COGS500/CmpE489: Cognitive Science

Reasoning

Adapted from Albert Ali Salah's slides
+ Atkinson&Hilgard's Introduction to Psychology Chapter 9,

Quiz

Difference between covert and overt attention.

Saliency map

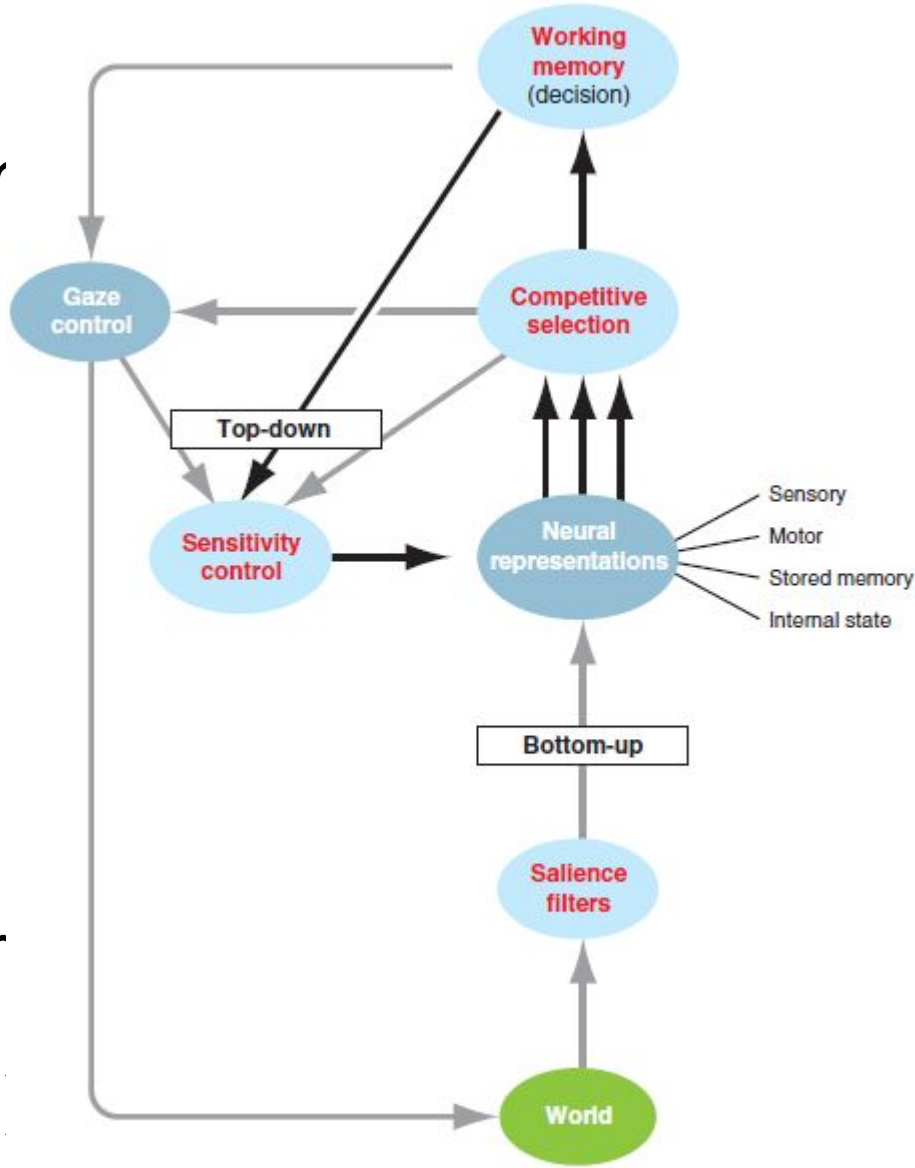
- The Saliency map integrates the normalized information from the individual feature maps into one global measure of conspicuity.
- Koch and Ullman, 1985 (p. 221), “Saliency at a given location is determined primarily by how different this location is from its surround in color, orientation, motion, depth etc.”
- Koch & Ullman hypothesized that a centralized retinotopic map can provide regulation of attention deployment (i.e. saliency map)
- Recent results show a number of retinotopically organized maps in the brain, particularly in midbrain, thalamus and occipital lobes, but also in temporal, parietal and even frontal cortices (Saygın & Sereno)

Attentional selection and inhibition of return

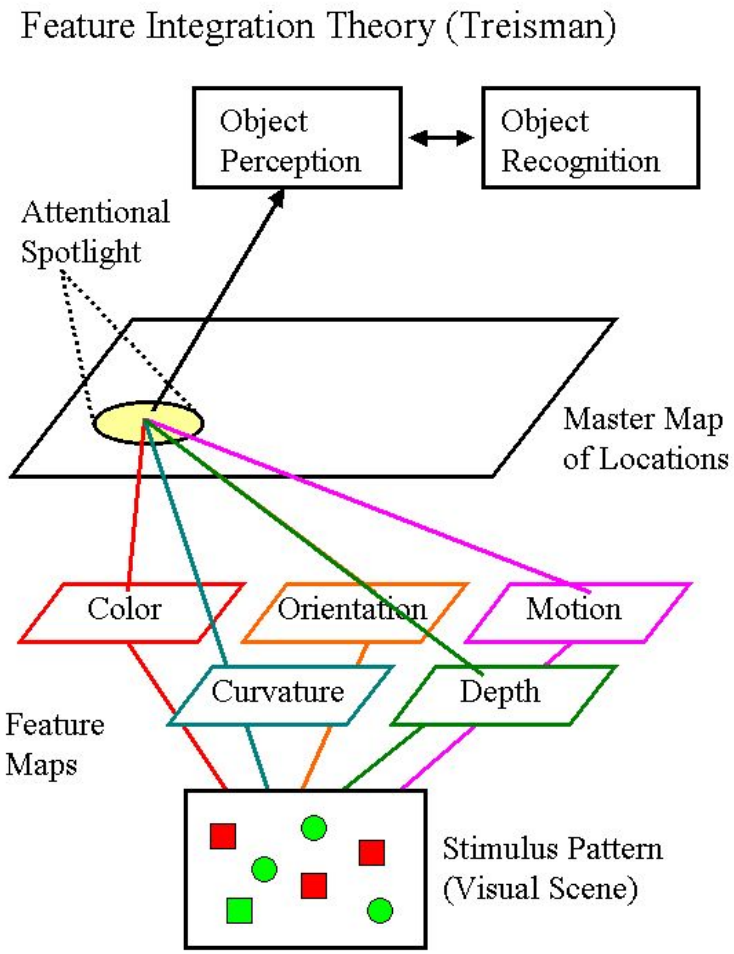
- Attentional spotlight attends to the most salient spot
- This spot is temporarily suppressed to allow the next salient spot to capture attention
  - Through “inhibition of return”
- Speeded local discriminations:
  - visual processing at recently attended locations might be slower
- Inhibition of return is shown to be object-bound

Components of attention

- According to Knudser (2007), attention requires four processes:
  - Working memory,
  - Competitive selection,
  - Top-down sensitivity control,
  - Filtering for stimuli that ar likely to be behaviorally important (salience filters

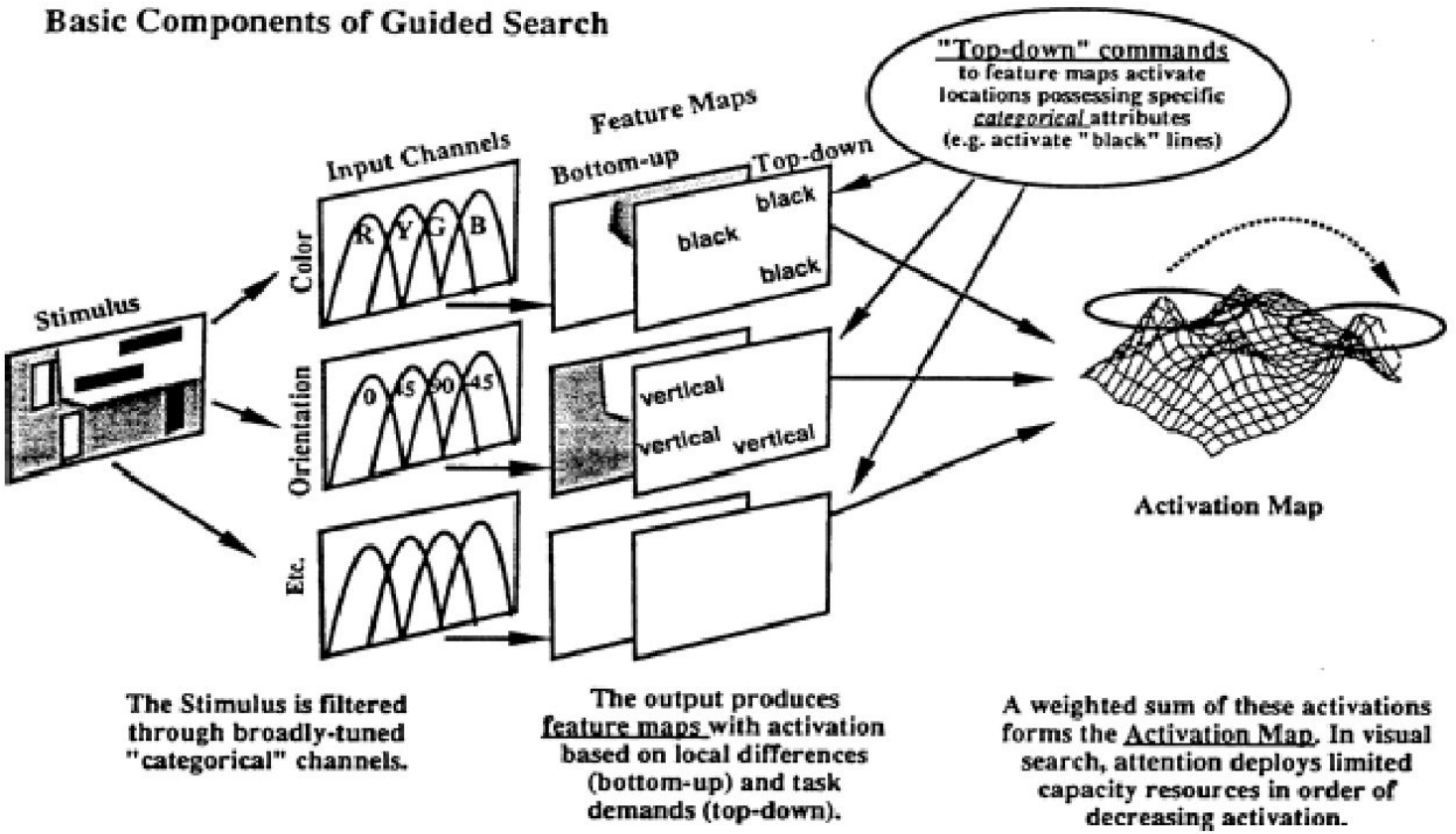


Feature Integration Theory



http://www.luc.edu/faculty/asutter/FIT.gif

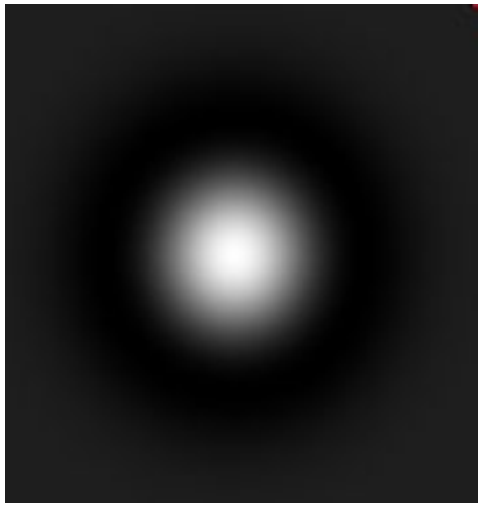
Guided Search



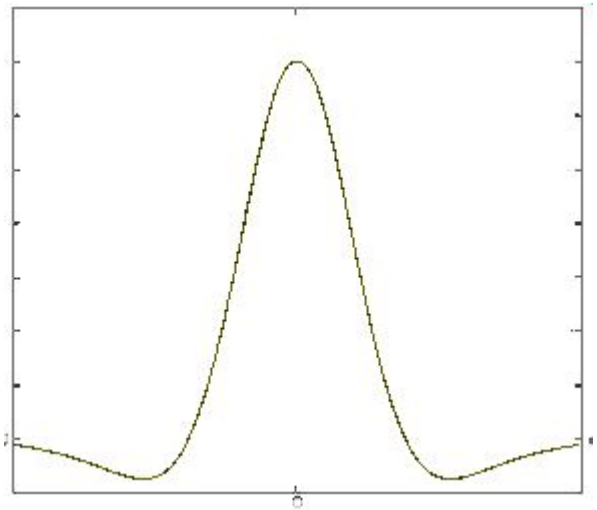
Wolfe, 1994

Centre-surround cell responses

$$\mathrm{DoG}_{\sigma,\gamma}(x,y) = A_c \frac{1}{\gamma^2} e^{-\frac{x^2+y^2}{2\gamma^2\sigma^2}} - A_s e^{-\frac{x^2+y^2}{2\sigma^2}}$$

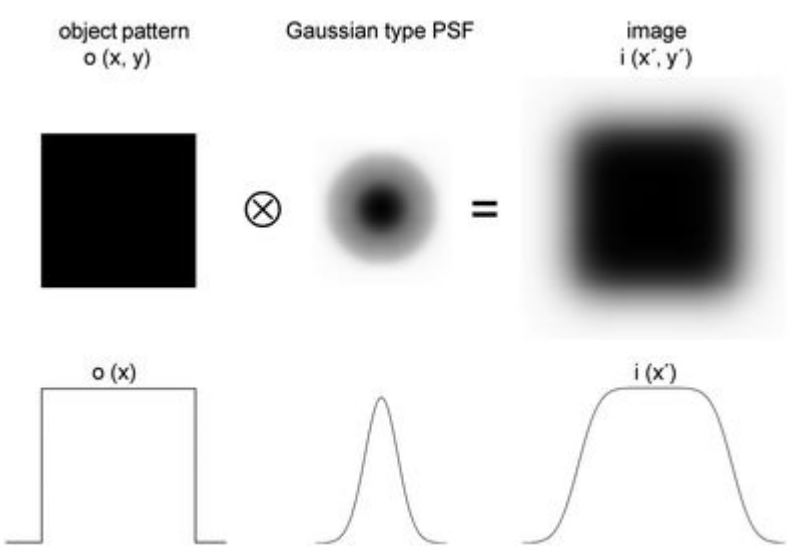


Intensity map of a
two-dimensional DoG function



One-dimensional DoG profile

CONVOLUTION

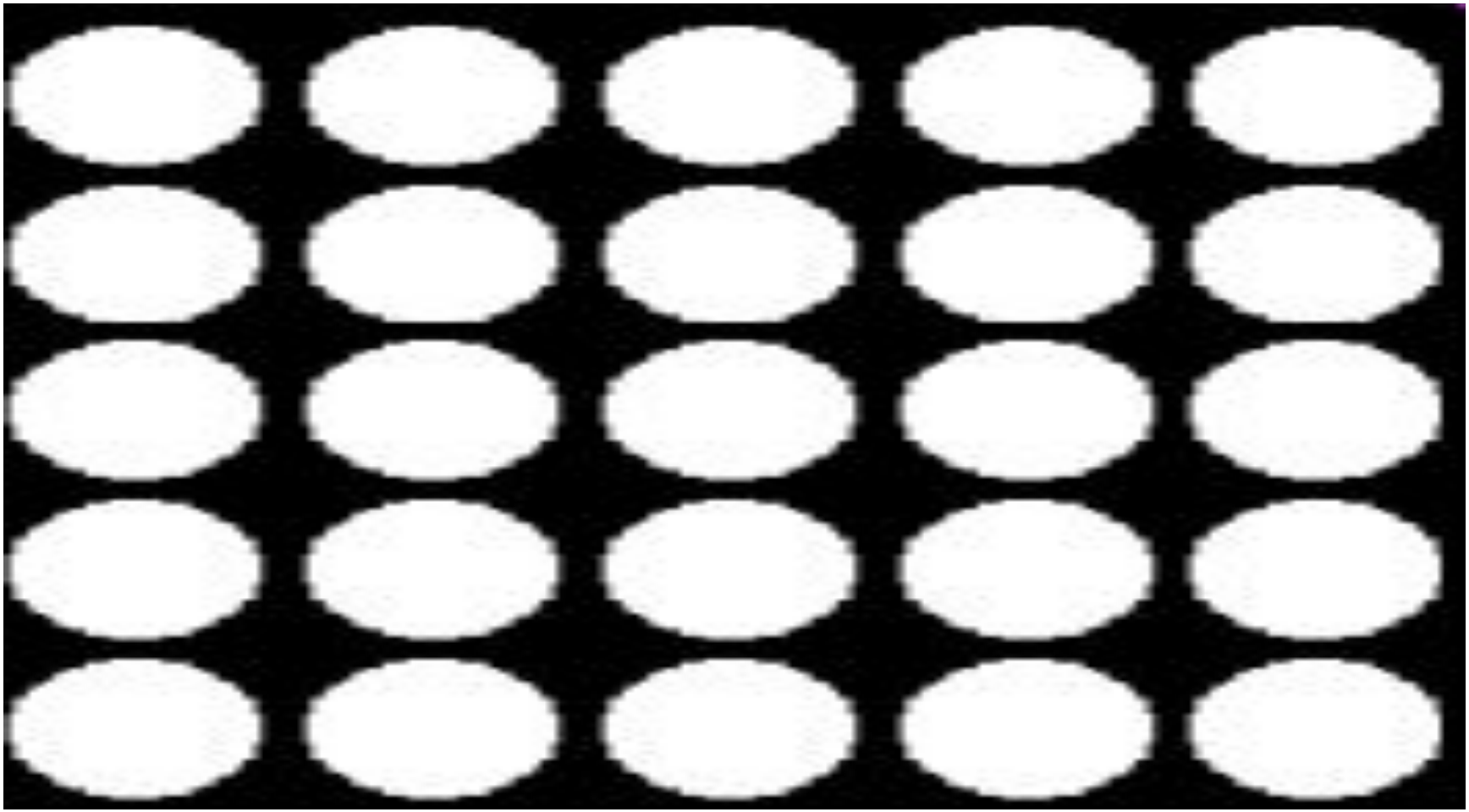


$\sigma = 1$

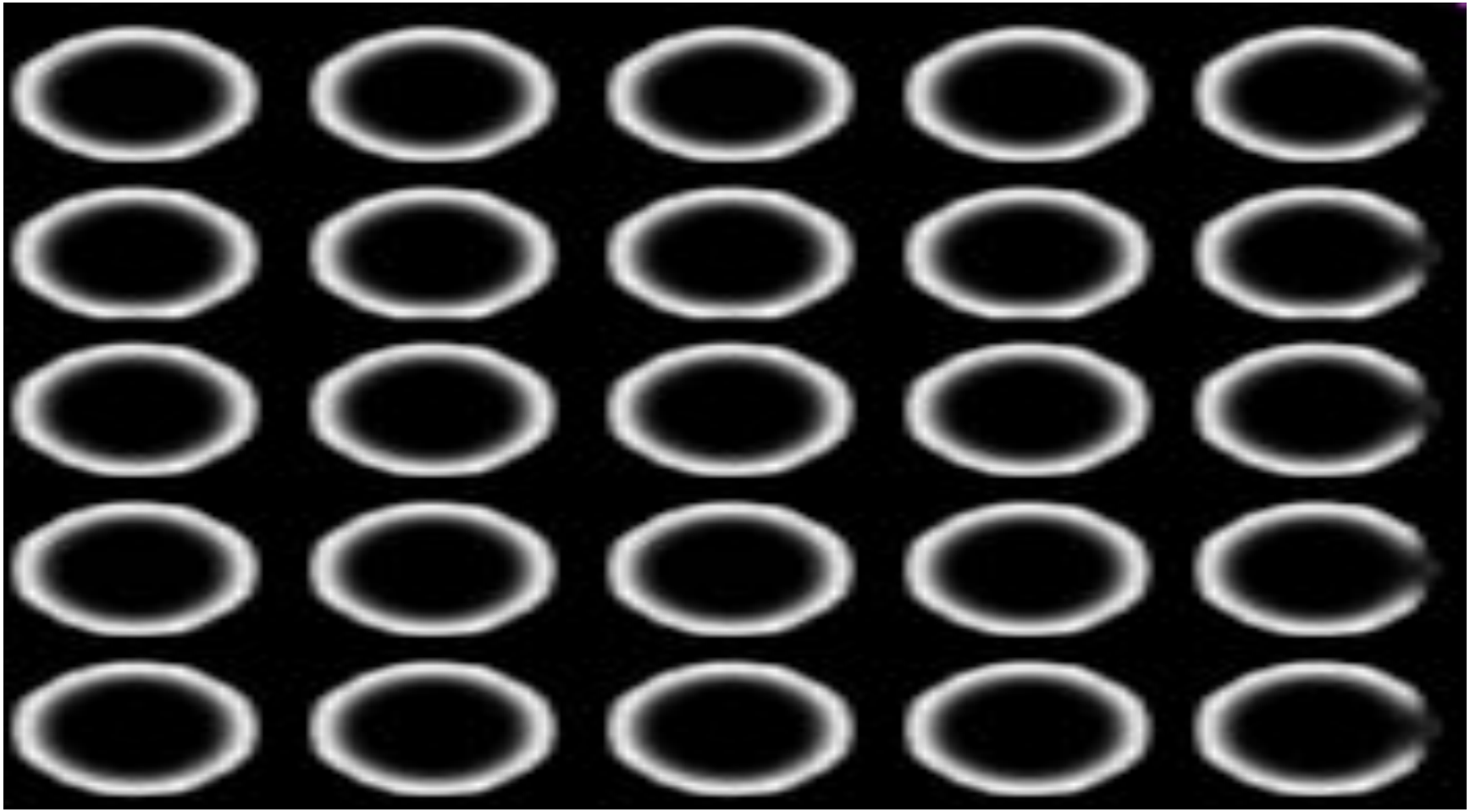
$\sigma = 2$

difference

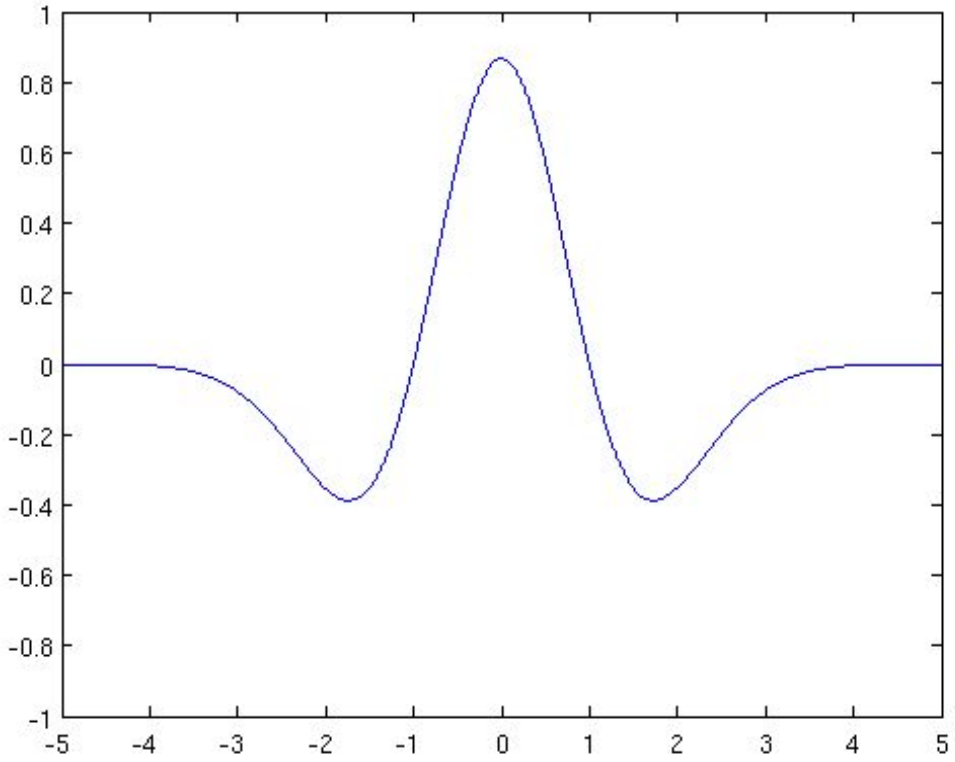
The effect of centre-surround



This will be the result

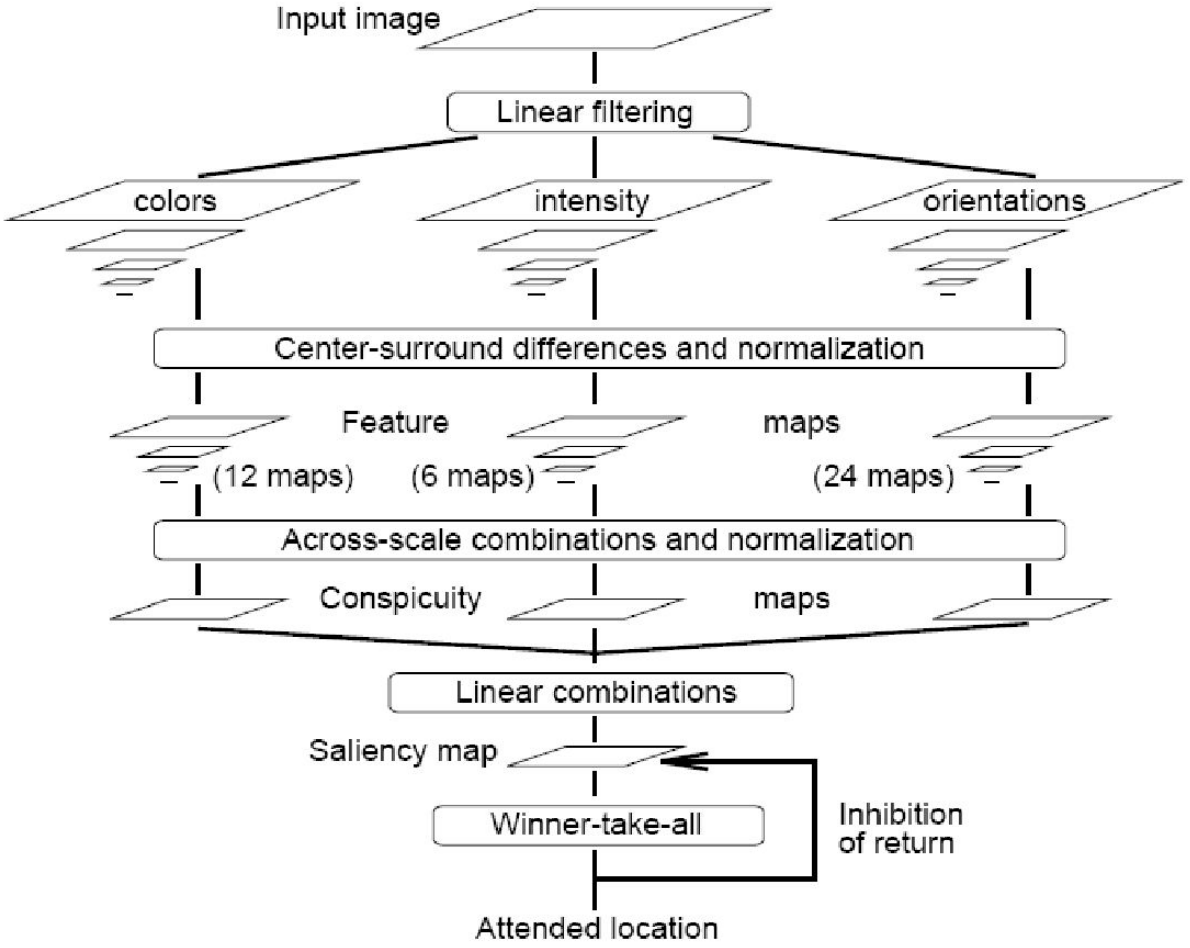


Mexican hat

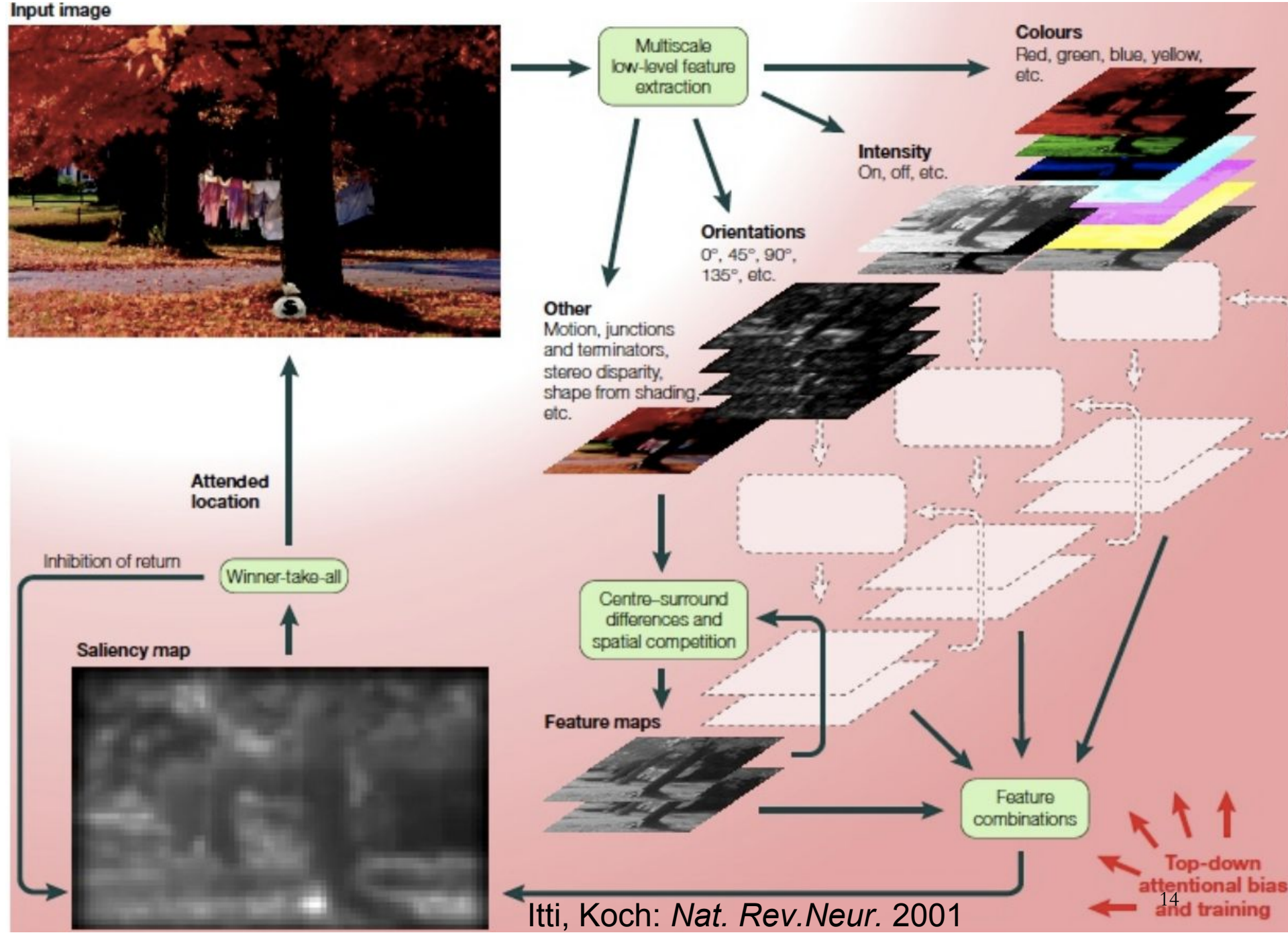


$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}\left(1 - \frac{t^2}{\sigma^2}\right)e^{\frac{-t^2}{2\sigma^2}}$$

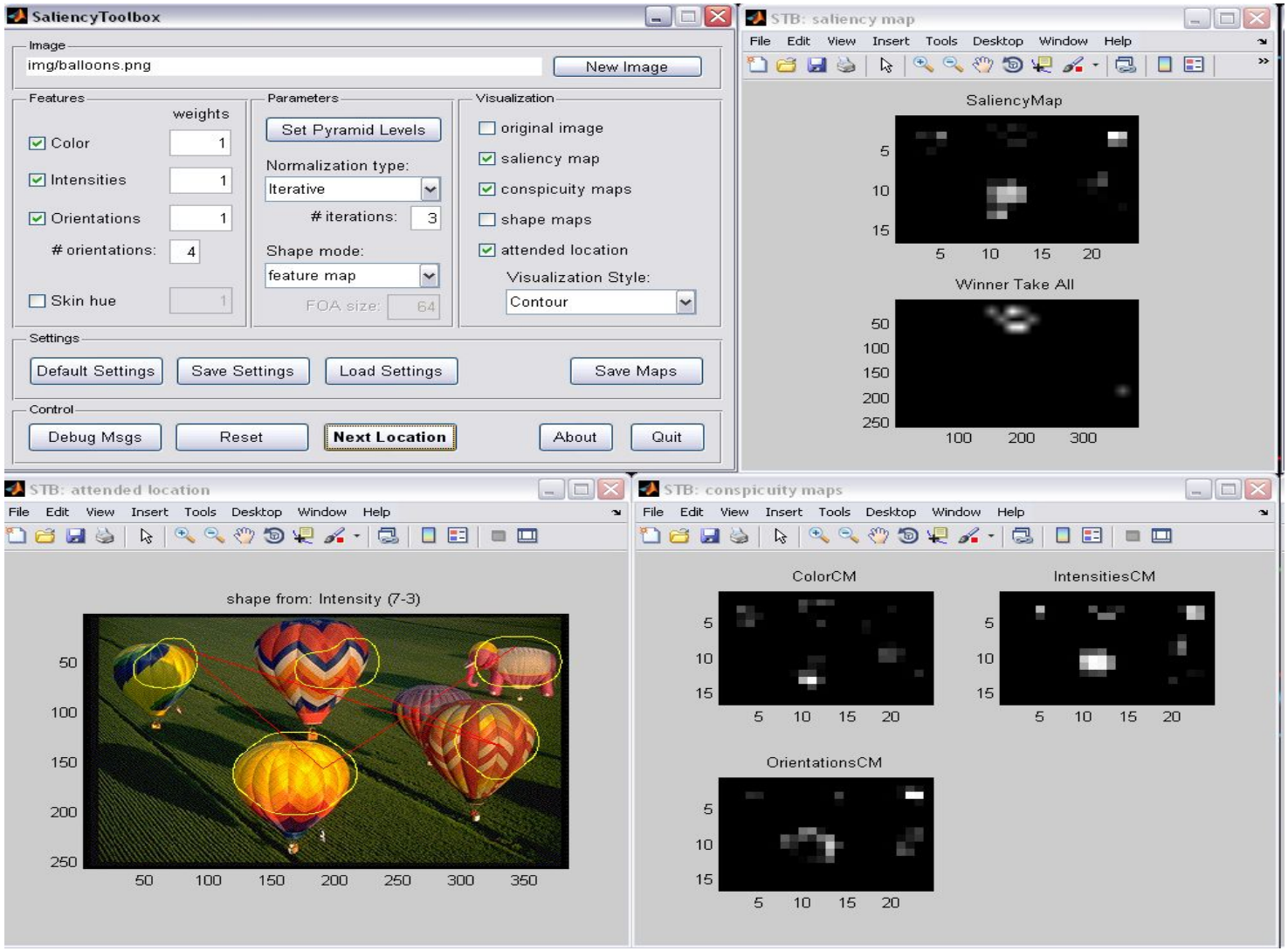
Computational Model for FIT



Itti, Koch, Niebur, IEEE Trans. PAMI, 1998.

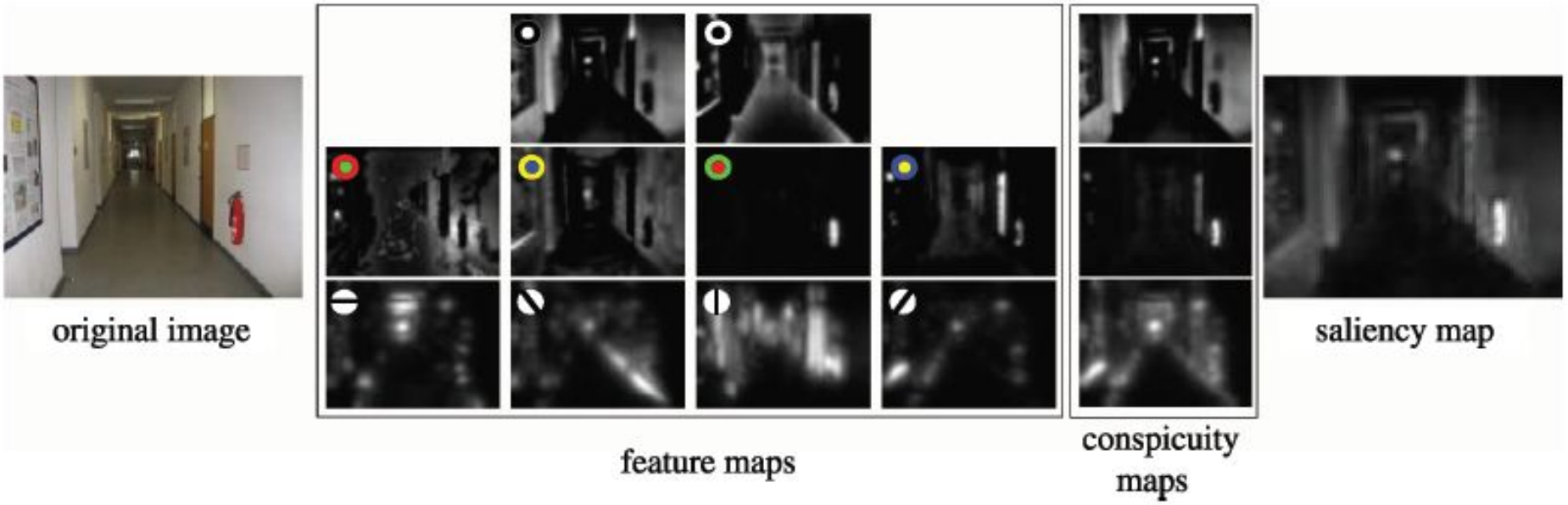


Itti, Koch: *Nat. Rev.Neur.* 2001



Dirk Walther, http://www.saliencytoolbox.net/

Saliency Map (VOCUS)



Frintrop, ACM TAP, 2010

Rational animals

- Aristotle defines man as a “rational animal”
- What does the definition of rationality involve?
  - Intentionality
  - Acting for reasons
  - Logic
  - Causal reasoning, tool use
  - Beliefs
  - Language
  - Products or processes

Questions on rationality

- To what degree is it a normative concept?
- Can we define rationality across species?
  - Dretske's birds: A bird avoids eating monarch butterflies (poisonous), it is minimally rational. It also avoids eating viceroy, which looks similar, but is not poisonous.

Definitions of rationality

- Webster's:
  - rational: having reason or understanding
  - irrational: not endowed with reason or understanding; lacking usual or normal mental clarity or coherence
- Greek roots
  - *logos*: reason
  - *ratio*: proportion
- Being logical: having "tacit knowledge of the fundamental semantic principle governing any inference" (Johnson-Laird – proposes "mental logic")
- Piaget: "reasoning is nothing more than the propositional calculus itself"
- Gardner: rational is an individual reaching conclusions "by shrewd intuition, lucky guessing or being programmed to issue only valid responses"

Definitions of rationality

- Newell: “If an agent has knowledge that one of his actions will lead to one of its goals, then the agent will select that action”
- Dennett: rationality is “a general purpose term of cognitive approval which requires maintaining only conditional and revisable allegiances between rationality, so considered, and the proposed (or even universally acclaimed) methods of getting ahead, cognitively in the world”
- Cohen: valid deductive or probabilistic reasoning
- Selten: full rationality requires unlimited cognitive resources

Artificial intelligence and rationality

■The exciting new effort to make computers thinks … *machine with minds,* in the full and literal sense" (Haugeland 1985)	■"The study of mental faculties through the use of computational models" (Charniak et al. 1985)
•"The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990)	• A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkol, 1990)

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

Thinking Rationally: Laws of Thought

- Aristotle (~ 450 B.C.) attempted to codify “right thinking”
  - What are correct arguments/thought processes?
- e.g., deductive reasoning: “Socrates is a man, all men are mortal; therefore Socrates is mortal”
- What about inductive reasoning? “Socrates is a man, Socrates is mortal; probably all men are mortal”
- Several Greek schools developed various forms of logic:
  - notation plus rules of derivation for thoughts

Thinking Rationally: Laws of Thought

- Problems:
  - Not all intelligent behavior has a logical basis.
  - Uncertainty: Not all facts are certain (e.g., *the flight might be delayed).*
  - Resource limitations:
    - Not enough time to compute/process
    - Insufficient memory/disk/etc
    - Etc.

Acting Rationally: The Rational Agent

- Rational behavior: Doing the right thing!
- The right thing: That which is expected to maximize the expected return
- Provides the most general view of AI because it includes:
  - Correct inference ("Laws of thought")
  - Uncertainty handling
  - Resource limitation considerations (e.g., reflex vs. deliberation)
  - Cognitive skills (NLP, AR, knowledge representation, ML, etc.)
- Advantages:
  - More general
  - Its goal of rationality is well defined

Rational Agents

- An **agent** is just something that acts (agent comes from the Latin *agere,* to do*).*
- Attributes that distinguish agents from mere “programs” include:
  - operating under autonomous control,
  - perceiving their environment,
  - persisting over a prolonged time period,
  - adapting to change, and
  - being capable of taking on another’s goals.
- The AI course is mostly about designing **rational agents**.
- A rational agent is one that **acts** so as **to achieve** the best outcome, or when there is uncertainty, **the best expected outcome**.

Value and utility

- Rationality in economics refers to optimality in face of expected value and/or expected utility
- Suppose I believe the following:
  - P(A25 gets me there on time) = 0:04
  - P(A90 gets me there on time) = 0:70
  - P(A120 gets me there on time) = 0:95
- Which action to choose depends on my preferences for missing flight vs. airport cuisine, etc. Utility theory is used to represent and infer preferences
- Decision theory = utility theory + probability theory

Utility

- There are also problems concerning people's calculations of utility.
- Some of these take the form of "thought experiments", where the argument depends on your intuitions (as the founding principles of normative theories generally do),
- and some depend on experimental work.
- We shall look at three classic paradoxes of preference and choice, then at preference reversals and framing effects.

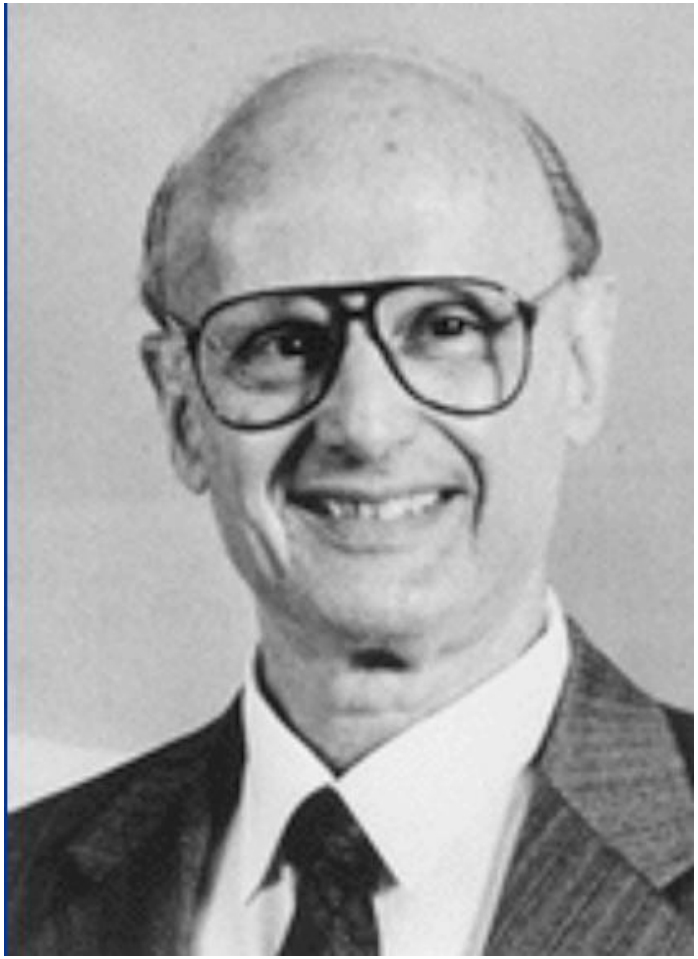
1. Allais paradox

(1a) $1000 with probability 1 (i.e. certain)

(1b) $1000 with probability .90
$5000 with probability .09
$0 with probability .01

(2a) $1000 with probability .10
$0 with probability .90

(2b) $5000 with probability .09
$0 with probability .91



2. Ellsberg paradox

30 red balls

60 black or yellow balls

A: receive $100 if a red ball

B: receive $100 if a black ball

C: receive $100 if a red or yellow ball

D: receive $100 if a black or yellow ball

Harry Markowitz

The ultimatum game

- The ultimatum game is a game in economic experiments. The first player (the proposer) conditionally receives a sum of money and proposes how to divide the sum between the proposer and the other player. The second player (the responder) chooses to either accept or reject this proposal. If the second player accepts, the money is split according to the proposal. If the second player rejects, neither player receives any money. The game is typically played only once so that reciprocation is not an issue. wikipedia
- Two players divide a sum given to them
- The first player divides, the second player accepts or rejects
- Played only once!
- In many cases, offers less than 20% are rejected!

Mutual decision making

- von Neumann & Morgenstern (Theory of Games, 1944)
  - Uncertainty also arising from mutual attempts of decision makers
  - Simon: There is no satisfactory definition of “optimal” rationality in the presence of opportunities for outguessing and outwitting.
- Muth (1961): rational expectations
  - all actors have the same economic model in mind
- Rappaport & Chammah (1965): Prisoner’s dilemma

Prisoner's Dilemma:

Two members of a criminal gan and imprisoned.

Prisoner's dilemma payoff matrix

A \ B	B stays silent	B betrays
A stays silent	-1 / -1	0 / -3
A betrays	-3 / 0	-2 / -2

- If A and B each betray the other, each of them serves 2 years in prison
- If A betrays B but B remains silent, A will be set free and B will serve 3 years in prison (and vice versa)
- If A and B both remain silent, both of them will only serve 1 year in prison (on the lesser charge)

Prisoner's Dilemma

- While cooperation is collectively rational, defection is individually rational.
- What changes when the game is iterated (played multiple times between the same players)?
- Tit-for-tat heuristic:
  - Unless provoked, the agent will always cooperate
  - If provoked, the agent will retaliate
  - The agent is quick to forgive
  - The agent must have a good chance of competing against the opponent more than once.

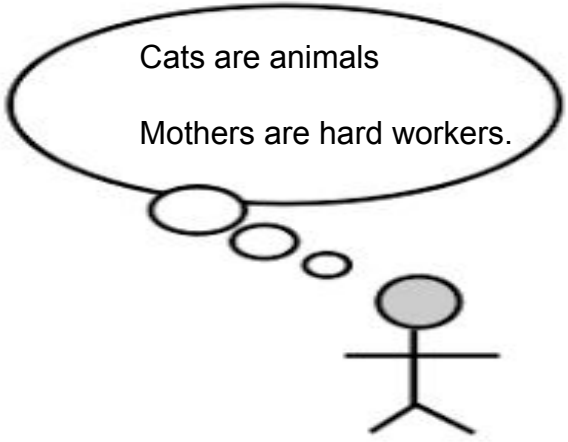
Iterated Prisoner's Dilemma

- If the game is played exactly N times and both players know this, then it is optimal to defect in all rounds. Why?
- Robert Aumann in a 1959 paper, rational players repeatedly interacting for indefinitely long games can sustain the cooperative outcome.
- Axelrod invited academic colleagues all over the world to devise computer strategies to compete in an IPD tournament
  - greedy strategies did poor in the long run
- The winning deterministic strategy was tit for tat, of Anatol Rapoport
- Nice, retaliating, forgiveness, non-envious

The building blocks of thought

- Language of the mind
  - Imaginal thought
  - Propositional thought
    - Consists of concepts
      - An entire class, reduce complexity
      - Allow prediction
      - Concepts not correspond to objects
      - Concepts on the spot

http://weclipart.com/



What is categorization?

Prototypes

- The properties that describe the best examples
- Grandmother
  - Prototype
  - Core
- Natural concepts are fuzzy. Why?
- Universality in prototypes of concepts
- Hierarchies of concepts

Different categorization processes

- Well-defined concepts:
  - Rule-based **If she's the female parent of a parent, she's a grandmother**
  - More properties, slower
- Fuzzy concepts
  - Defining properties?
  - Similarity to Prototypes, stored examplars
    - A typical experiment:
      - determine properties of the prototype of a concept
      - determine similarity between each instance and prototype based on shared properties
      - correlation, similarity-to-prototype vs accuracy speed

The neural basis

- Fuzzy vs. well-defined
  - Animals vs artifacts
  - Recognize pictures of
  - Is this sufficient?
- Prototype vs. exemplar

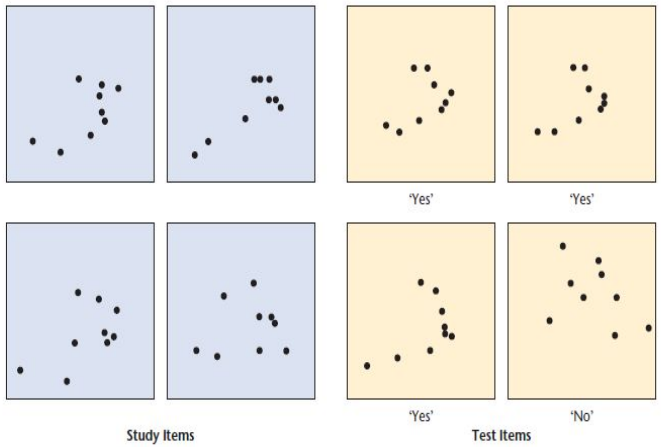


Figure 9.5 Examples of Dot Patterns Used to Study Categorization in Amnesiac Patients. *Individuals learned that the study items all belonged to one category and then had to decide whether each of the test items belonged to that category. The test items that belong to the category (the ones labeled 'yes') do not match the study items directly. Rather, the test items that belong to the category are*

medial-temporal lobe damage

The dissolution of semantics.

EXPORT ★ Add To My List Database: PsycINFO Journal Article

McCarthy, Rosaleen Warrington, Elizabeth K.

Citation

McCarthy, R., & Warrington, E. K. (1990). The dissolution of semantics. *Nature, 343*(6259), 599. http://dx.doi.org/10.1038/343599a0

Abstract

Describes further degeneration of semantic knowledge in a patient previously described by R. A. McCarthy and E. K. Worthington (see record 1989-16167-001). The S had an impairment of comprehension that was apparently restricted to certain semantic categories in the auditory-verbal domain. The different rates of degeneration in the S's verbal and visual knowledge strengthen the hypothesis of modality- and category-specific cerebral meaning systems. (PsycINFO Database Record (c) 2016 APA, all rights reserved)

Neurocase (1998) Vol. 4, pp. 311–338 © Oxford University Press 1998

Are Living and Non-living Category-specific Deficits Causally Linked to Impaired Perceptual or Associative Knowledge? Evidence From a Category-specific Double Dissociation

Matthew A. Lambon Ralph, David Howard[1], Gemma Nightingale[2] and Andrew W. Ellis[3]

MRC Cognition and Brain Sciences Unit, Cambridge, [1]Department of Speech, University of Newcastle, [2]Department of Linguistics, University of York and [3]Department of Psychology, University of York, UK

Abstract

Perhaps the most influential view of category-specific deficits is one in which the dissociation between living and non-living kinds reflects differential reliance on, or weighting of visual or associative–functional attributes. We present data collected from two patients, which question the apparent relationship between category-specific deficits and loss of specific attribute types. One patient with dementia of Alzheimer's type presented with relatively poor performance on living things but failed to show a difference between knowledge of visual and associative–functional information. The other patient with semantic dementia demonstrated relatively poor knowledge of visual attributes but failed to exhibit a category-specific impairment for animate kinds. In fact her comprehension and naming were slightly but significantly better for living things. The data are discussed with reference to various theories of category-specific impairment. We suggest that category-specific deficits for living things probably results from a combination of atrophy to medial and neocortical temporal structures, including the inferior temporal lobe. It is proposed that at the behavioural level, category-specific deficits arise when both critical identifying attributes of knowledge are lost and the intercorrelation between features causes disintegration of the category such that each exemplar 'regresses' towards a category prototype.

Cognitive, Affective, & Behavioral Neuroscience
2001, 1 (4), 360-370

PET evidence for multiple strategies of categorization

ANDREA L. PATALANO, EDWARD E. SMITH, JOHN JONIDES, and ROBERT A. KOEPPE
University of Michigan, Ann Arbor, Michigan

It is widely held that people use multiple strategies to categorize their experiences in the world. We conducted a pair of neuroimaging experiments to identify the neural correlates of two of these strategies—rule application and exemplar similarity. Participants were instructed to perform either a rule- or an exemplar-based categorization task while changes in cerebral blood flow were measured using positron emission tomography. Patterns of neural activity were consistent with the predictions of cognitive models of rule- versus exemplar-based categorization and with existing neuroscience data. The identification of strategy-specific neural patterns offers future researchers a diagnostic tool for assessing strategy use in other situations.

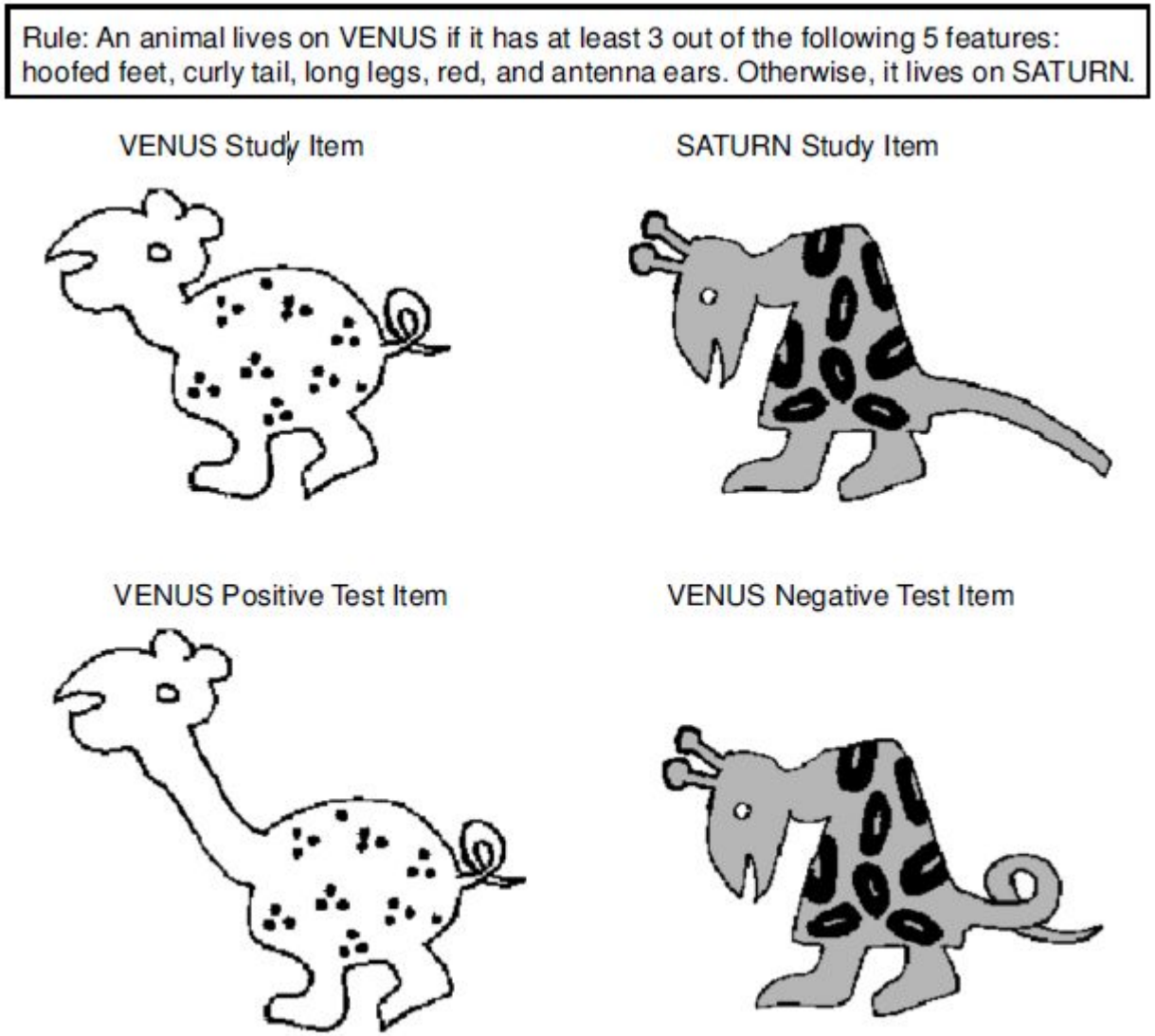


Figure 1. Example animal stimuli used in Experiments 1 and 2. In the original, the two left-hand images are red, and the two right-hand images are blue.

In the prefrontal cortex, exemplar group activationwas not different from 0% [M 5 3.4%; t (11) 5 1.49, p 5 .100], suggesting a qualitative difference between the rule and the exemplar groups. In both left (M 5 3.4%) and right (M 5 5.6%) posterior parietal areas, respectively, exemplar activation was greater than 0% (

The Neural Basis for Categorization in Semantic Memory[1]

Murray Grossman,*[,2] Edward E. Smith,† Phyllis Koenig,* Guila Glosser,*
Chris DeVita,* Peachie Moore,* and Corey McMillan*

*Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania, 19104; and
†Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109

Received March 11, 2002

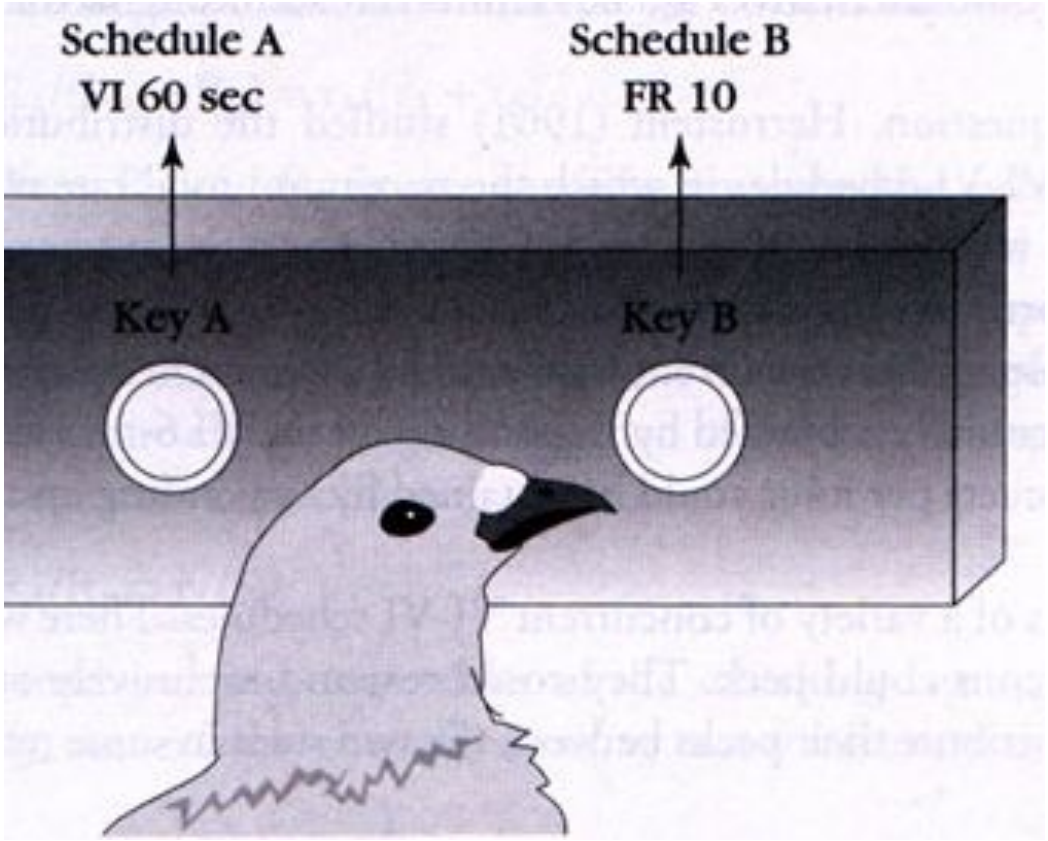
We asked young adults to categorize written object descriptions into one of two categories, based on a rule or on overall similarity, while we monitored regional brain activity with functional magnetic resonance imaging (fMRI). We found significantly greater recruitment of left dorsolateral prefrontal cortex for rule-based categorization in direct comparison with similarity-based categorization. Recruitment of right ventral frontal cortex and thalamus was uniquely associated with rule-based categorization as well. These observations lend support to the claim that executive functions such as working memory, inhibitory control, and selective attention contribute to rule-based categorization. Right inferior parietal activation was uniquely associated with similarity-based categorization. This region may play an important role in overall feature configuration that is important for this form of categorization. We found other brain regions recruited for both rule-based and similarity-based categorization: Anterior cingulate cortex may support the implementation of executive functions during situations with competing response alternatives; and left inferior parietal cortex may be related to the integration of feature knowledge about objects represented in modality-specific association cortices. We also administered a degraded-similarity condition where the task of categorizing a written object description was made more difficult by perceptually degrading the stimulus materials. The degraded condition and the rule-based condition, but not the similarity-based condition, were associated with caudate activation. The caudate may support resource demands that are not specific for a particular categorization process. These findings associate partially distinct large-scale neural networks with different forms of categorization in semantic memory. ◦ 2002 Elsevier Science (USA)

INTRODUCTION

Semantic memory is our accumulated knowledge of basic meanings and facts (Tulving, 1972). Categorization is fundamental to understanding and using the concepts in semantic memory, since this process helps organize our knowledge and relate a test object to other known objects in the world. Categorization also allows us to engage in activities such as understanding unfamiliar objects and learning about novel objects. In the present study, we use functional magnetic resonance imaging (fMRI) to study the neural basis for categorization during semantic memory processing.

At least two kinds of categorization processes are thought to apply to semantic memory. The first process involves a global comparison of a test object with remembered instances of the category (Medin and Schaffer, 1978; Medin *et al.,* 1993), or possibly a comparison with a mental prototype representing category members (Rosch and Mervis, 1975; Smith and Medin, 1981). This "similarity-based" approach appears to depend in part on the implicit organization of perceptual features contributing to the overall appearance of an object (Ashby *et al.,* 1998) and a comparison of the test object with previously encountered exemplars or a prototype of the category (Medin and Schaffer, 1978; Smith and Medin 1981). The second kind of categorization process involves a "rule-based" approach, where a test object is evaluated with reference to a set of propositions or rules representing category membership criteria (Bruner *et al.,* 1956; Murphy and Medin, 1985; Smith

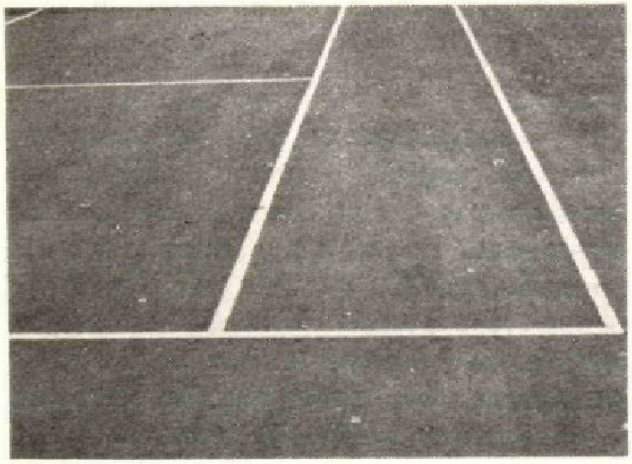
Herrnstein's Pidgeons



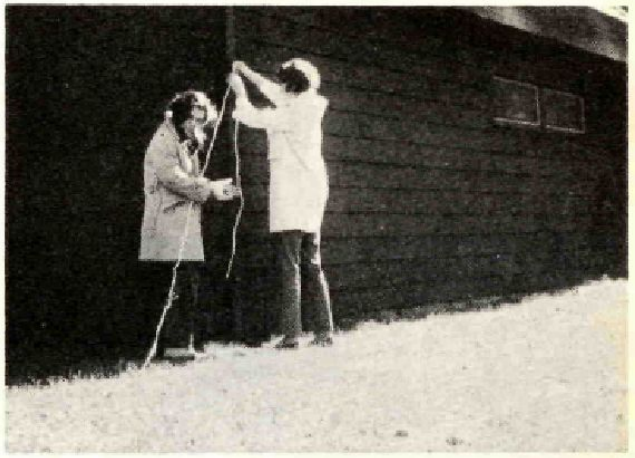
Herrnstein, 1976, Natural Concepts in Pidgeons

Pidgeons correctly classified trees...

... and water



... and specific persons



Reasoning

- Propositions
  - Conclude
  - From reasons
- Deductive reasoning
  - If premises are true, conclusion should be true
  - More rules,
    - longer
    - Likely to make mistakes
    - Specific conditions

a If it's raining, I'll take an umbrella.

b It's raining.

c Therefore, I'll take an umbrella.

a If it's raining, I'll take an umbrella.

b If I take an umbrella, I'll lose it.

c It's raining.

d Therefore, I'll lose my umbrella.

Reasoning

a No addictive things are inexpensive.

b Some cigarettes are inexpensive.

c Therefore, some addictive things are cigarettes.

a No addictive things are inexpensive.

b Some cigarettes are inexpensive.

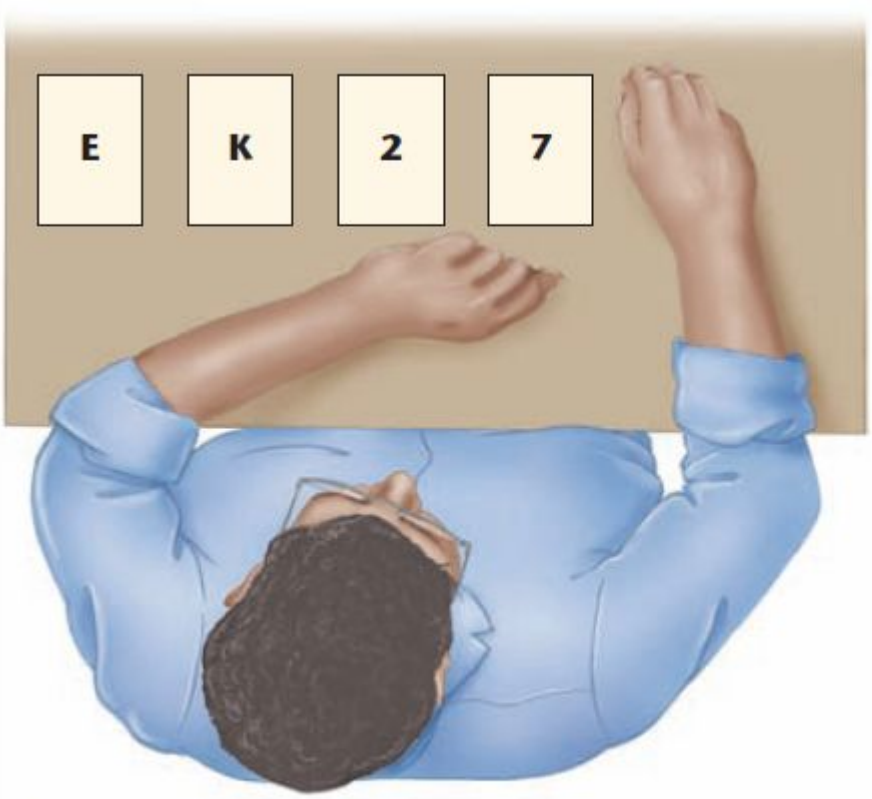
c Therefore, some cigarettes are not addictive.

Effects of content

- Evaluation depends on the content
- Wason selection task
  - Whether the following claim is correct:

    If a card has a vowel on one side, then it has an even number on the other side'

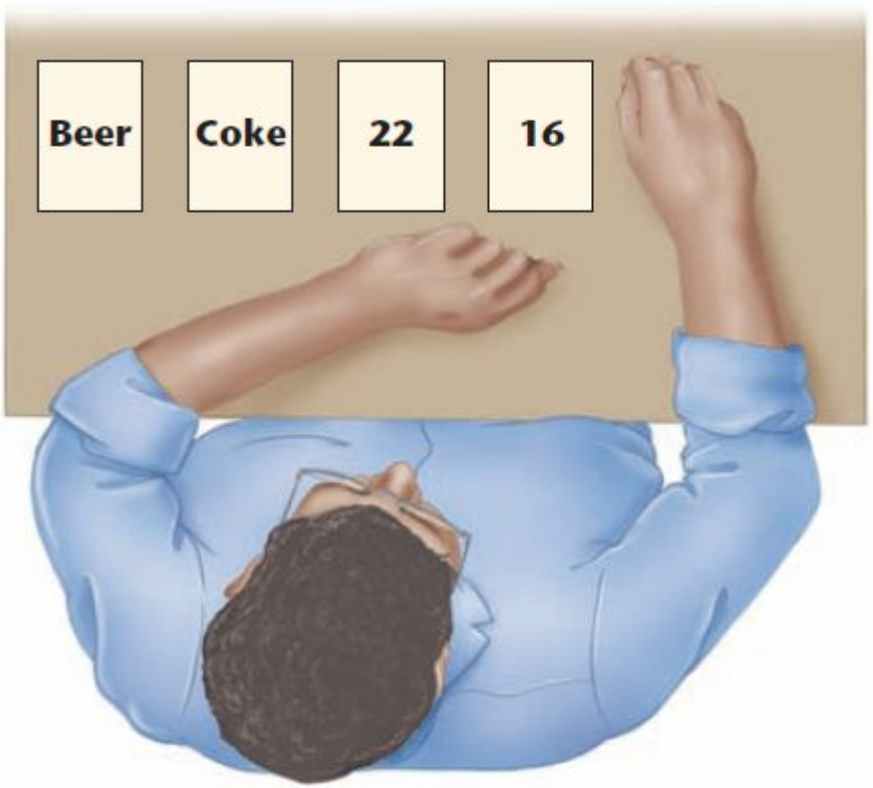
Wason selection task



a) Hypothesis: If a card has a vowel on one side, it has an even number on the other side.

Kartın bir tarafında sesli harf varsa diğer tarafında çift sayı vardır

Wason selection task



b) Hypothesis: If a person is drinking beer, he or she must be over 19.

- Pragmatic rules
  - Permission rule:

if a particular action is taken, often a

Inductive reasoning

- An argument can be good even if it is not deductively valid.
- Inductive strength:
  - Matter of probabilities

- Base-rate rule: membership
- Conjunction rule: combination of propositions

a Mitch majored in accounting in college.

b Mitch now works for an accounting firm.

c Therefore, Mitch is an accountant.

Heuristics

- A short-cut
  - Relatively easy to apply
  - Often correct answer
  - Not inevitably so!
- May violate basic probability rules
- Ignore info about base rates

In one experiment, one group of participants was told that a panel of psychologists had interviewed 100 people – 30 engineers and 70 lawyers – and written personality descriptions of them. These participants were then given a few descriptions and asked to indicate the probability that the person described was an engineer. Some descriptions were prototypical of an engineer (for example, 'Jack shows no interest in political issues and spends his free time on home carpentry'), and others were neutral (for example, 'Dick is a man of high ability and promises to be quite successful'). Not surprisingly, these participants rated the prototypical description as more likely to be that of an engineer.

Tversky & Kahneman, 1973

Heuristics

Linda is 31 years old, single, outspoken, and very bright. In college, she majored in philosophy . . . and was deeply concerned with issues of discrimination.

1. Linda is a bank teller.
2. Linda is a bank teller and is active in the feminist movement.

Conjunction rule.

The probability of a proposition cannot be less than the probability of that proposition combined with another proposition

Similarity heuristics?

Similarity heuristic

1. a All robins have sesamoid bones.
   b Therefore all sparrows have sesamoid bones.

versus

2. a All robins have sesamoid bones.
   b Therefore all ostriches have sesamoid bones.

Causality heuristic

1. Sometime during the year 2010, there will be a massive flood in California in which more than 1,000 people will drown.

2. Sometime during the year 2010, there will be an earthquake in California, causing a massive flood in which more than 1,000 people will drown.

Availability heuristic

estimated the frequency of words starting with the letter r (like rose) as higher than the frequency of words with the letter r in the third position (such as care).

Representativeness heuristic

Each case is representative of its category.

why subjects overestimate the number of fatalities caused by floods or murder (which get high press coverage, and are easily remembered), while they underestimate the number of fatalities caused by specific diseases (Slovic, Fischhoff, & Lichtenstein, 1982)

Confirmation bias. We give more credence to evidence that is in line with our previous beliefs than to evidence that contradicts it

Heuristics

Do not be too pessimistic about our level of rationality!