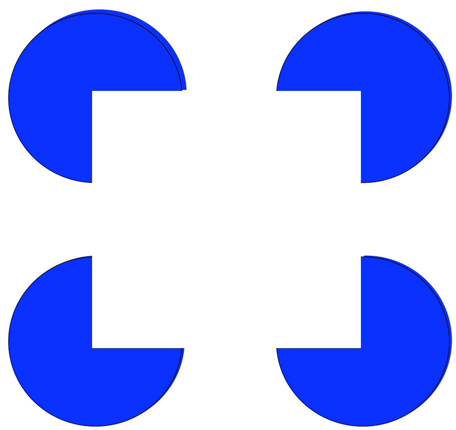
Interactions between streams overcomes their complementary weaknesses and support intelligent and creative behaviors

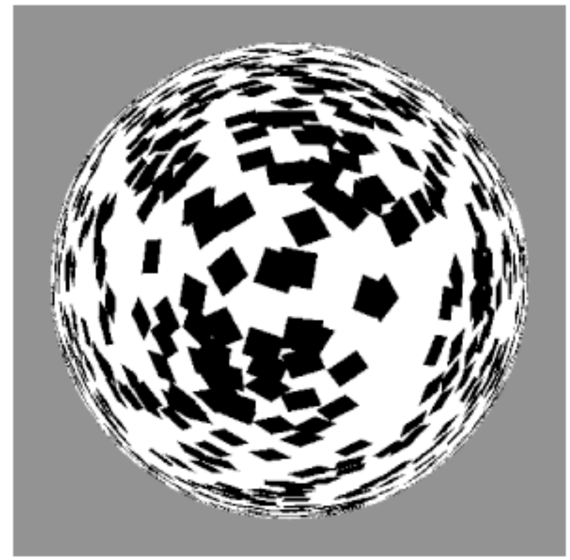
Not just one learning law!

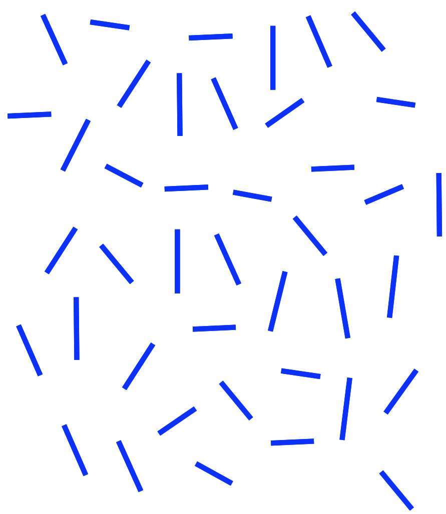
SOME COMPLEMENTARY PROCESSES

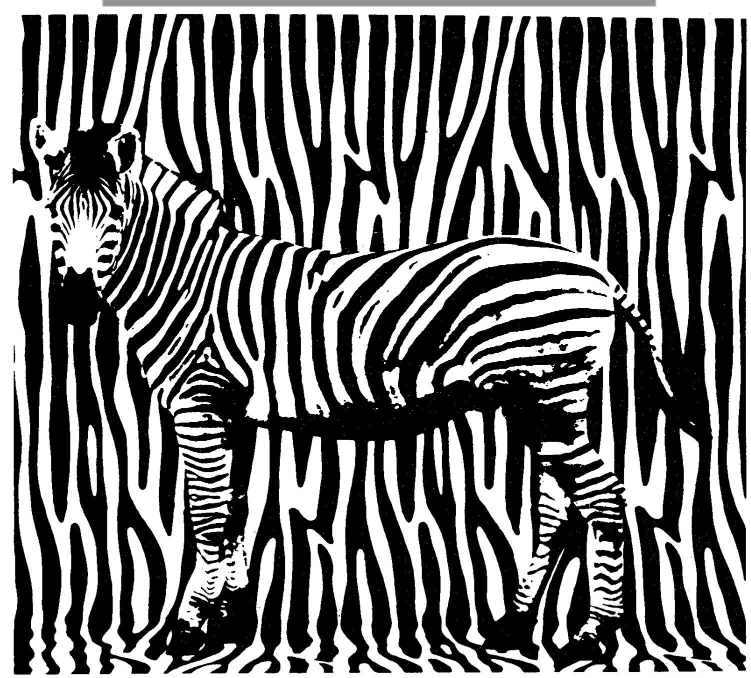
Visual Boundary Interbob Stream V1-V4	Visual Surface Blob Stream V1-V4
Visual Boundary Interbob Stream V1-V4	Visual Motion Magno Stream V1-MT
WHAT learning/ Matching Inferotemporal and Prefrontal areas	WHERE learning/ Matching Parietal and Prefrontal areas
Object Tracking MT Interbands and MSTv	Optic Flow Navigation MT Bands and MSTd
Motor Target Position Motor and Parietal Cortex	Volitional Speed Basal Ganglia

VISUAL BOUNDARIES OR GROUPINGS



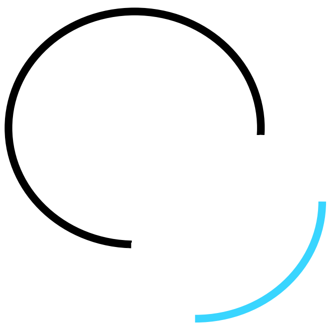


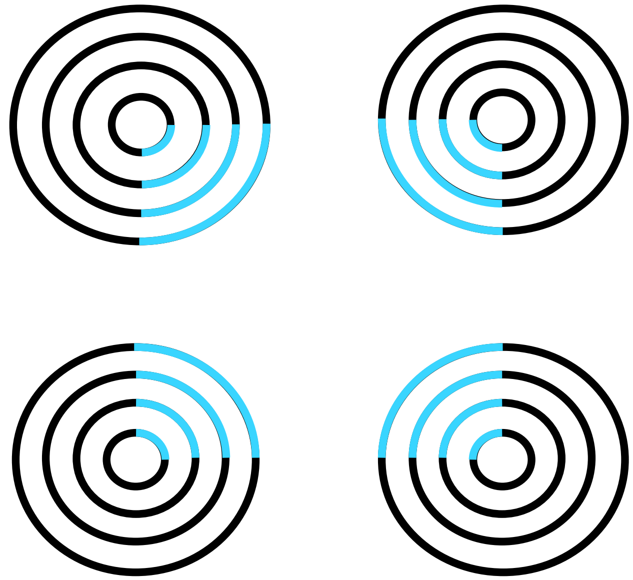




VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY

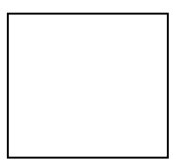
Grossberg (1984)





Neon color spreading

BOUNDARY COMPLETION



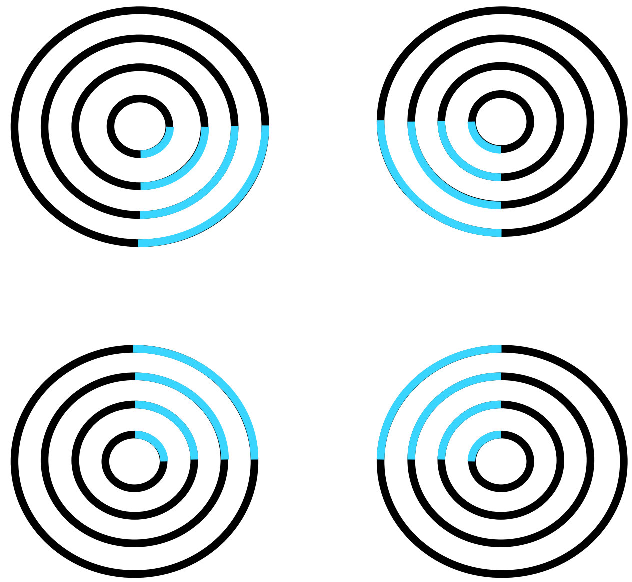
oriented
inward
insensitive to
direction-of-contrast

SURFACE FILLING-IN

unoriented
outward
sensitive to
direction-of-contrast

Grossberg Plenary
IJCNN'07

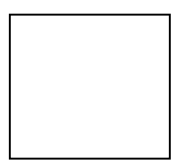
VISUAL BOUNDARY AND SURFACE COMPUTATIONS ARE COMPLEMENTARY



Neon color spreading

All Boundaries
Are
Invisible!

BOUNDARY COMPLETION



oriented
inward
insensitive to
direction-of-contrast

SURFACE FILLING-IN

unoriented
outward
sensitive to
direction-of-contrast

SEEING vs. KNOWING

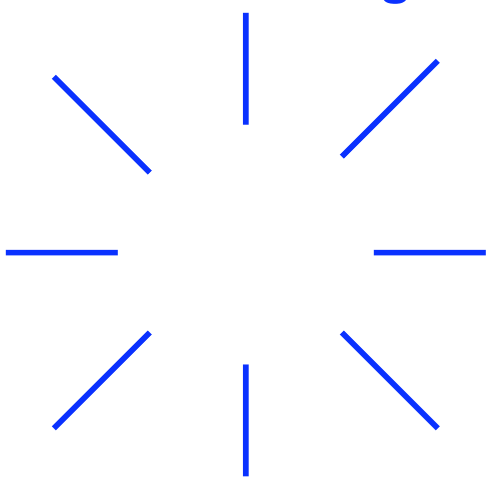
SEEING vs. **KNOWING**

an object / **what it is**

Epstein, Gregory, Helmholtz, Kanizsa, Kellman, Michotte,…

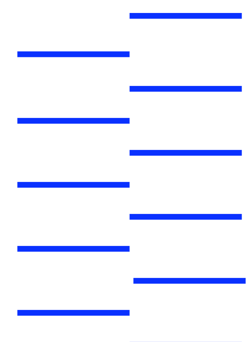
SEEING vs. **RECOGNIZING**

Ehrenstein Figure



See
Recognize

Offset Grating



Some boundaries are invisible, or amodal

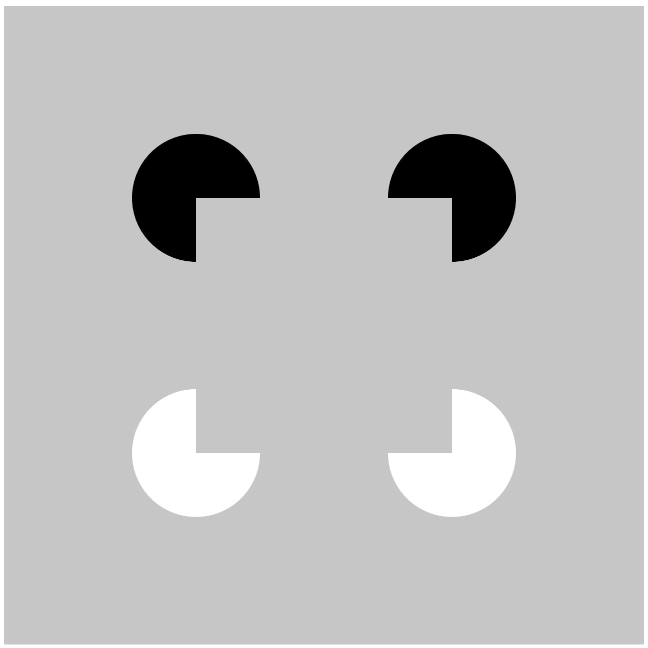
Do not see
Recognize

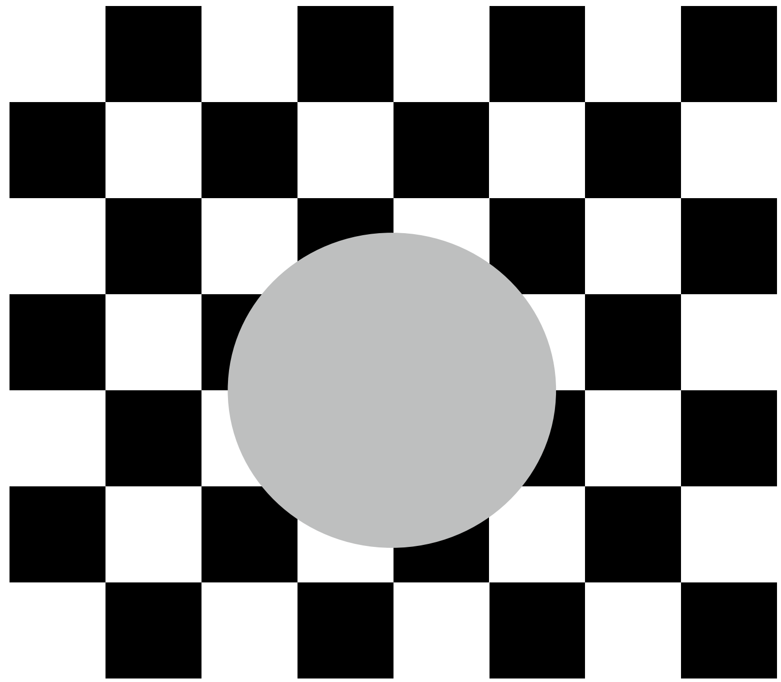
ALL BOUNDARIES ARE INVISIBLE!

Within the Boundary System

Grossberg (1984)

WHY? To recognize object boundaries in front of textured backgrounds

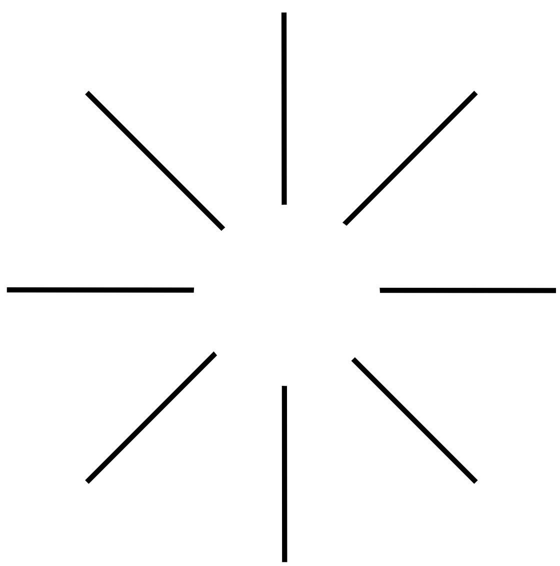




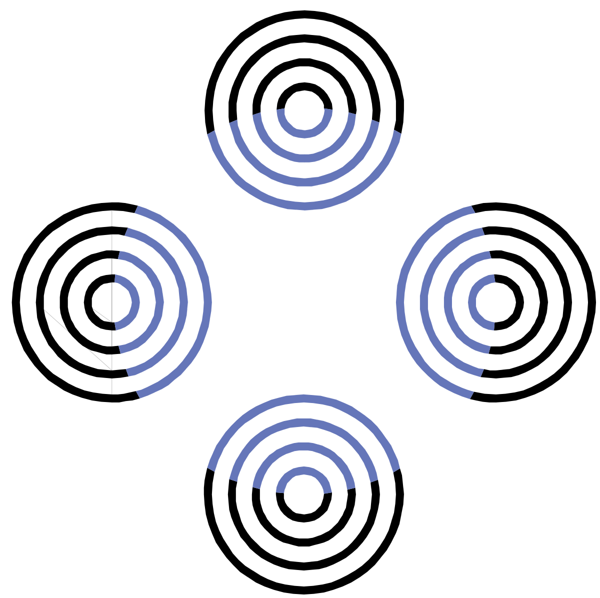
IF BOUNDARIES ARE INVISIBLE, HOW DO WE SEE?

Filling-In of Surface Color

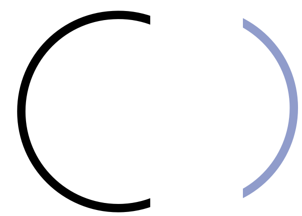
Boundaries define the compartments
within which lightness and color spread



Ehrenstein (1941)

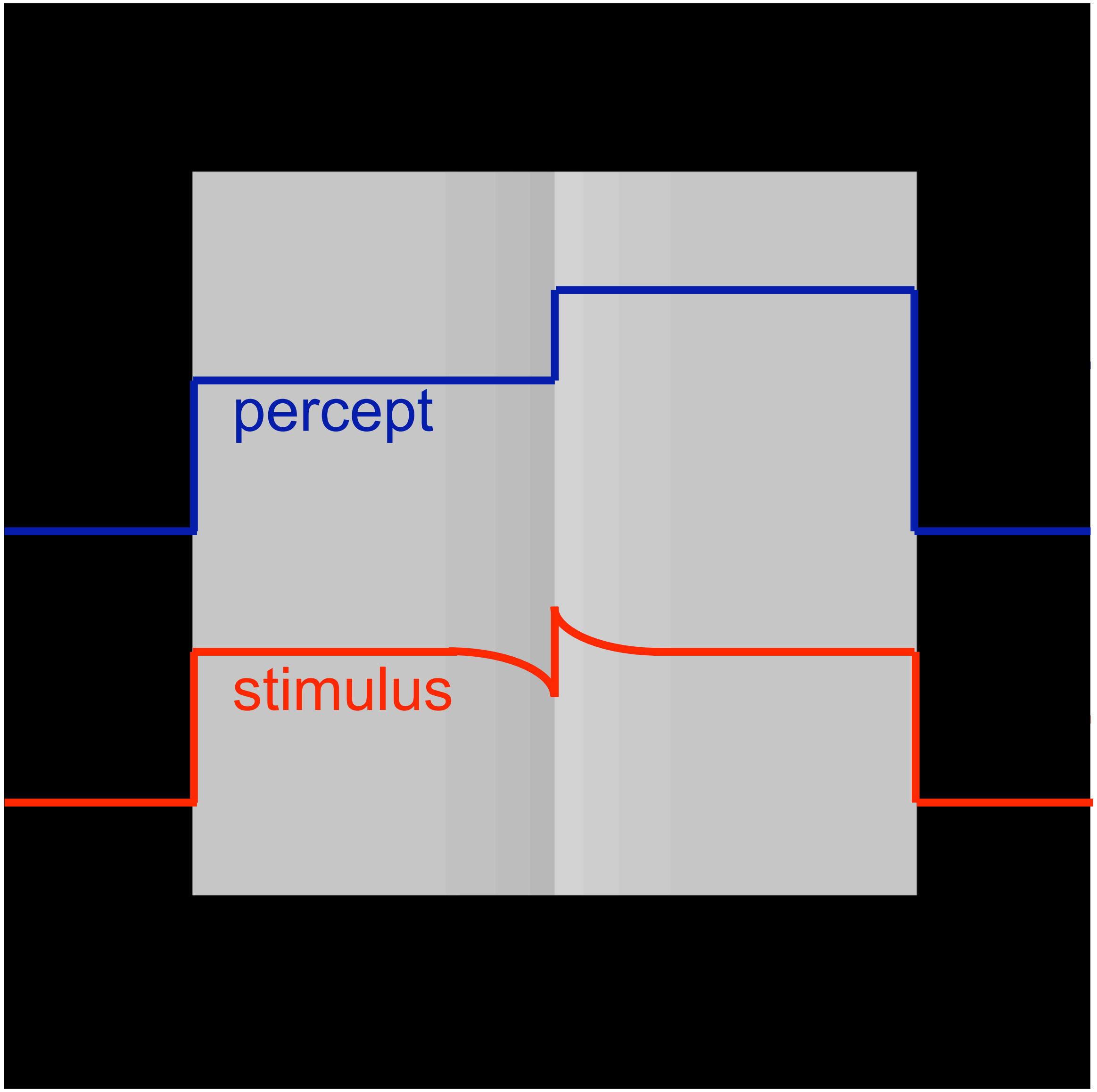


Varin (1971)



Neon color
spreading

Craik-O'Brien-Cornsweet Effect

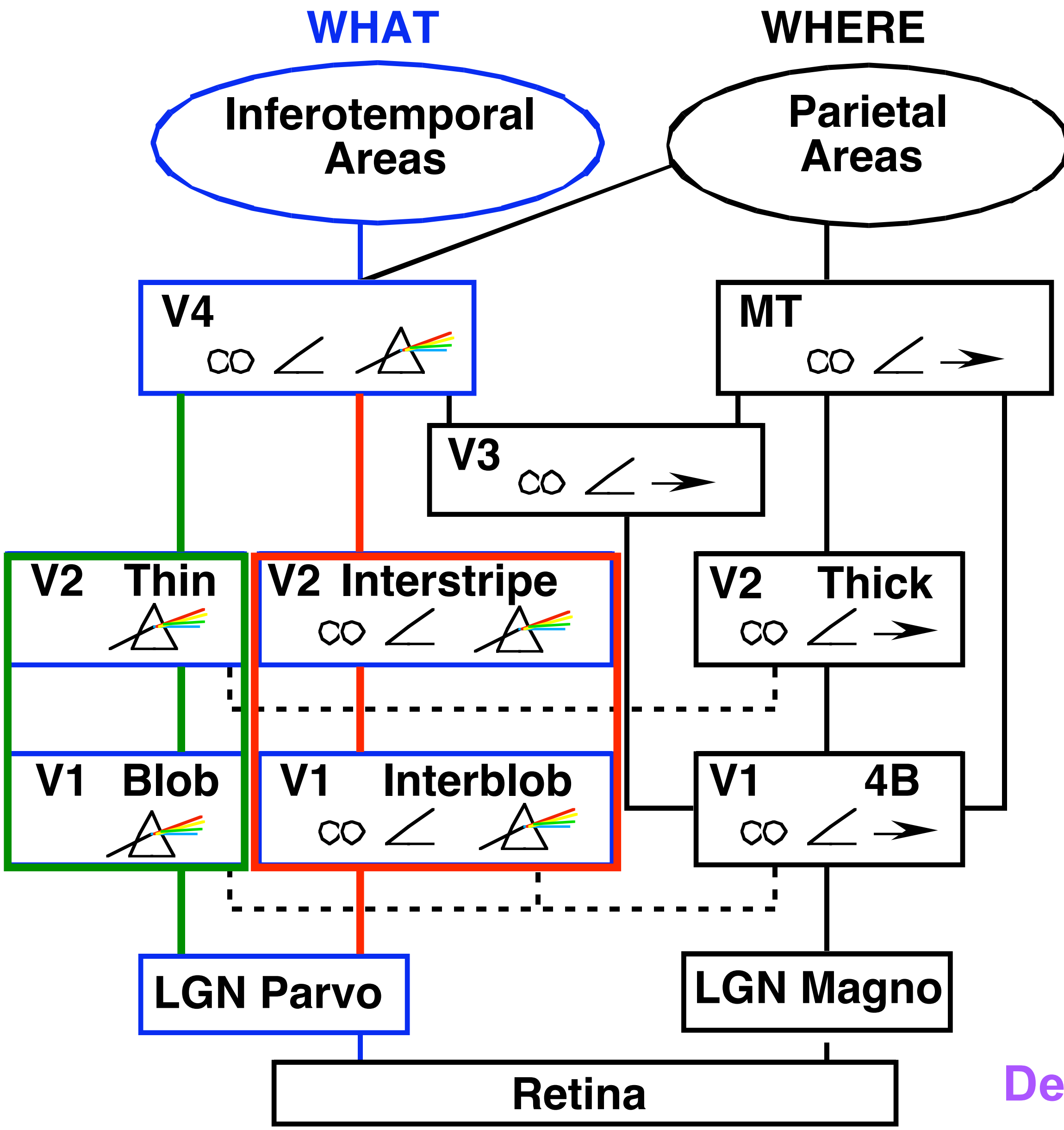


Boundary completion defines filling-in compartments

Filling-in determines what we see in each compartment

Grossberǵ (1984)
Todorovic (1987)

BOUNDARY AND SURFACE CORTICAL STREAMS



Boundaries
interblob stream

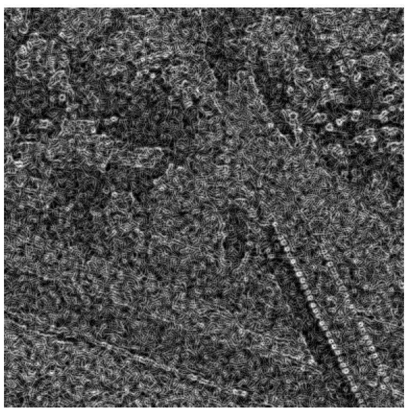
Surfaces
blob stream

DeYoe and van Essen (1988)

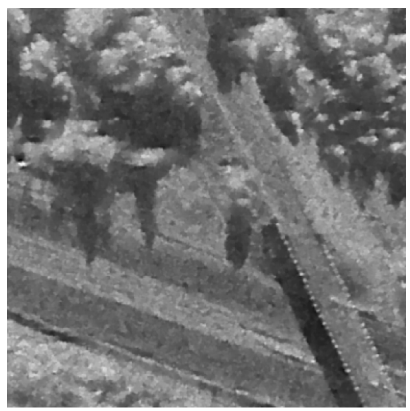
DO THESE IDEAS WORK ON HARD PROBLEMS?

input

feature



boundary



filling-in

Mingolla, Ross, and Grossberg (1999)

Application:
Image Enhancement

Synthetic aperture radar

signal: 5 orders of magnitude of power In radar return

multiplicative noise

sparse high-intensity pixels

Cf. Impressionist paintings
Monet

LIGHT ADAPTATION

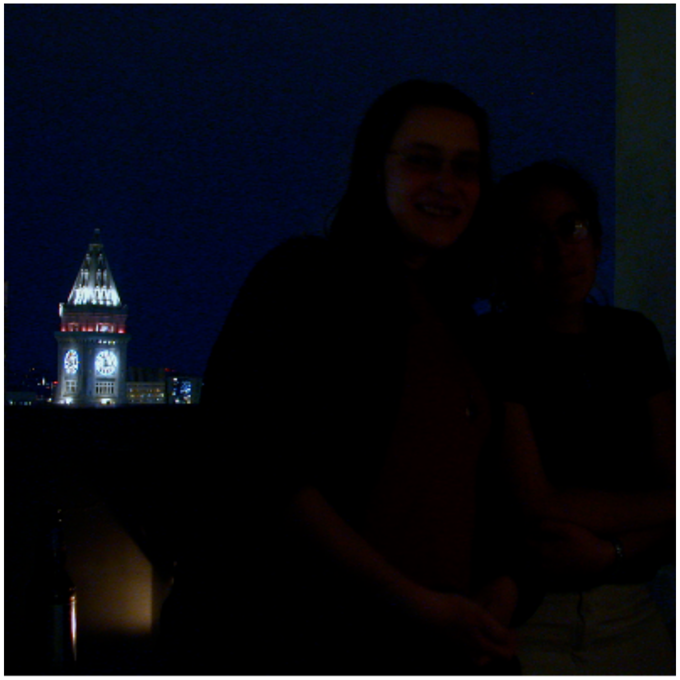
Ten orders of magnitude of daily variations of ambient illumination

Martin (1983)

Habituative transmitters + shunting on-center off-surround nets + boundaries + surfaces

Grossberg and Hong (2006)

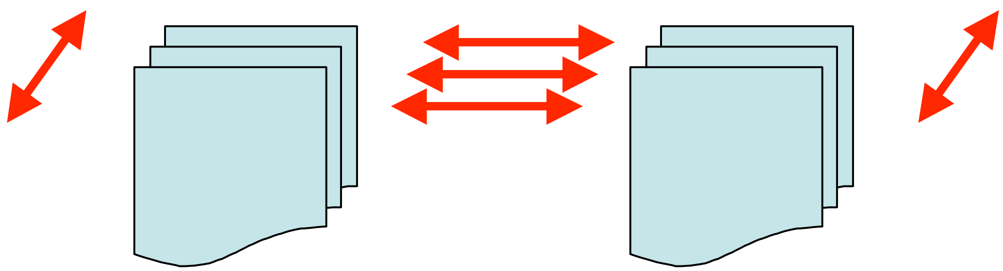
INPUT



MODEL SIMULATION

A VERY LARGE FUNCTIONAL UNIT!

Parallel and hierarchical interactions within
PAIRS OF
COMPLEMENTARY CORTICAL STREAMS
are needed to compute
COMPLETE INFORMATION
about a changing world



Clarifies why understanding how brains work is so difficult

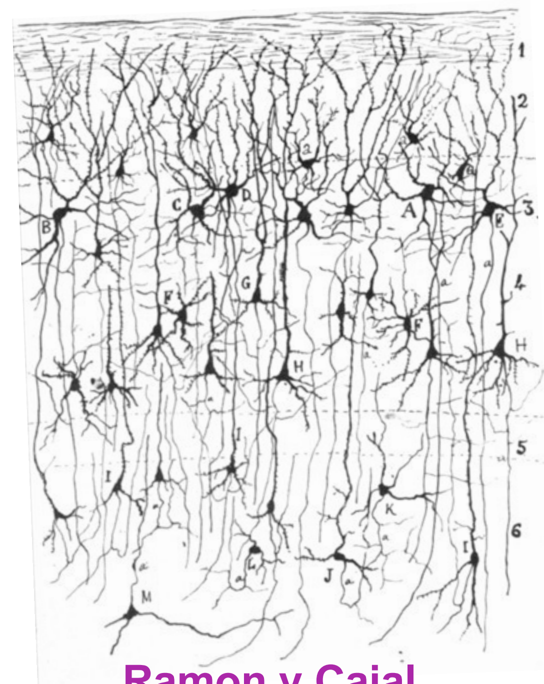
HOW DOES THE CEREBRAL CORTEX WORK?

It supports the highest levels of
biological intelligence in all modalities

VISION, SPEECH, COGNITION, ACTION

Why does the cortex have LAYERS?

How does LAMINAR COMPUTING
give rise to biological intelligence?



Ramon y Cajal

New modeling paradigm: show how variations of the same cortical design carry out all higher intelligent processes

Today, illustrate this with 2 examples:

VISION: unify perceptual learning, grouping, and attention;
also 3D vision and figure-ground perception (spatial)

COGNITION: unify working memory and sequence learning (temporal)

SHARED LAMINAR DESIGN AND CONNECTIVITY ACROSS MULTIPLE AREAS OF GRANULAR NEOCORTEX

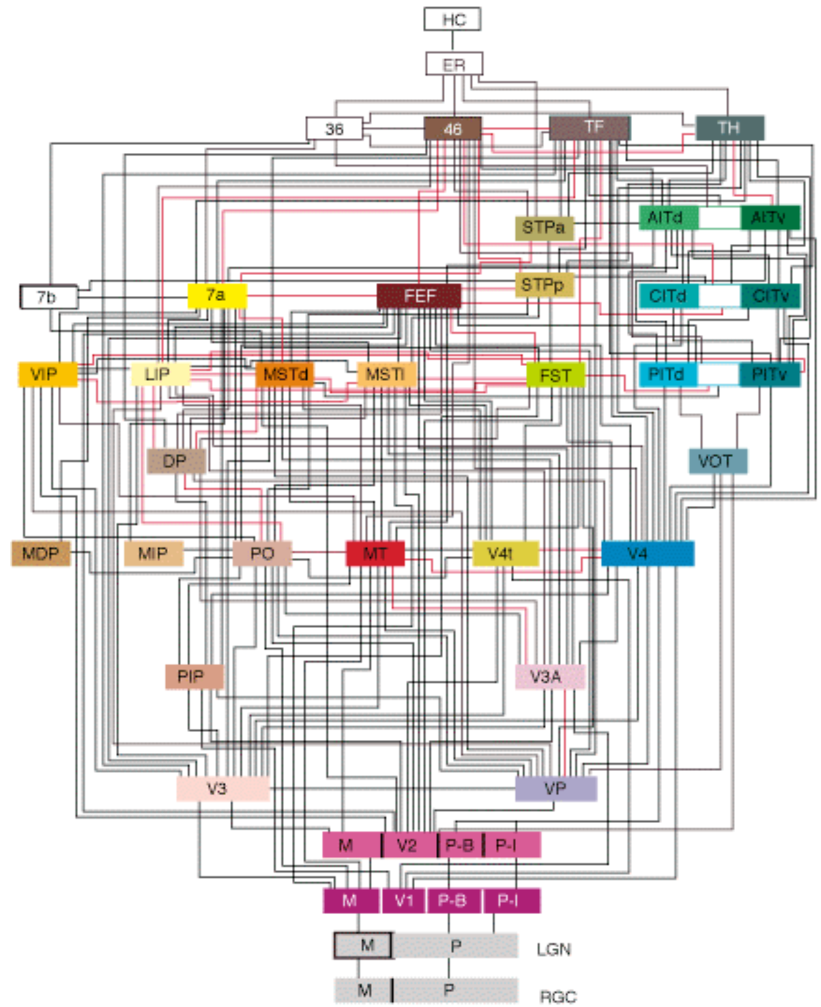
Feedforward

Superficial layers in one area to layers 4 and 6 of the next

Feedback

Deep layers in one area to mainly outside layer 4 of another

Can this known anatomical similarity be elaborated to dynamically explain very different behaviors?



Van Essen et al

VISUAL LAMINAR MODEL: 3D LAMINART

CORTICAL AREAS V1 AND V2

Deep layers (4-6)

Item Storage

Normalization

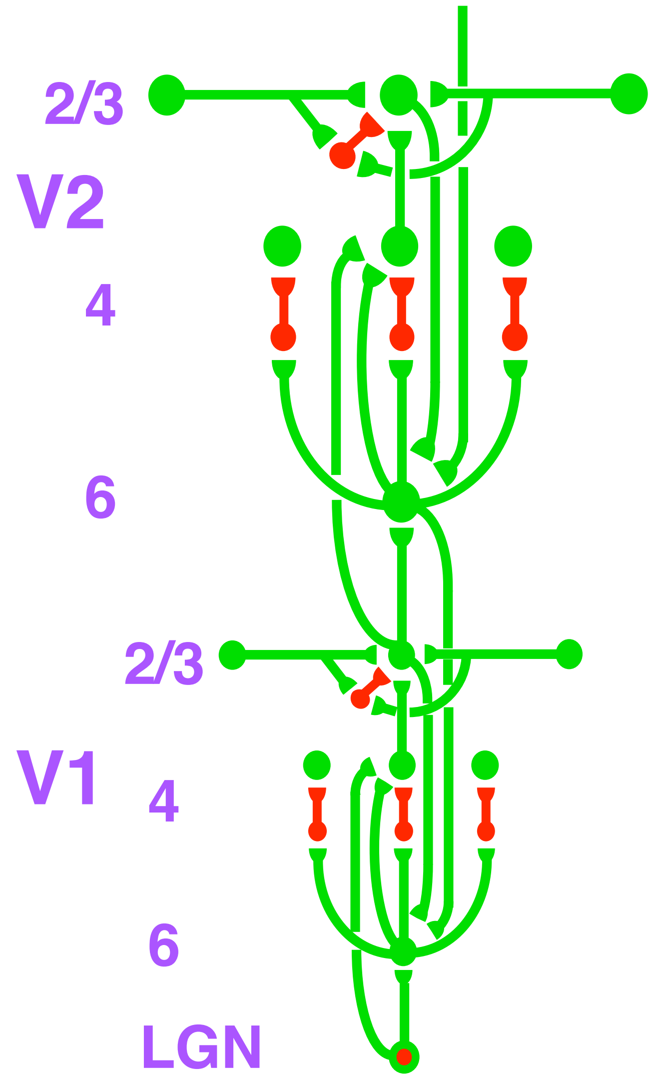
Contrast enhancement

ORIENTED FILTERING of image contrasts

Superficial layers (2/3)

Grouping across processing channels

BINOCULAR MATCHING and PERCEPTUAL GROUPING of oriented image features

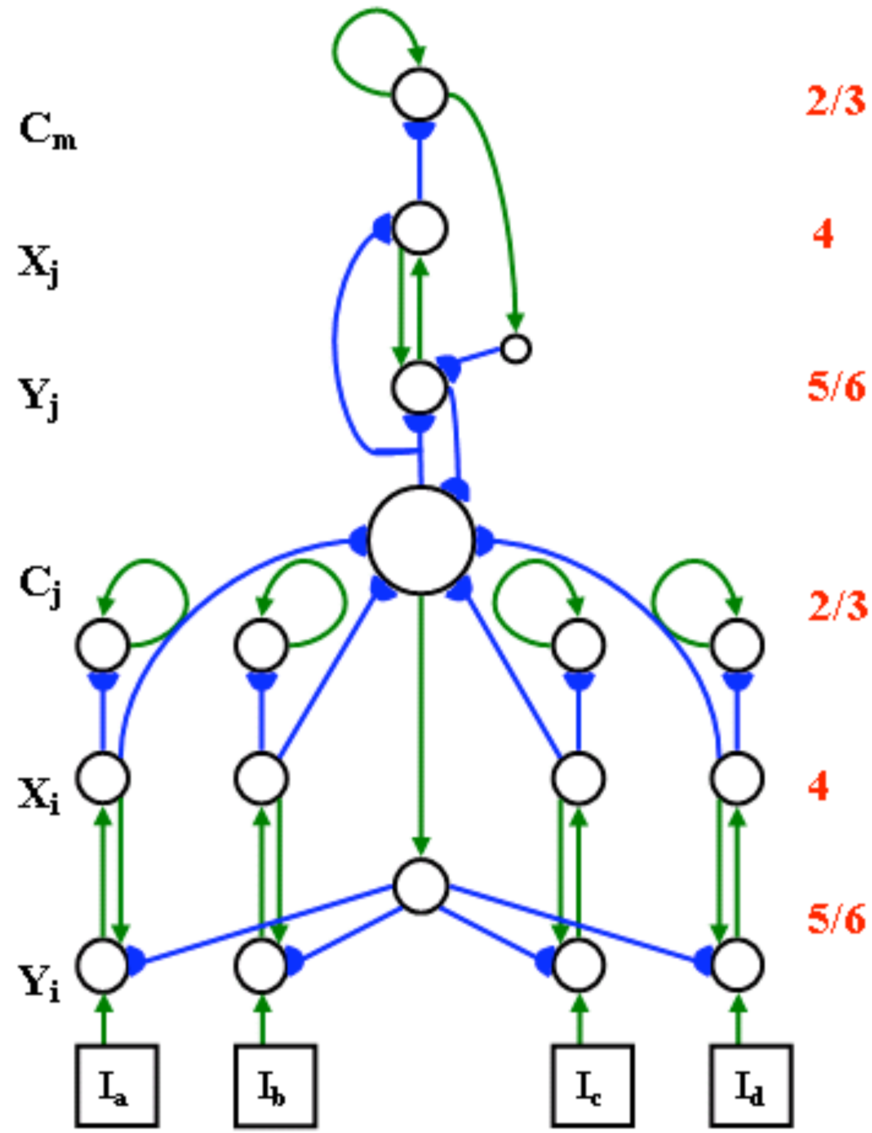


COGNITIVE LAMINAR MODEL: LIST PARSE

LATERAL PREFRONTAL CORTEX

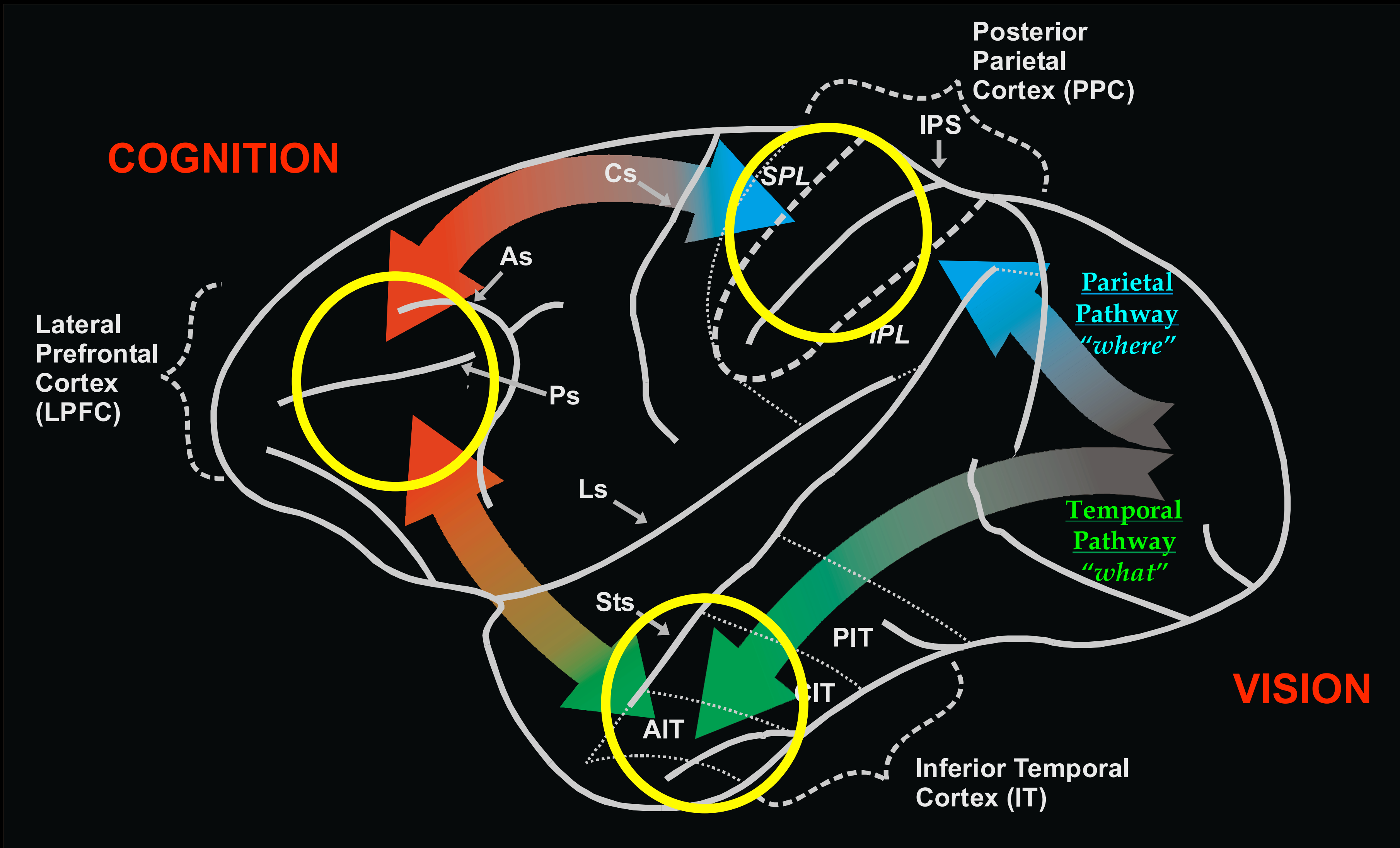
Deep layers (4-6)
Item storage
Normalization
Contrast enhancement
WORKING MEMORY
for short-term storage of
event sequences

Superficial layers (2/3)
Grouping across processing channels
SEQUENCE CHUNKING NETWORK
for long-term coding of
familiar event sequences



Grossberg Plenary
IJCNN'07

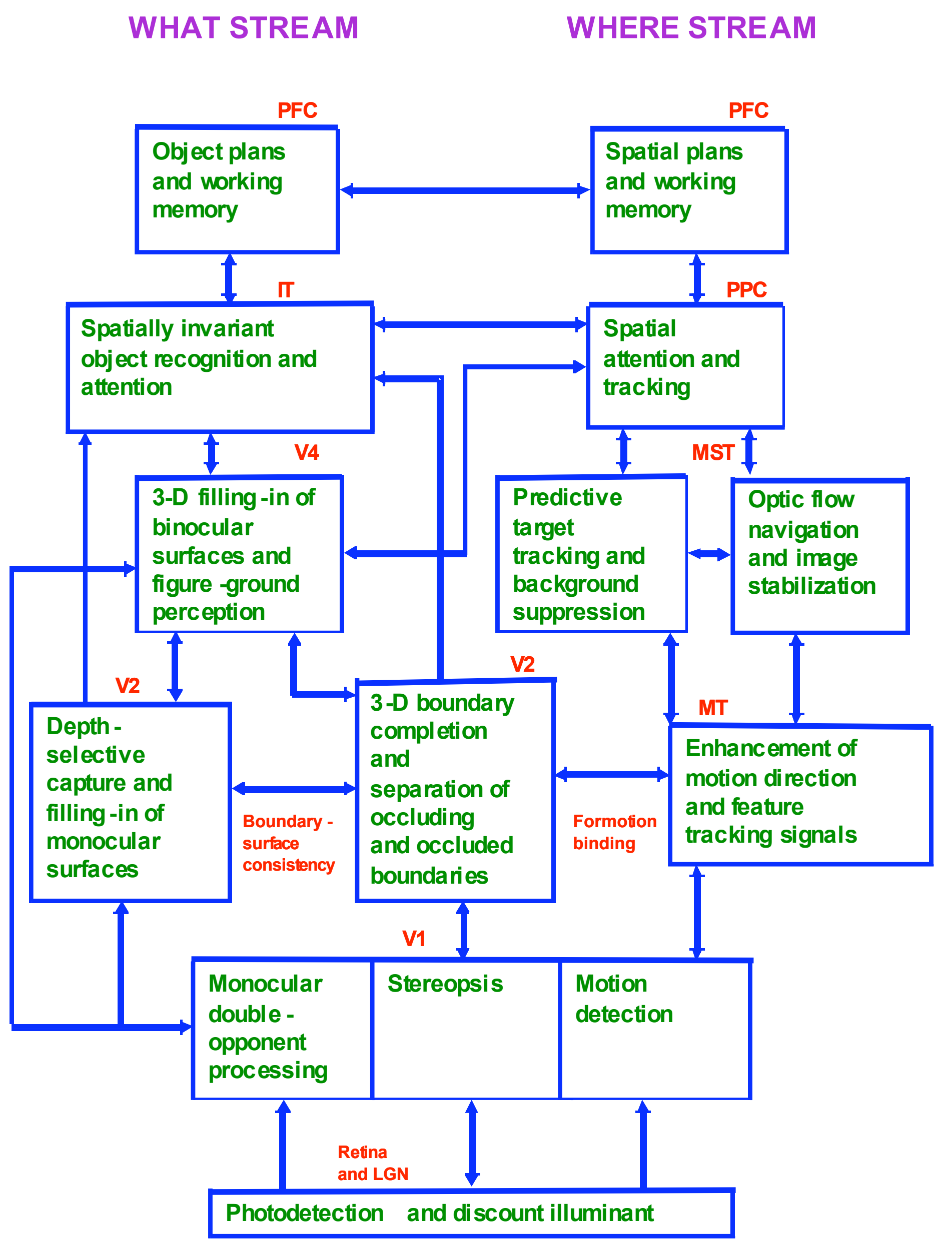
TWO ENDS OF THE CORTICAL HIERARCHY



We've come a long way, baby…
CELEST projects towards a theory of visual intelligence

BOTTOM-UP
TOP-DOWN
HORIZONTAL
interactions everywhere to overcome COMPLEMENTARY WEAKNESSES

Not independent modules



Different projects study different combinations of processes

Together they put much more conceptual pressure on the design of each process than any single project could

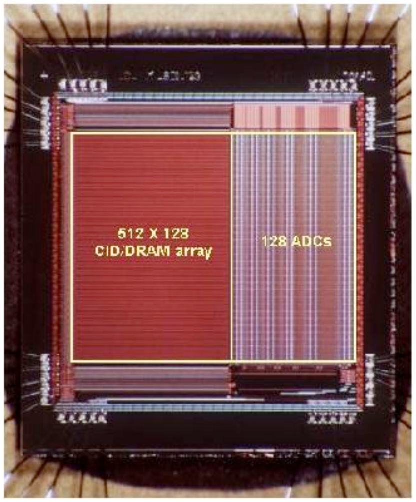
A KEY RESEARCH GOAL

Develop a unified theory of how laminar neocortical circuits are specialized for different types of intelligence

Show how these cortical circuits learn from different environments

A potentially huge technological impact:

A self-organizing VLSI chip set for multiple intelligent tasks



HOW DOES THE CEREBRAL CORTEX WORK?
VISION

Consider 3 basic problems:

1. How does visual cortex stably DEVELOP and LEARN to optimize its structure to process different environments?

2. How does visual cortex GROUP distributed information into emergent object representations?

3. How does top-down ATTENTION bias visual processing to selectively process interesting data?

Breakthrough:

In laminar neocortex, this is really one problem, not three:
show how 1 implies 2 and 3!

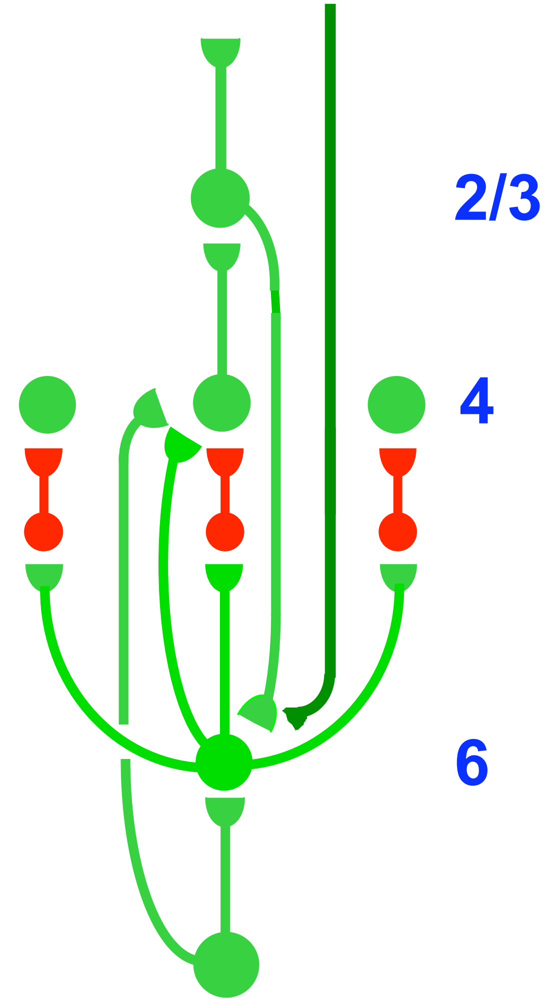
e.g., Grossberg et al. (1997, TINS), Grossberg (2003, Beh&Cog Neurosci Reviews)

WHAT DOES LAMINAR COMPUTING ACHIEVE?

1. Self-stabilizing
   development and learning

2. Seamless fusion of
   pre-attentive automatic
   bottom-up processing
   and
   attentive task-selective
   top-down processing

3. ANALOG COHERENCE: Solution of BINDING PROBLEM for perceptual grouping without a loss of analog sensitivity



Even the earliest visual cortical stages carry out active adaptive information processing:
LEARNING, GROUPING, ATTENTION

LAMINAR COMPUTING: A NEW WAY TO COMPUTE

1. FEEDFORWARD AND FEEDBACK

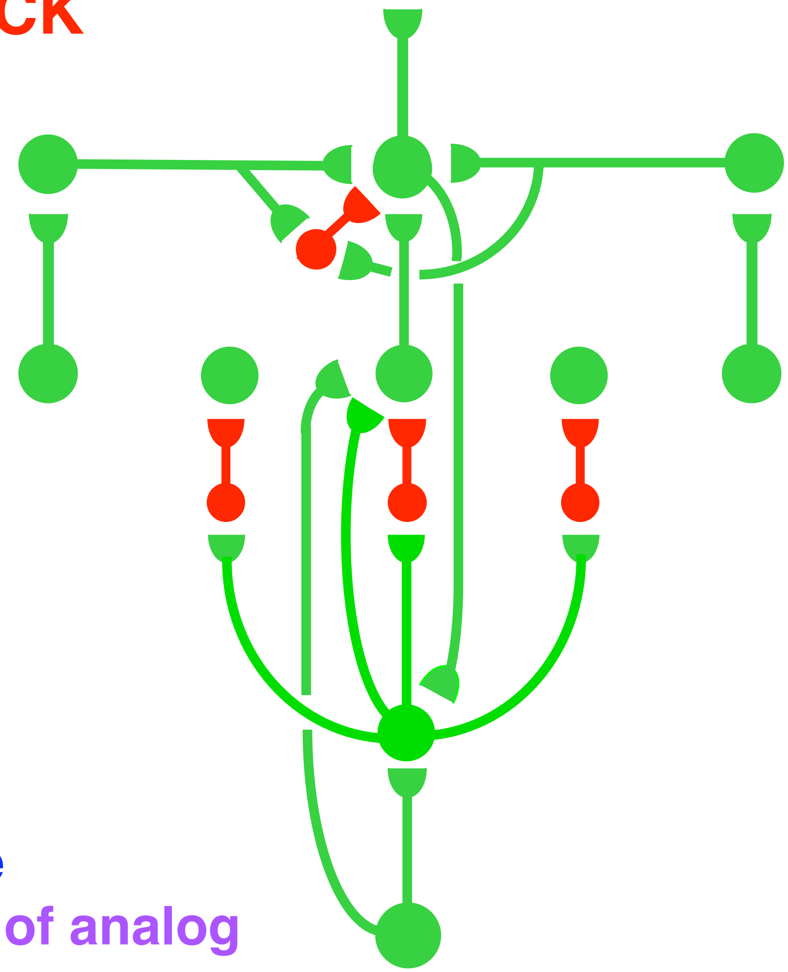
Fast feedforward processing when data are unambiguous

e.g., Thorpe et al

Slower feedback chooses among ambiguous alternatives:

self-normalizing competition

"real-time probability theory"



A self-organizing system that trades certainty against speed

Goes beyond Bayesian models!

2. ANALOG AND DIGITAL

ANALOG COHERENCE combines the stability of digital with the sensitivity of analog

3. PRE-ATTENTIVE AND ATTENTIVE LEARNING

Reconciles the differences of (e.g.) Helmholtz and Kanizsa

"A pre-attentive grouping is its own 'attentional' prime"

HOW DOES THE CEREBRAL CORTEX WORK?
VISION

Started with 3 basic problems:

1. **How does visual cortex stably DEVELOP and LEARN to optimize its structure to process different environments?**

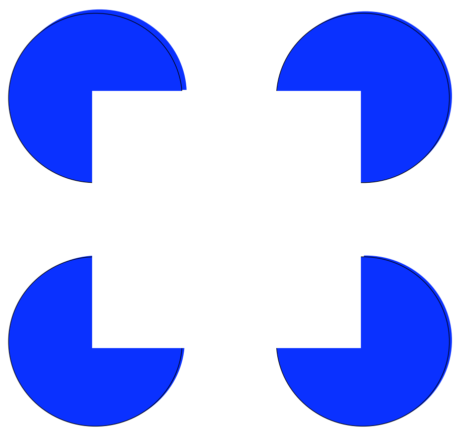
2. **How does visual cortex GROUP distributed information into emergent object representations?**

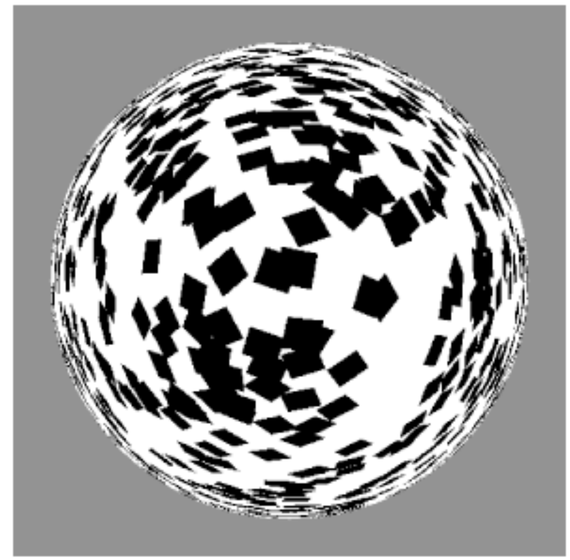
3. **How does top-down ATTENTION bias visual processing to selectively process interesting data?**

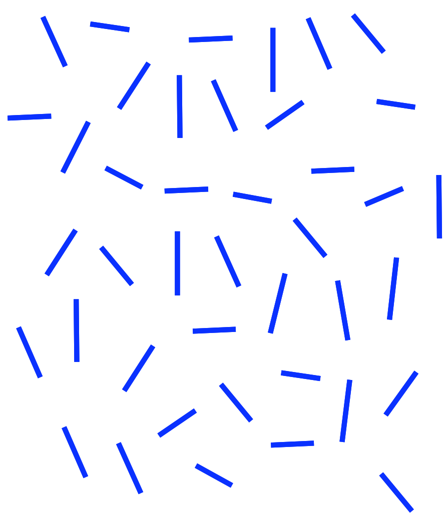
What is the relationship between GROUPING and ATTENTION?

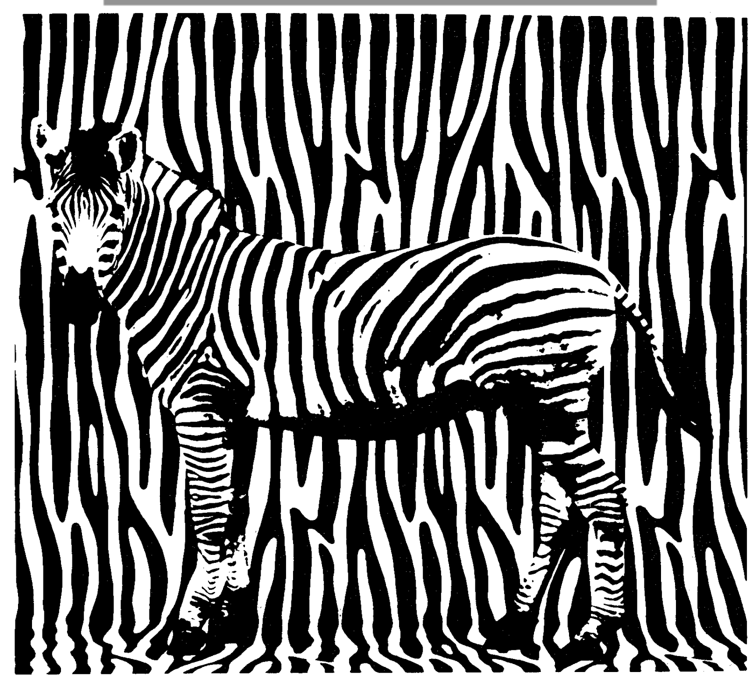
I will discuss circuits, but the work always starts with psychological data

VISUAL BOUNDARIES OR GROUPINGS

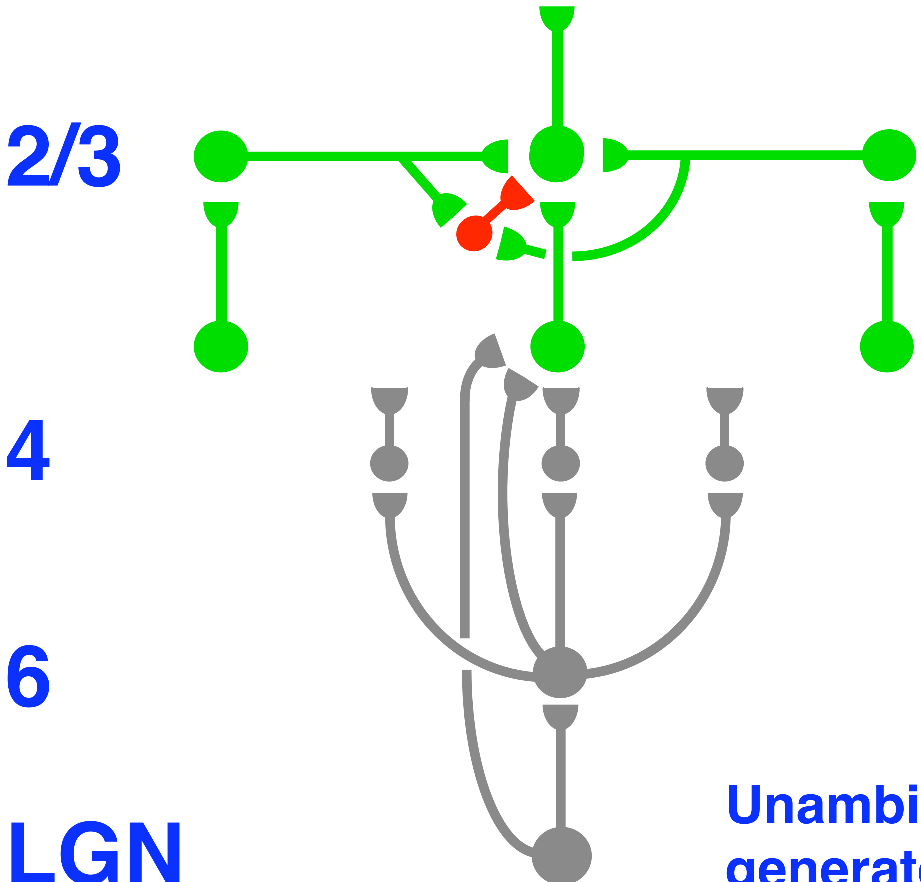








GROUPING STARTS IN LAYER 2/3



Long-range
horizontal excitation links
collinear, coaxial receptive fields

Gilbert & Wiesel, 1989
Bosking et al., 1997
Schmidt et al, 1997

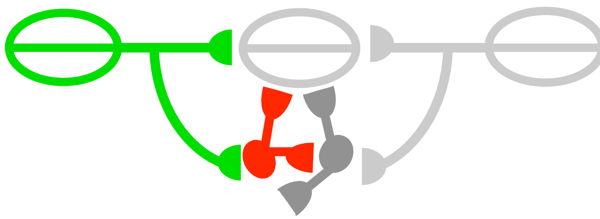
Short-range
disynaptic inhibition of target pyramidal via pool of interneurons

Hirsch & Gilbert, 1991

Unambiguous groupings can form and generate feedforward outputs quickly

Thorpe et al, 1996

1984 PREDICTION:
BIPOLE PROPERTY CONTROLS PERCEPTUAL GROUPING



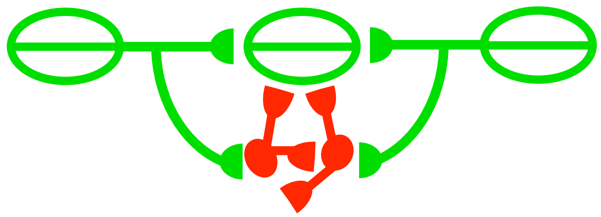
Input on just one side

ONE-AGAINST-ONE:
Balanced Excitation and Inhibition

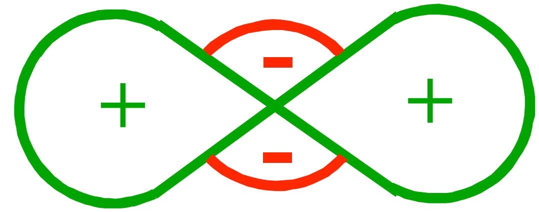
Cell not excited

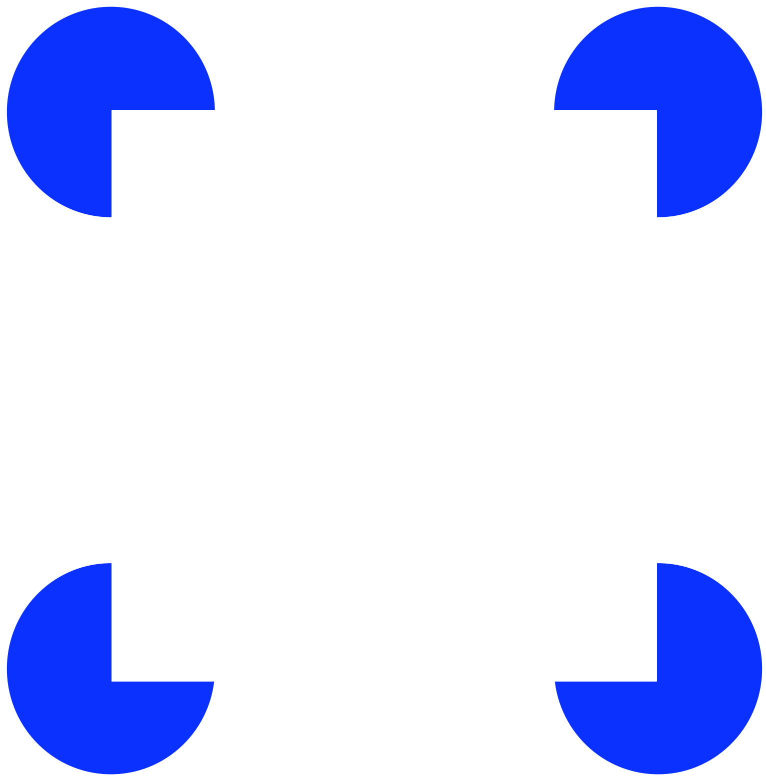
Grossberg, 1984
Grossberg & Mingolla, 1985
Laminar: Grossberg, Mingolla & Ross, 1997

BIPOLE PROPERTY CONTROLS PERCEPTUAL GROUPING



vs.





Collinear input on both sides

Excitatory inputs summate

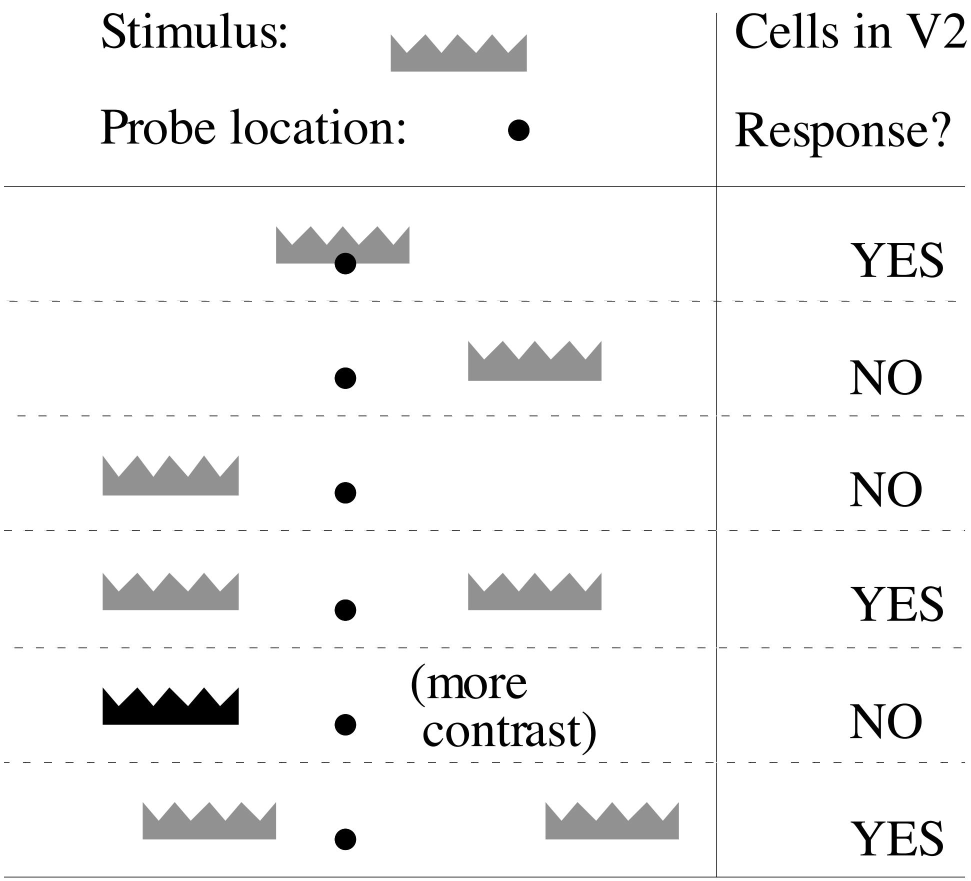
Inhibitory inputs normalize

Shunting inhibition!

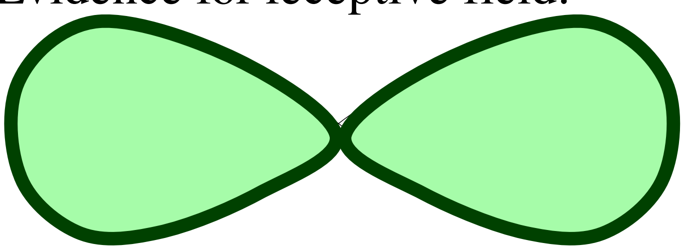
TWO-AGAINST-ONE

Cell is excited

BIPOLES: FIRST NEUROPHYSIOLOGICAL EVIDENCE (V2)



Evidence for receptive field:

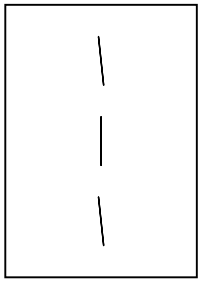


von der Heydt, Peterhans, and Baumgartner, 1984

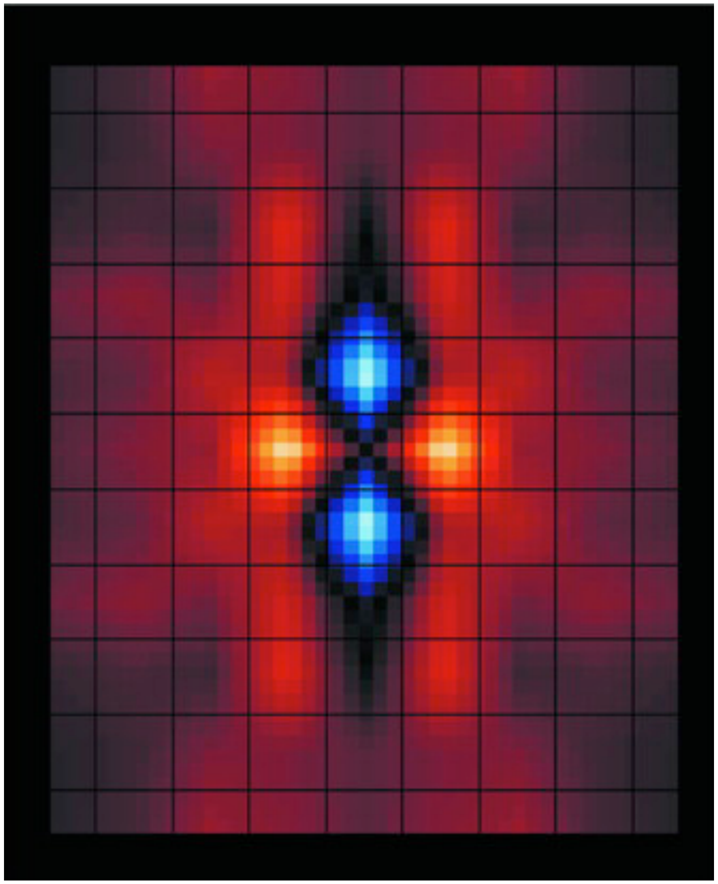
Peterhans and von der Heydt, 1988

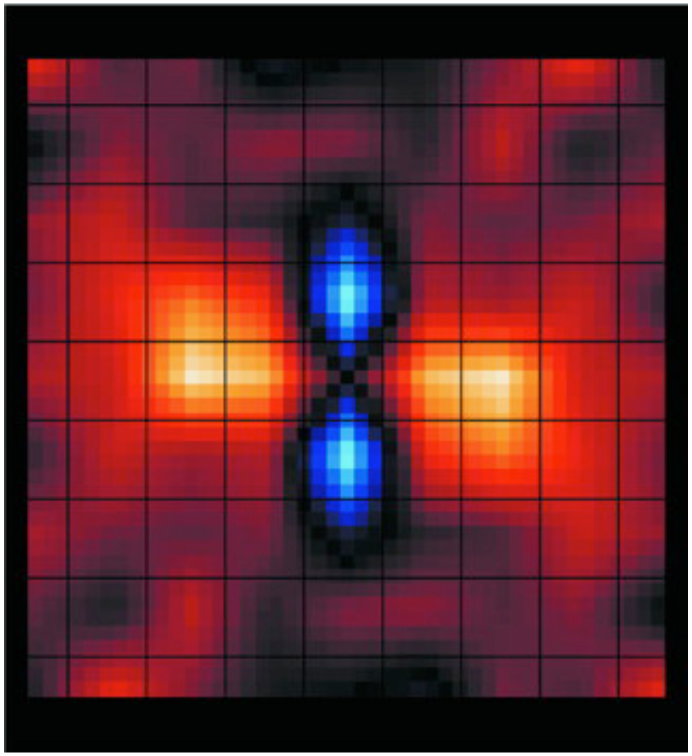
KAPADIA, ITO, GILBERT & WESTHEIMER (1995)

Psychophysics

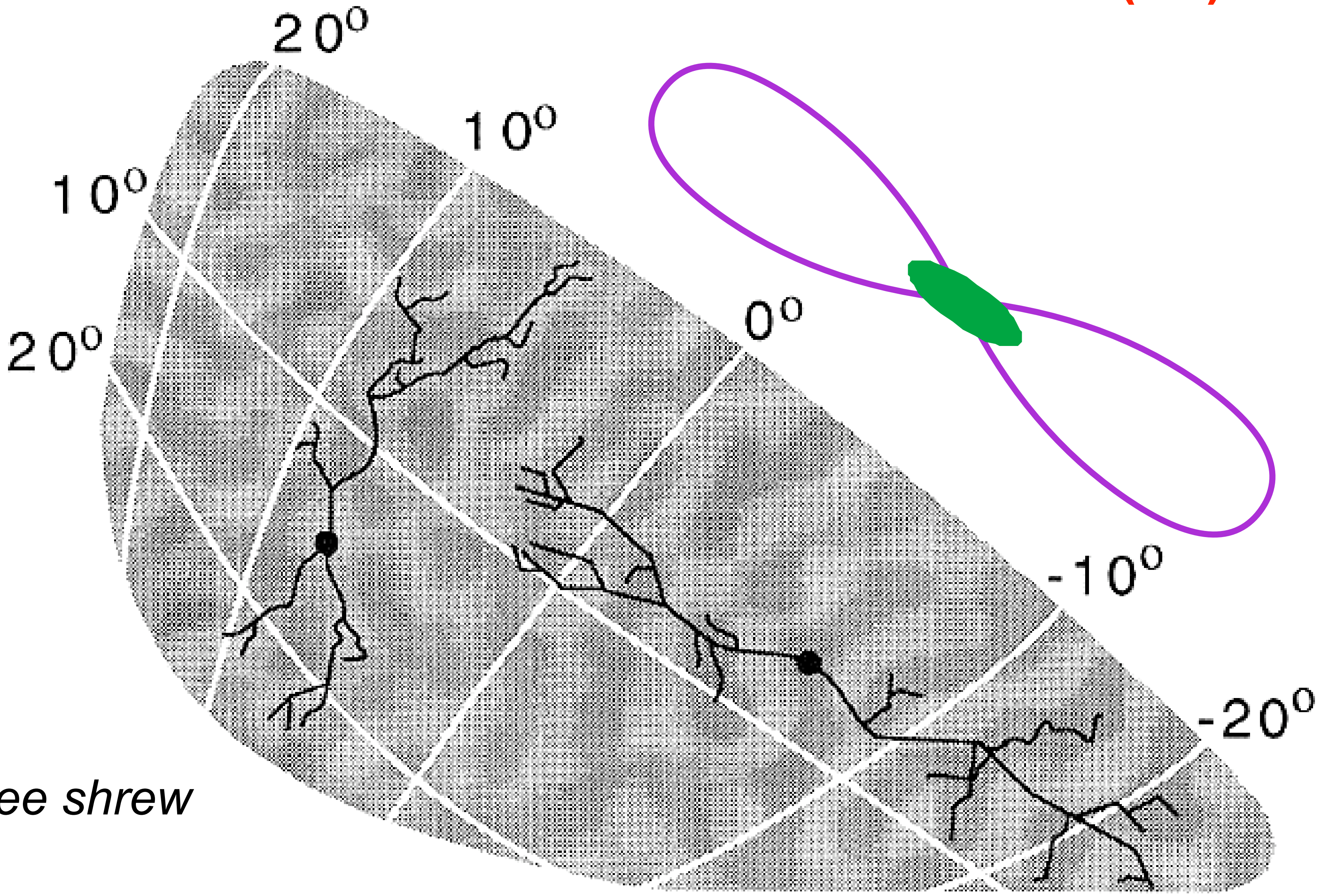


Neurophysiology
V1





ANATOMY: HORIZONTAL CONNECTIONS (V1)

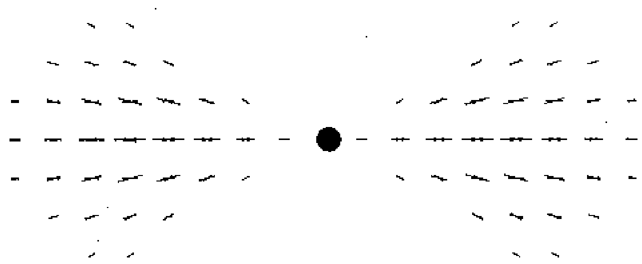


tree shrew

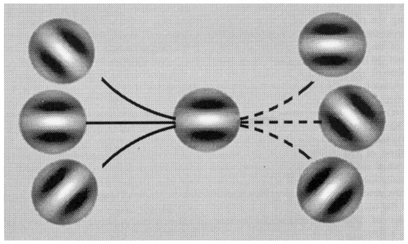
Bosking, *et al.*, 1997

BIPOLES THROUGH THE AGES

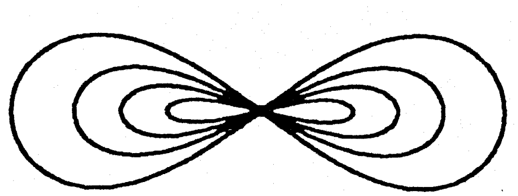
Grossberg and Mingolla, 1985



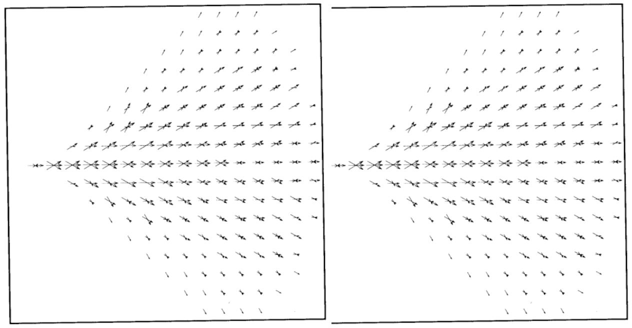
Field, Hayes, and Hess, 1993
"association field"



Heitger and von der Heydt, 1993

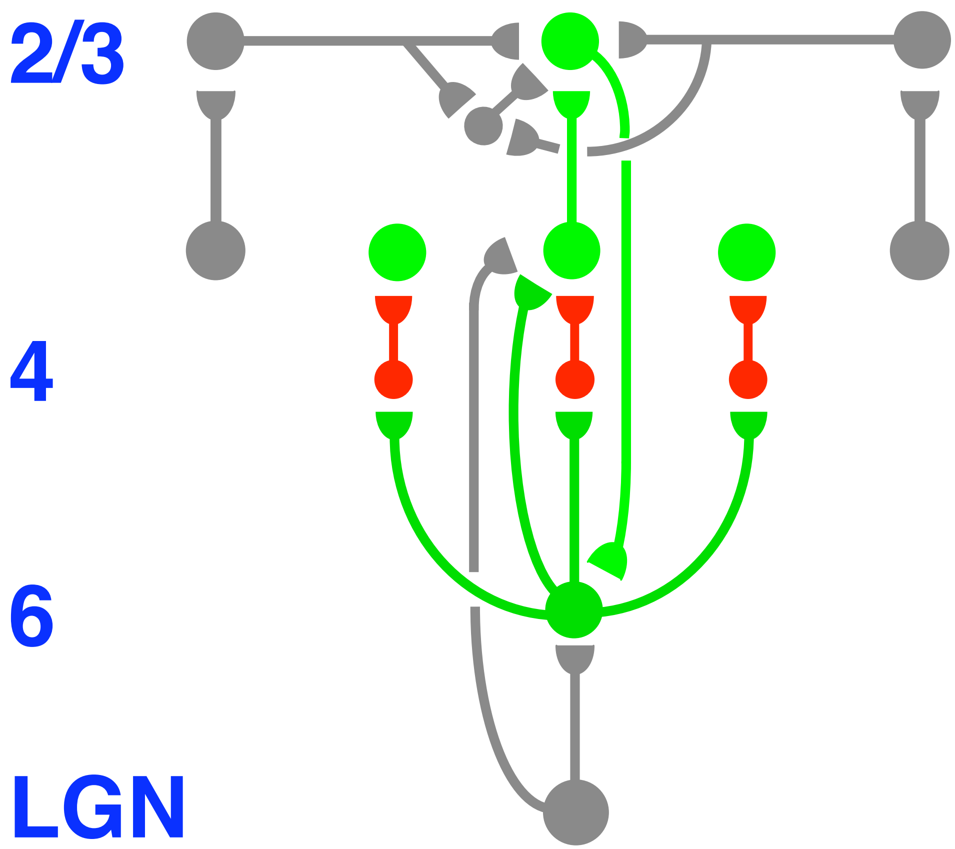


Williams and Jacobs, 1997



Cf. "relatability" geometric constraints on which contours get to group with which Kellman & Shipley, 1991
Also, Ullman, Zucker, Mumford, Guy & Medione "tensor voting"

HOW IS THE FINAL GROUPING SELECTED?
FOLDED FEEDBACK



Layer 2/3 groupings feed back into

6-to-4 on-center off-surround:

Direct layer 2/3-to-6 path

Can also go via layer 5

Blasdel et al., 1985
Kisvarday et al., 1989

Strongest grouping enhanced by its on-center

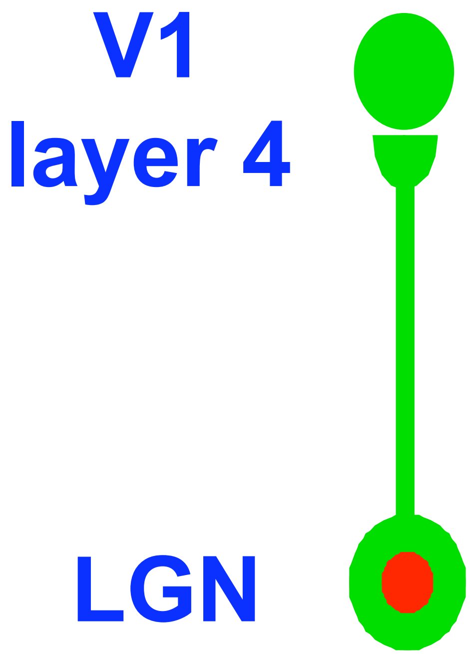
Inputs to weaker groupings suppressed by off-surround

Interlaminar feedback creates functional columns

Activities of conflicting groupings are reduced by self-normalizing inhibition, slowing processing; intracortical feedback selects and contrast-enhances the winning grouping, speeding processing

HOW ARE LAYER 2/3 BIPOLE CELLS ACTIVATED?

DIRECT BOTTOM-UP ACTIVATION OF LAYER 4



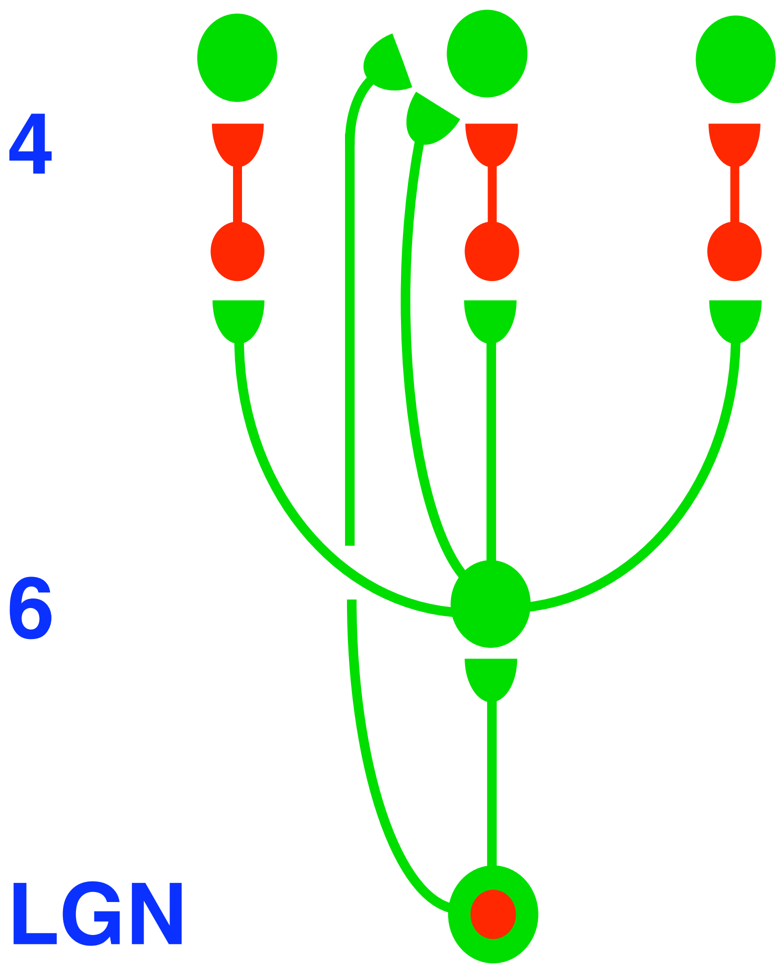
Strong bottom-up LGN input to layer 4

Stratford et al. (1996)
Chung & Ferster (1998)

(Many details omitted!)

ANOTHER BOTTOM-UP INPUT TO LAYER 4: *WHY?*

LAYER 6-TO-4 ON-CENTER OFF-SURROUND



LGN projects to layers 6 and 4

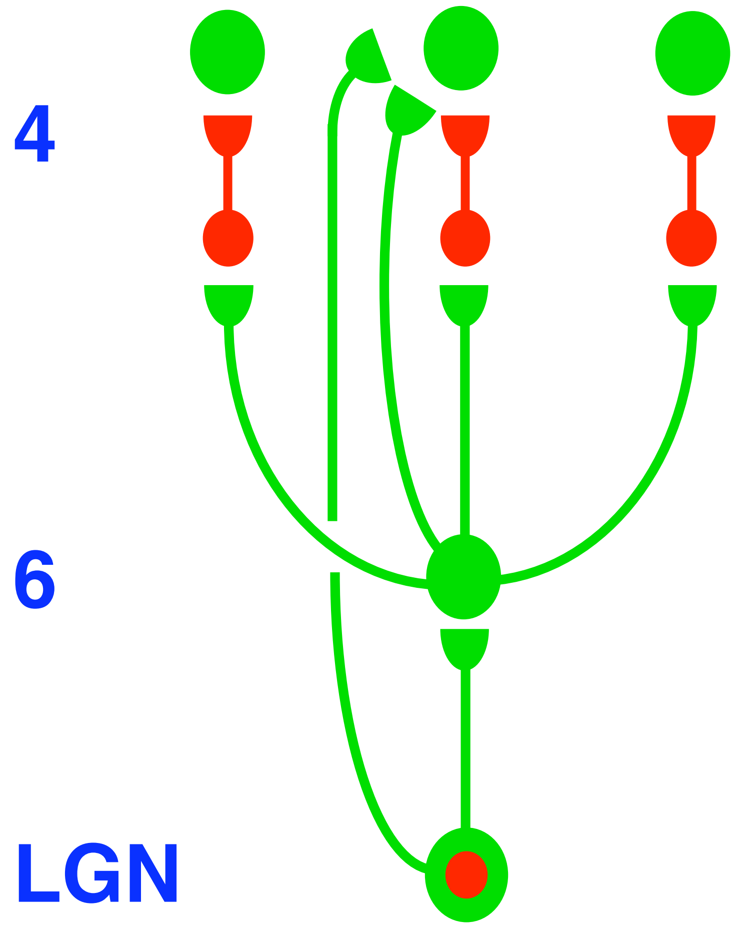
Layer 6 excites spiny stellates in column above it

Medium-range connections onto inhibitory interneurons

6-to-4 path acts as on-center off-surround

Grieve & Sillito, 1991, 1995
Ahmed et al., 1994, 1997

BOTTOM-UP CONTRAST NORMALIZATION



Together, direct LGN-to-4 path and 6-to-4 on-center off-surround provide contrast normalization if cells obey shunting or membrane equation dynamics

Grossberg, 1968, 1973

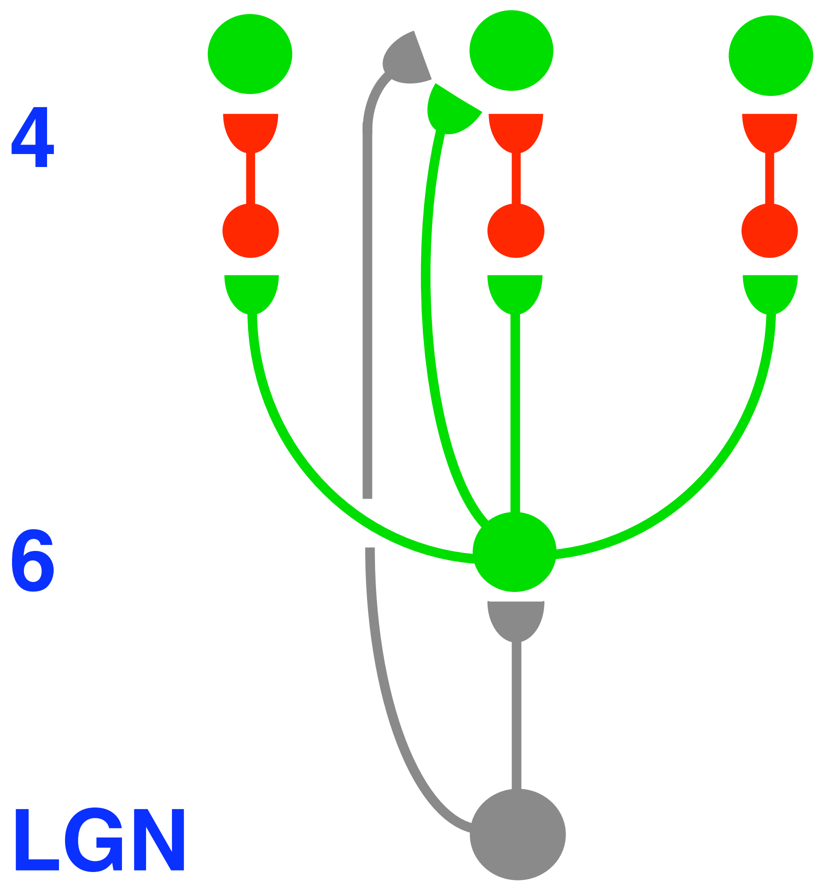
Sperling and Sondhi, 1968

Heeger, 1992

Douglas et al., 1995

Shapley et al., 2004

MODULATION OR PRIMING BY 6-TO-4 ON-CENTER



On-center 6-to-4 excitation is inhibited down to being modulatory
(priming, subthreshold)

Stratford et. al, 1996

Callaway, 1998

On-center 6-to-4 excitation cannot activate layer 4 on its own

Clarifies need for direct path

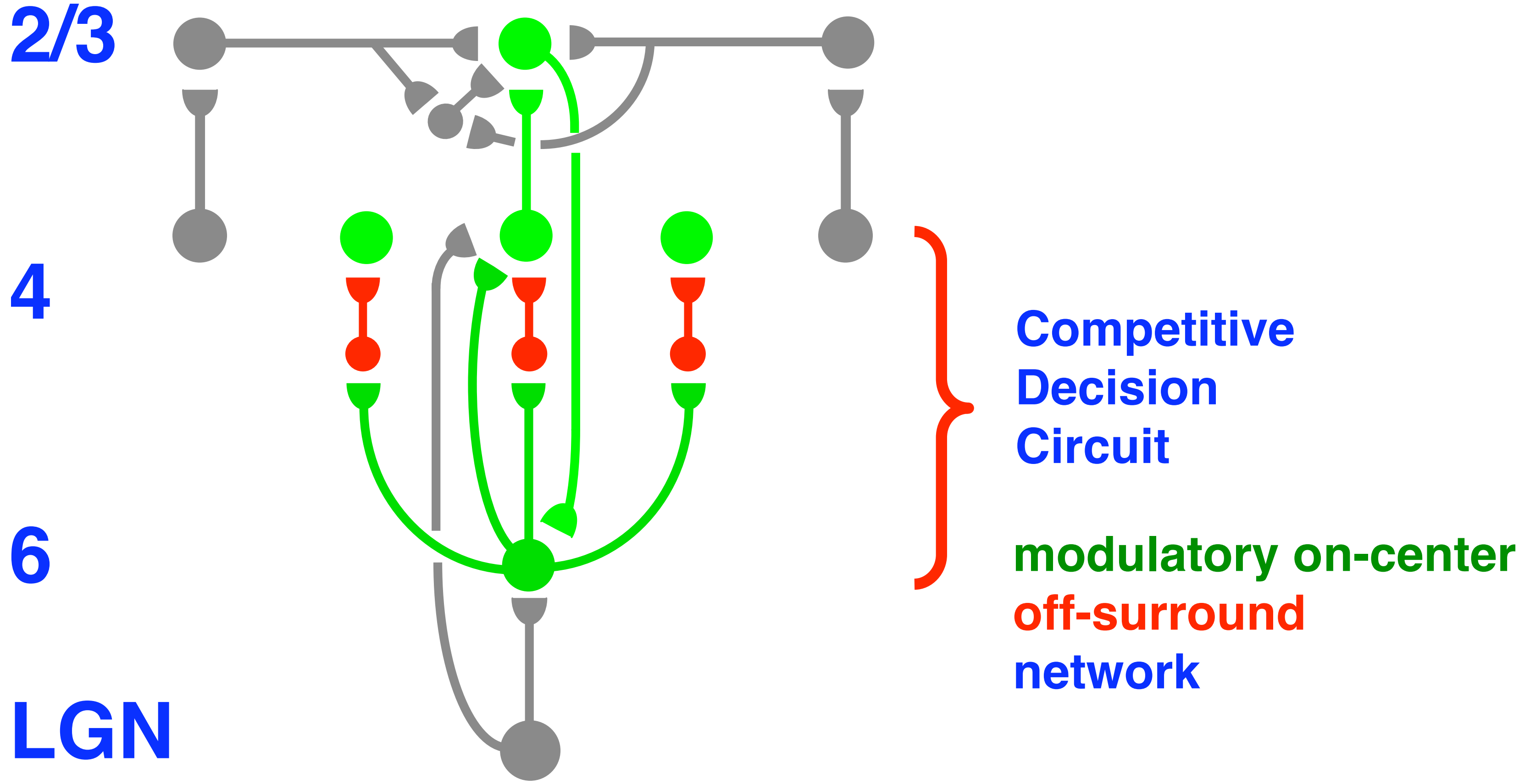
Predictions:

Test modulatory property directly

Plays key role in stable grouping, development and learning

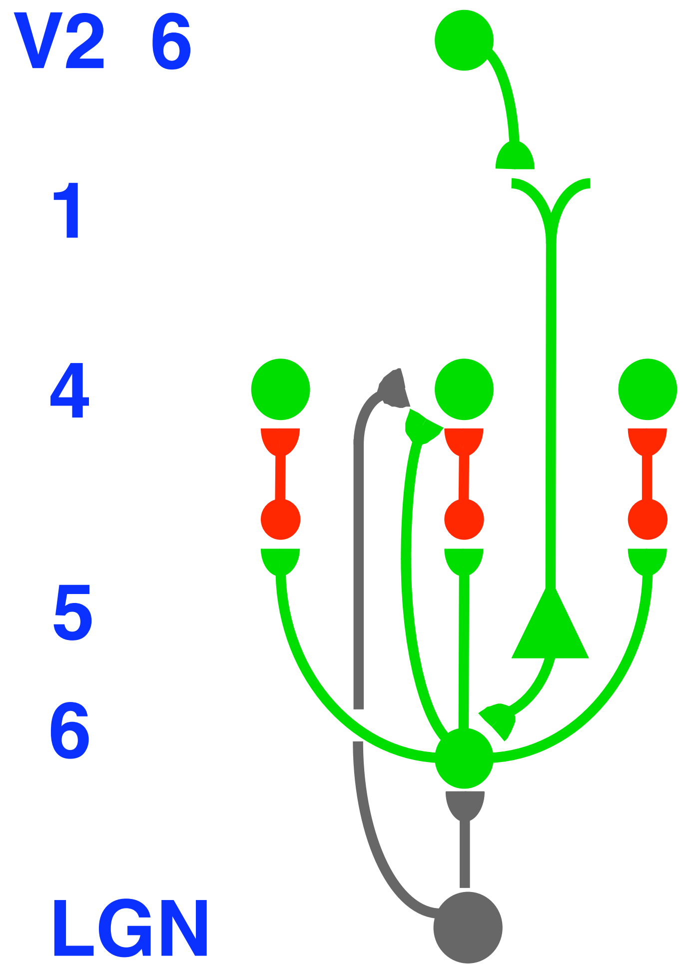
ART MATCHING RULE!

BOTTOM-UP FILTERS AND INTRACORTICAL GROUPING FEEDBACK USE THE SAME 6-TO-4 DECISION CIRCUIT



TOP-DOWN INTERCORTICAL ATTENTION ALSO USES THE SAME 6-TO-4 DECISION CIRCUIT!

TOP-DOWN ATTENTION AND FOLDED FEEDBACK



Attentional signals also feed back into 6-to-4 on-center off-surround

1-to-5-to-6 feedback path

Macaque: Lund & Boothe, 1975
Cat: Gilbert & Wiesel, 1979

V2-to-V1 feedback is on-center off-surround and affects layer 6 of V1 the most

Bullier et al., 1996
Sandell & Schiller, 1982

Attended stimuli enhanced
Ignored stimuli suppressed

This circuit supports the predicted

ART MATCHING RULE!

ART: LINK BETWEEN COMPETITION AND ATTENTION

Neurophysiological Data

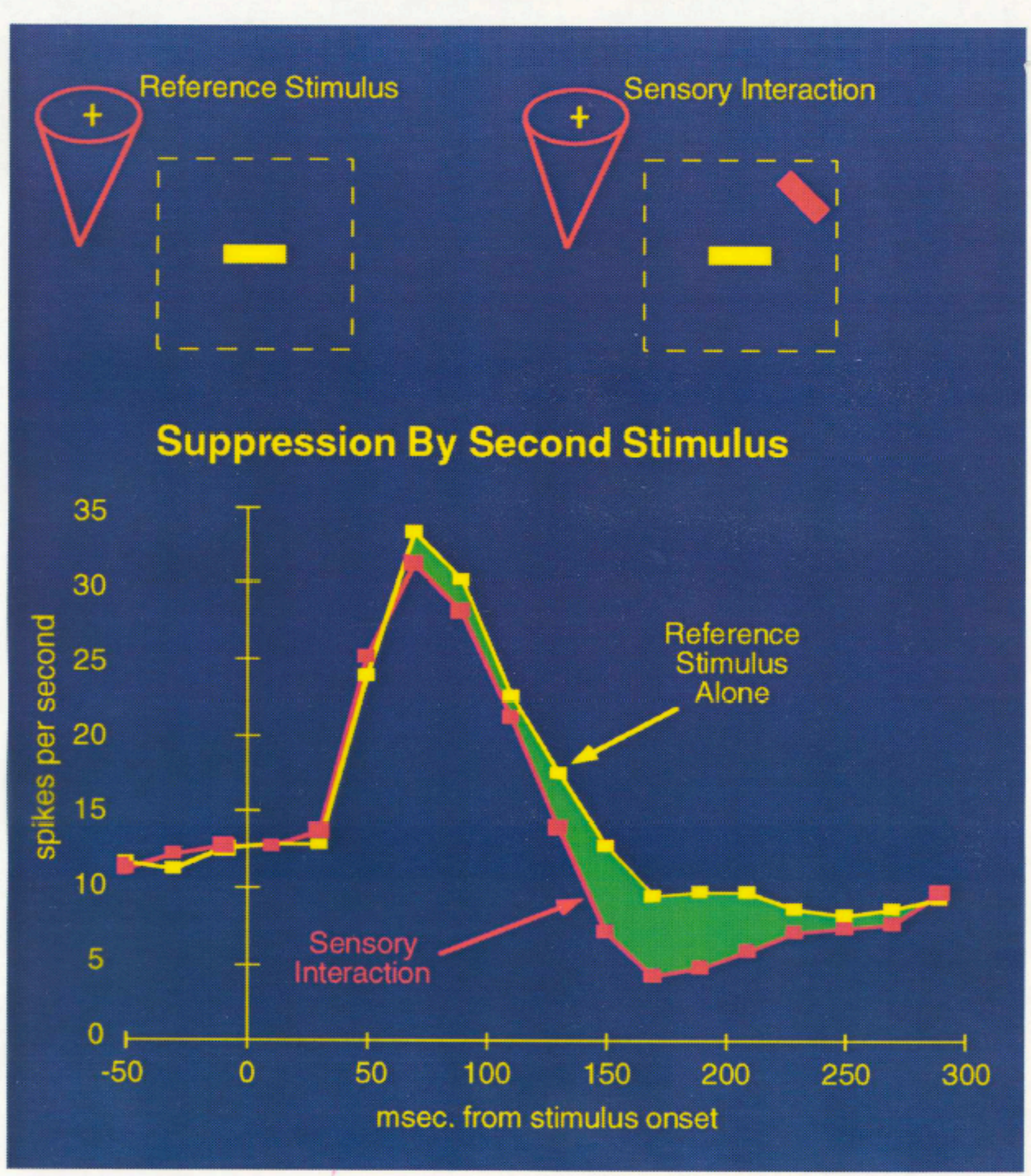
Reynolds, J., Nicholas, J., Chelazzi, L., & Desimone, R. (1995)
Spatial attention protects macaque V2 and V4 cells from the influence of non-attended stimuli
Society for Neuroscience Abstracts, 693.1, page 356

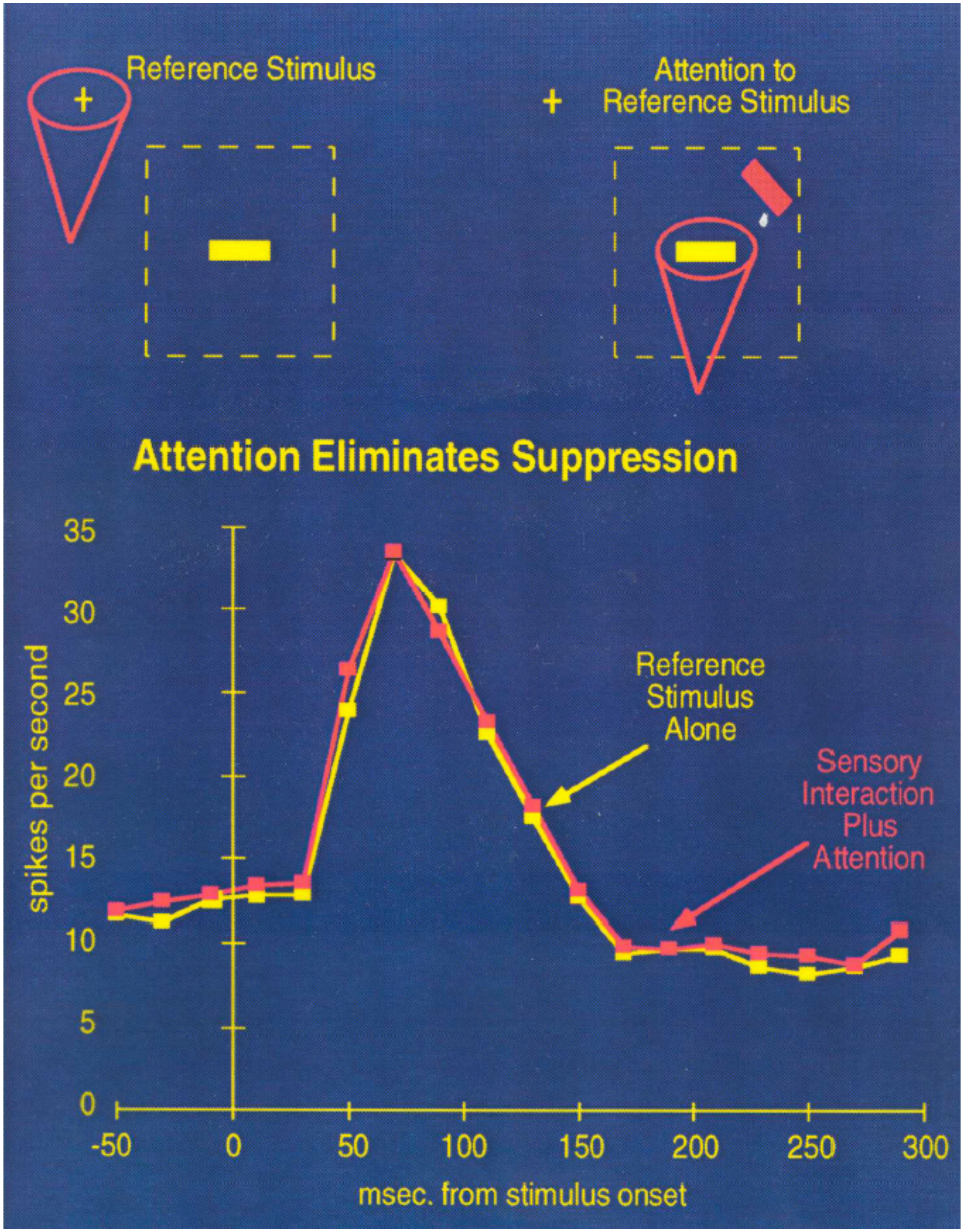
Reynolds, J., Chelazzi, L., & Desimone, R. (1999)
Competitive mechanisms subserve attention in Macaque areas V2 and V4
Journal of Neuroscience, 19, 1736 - 1753

****

Carpenter, G.A., Grossberg, S., and Reynolds, J.H. (1991)
ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network
Neural Networks, 4, 565-588

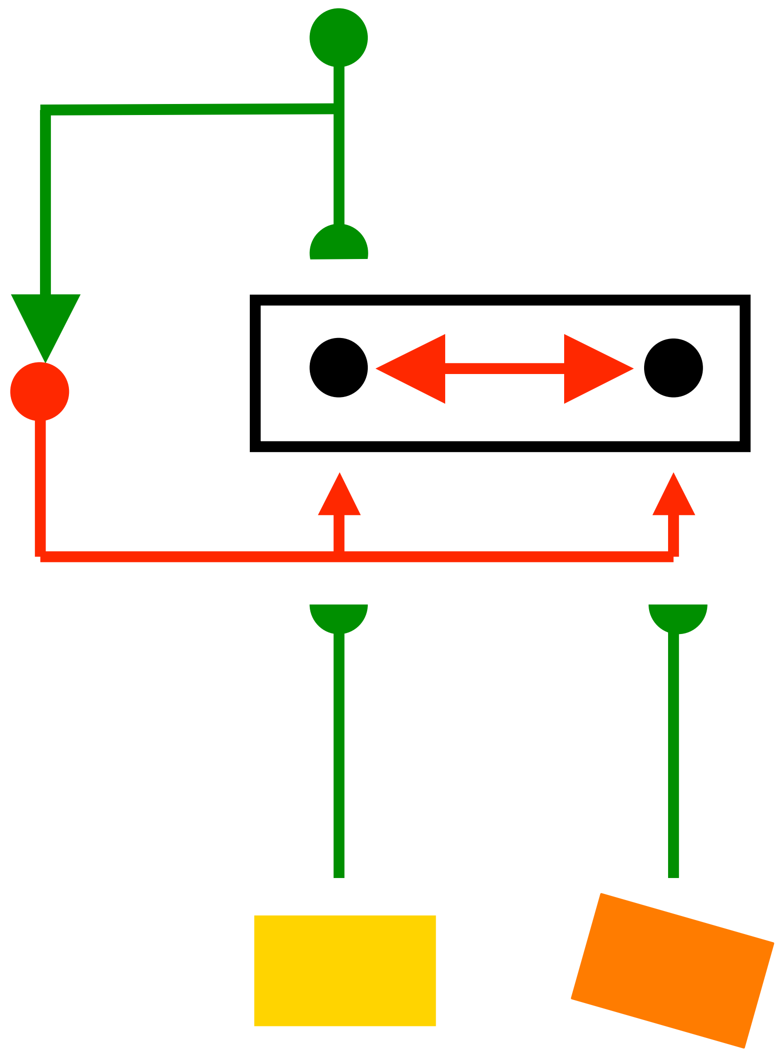
Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H., and Rosen, D.B. (1992)
Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps
IEEE Transactions on Neural Networks, 3, 698-713


Reference Stimulus
Sensory Interaction
+
+
Suppression By Second Stimulus
spikes per second
0
5
10
15
20
25
30
35
-50
0
50
100
150
200
250
300
msec. from stimulus onset
Reference
Stimulus
Alone
Sensory
Interaction


Reference Stimulus
Attention to
Reference Stimulus
Attention Eliminates Suppression
35
30
25
20
15
10
5
0
spikes per second
-50
0
50
100
150
200
250
300
msec. from stimulus onset
Reference
Stimulus
Alone
Sensory
Interaction
Plus
Attention

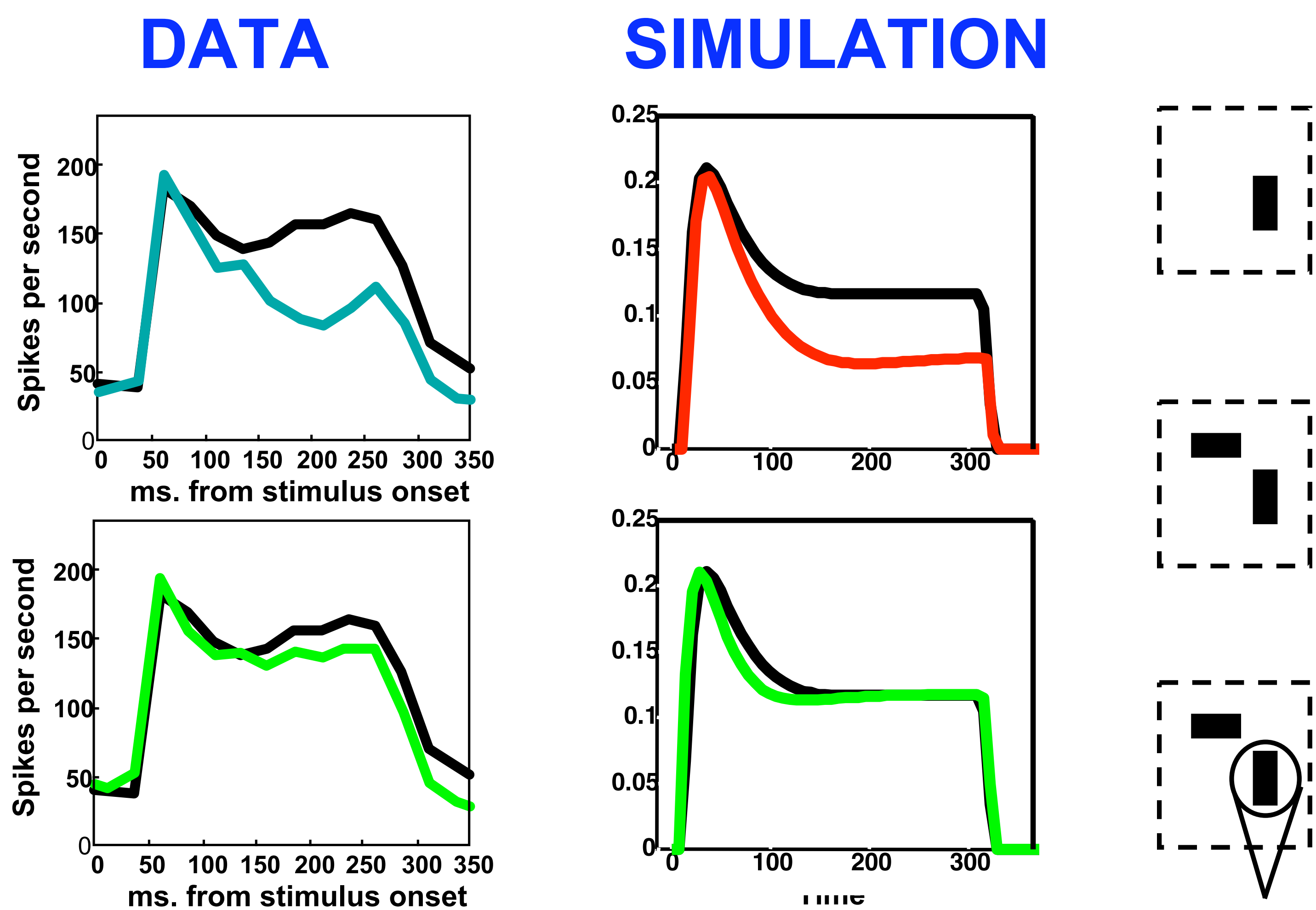
COMPETITIVE MATCHING AND ATTENTION

ART Matching Rule



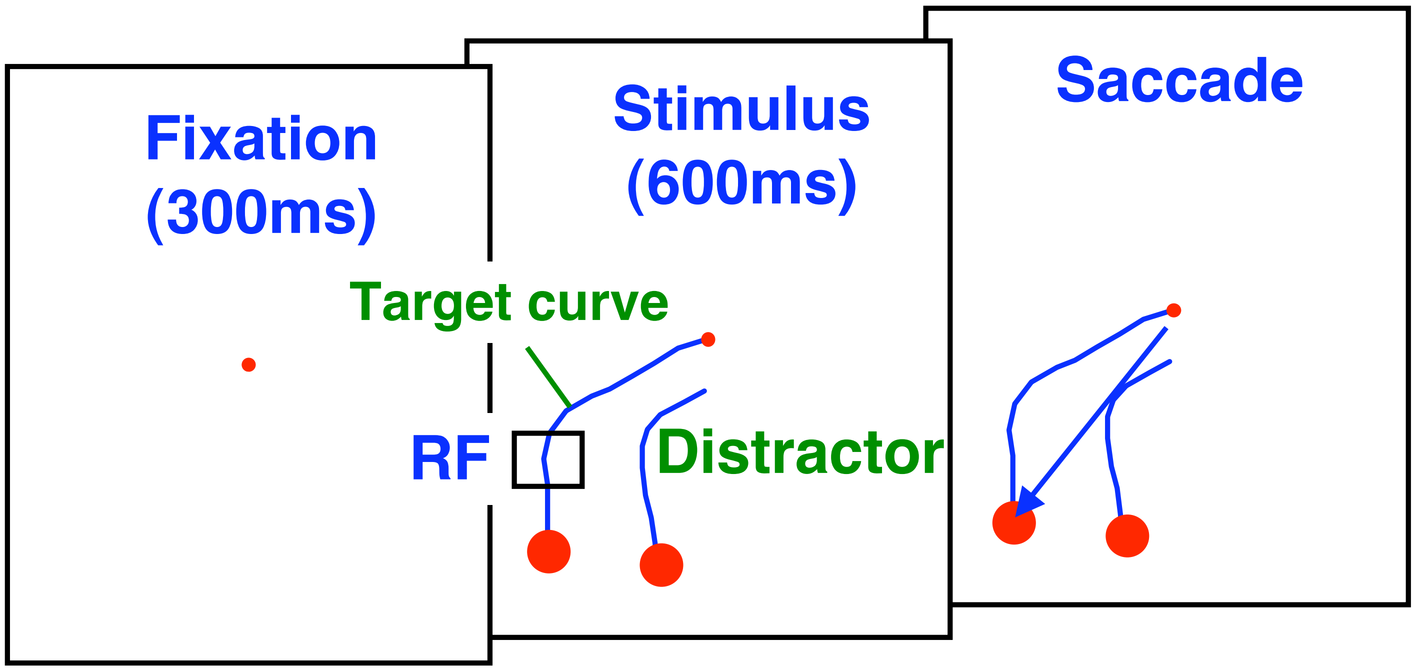
SIMULATION OF REYNOLDS ET AL. (1995)

Grossberg and Raizada (2000, Vision Research)

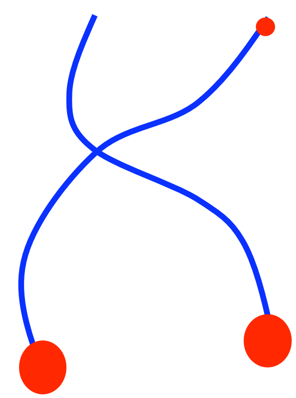


Data plots adapted with permission from Reynolds et al.

ATTENTION FLOWS ALONG CURVES: ROELFSEMA ET AL. (1998): MACAQUE V1

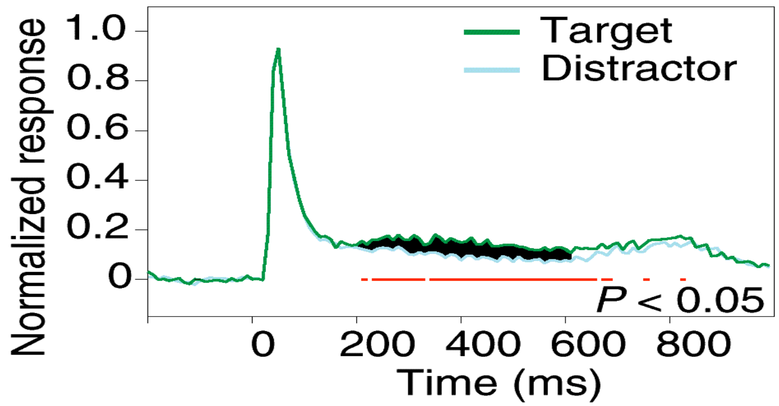


Crossed-curve condition:
Attention flows across junction between smoothly connected curve segments
(Good Continuation)

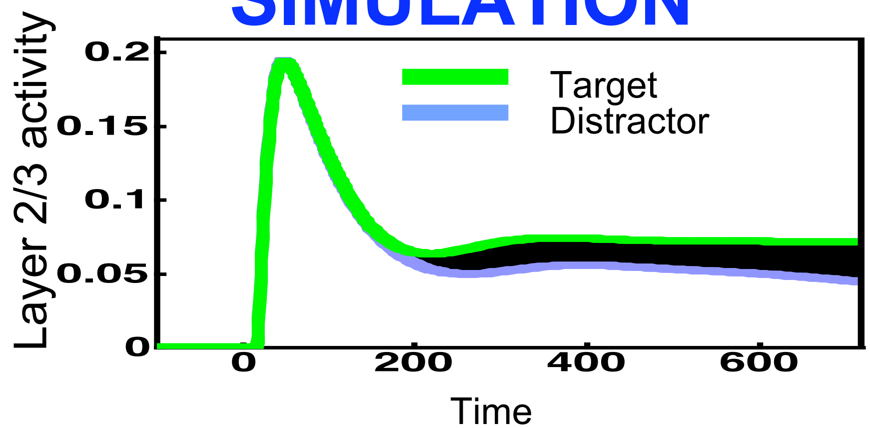


SIMULATION OF ROELFSEMA ET AL. (1998)

DATA



SIMULATION

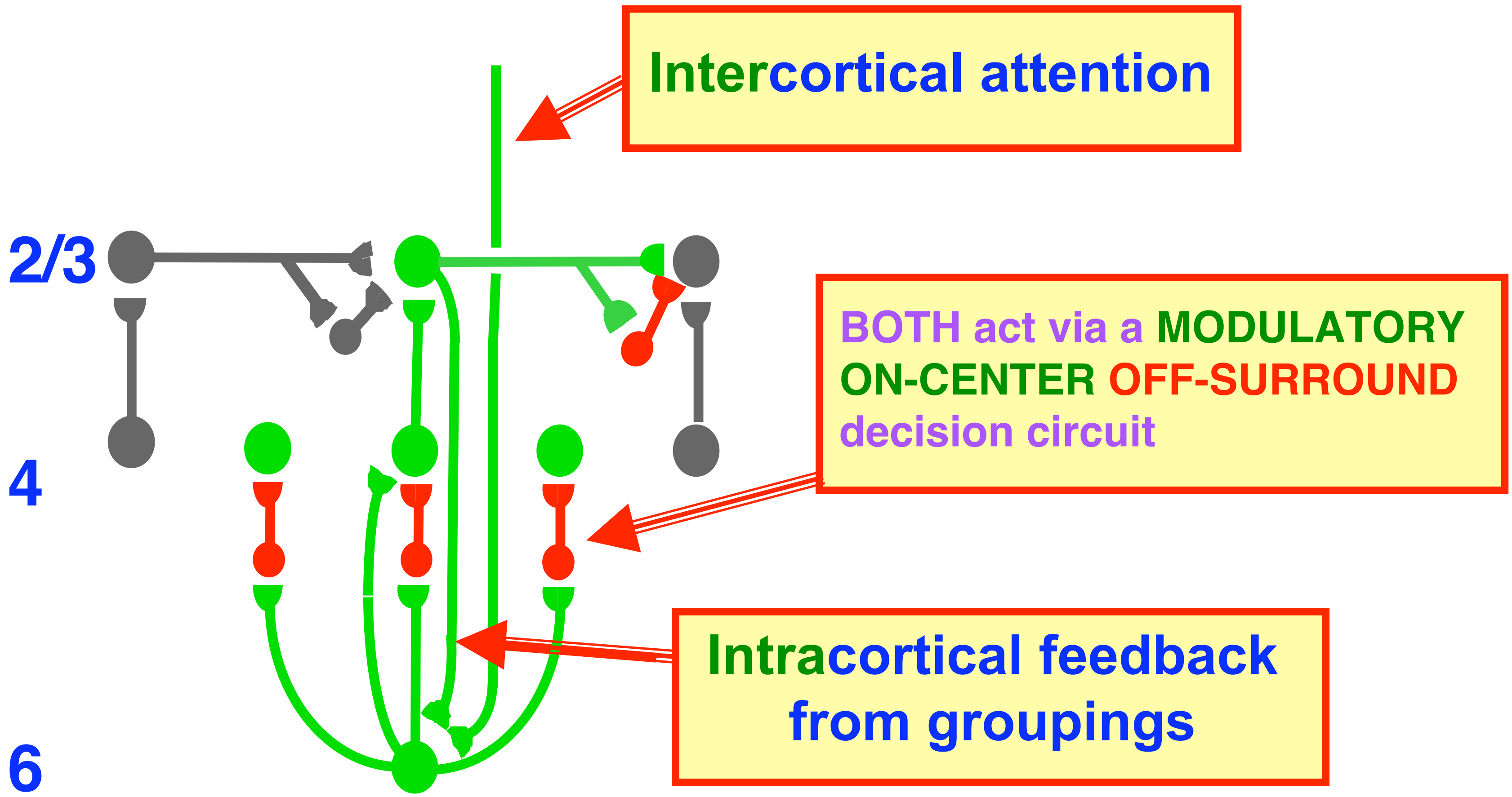


Attention directed only to far end of curve

Propagates along active layer 2/3 grouping to distal neurons

Grossberg and Raizada (2000, Vision Research)

EXPLANATION: GROUPING AND ATTENTION SHARE THE SAME MODULATORY DECISION CIRCUIT



HOW ATTENTION CAN SELECT AN ENTIRE OBJECT

"OBJECT ATTENTION"

NOT the only kind of object-based attention!

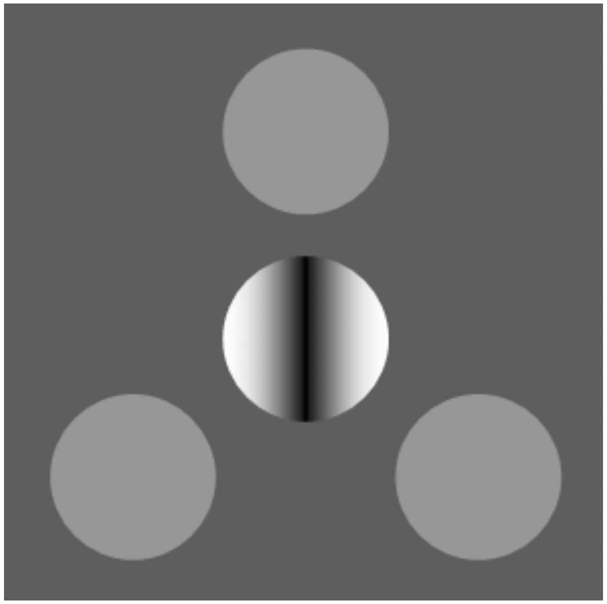
Boundary-mediated attention

Surface-mediated attention (Mingolla talk)

Prototype-mediated attention (ART)

…the model simulates lots of other behavioral and brain data about attention

DE WEERD ET AL. (1999): ATTENTIONAL FEEDBACK IS NECESSARY TO SELECT WEAK TARGETS



Stimulus

Variable-contrast grating + distractors

Task

Discriminate orientation of grating

Conditions

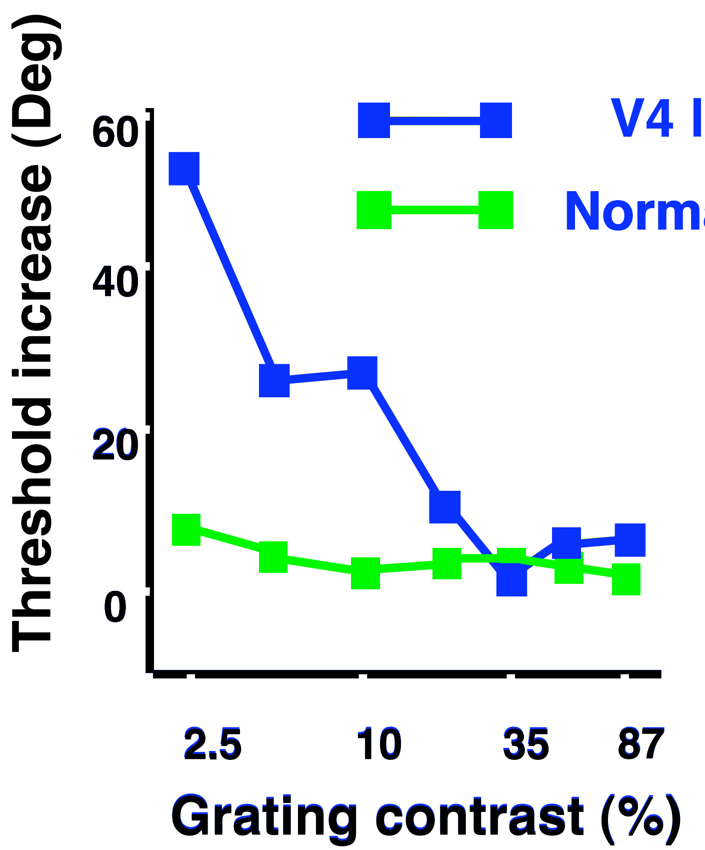
Stimulus either in normal visual quadrant, or quadrant with lesioned attentional regions (V4, TEO)

Result

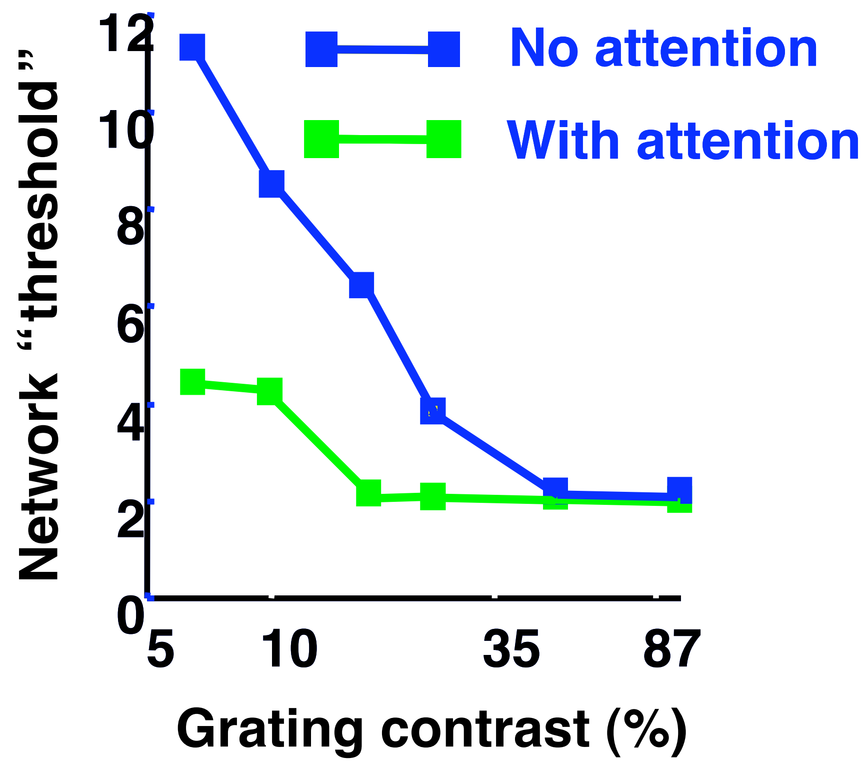
Attention needed only for low-contrast gratings

ATTENTION HAS GREATER EFFECT ON LOW-CONTRAST TARGETS

MACAQUE DATA



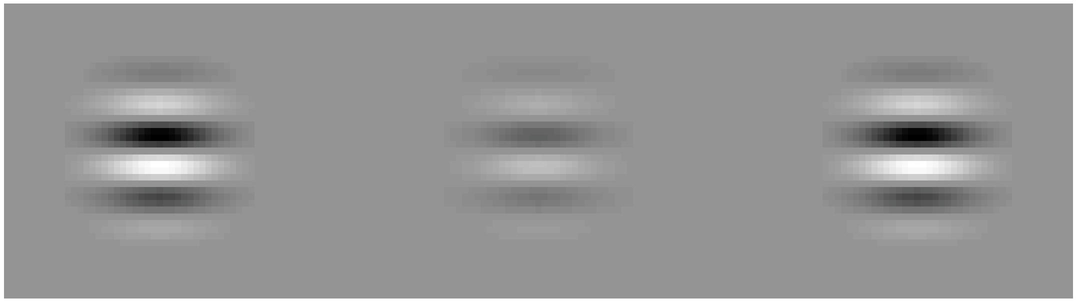
MODEL SIMULATION



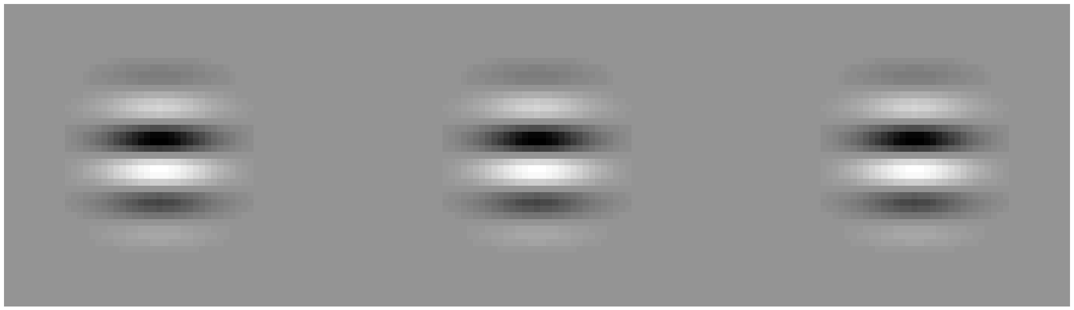
POLAT ET AL. (1998): CAT AREA 17 (V1) CONTRAST-SENSITIVE GROUPING

TARGET: Variable-contrast Gabor in neuron's Classical RF
FLANKERS: Constant-contrast collinear Gabors outside RF

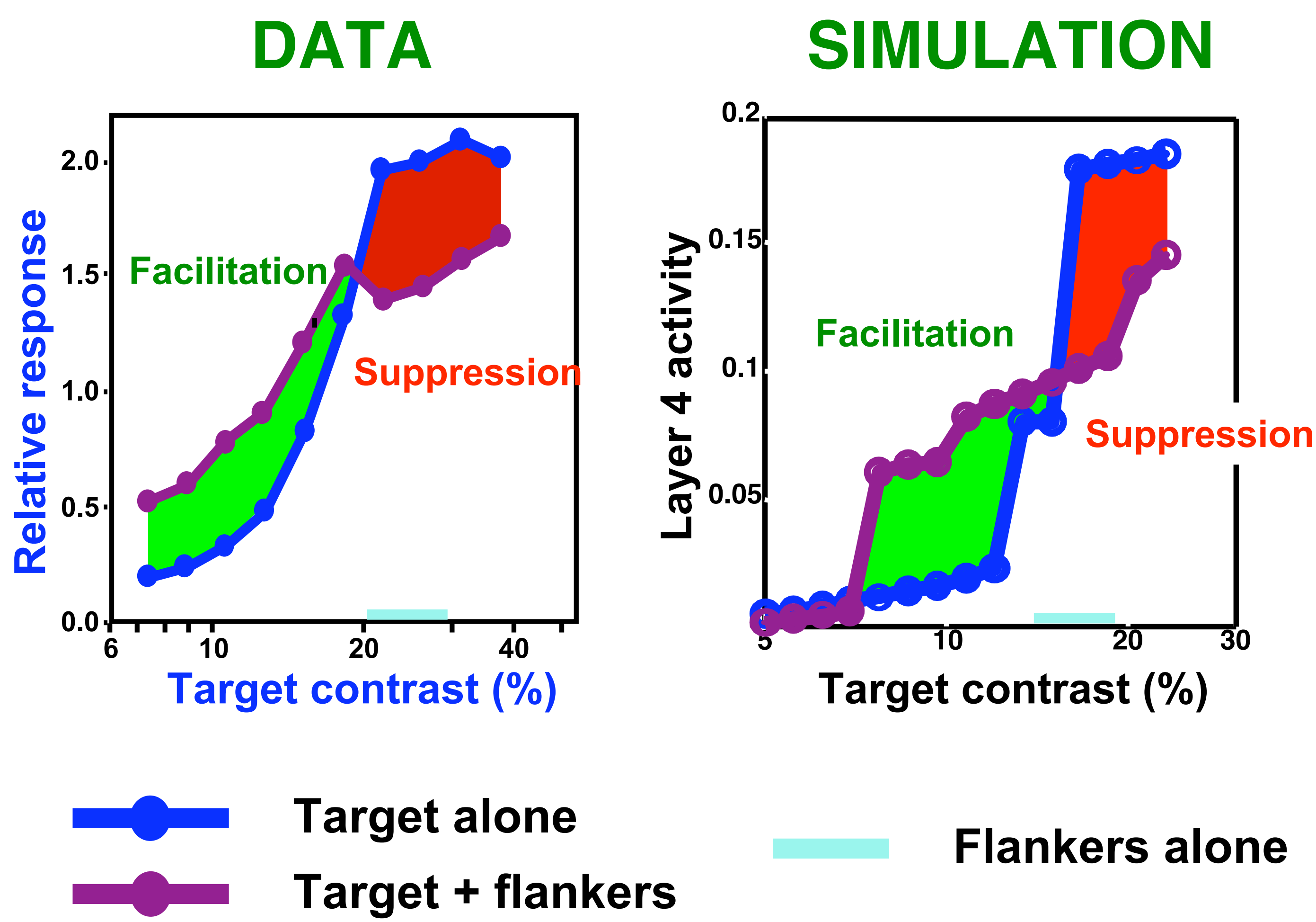
Collinear flankers **ENHANCE** response to near-threshold target



Flankers **SUPPRESS** response to high contrast target



SIMULATION OF POLAT ET AL. (1998)

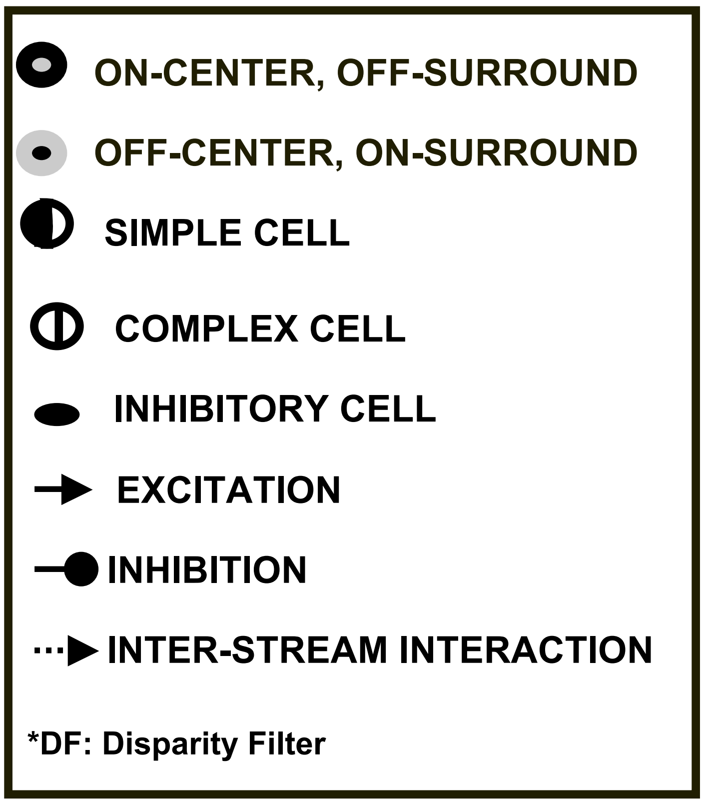


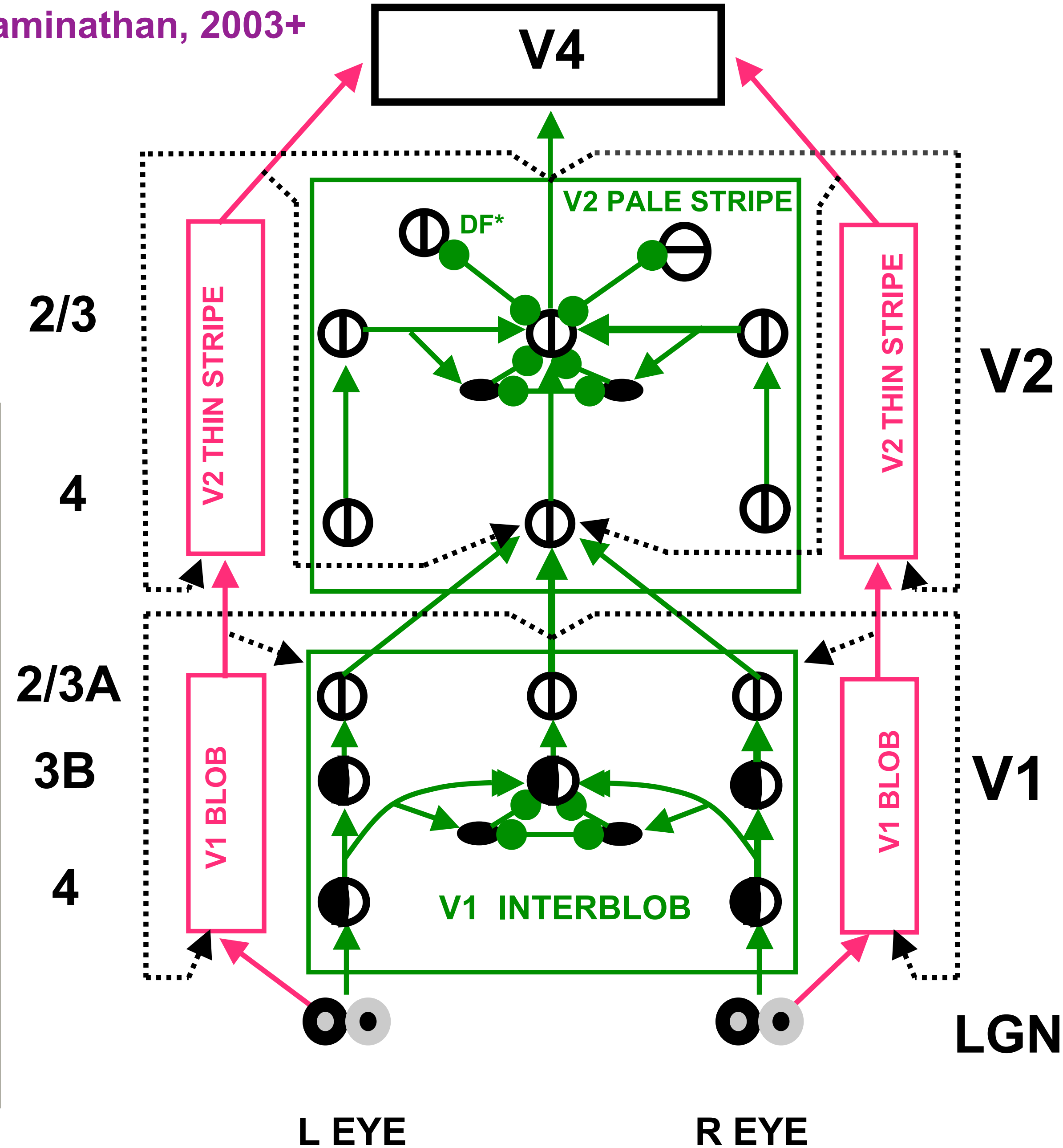
3D LAMINART: 3D VISION & FIGURE-GROUND PERCEPTION

Grossberg, Cao, Fang, Howe, Swaminathan, 2003+

Boundary stream (green)

Surface stream (red)





WHY IS THE MODEL CALLED LAMINART?

LAMINART = LAMINAR ART

ART = ADAPTIVE RESONANCE THEORY

Grossberg (1976, 1980), Carpenter and Grossberg (1987),…

ART proposes how
stable development and learning occur throughout life
using top-down attention

ART MATCHING RULE

ART predicted in the 1970's-1980's that
attention is realized by a top-down
modulatory on-center off-surround network!

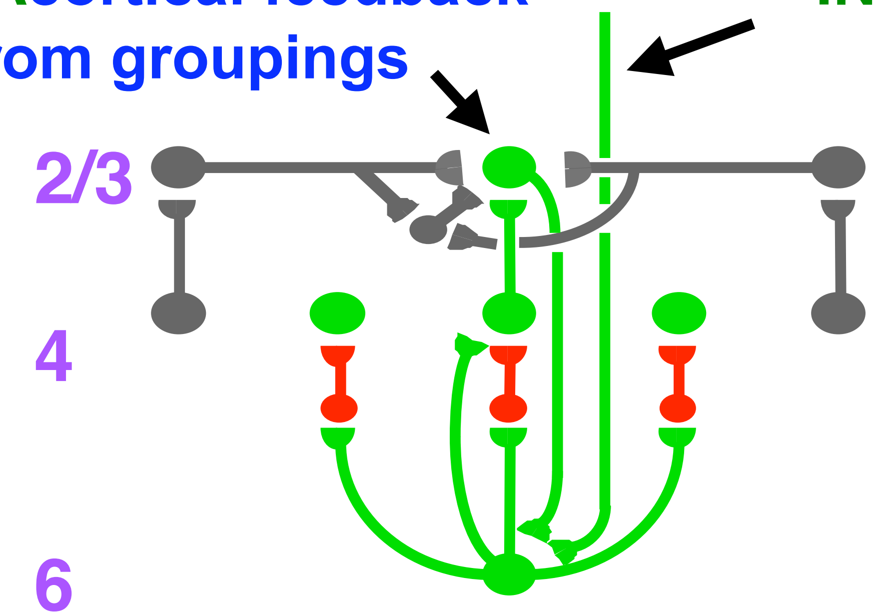
Attention matching helps to dynamically stabilize learning

The LAMINART model predicts what laminar cortical circuit embodies the ART Matching Rule

Grossberg (1999, Spatial Vision)

INTRAcortical feedback from groupings

INTERcortical attention



Attention acts via a
TOP-DOWN
MODULATORY ON-CENTER
OFF-SURROUND NETWORK

INTERcortical loop
attentively stabilizes learning
INTRAcortical loop
pre-attentively stabilizes learning

Cf., Watanabe et al

CREATIVE DISCOVERY OF CAUSAL RELATIONS IN A CHANGING WORLD

ART mechanistically predicts why we are

symbol forming

intentional

attentional

learning

beings

ADAPTIVE RESONANCE THEORY

ART

Grossberg (1976)

A unifying theme:

Stability-Plasticity Dilemma

How can learning continue into adulthood without causing catastrophic forgetting?

How can we LEARN quickly without being forced to FORGET just as quickly?

e.g., why learning your faces does not force me to forget faces of my family and friends!

USES OF ART IN ENGINEERING AND TECHNOLOGY

Boeing parts design retrieval; used in Boeing 777 design
satellite remote sensing
radar identification
robot sensory-motor control and navigation
machine vision
3D object and face recognition
Macintosh operating system software
automatic target recognition
ECG wave recognition
protein secondary structure identification
character classification
musical analysis
air quality monitoring and weather prediction
medical imaging and database analysis
multi-sensor chemical analysis
strength prediction for concrete mixes
signature verification
decision making and intelligent agents
machine condition monitoring and failure forecasting
chemical analysis
electromagnetic and digital circuit design…

ART MATCHING AND RESONANCE RULES

help to solve the Stability-Plasticity Dilemma

BOTTOM-UP ACTIVATION

by itself can activate target cells

(automatic activation)

TOP-DOWN EXPECTATIONS

learn prototypes that

- select consistent bottom-up signals (hypothesis testing)
- suppress inconsistent bottom-up signals (attentional focusing)
- cannot by themselves fully activate target cells (modulation, priming)

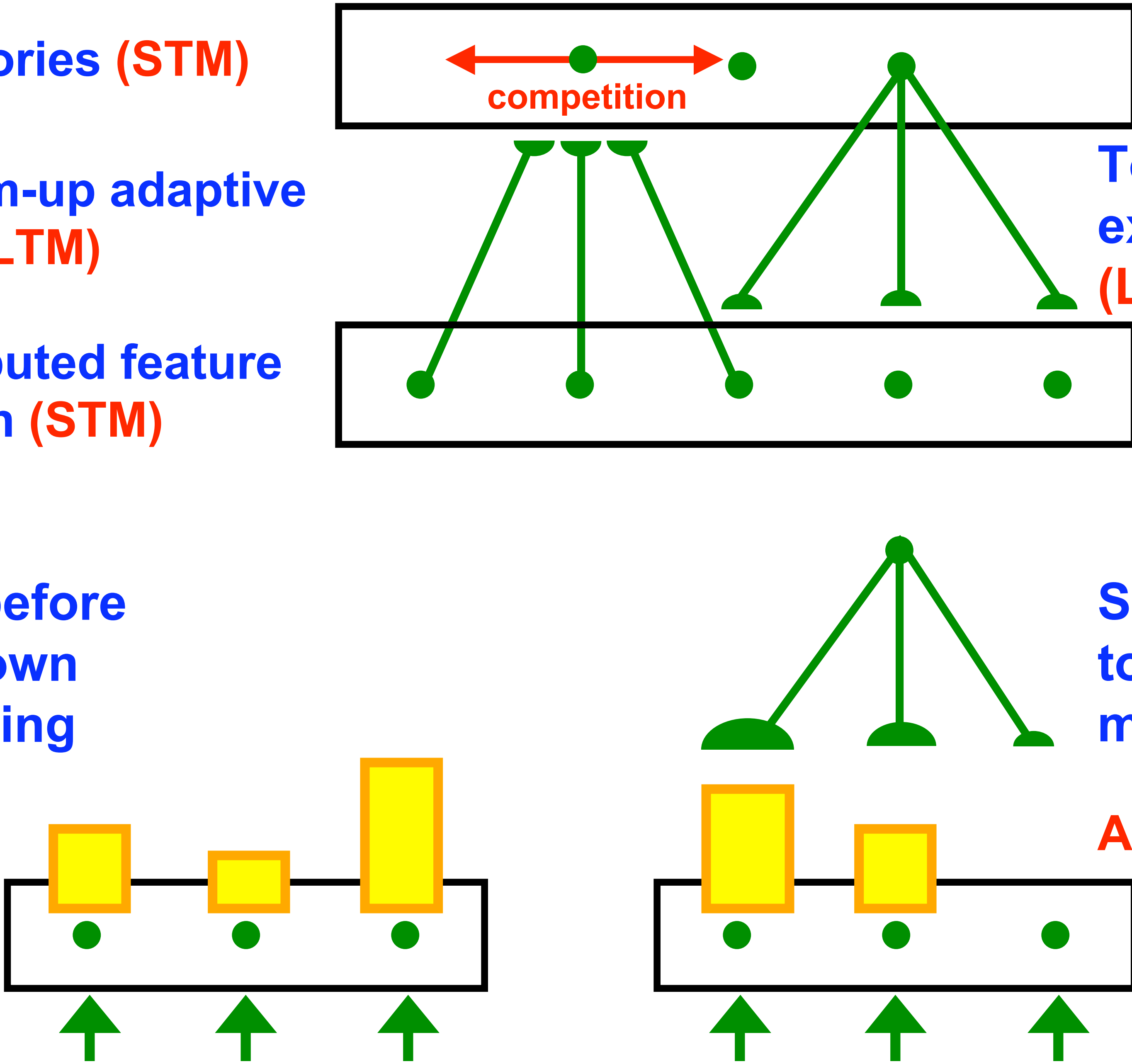
EXPECTATIONS FOCUS ATTENTION

Categories (STM)

Bottom-up adaptive filter (LTM)

Distributed feature pattern (STM)

Top-down expectations (LTM)

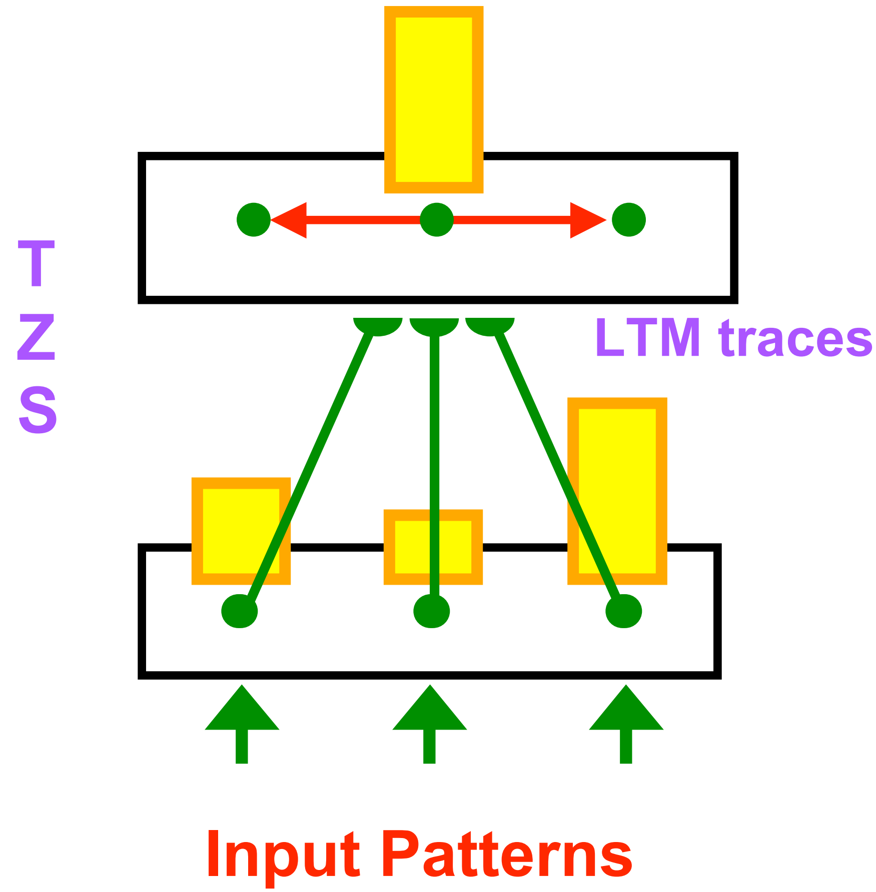


STM before top-down matching

STM after top-down matching

Attention!

COMPETITIVE LEARNING AND SELF-ORGANIZING MAPS



Categories

Compressed STM representation

competition

Adaptive Filter T=ZS

Features

Distributed STM representation

Grossberg, 1972, 1976; von der Malsburg, 1973; Kohonen, 1984

STABLE SPARSE LEARNING THEOREM

Grossberg (1976)

If a sequence of feature patterns does not form too many clusters relative to the number of category coding cells, then learning is

stable

self-normalizing

tracks input statistics

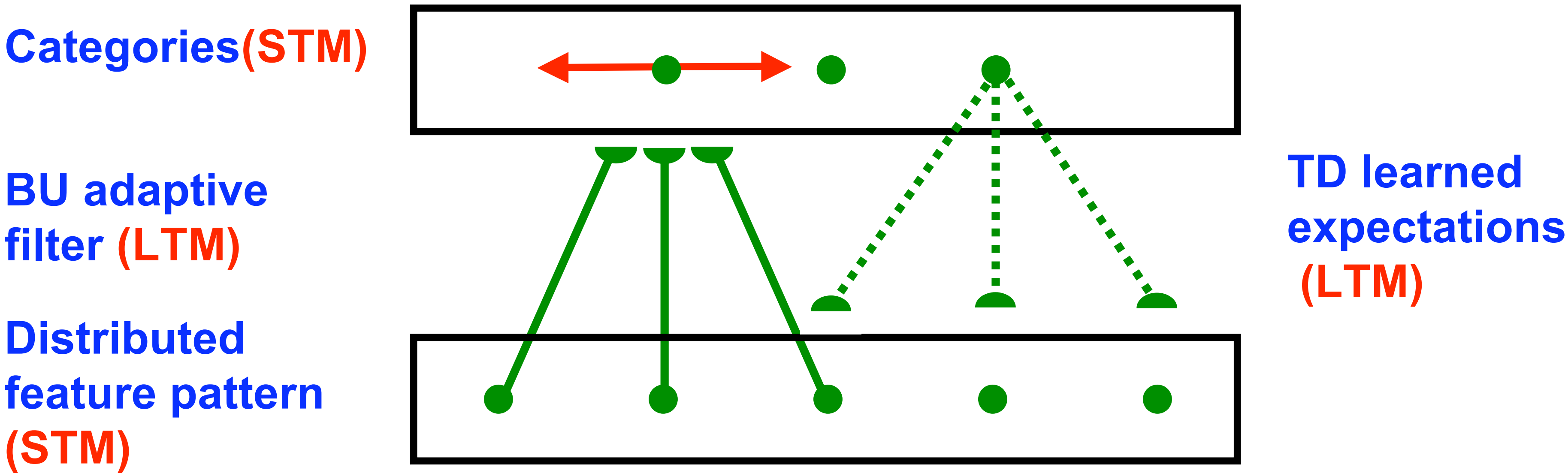
Bayesian

In general, learning is unstable in response to a dense series of inputs whose statistics change through time:

Recent learning can force unselective forgetting or catastrophic forgetting of older learning

FROM COMPETITIVE LEARNING AND SELF-ORGANIZING MAPS TO ADAPTIVE RESONANCE THEORY

ART was introduced to dynamically stabilize recognition learning using top-down EXPECTATIONS and ATTENTION



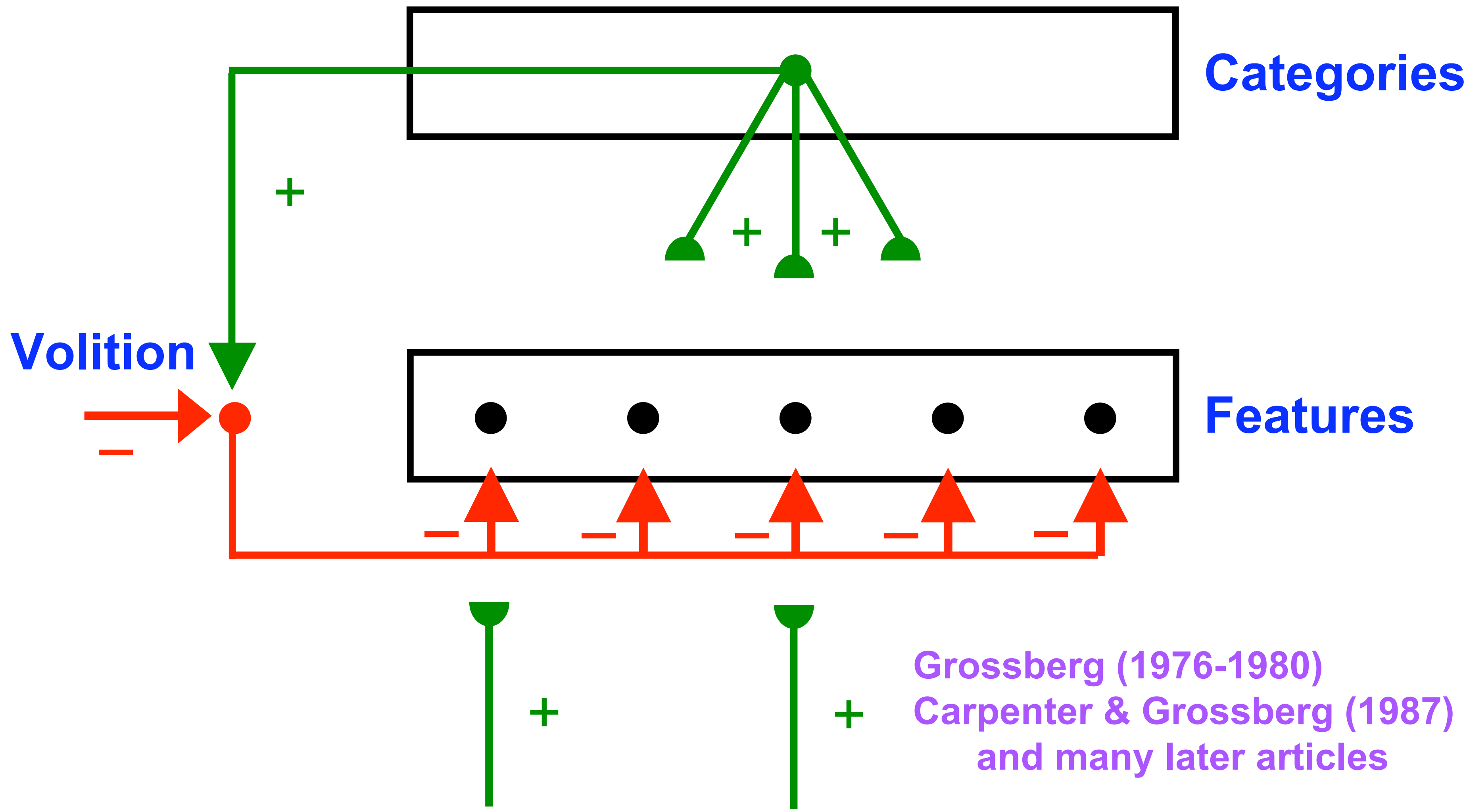
Cf. experiments of C. Gilbert, E. Kandel, M. Merzenich, etc.

HOW do expectations focus attention and stabilize learning?

ART MATCHING RULE

Stabilizes Learning

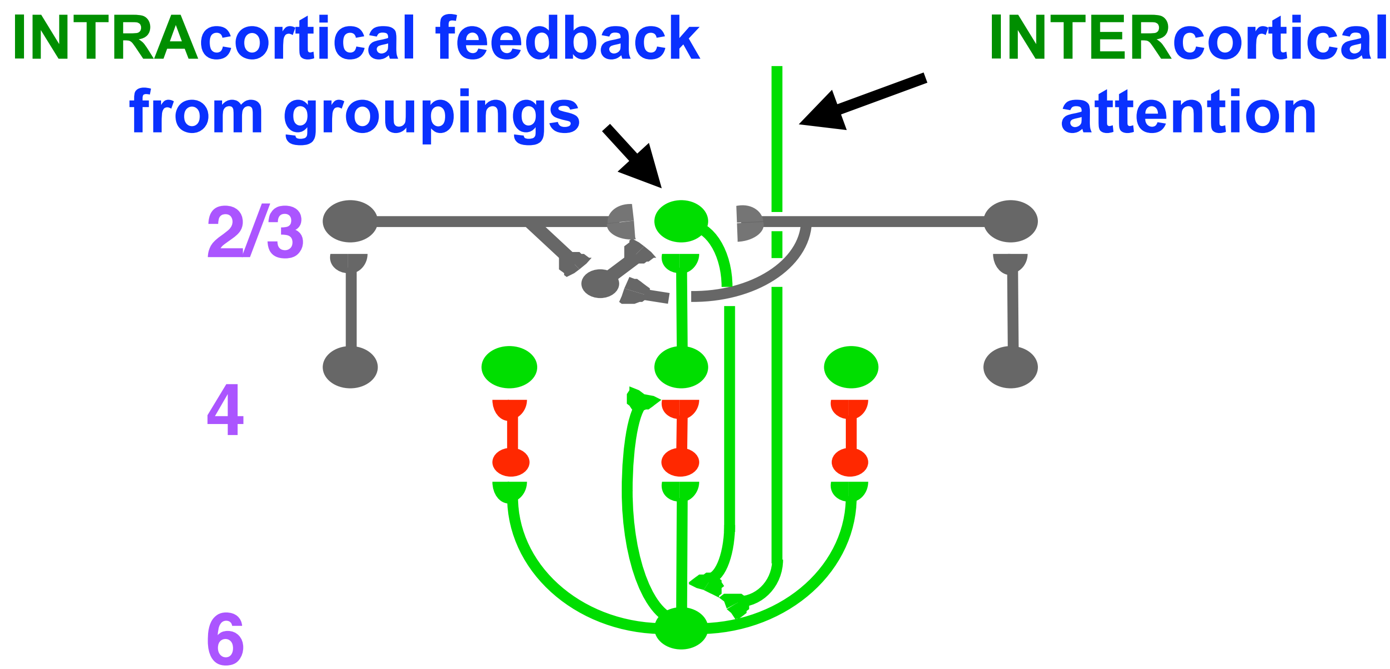
Top-down, modulatory on-center, off-surround network



Grossberg (1976-1980)
Carpenter & Grossberg (1987)
and many later articles

LAMINAR CORTICAL CIRCUIT FOR ATTENTION

Grossberg (1999, Spatial Vision)



Attention acts via a
TOP-DOWN
MODULATORY ON-CENTER
OFF-SURROUND NETWORK

INTERcortical loop
attentively stabilizes learning
INTRAcortical loop
pre-attentively stabilizes learning

Cf., Watanabe et al

NOT ALL BRAIN MODELS ARE CREATED EQUAL

All the major ART predictions have received experimental confirmation

COMPETITIVE MATCHING AND ATTENTION

ART predicted a link between

TOP-DOWN EXPECTATION

COOPERATIVE-COMPETITIVE MATCHING

ATTENTION

GROWING EXPERIMENTAL SUPPORT FOR ALL THE MAJOR ART PREDICTIONS

ATTENTION HAS AN ON-CENTER OFF-SURROUND

Bullier, Jupe, James, and Girard, 1996

Caputo and Guerra, 1998

Downing, 1988

Mounts, 2000

Reynolds, Chelazzi, and Desimone, 1999

Smith, Singh, and Greenlee, 2000

Somers, Dale, Seiffert, and Tootell, 1999

Sillito, Jones, Gerstein, and West, 1994

Steinman, Steinman, and Lehmkuhne, 1995

Vanduffell, Tootell, and Orban, 2000

"BIASED COMPETITION"

Desimone, 1998

Kastner and Ungerleider, 2001

GROWING EXPERIMENTAL SUPPORT FOR ALL THE MAJOR ART PREDICTIONS

ATTENTION CAN FACILITATE MATCHED BOTTOM-UP SIGNALS

Hupe, James, Girard, and Bullier, 1997
Luck, Chellazi, Hillyard, and Desimone, 1997
Roelfsema, Lamme, and Spekreijse, 1998
Sillito, Jones, Gerstein, and West, 1994
and many more…

INCONSISTENT WITH MODELS WHERE TOP-DOWN MATCH IS SUPPRESSIVE

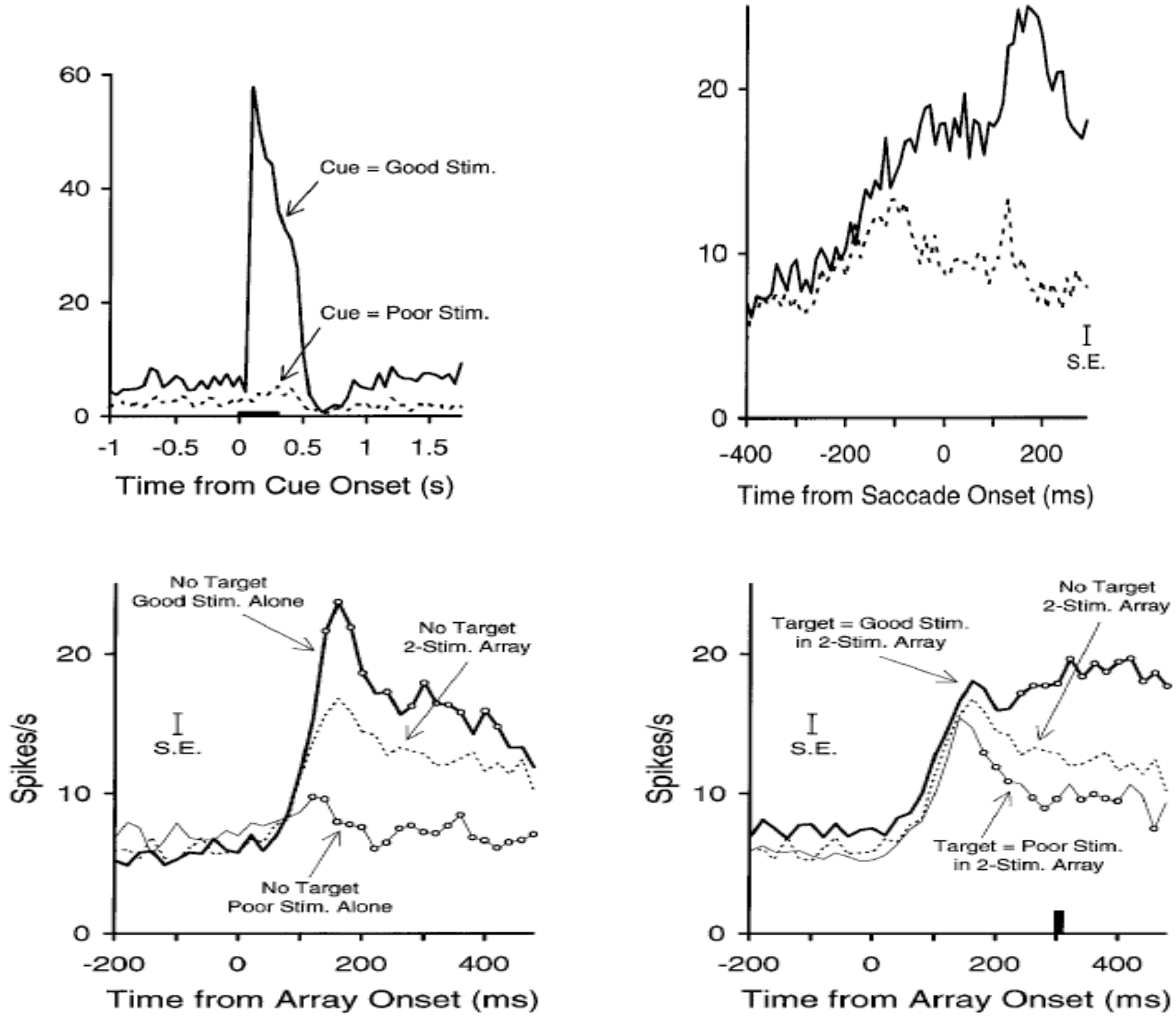
Mumford, 1992
Rao and Ballard, 1999

Bayesian Explaining Away

IT CELLS DURING MEMORY-GUIDED SEARCH

Priming and Competition

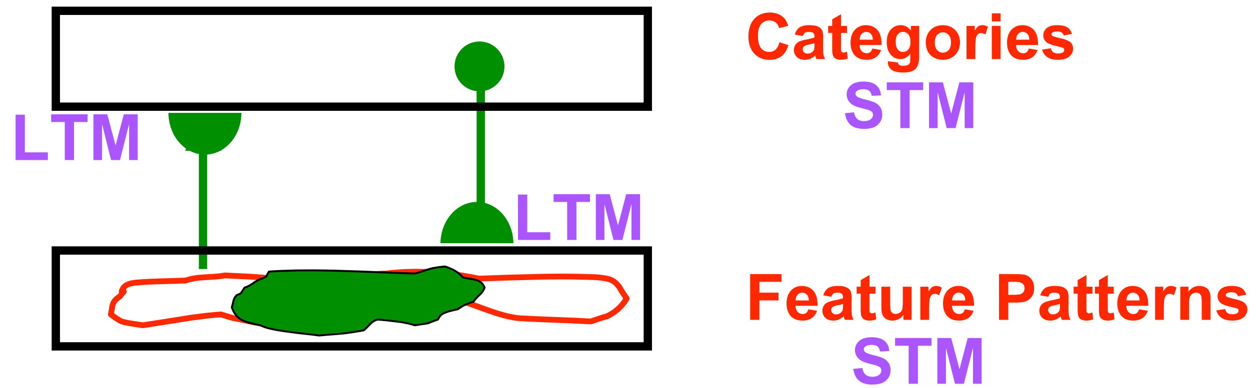
Chelazzi, Duncan, Miller, and Desimone, 1998



ADAPTIVE RESONANCE

Attended featured clusters reactivate bottom-up pathways

Activated categories reactivate their top-down pathways



Resonance synchronizes
amplifies
prolongs system response

Resonance triggers learning in bottom-up and top-down adaptive weights: *adaptive* resonance!

KEY ART PREDICTION

ALL CONSCIOUS STATES ARE RESONANT STATES

Grossberg, 1976

Growing neurophysiological support during the past several years for the predicted connection between:

Consciousness
Learning
Expectation
Attention
Resonance
Synchrony

e.g., experiments by J. Bullier, R. Desimone, C. Gilbert, V. Lamme, J. Reynolds, P. Roelfsema, W. Singer, N. Suga, etc.

SUPPORT FOR ART CLEARS PREDICTIONS

LINK BETWEEN ATTENTION AND LEARNING

VISUAL LEARNING

Ahissar and Hochstein, 1993

AUDITORY LEARNING

Gao and Suga, 1998

SOMATOSENSORY LEARNING

Krupa, Ghazanfar, and Nicolelis, 1999
Parker and Dostrovsky, 1999

SUPPORT FOR ART CLEARS PREDICTIONS

LINK BETWEEN ATTENTION AND SYNCHRONY

Engel, Fries, and Singer, 2001

Fries, Reynolds, Rorie, and Desimone, 2001

Pollen, 1999

A match can support learning that refines a learned prototype

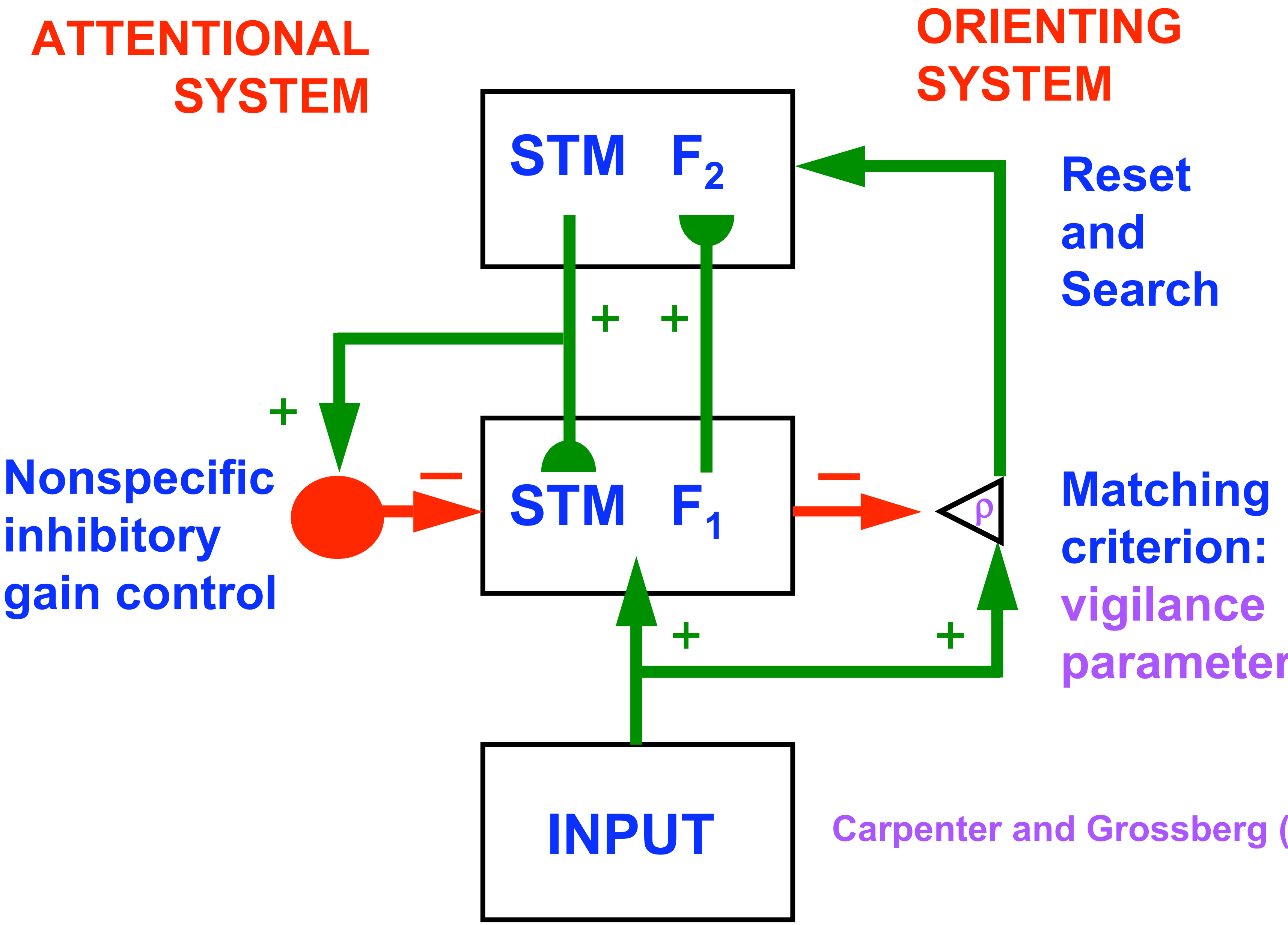
How are new categories discovered when there is a big enough mismatch?

ART

Interactions between COMPLEMENTARY SYSTEMS

Attentional System ⟷	Orienting System
Expected Events	Unexpected Events
Familiar Events	Unfamiliar Events
Resonance	Reset
Attention	Memory Search
Learning	Hypothesis Testing
Recognition	
Temporal cortex Prefrontal cortex	Nonspecific thalamus Hippocampal system

ART 1 MODEL



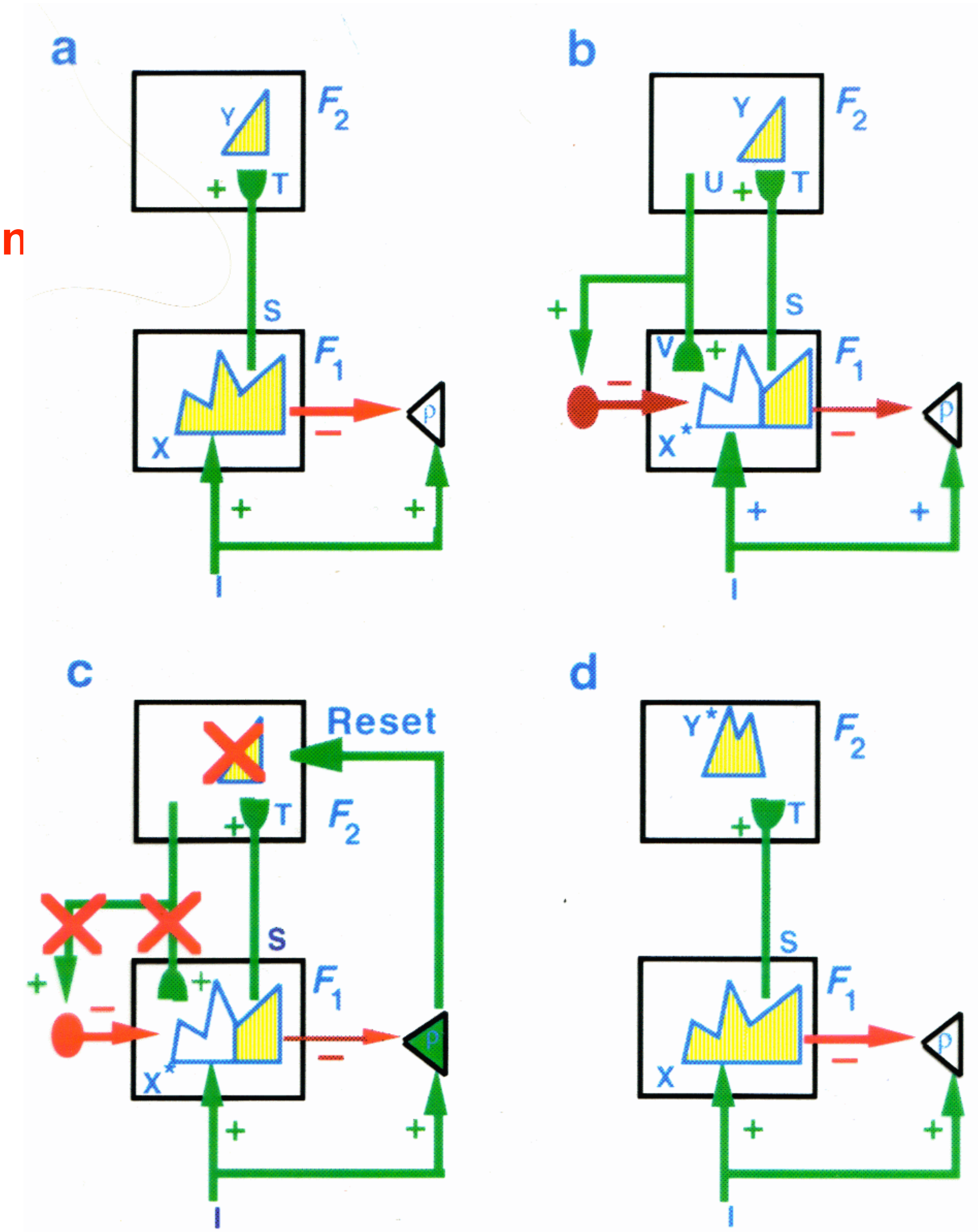
Carpenter and Grossberg (1987)

ART HYPOTHESIS TESTING AND LEARNING CYCLE

Choose category, or symbolic representation

Test hypothesis

VIGILANCE
How big a mismatch causes reset?



Mismatch reset

Choose another category

ERP SUPPORT FOR HYPOTHESIS TESTING CYCLE

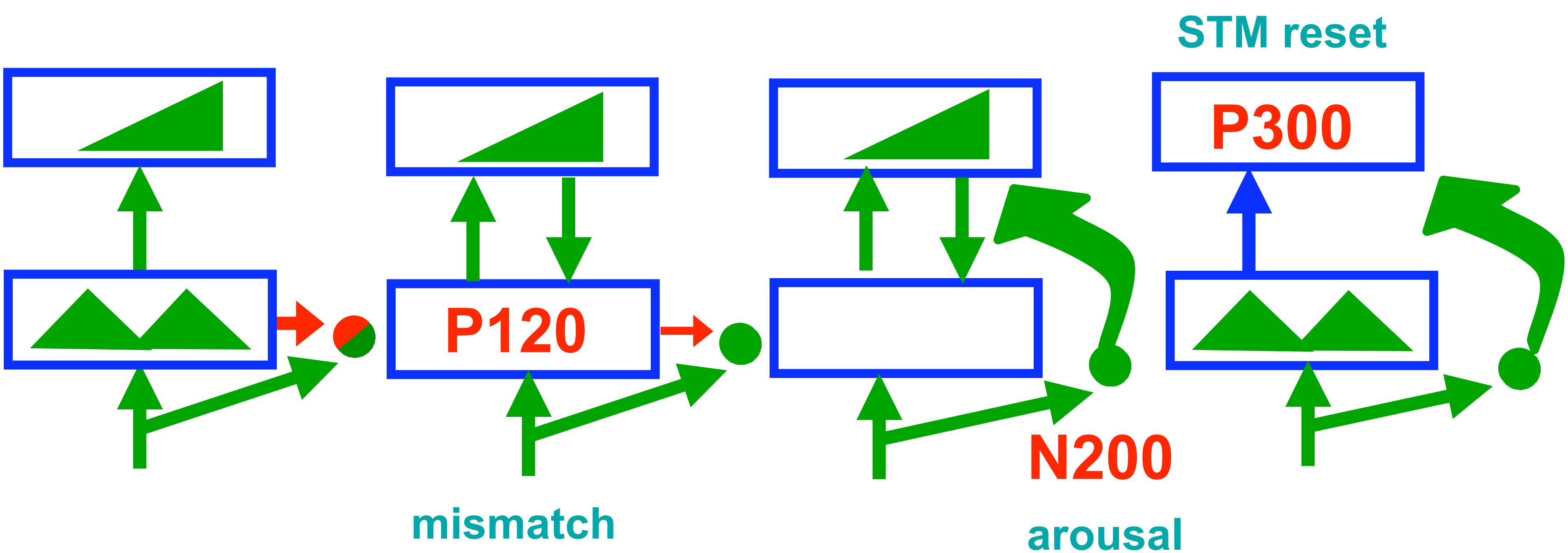
Event-Related Potentials: Human Scalp Potentials

ART predicted correlated sequences of P120-N200-P300

Event Related Potentials during oddball learning

P120 - mismatch; N200 - arousal/novelty; P300 - STM reset

Confirmed in: Banquet and Grossberg (1987)



NEUROPHYSIOLOGICAL SUPPORT FOR MATCHING AND RESET IN INFEROTEMPORAL CORTEX

Cells in inferotemporal cortex learn to categorize visual events in the world

These cells are actively reset during working memory tasks

There is an "active matching process that was reset between trials."

Miller, Li, Desimone (1991)

COGNITIVE LEARNING CYCLE

A dynamic cycle of

RESONANCE

and

RESET

As inputs are learned, search automatically disengages and direct access to globally best-matching category occurs

Mathematical proof in: Carpenter & Grossberg, *CVGIP*, 1987

Explains how we can quickly recognize familiar objects and events even if, as we get older, we store enormous numbers of memories

Are the predicted CLEARS dynamics realized within laminar cortical and thalamic circuits?

YES!

SMART model

Synchronous Matching ART

Grossberg and Versace (2006+)

MAIN QUESTIONS:

How are multiple levels of brain organization

- spikes
- local field potentials
- inter-areal synchronous oscillations
- spike-timing dependent plasticity

coordinated to

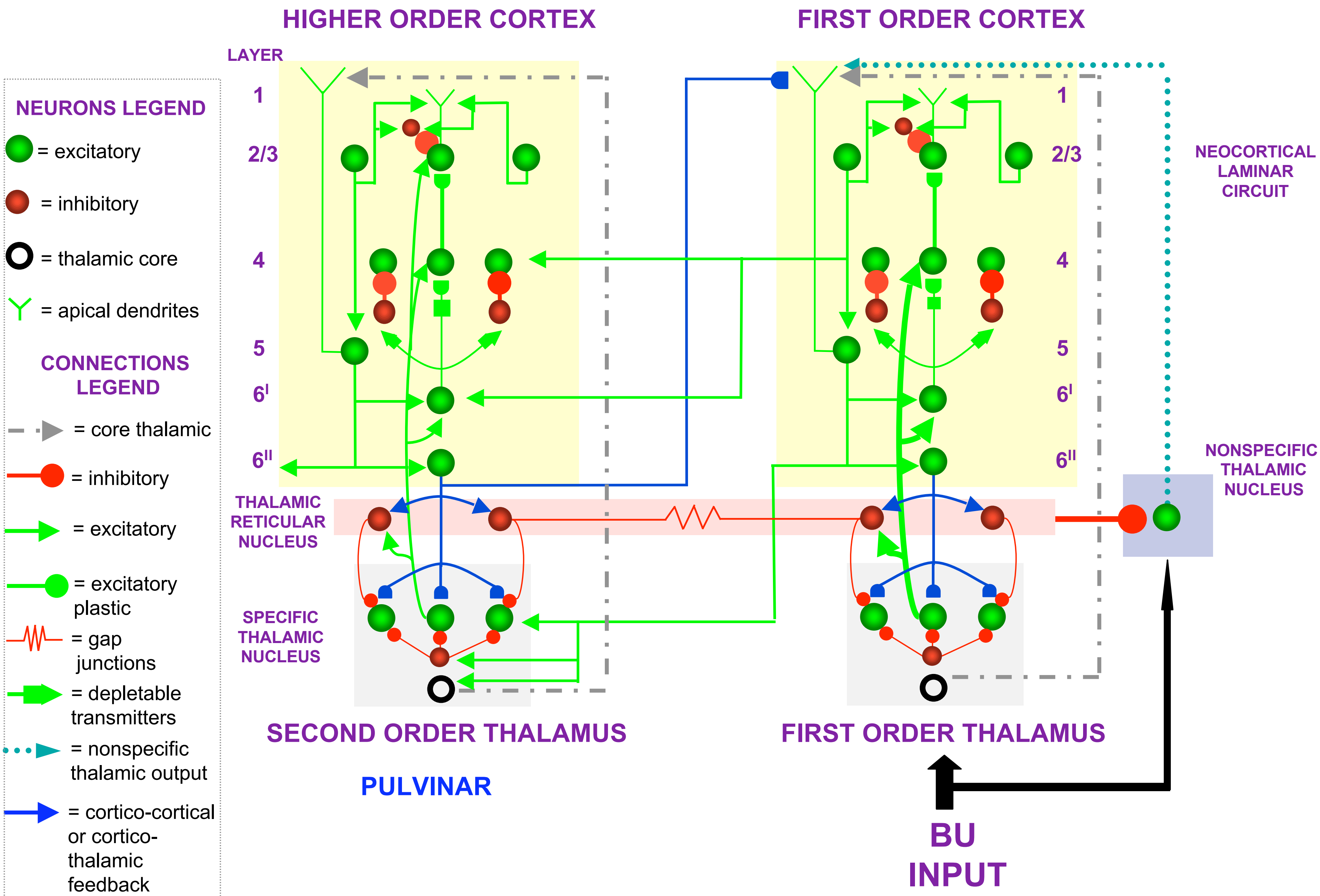
regulate stable category learning and attention

during cognitive information processing via

- laminar cortical circuits
- specific and nonspecific thalamic nuclei?

SMART: MODEL MACROCIRCUIT

Grossberg Plenary
IJCNN'07



THE MODEL FUNCTIONALLY EXPLAINS LOTS OF ANATOMICAL DATA

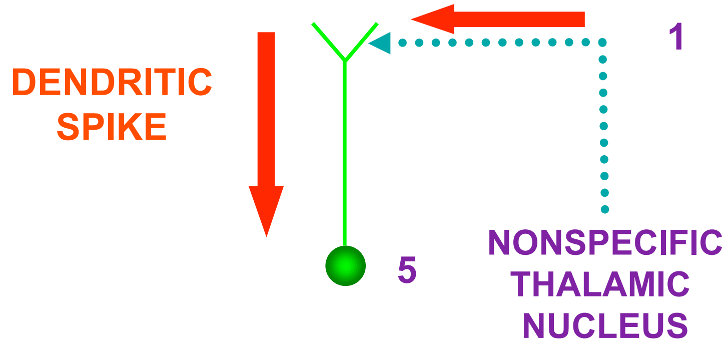
Connections	Type	Functional interpretation	References
thalamic core A → 4 A	D	Primary thalamic relay cells drive layer 4.	Blasdel and Lund (1983)
thalamic core A → 6^{I} A	D	Primary thalamic relay cells prime layer 4 via the 6 → 4 modulatory circuit.	Blasdel and Lund (1983) for LGN → 6; Callaway (1998) LGN input to 6 is weak and Layer 5 projections to 6 **[Note 1]**
thalamic core A → RE A	D	Recurrent inhibition to primary and secondary thalamic relay cells.	Sherman and Guillery (2001); Jones (2002)
RE A → thalamic core A	I	Off-surround to primary and secondary thalamic relay cells, synchronization of thalamic relay cells.	Cox *et al.* (1997); Pinault and Deschenes (1998); Sherman and Guillery (2001)
RE A → RE A	I	Normalization of inhibition.	Jones (2002); Sohal and Huguenard (2003)
RE A (B) → RE B(A)	GJ	Synchronize RE and thalamic relay cells.	Landisman *et al.* (2002)
RE A → nonspecific thalamic A	I	Inhibition of nonspecific thalamic cells, participates in the reset mechanism.	Kolmac and Mitrofanis (1997); Van der Werf *et al.* (2002)
nonspecific thalamic A → 5 A	M	To 5 through apical dendrites in 1, participates in the reset mechanism.	Van der Werf *et al.* (2002)
4 A → 4 inh. A	D	Lateral inhibition in layer 4.	Markram *et al.* (2004)
4 inh. A → 4 A	I	Lateral inhibition in layer 4.	Markram *et al.* (2004)
4 inh. A → 4 inh. A	I	Normalization of inhibition in layer 4.	Ahmed *et al.* (1997); Markram *et al.* (2004)
4 A → 2/3 A	D	Feedforward driving output from 4 to 2/3.	Fitzpatrick *at al.* (1985); Callaway and Wiser (1996)
2/3 A → 2/3 A	D	Recurrent connections (grouping) in 2/3.	Bosking *et al.* (1997); Schmidt *et al.* (1997); Grossberg and Raizada (2003)
2/3 A → 2/3 inh. A	D	Avoid outward spreading (bipole) in 2/3.	McGuire *et al.* (1991); Grossberg and Raizada (2003)
2/3 inh. A → 2/3 inh. A	I	Normalization of inhibition.	Tamas *et al.* (1998); Grossberg and Raizada(2003)
2/3 A → 4 B	D	Feedforward output from Area A to Area B.	Van Essen *et al.* (1986)
2/3 A → 6^{II} B	D	Feedforward output from Area A to Area B.	Van Essen *et al.* (1986)
2/3 A → 5 A	D	Conveys layer 2/3 output to layer 5.	Callaway and Wiser (1996)
2/3 A → 6^{II} A	D	Conveys layer 2/3 output to layer 6^{II}.	Callaway (1998)

THE MODEL FUNCTIONALLY EXPLAINS LOTS OF ANATOMICAL DATA

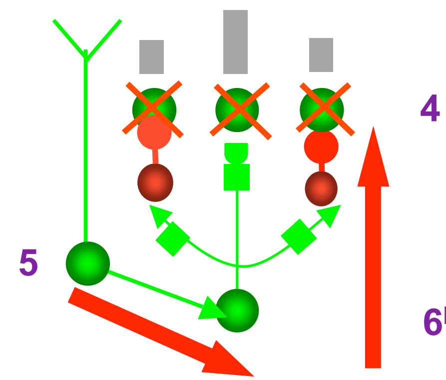
Connections	Type	Functional interpretation	References
5 A → thalamic core B	D	Feedforward connections from Area A to Area B through secondary thalamic relay neurons.	Rockland (1999); Sherman and Guillery (2001)
5 A → 6^{I} A	D	Delivers feedback to the 6 → 4 circuit from higher cortical areas, sensed at the apical dendrites of 5 branching in 1.	Callaway (1998); Callaway and Wiser (1996), class B" cells **[Note 2]**
6^{I} A → 4 A	M	On-center to 4. Mediated by habituative gates.	Stratford *et al.* (1996); Callaway (1998); Grossberg and Raizada (2003)
6^{I} A → 4 int. A	D	Off-surround to 4.	McGuire *et al.* (1984); Ahmed *et al.*, (1997); Callaway (1998)
6^{II} A → thalamic Core A	M	On-center to primary thalamic relay cells.	Sillito *et al.* (1994); Callaway (1998);
6^{II} A → RE A	D	Off-surround to primary thalamic relay cells mediated by thalamic RE.	Guillery and Harting (2003); Sherman and Guillery (2001)
6^{II} B → 2/3, 2/3 inh., 5 A	M	Intercortical feedback from 6^{II} area B to 1 area A, where it synapses on 2/3 excitatory and inhibitory neurons, as well as 5 apical dendrites branching in 1	Rockland and Virga (1989); Rockland (1994); Salin and Bullier (1995)

Abbreviations: **inh. =** inhibitory neurons; **RE** = reticular nucleus**;** **A** = primary (thalamic, cortical) loop; **B** = secondary (thalamic, cortical) loop; **D** = driving excitatory connections; **M** = modulatory connections; **I** = inhibitory connections; **GJ** = gap junctions; **int.** = inhibitory interneuron. **[Note 1]:** Callaway (1998) subdivides Layer 6 neurons in 3 classes: *Class I:* provide feedback to 4C, receive input from LGN, and project back to LGN; *Class IIa:* dendrites in 6, axons from 2/3, project back to 2/3 with modulatory connections; *Class IIb:* dendrites in 5, project exclusively to deep layers (5 & 6) and claustrum. In the model, these populations are clustered in 2 classes, layer 6I and 6II, which provide feedback to thalamic relay cells and layer 4, respectively. **[Note 2]:** Callaway (1998) subdivides Layer 5 neurons in 3 classes: *Class A:* dendrites in 5, axons from 2/3, project back to 2/3 with modulatory connections; *Class B:* dendrites in 5, axons from 2/3, project laterally to 5 and PULVINAR; *Class C:* dendrites in 1, project to superior colliculus. In the model, these differences are ignored, and it is assumed that the model layer 5 neuron receives input from 2/3 (Classes A and B), as well modulatory input from the nonspecific thalamic nuclei (Class C, apical dendrites in layer 1), and provide output to 6^{I} and second-order thalamic nuclei. The inner, recurrent loop with 2/3 has also been ignored.

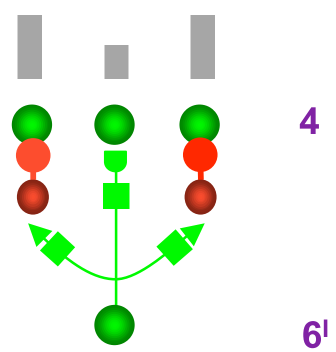
MISMATCH CAUSES LAYER 5 DENDRITIC SPIKES THAT TRIGGER RESET IN DEEP LAYERS (6-4)



(a) **Arousal causes increase in nonspecific thalamic nuclei firing rate and layer 5 dendritic and later somatic spikes**
Larkum and Zhu (2002), Williams and Stuart (1999)



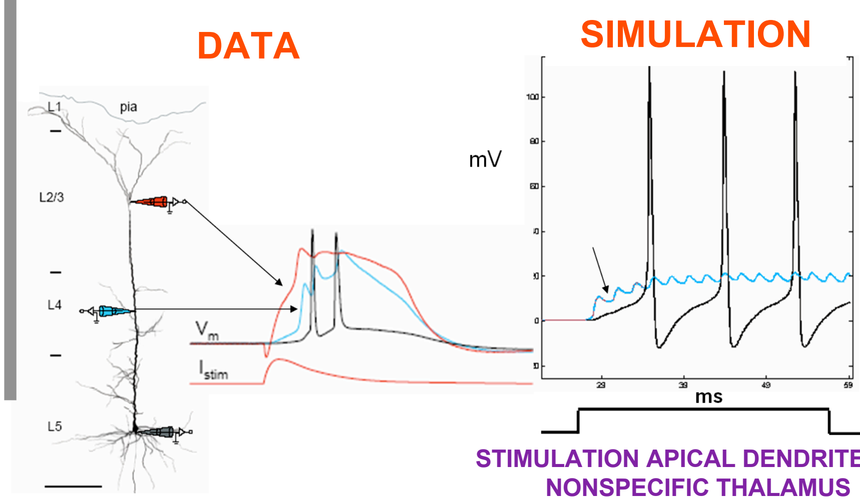
(b) **Layer 5 spikes reach layer 4 via layer 6^I and inhibitory interneurons**
Lund and Boothe (1975), Gilbert and Wiesel (1979)



(c) **Habituative neurotransmitters in layer 6^I shift the balance of active cells in layer 4**
Grossberg (1972, 1976)

Dendritic stimulation fires layer 5

Larkum and Zhu (2002)

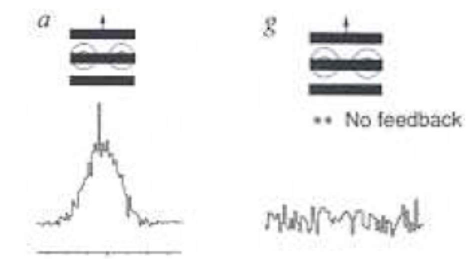


STIMULATION APICAL DENDRITES OF NONSPECIFIC THALAMUS

BRAIN OSCILLATIONS DURING MATCH/MISMATCH

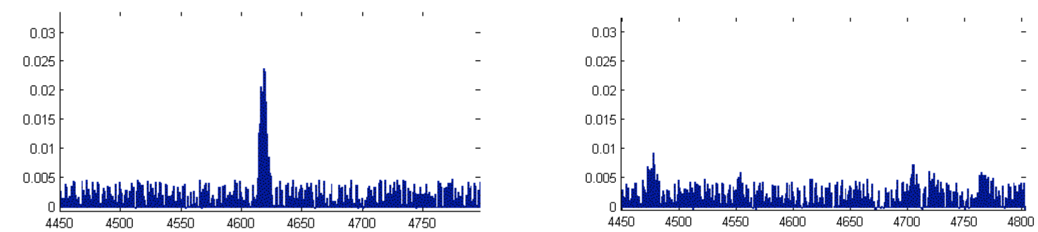
DATA

FB ON **FB OFF**

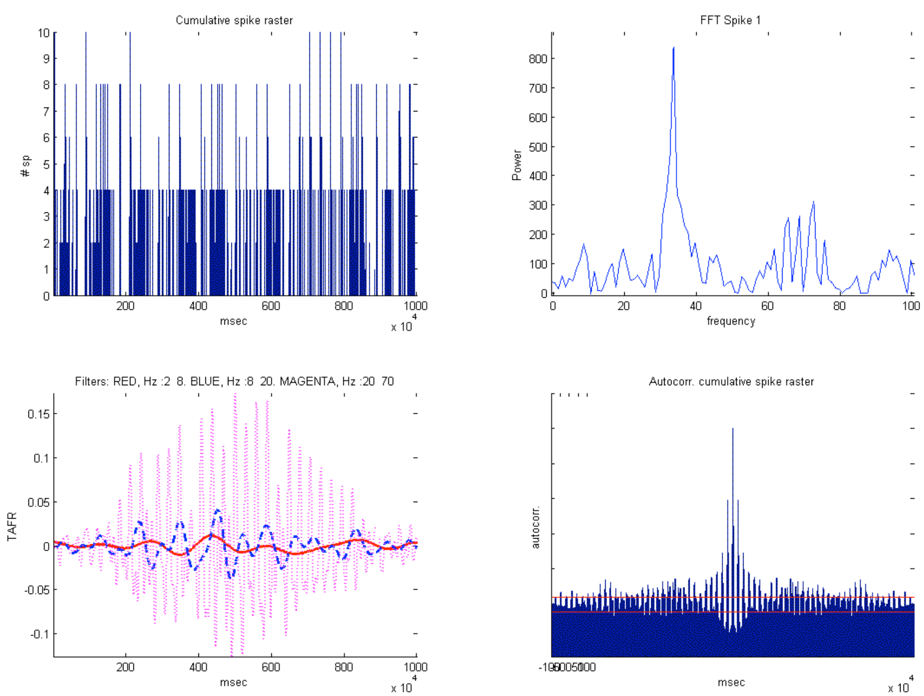


SIMULATION

FB ON **FB OFF**

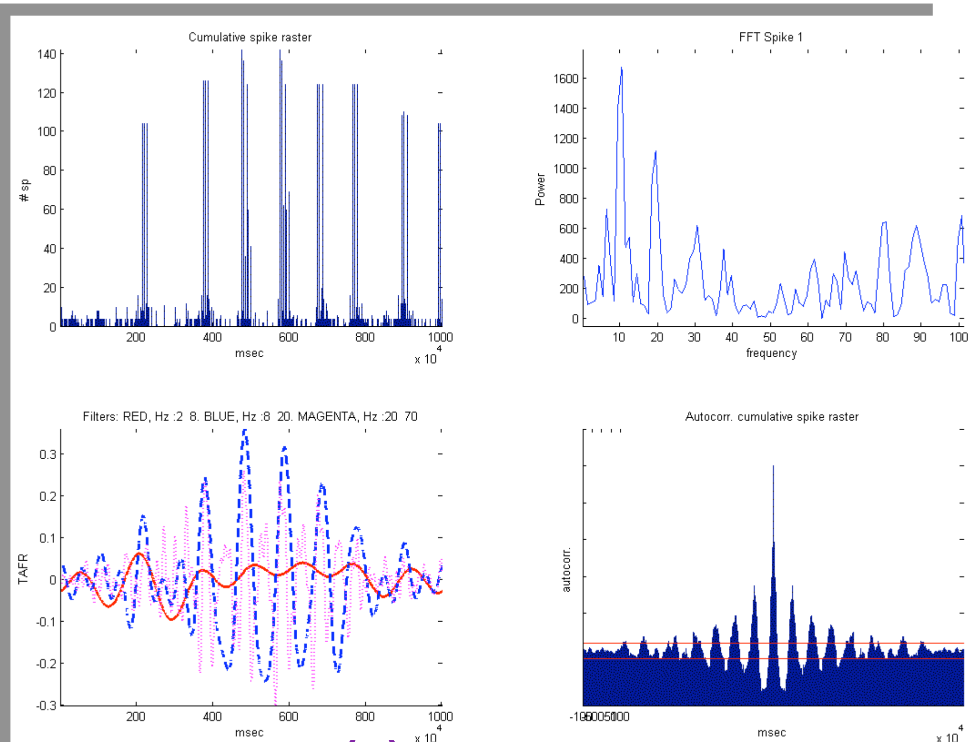


(a) TD CORTICOTHALAMIC FEEDBACK increases SYNCHRONY Sillito *et al.* (1994)



(b) MATCH

increases γ oscillations

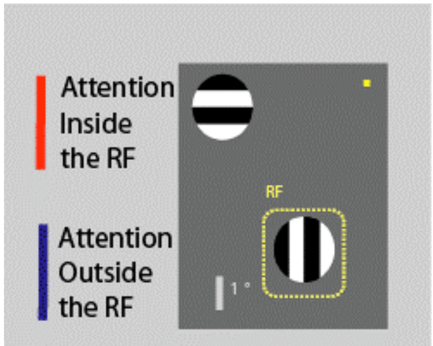


(c) MISMATCH

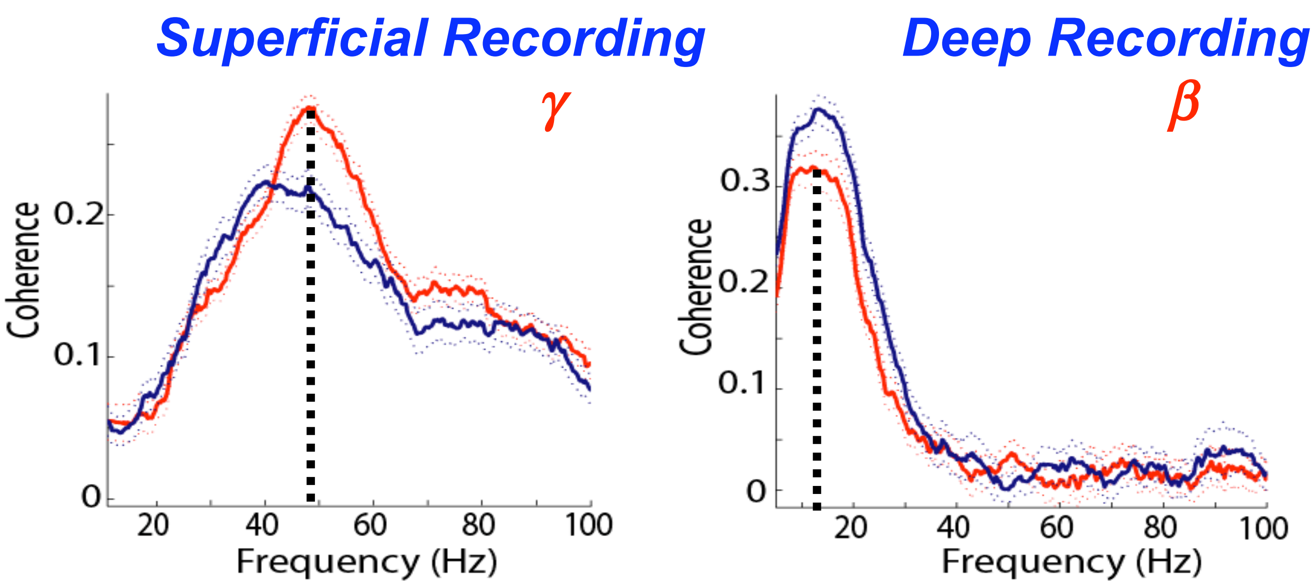
increases θ, β oscillations

DIFFERENT AVERAGE OSCILLATION FREQUENCIES IN SUPERFICIAL AND DEEP CORTICAL LAYERS

Buffalo, E.A., Fries, P., and Desimone, R. (2004). Layer-specific attentional modulation in early visual areas. Society for Neuroscience Abstract. 30, 717–6.



PREDICTION: Does this difference in average oscillation frequencies in the superficial and deep layers reflect layer 4 RESET dynamics?



HOW DOES THE CEREBRAL CORTEX WORK?

COGNITION

Are cognitive circuits variations on vision circuits?

How do laminar cortical circuits store and learn about sequences of events through time?

LIST PARSE MODEL

LAMINAR INTEGRATED STORAGE OF TEMPORAL PATTERNS FOR ASSOCIATIVE RETRIEVAL, SEQUENCING AND EXECUTION

Grossberg and Pearson (2006+)

EXCITING PREDICTION: How prefrontal cortex may use a variant of visual cortical circuits to carry out cognitive temporal functions:

MODEL FUNCTIONS

Working memory storage of incoming TEMPORAL series of events as a parallel SPATIAL activation gradient

Learning of list, or event sequence, categories

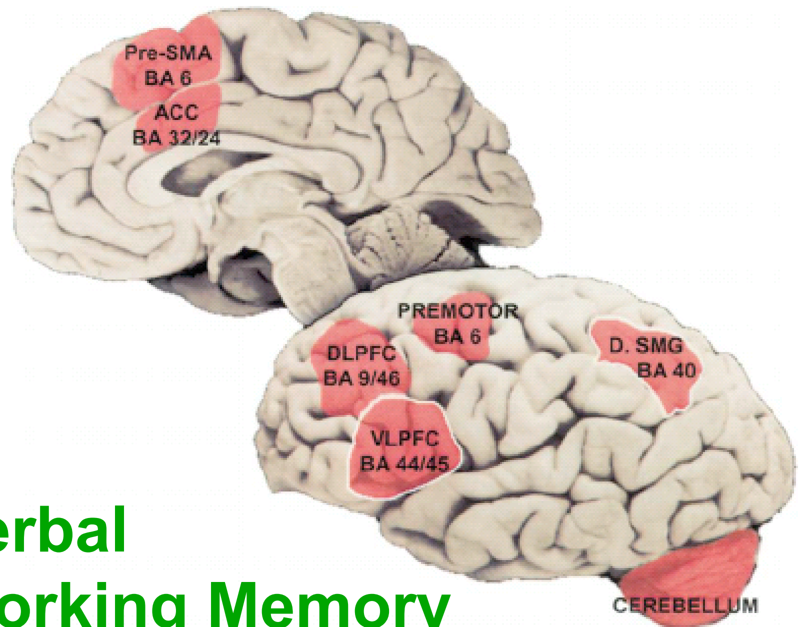
Voitional performance of event sequences at variable rates

MODEL MECHANISMS

Predict how layered circuits in prefrontal and motor cortex accomplish this

Grossberg Plenary
IJCNN'07

LIST PARSE MACROCIRCUIT



Verbal Working Memory

Chien, Ravizza, Fiez (2003)

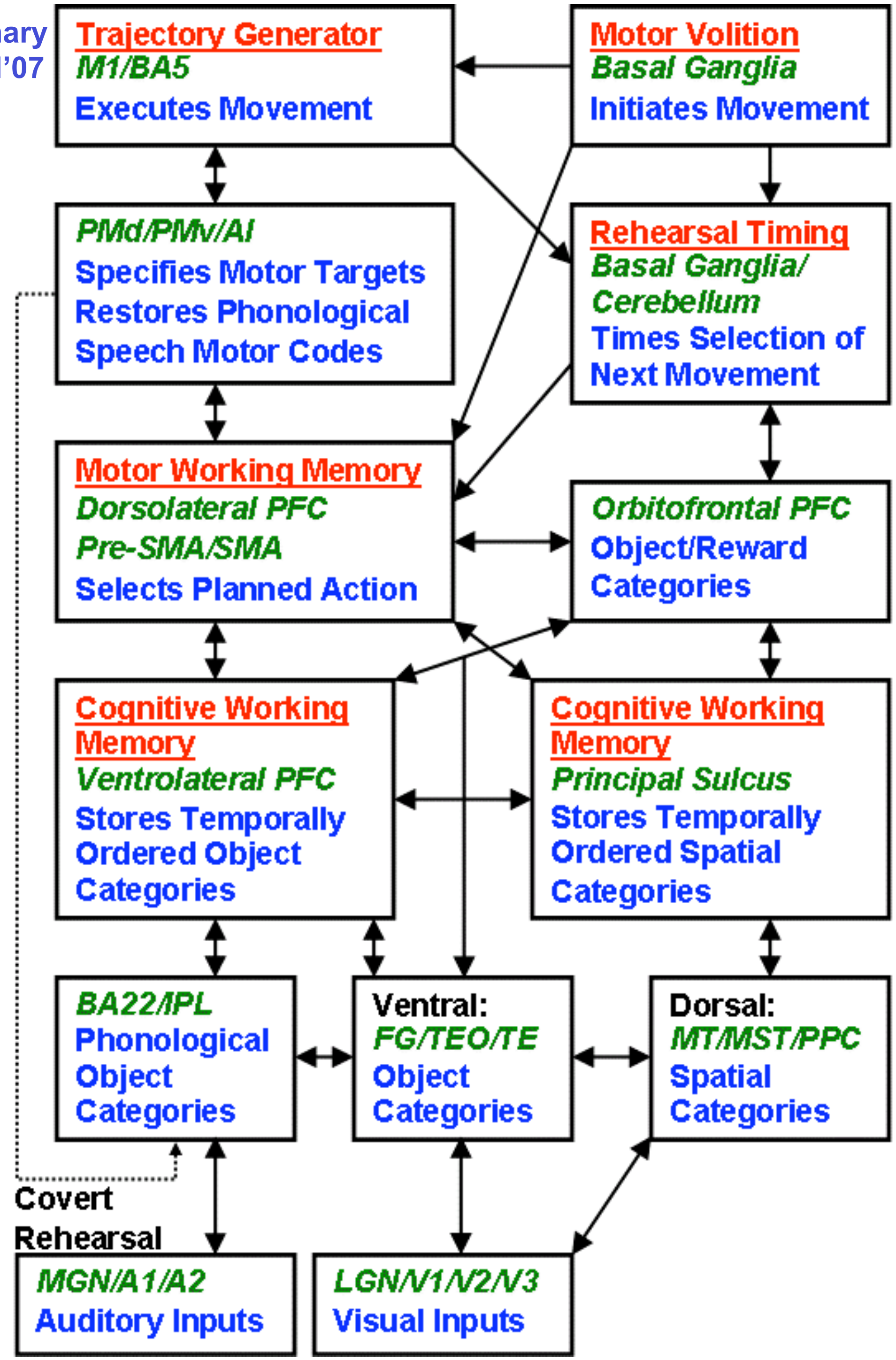


Diagram Convention

Model Components

Modeled

Proposed Localization

Proposed Functionality

Grossberg Plenary
IJCNN'07

LIST PARSE CIRCUIT DIAGRAM

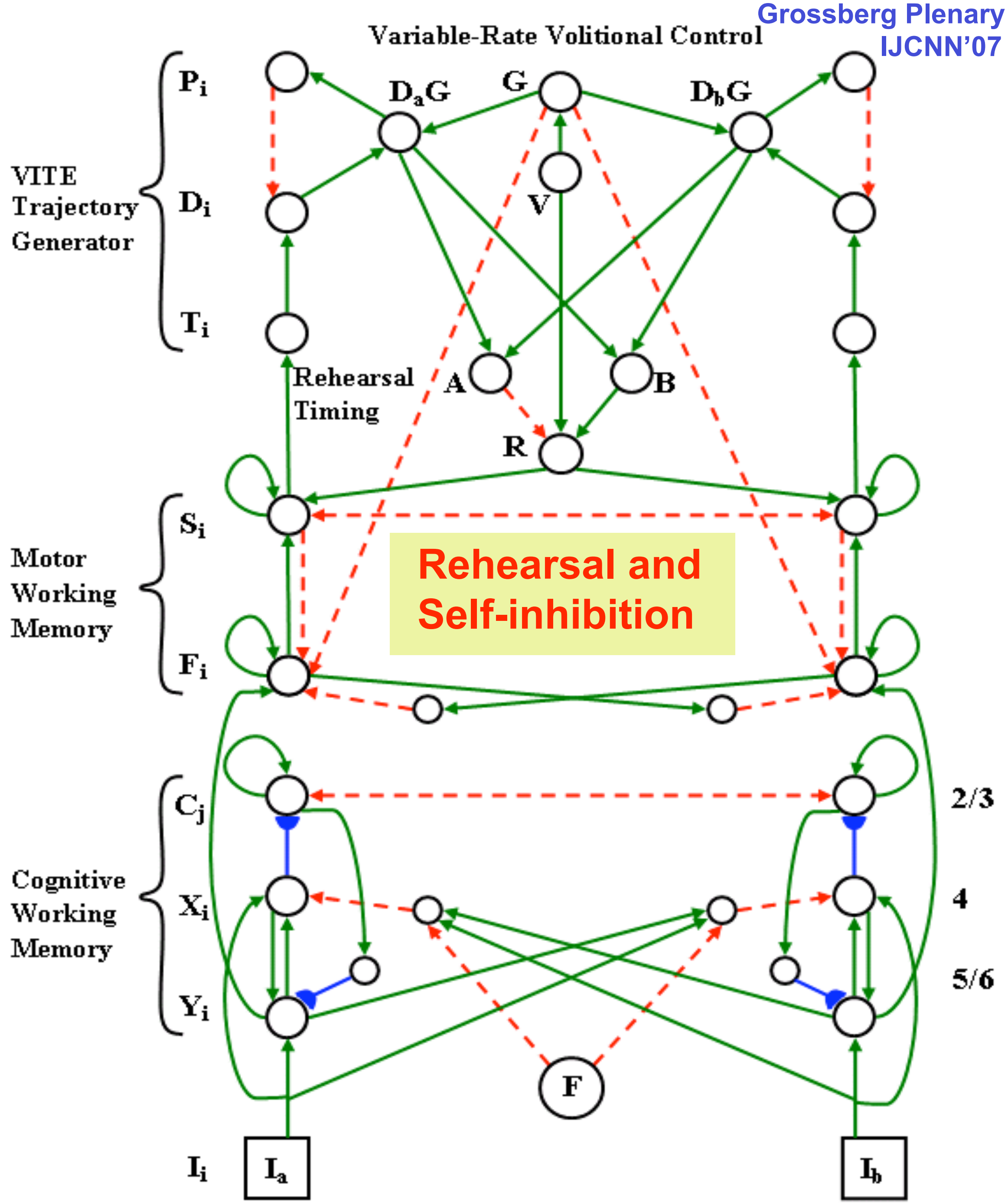
Connectivity Convention

Connection Type	Fixed	Modifiable
Excitatory		
Inhibitory		

Sequence chunks

BU Filter TD Expectation

Working memory



WORKING MEMORY MODELS: ITEM AND ORDER, or COMPETITIVE QUEUING, models

A Type of Temporal ART Dynamics

In 1978 this was a new paradigm when the Atkinson and Shiffrin model was popular

Content-addressable cells code both:

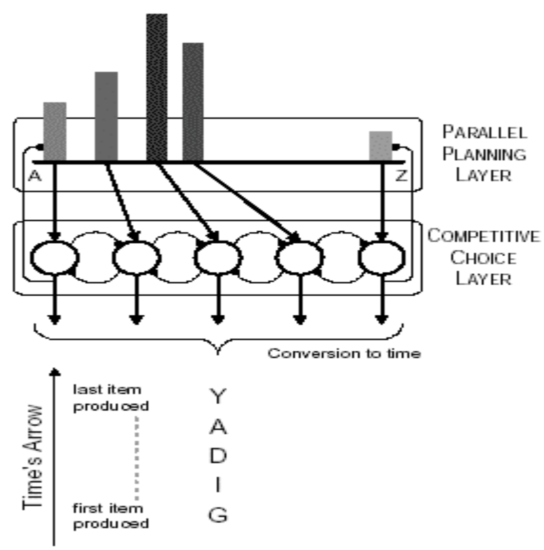
ITEM information (WHAT event occurred)
as well as
ORDER information (WHEN it occurred)

What constraints govern the design of such a working memory?

How can evolution discover a design for a process as sophisticated as a working memory?

I proposed an answer in a 1978 J. Math. Psychol. article, but with no laminar understanding: recurrent shunting on-center off-surround network

Grossberg (1978)
Houghton (1990)
Page & Norris (1998)
Farrell & Lewandowsky (2004)

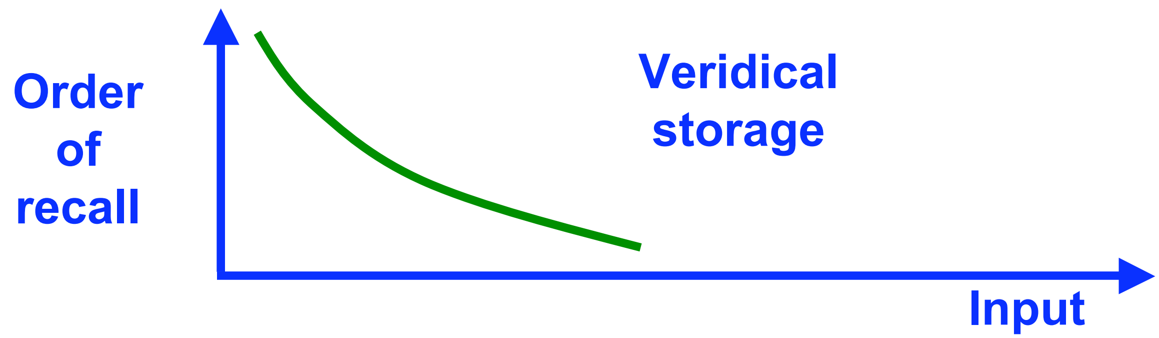


SURPRISING PREDICTION

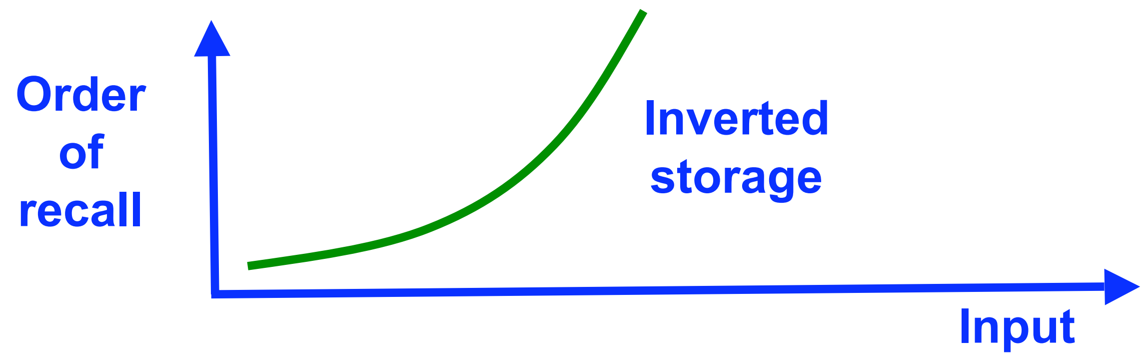
The ability to stably categorize temporal order information implies that working memory storage is not always veridical!

Three Cases:

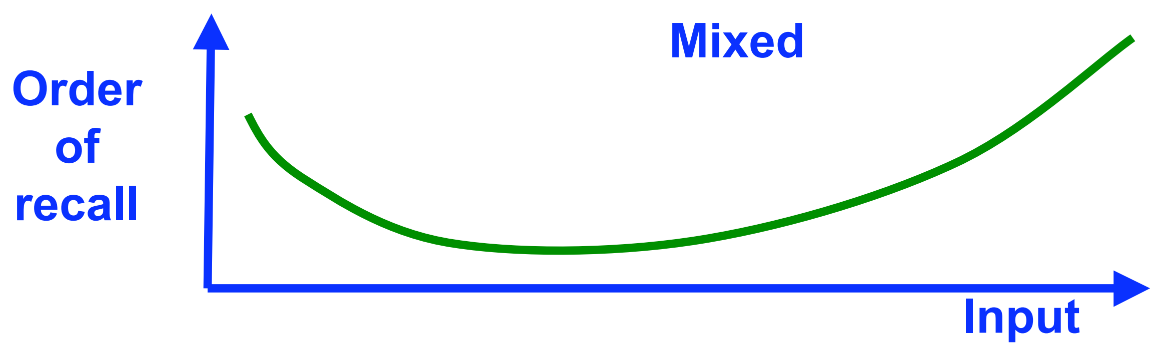
Primacy Gradient



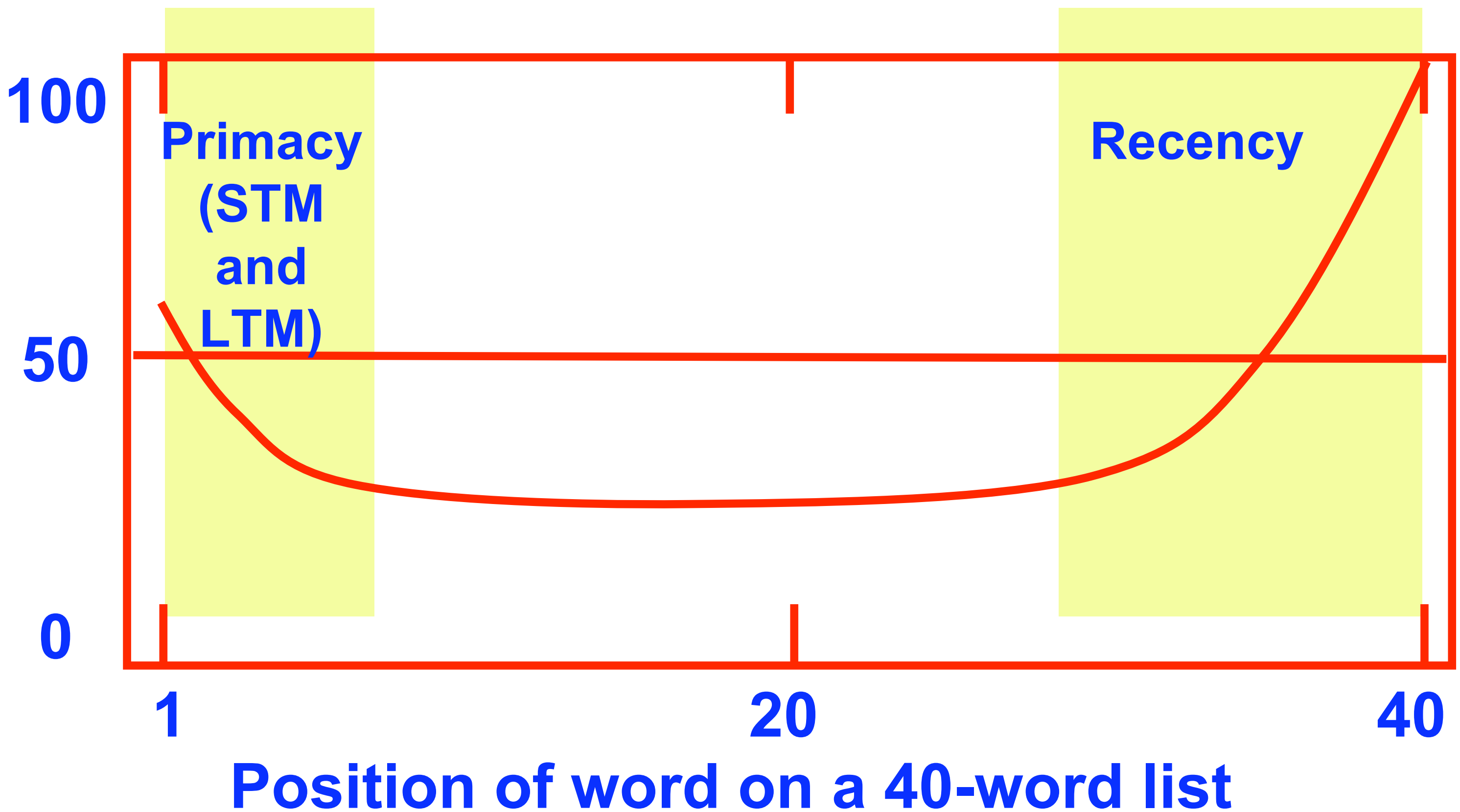
Recency Gradient



Bow



SERIAL-POSITION FUNCTION FOR FREE RECALL



Data of Murdock, B.B., J. Experimental Psychology, 1962, 64, 482-488

Grossberg (1982) Studies of Mind and Brain, Kluwer/Reidel

Grossberg and Stone (1986) Memory and Cognition

ANOTHER CONFIRMED PREDICTION

Call the position where the bow occurs the

TRANSIENT MEMORY SPAN

A pure STM effect

$TMS \approx 4$ Cf., N. Cowan (2001) Magical # 4

IMMEDIATE MEMORY SPAN

STM plus LTM readout

$IMS \approx 7$ Cf., G. Miller (1956) Magical # 7

$$\boxed{IMS > TMS}$$

Grossberg (1978)

FARRELL AND LEWANDOWSKY (2004)
SUMMARY OF CQ PREDICTIVE POWER

Abstract: "Several competing theories of short-term memory can explain serial recall performance at a quantitative level. However, most theories to date have not been applied to the accompanying pattern of response latencies, thus ignoring a rich and highly diagnostic aspect of performance. This article explores and tests the error latency predictions of four alternative mechanisms for the representation of serial order. Data from three experiments show that latency is a negative function of transposition displacement, such that list items that are reported too soon (ahead of their correct serial position) are recalled more slowly than items that are reported too late. We show by simulation that these data rule out three of the four representational mechanisms. The data support the notion that serial order is represented by a primacy gradient that is accompanied by suppression of recalled items."

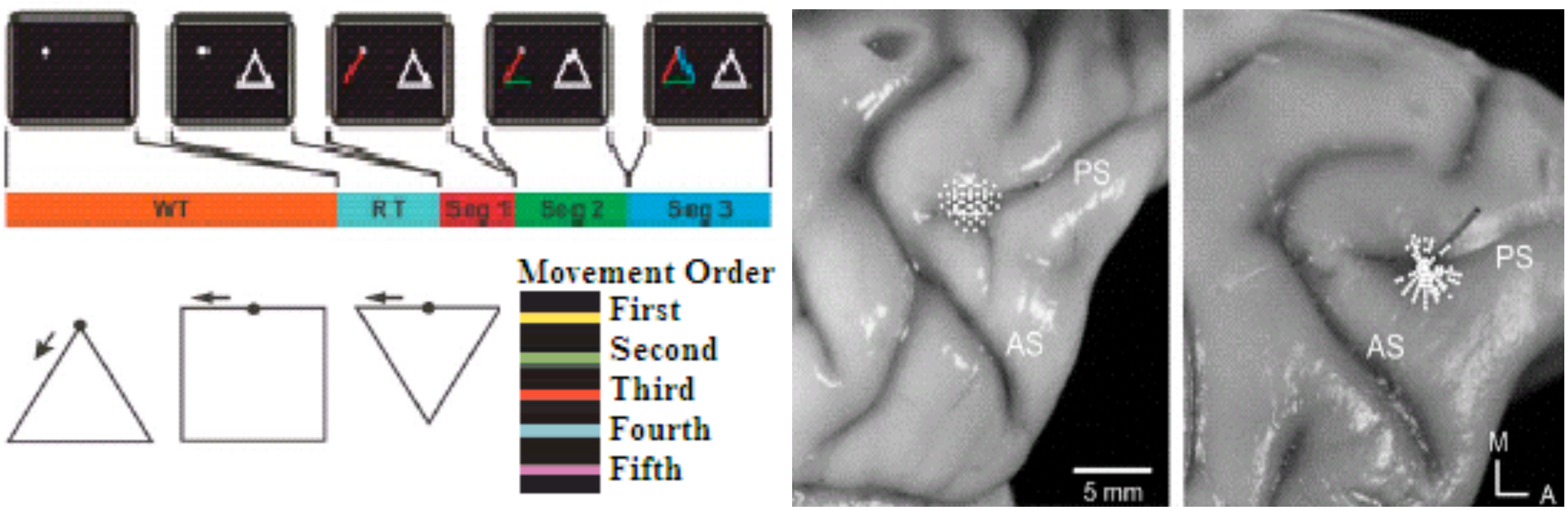
Farrell, S. and Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, 51: 115-135.

NEUROPHYSIOLOGY OF SEQUENTIAL COPYING

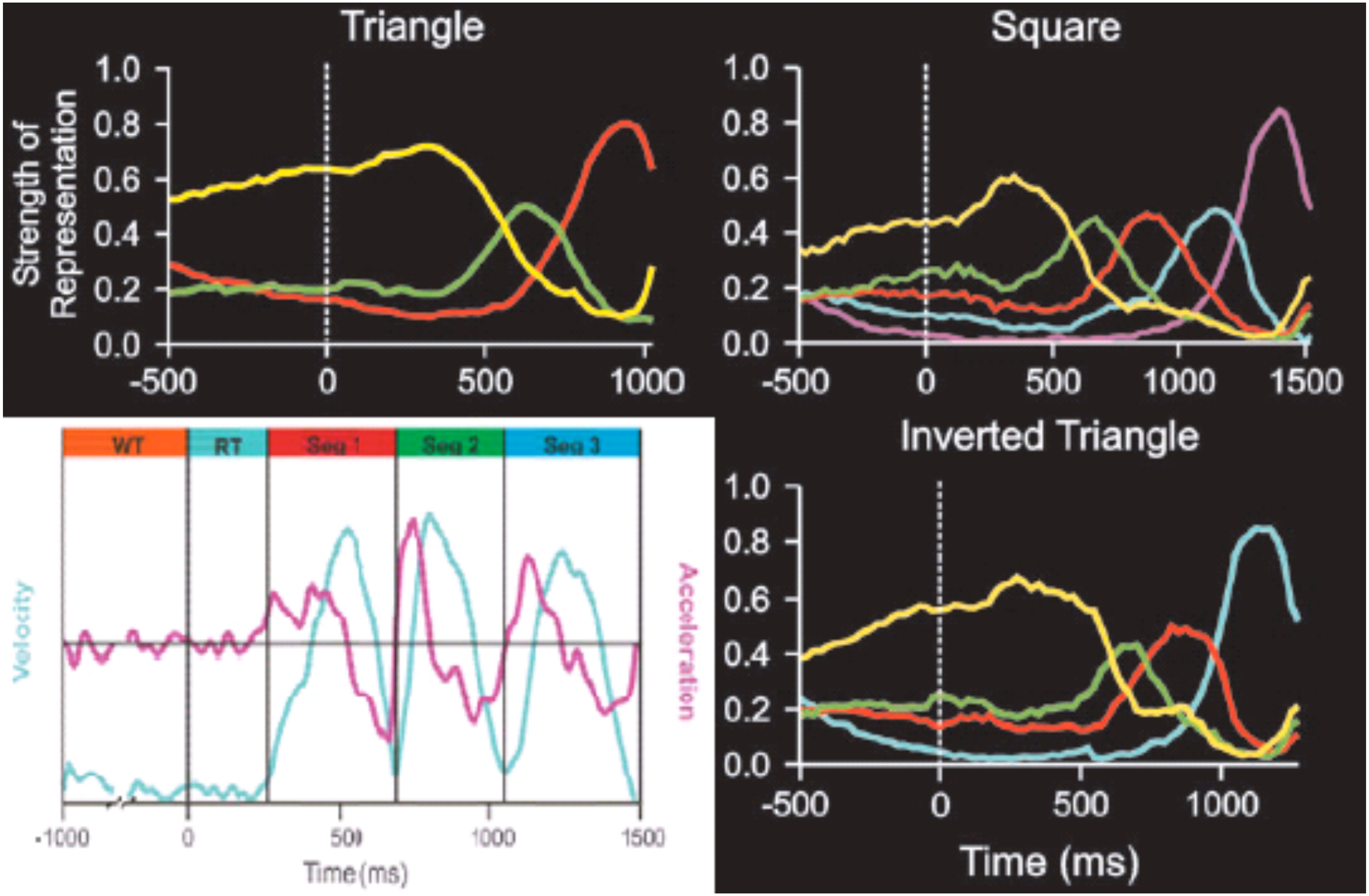
Strong neurophysiological support for 1978 working memory prediction

Extra-cellular recording in macaque peri-principalis region during joystick controlled copying

Averbeck, Chafee, Crowe & Georgopoulos (2002, 2003a, 2003b)



NEUROPHYSIOLOGY OF SEQUENTIAL COPYING



ERROR-TYPE DISTRIBUTIONS DURING IMMEDIATE SERIAL RECALL

Six-letter visual ISR

Order errors

Transpositions of neighboring items are the most common

Model Explanation

Noisy activation levels change relative order in primacy gradient

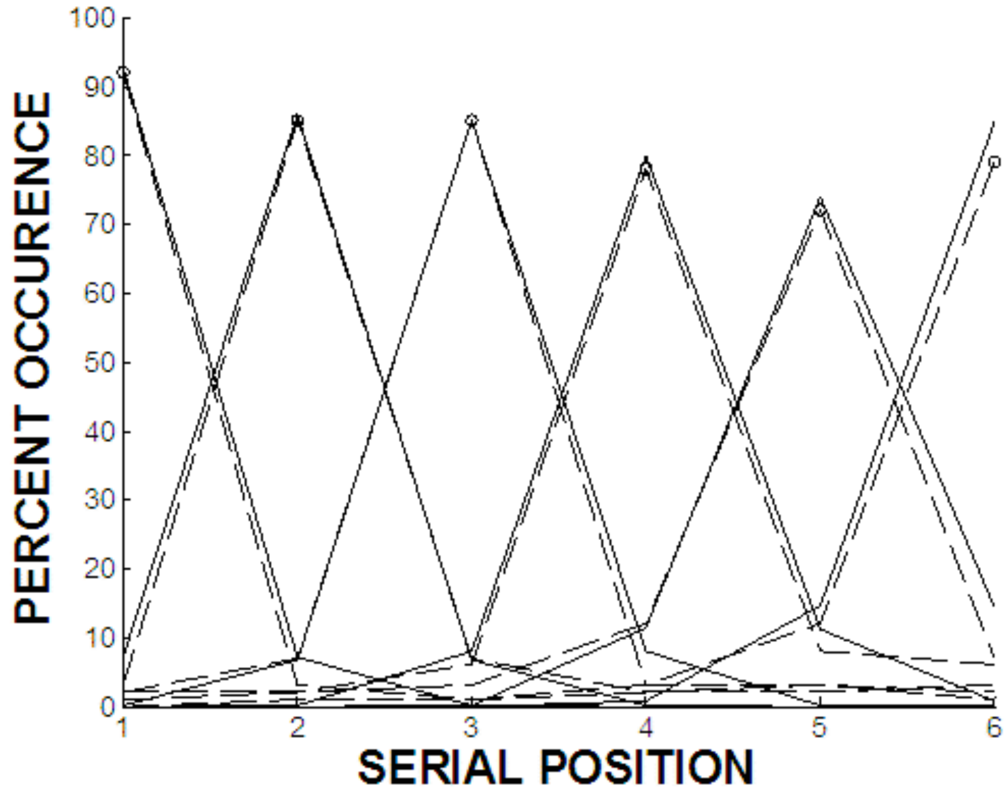
Similar activation of neighboring items; most susceptible to noise

Model parameters fitted on these data

Graph Convention

Data: Dashed Lines

Simulations: Solid Lines



Henson et al. (1996)

BOWING OF SERIAL POSITION CURVE DURING IMMEDIATE SERIAL RECALL

Auditory ISR with Various List Lengths
- **(graphs shifted rightward)**

For span and sub-span length lists:
- **Extended primacy, with one (or two) item recency**

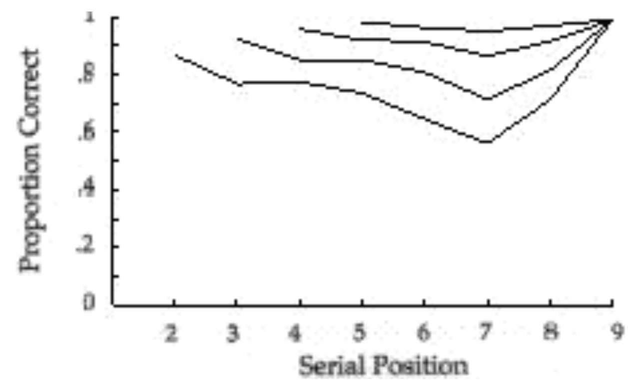
Auditory presentation: Enhanced performance for last items

LIST PARSE

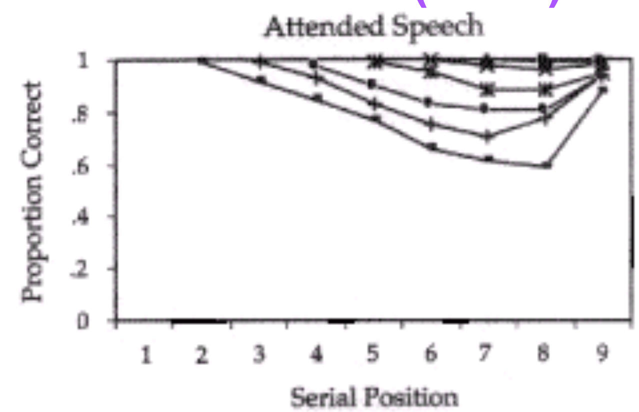
End effects: First and last items half as many neighbors

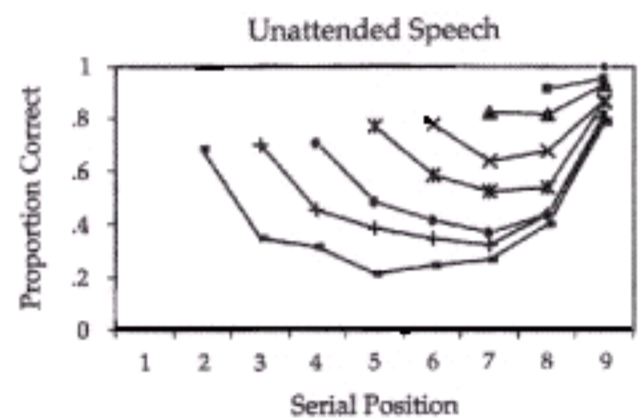
Echoic Memory: Last presented item retained in separate store

Simulation



Cowan et al. (1999)





LIST LENGTH EFFECTS DURING IMMEDIATE SERIAL RECALL

Variable List Length ISR

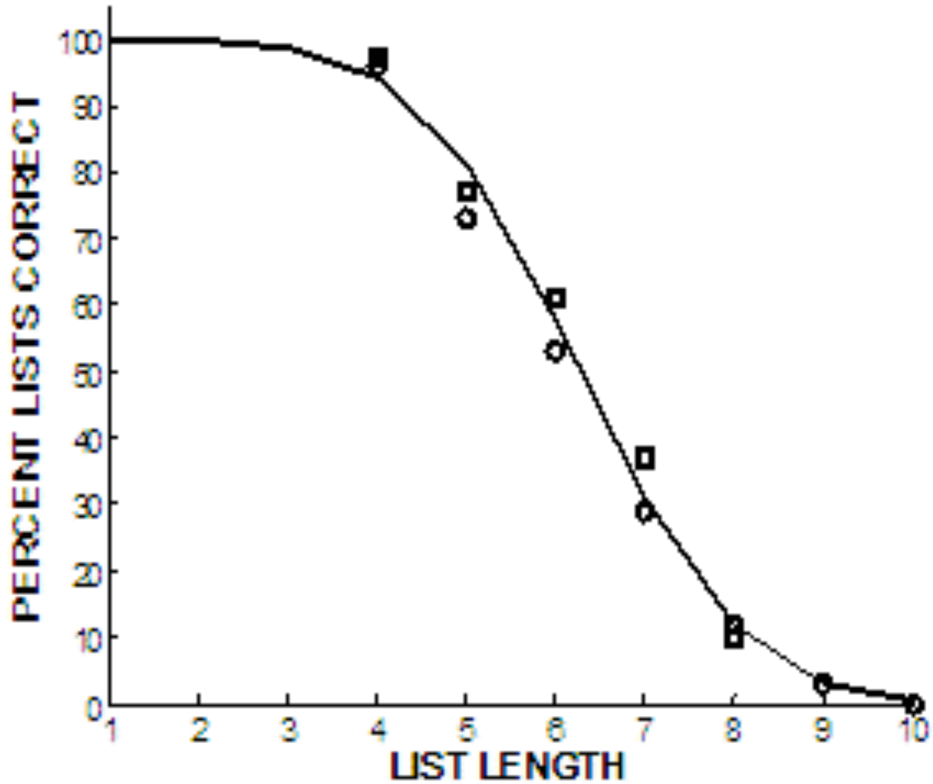
Longer lists are more difficult to recall

LIST PARSE

More items: Closer activation levels and lower absolute activity level with enough inputs

Noise is more likely to produce order errors

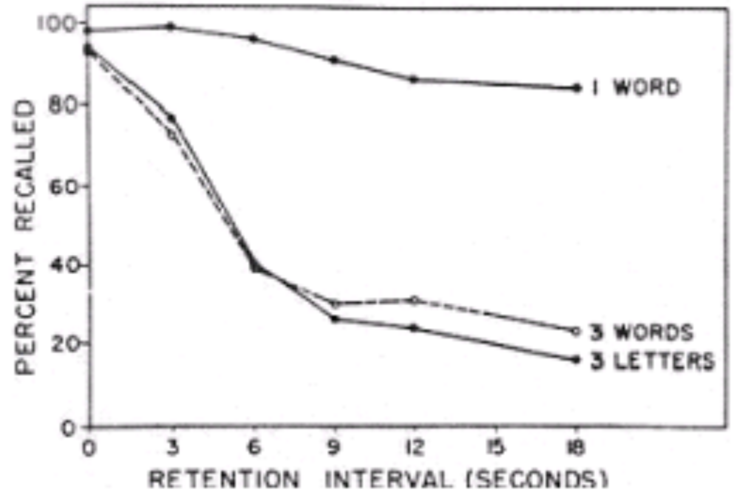
Activity levels more likely to drop below threshold



Circles: Crannell and Parrish (1968)
Squares: Baddeley and Hitch (1975)
Solid Line: Simulation

LIMITED TEMPORAL EXTENT FOR RECALL DURING IMMEDIATE SERIAL RECALL

Murdoch (1961)



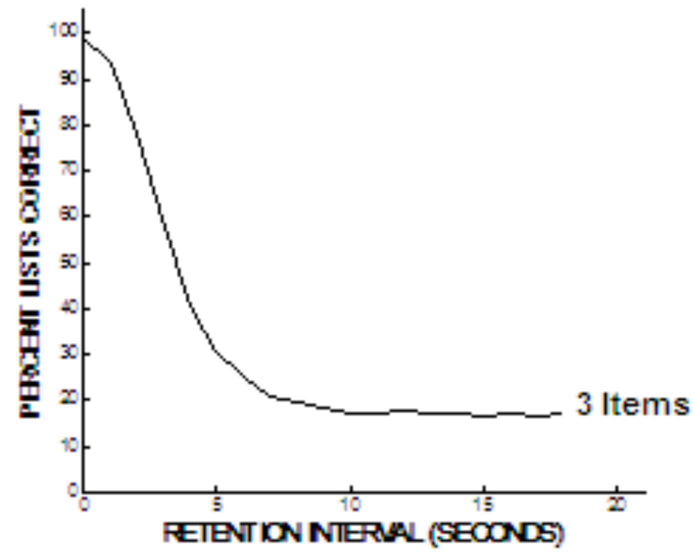
ISR task with distracter-filled retention intervals
(to prevent rehearsal)

Increasing Retention Interval:
Decreases Probability of Recalling List Correctly

Load Dependence: Longer lists more effected by delays

Performance Plateaus: Subjects reach apparent asymptote

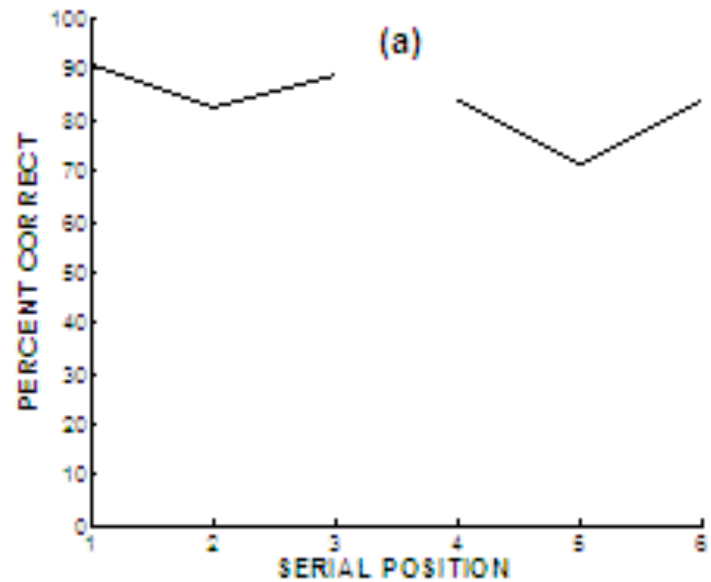
LIST PARSE
Increased convergence of activities with time; loss of order information



Simulation

TEMPORAL GROUPING AND PRESENTATION VARIABILITY

Simulations



LIST PARSE (Unique)

Temporal Grouping

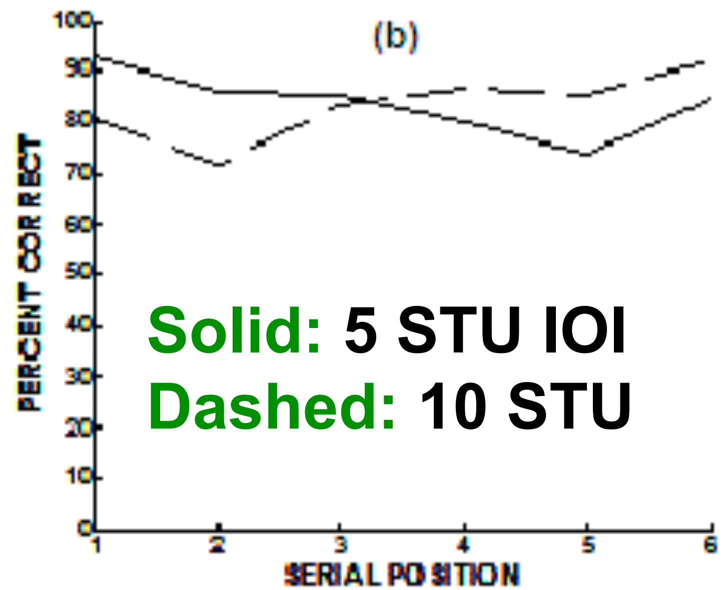
Inserting an extended pause leads to inter-group bowing

Significantly different times of integration and activity levels across pause; fewer interchanges

Prediction

Increasing IOIs while effectively preventing rehearsal

Enhances performance of recency items, weakens primacy?



IMMEDIATE FREE RECALL

Overt Rehearsal IFR Task with Super-span (i.e. 20 Item) Lists

Extended recency; even more extended with shorter ISIs

Increased probability of recall with diminished time from last rehearsal

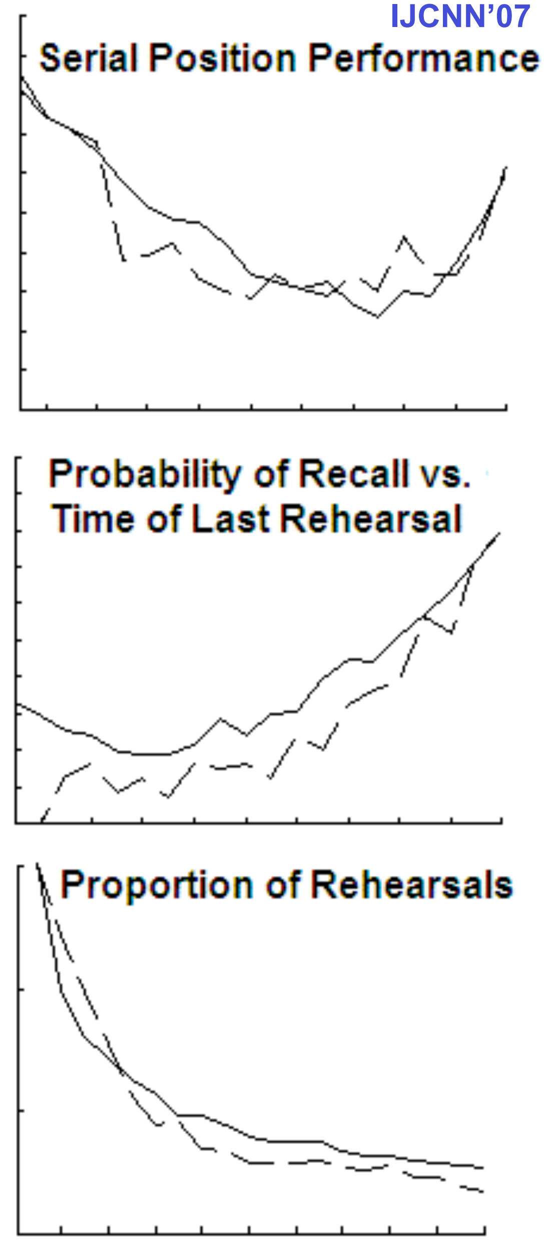
Early items in list rehearsed most

LIST PARSE (Unique)

For long lists:

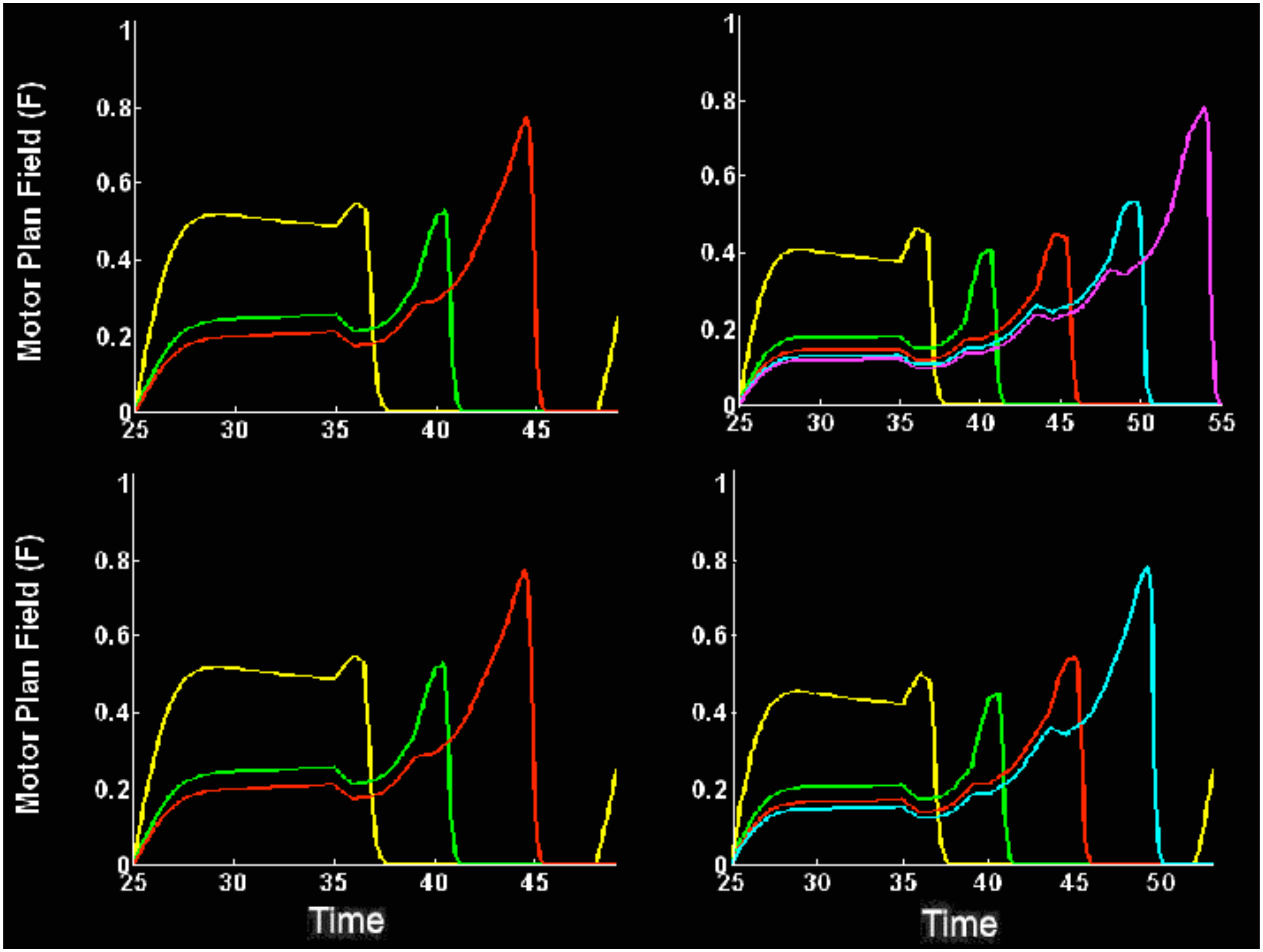
Incoming items form recency gradient

Rehearsal (re-presentation) based upon level of activity



Dashed: Data
Solid: Simulation

SIMULATES MONKEY SEQUENTIAL COPYING DATA



Simulations of neural activity in the motor plan field (F) vs. time for 3, 4 and 5 item sequences

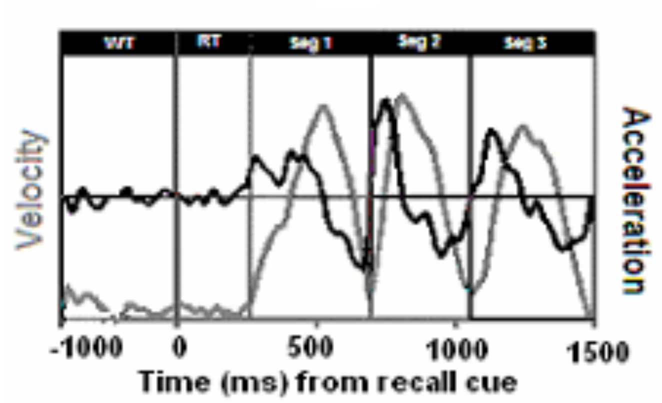
SIMULATES MOVEMENT KINEMATICS AND THEIR INTERNAL REPRESENTATION

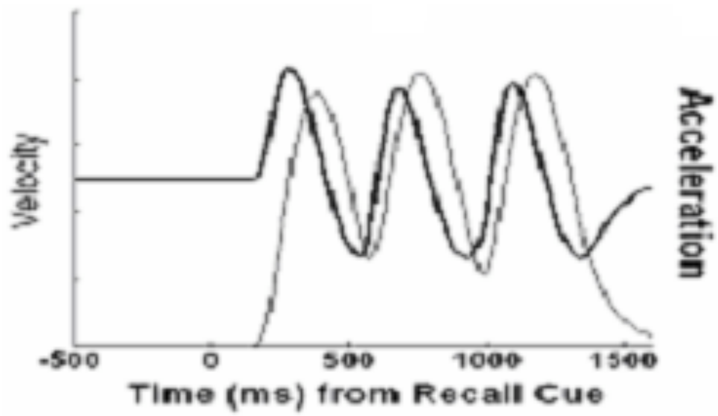
VITE Trajectory Generator
Produces bell-shaped velocity outflow signals (DG)

At start of movement, fast time-averaging cells (A) closely track the increasing outflow signals, slow cells (B) lag. B-A<0

Near completion of movement, fast cells (A) closely track decreasing outflow signals, slow cells (B) lag this decrease. B-A>0

Top: Observed Movement Kinematics





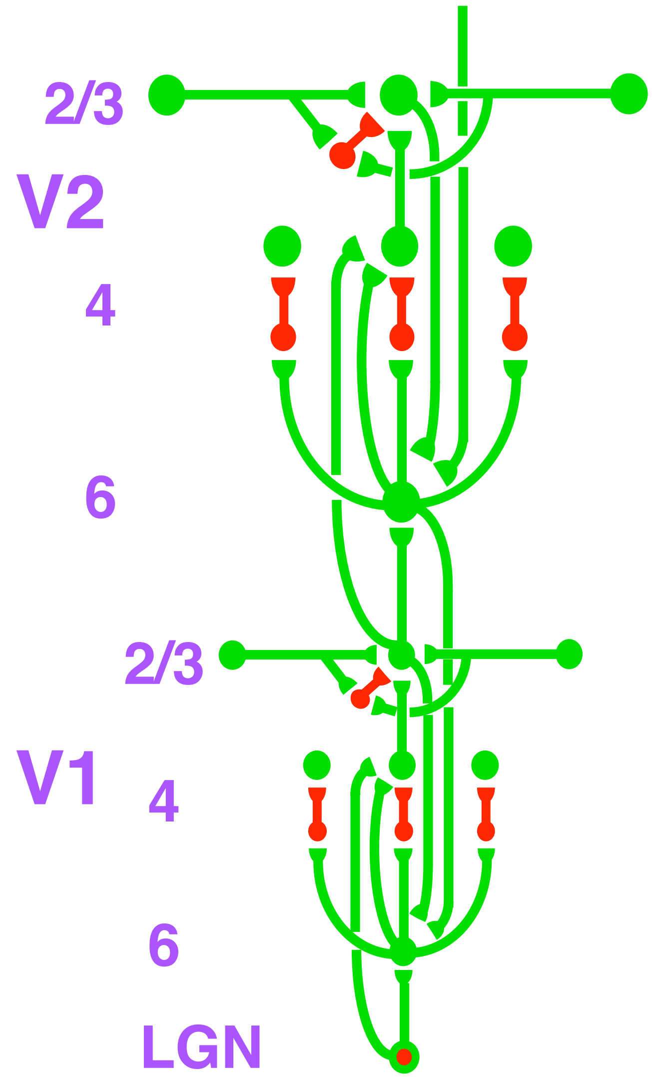
Bottom: Internal Estimates of Velocity (B) and Acceleration (A-B)

VISUAL LAMINAR MODEL: 3D LAMINART

CORTICAL AREAS V1 AND V2

Deep layers (4-6)
Item Storage
Normalization
Contrast enhancement
ORIENTED FILTERING of image contrasts

Superficial layers (2/3)
Grouping across processing channels
BINOCULAR MATCHING and PERCEPTUAL GROUPING of oriented image features

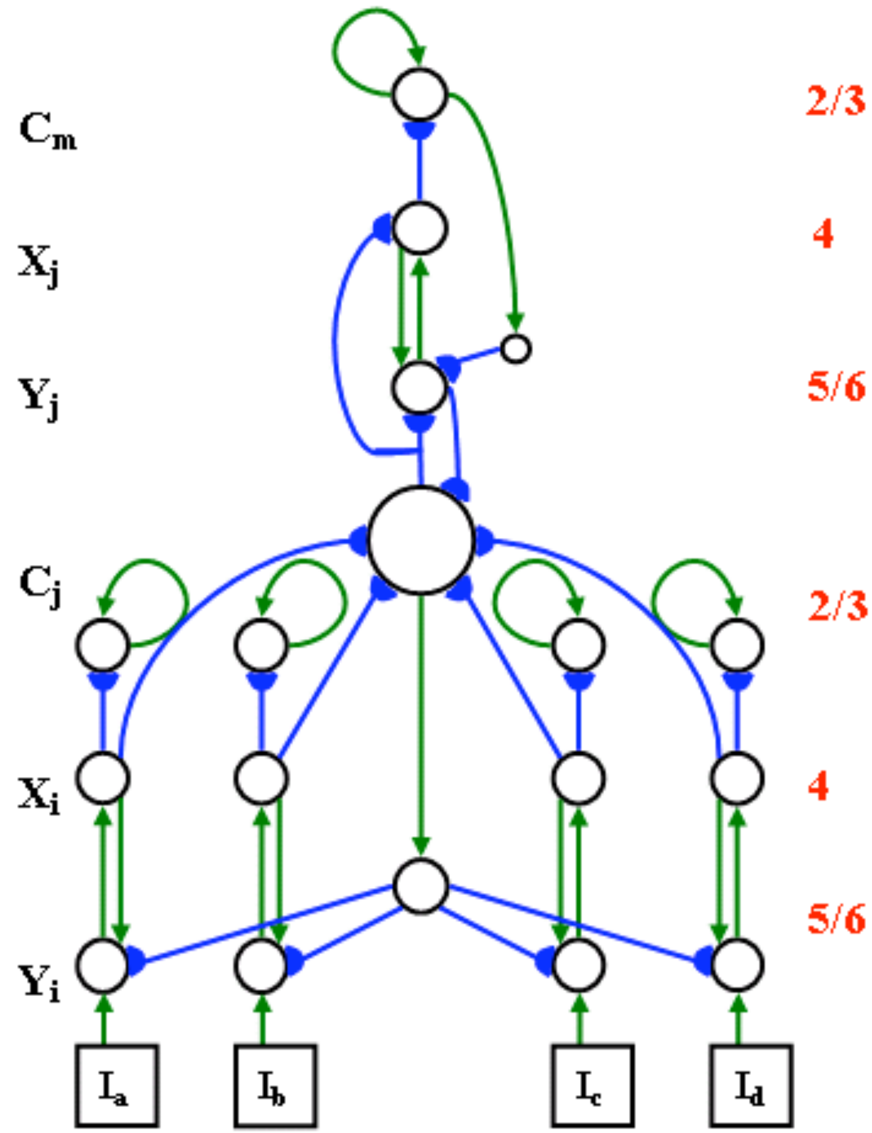


COGNITIVE LAMINAR MODEL: LIST PARSE

LATERAL PREFRONTAL CORTEX

Deep layers (4-6)
Item storage
Normalization
Contrast enhancement
WORKING MEMORY
for short-term storage of
event sequences

Superficial layers (2/3)
Grouping across processing channels
SEQUENCE CHUNKING NETWORK
for long-term coding of
familiar event sequences



TOWARDS A UNIFIED THEORY OF NEOCORTEX

Existence Proof:

Visual form perception: learning, grouping, attention

more spatial

Working memory and sequence learning

more temporal

They are variations on the same circuitry!

PREDICTION:

The volitional mechanism that controls
visual imagery and fantasy
is homologous to the mechanism that controls
storing a telephone number in working memory

WE ARE PART OF A MAJOR SCIENTIFIC REVOLUTION!

A big step forward in the physical theory of MEASUREMENT

Newton, Einstein, Heisenberg,…

Fast autonomous SELF-ORGANIZATION of a Measurement System in a Non-Stationary World

Not just small incremental steps based on known physical theories

NEW PARADIGMS
to understand
AUTONOMOUS ADAPTATION IN A CHANGING WORLD

COMPLEMENTARY COMPUTING

LAMINAR COMPUTING

Especially to those of you who are “young at heart”, I say:

JUMP ON. IT IS A GREAT RIDE!