Questions

- How is [effortless recognition across viewing conditions] accomplished by the brain?
- What kind of information does the visual system derive from the retinal image to construct descriptions of sets of object features that capture the invariant properties of objects?
- How are such descriptions stored, and how are they activated by the viewed object?
- Are object representations general, or are they specific to an action or to a cognitive process, such as learning, planning, or reasoning?

Categorization: Basic level

Perceptual categorizations...reflect the redundant, correlational structure of the environment and occur most often at the level at which individual members of categories are most similar to each other and maximally different from members of other categories.

- **Basic level** (Rosch): The most inclusive category within which attributes are common to most category members. Across levels, basic level exhibits the greatest increase in
  - number of characteristic features
  - shape similarity
  - similarity in manner of interaction
- More rapid recognition/naming (but confounded w/ frequency)
- Experts: subordinate categories become "basic" (e.g., birdwatchers)

Central claim

A multipurpose general recognition system does not actually exist. Instead, in the process of biological recognition, multiple representations of an object are formed, each specific to the transformations required by either perception or action.

Issue: Relevance of different tasks
- Categorization
- Individuation
- Use/interaction
Categorization is not (entirely?) due to verbal labels

Three sources of evidence
- Adult learning of novel categories (with no labels)
- Categorization by (preverbal) infants
- Categorization by (nonverbal) animals

Adult learning of novel categories

Posner & Keele dot-based prototypes

Categorizations are not based upon a recognition threshold that, once exceeded, definitively endows a stimulus with class membership. Instead, a familiarity continuum exists, according to which the probability of a correct classification depends on the structural typicality of the stimulus, determined by the closeness to the class prototype. Atypical exemplars greatly differing from the prototype are recognized as individual entities rather than as class members—that is, they themselves become the entry point of recognition.

Categorization by infants

- Infants as young as three or four months old can form basic-level categorical representations based on visual and auditory stimuli.
  - horses / cats / zebras / giraffes
- Some types of subordinate recognition (e.g., faces) appear to begin extremely early in life
  - neonates can visually discriminate their mother’s face from that of a stranger
- Superordinate-level classifications of object pictures appear later and improve with age, usually in close relation to linguistic developments
- Posner & Keele dot patterns
  - few exemplars (6): exemplar learning, no extraction of prototype
  - many exemplars (12): extraction of prototype, no exemplar learning

Categorization by animals

- Monkeys...are obviously capable of making basic-level categorizations, but they can also easily learn to discriminate individual human or monkey faces...and novel artificial object classes, even generalizing learning across basic image transformations.
- The ability to discriminate between basic classes has been demonstrated in the goldfish...and in many different bird species.
- Pigeons can learn to discriminate (1) people vs. non-people; (2) trees vs. water vs. person; (3) oak leaves vs. other leaves (but not an individual oak leaf)
- Chickens can discriminate slides of one bird in a variety of poses from slides of other birds, and they can transfer this discrimination to novel sets of slides.
Exemplars vs. prototypes

**Question:** Do categorizations at different abstraction levels rely on the same type of stored representations?

The accuracy of classification of new instances depends on their similarity to old exemplars, but the effect is much greater for small categories (5 items) than for large ones (20 items), implying that for large categories, individual members are not specifically encoded.

**Dual-system account**
Store both exemplars and prototypes separately (e.g., Marsolek)

**Single-system account**
Superposition (e.g., McClelland & Rumelhart distributed memory model)

degree of viewpoint invariance in visual recognition

**Image transformations, even simple scaling and translation, can sometimes affect recognition. Moreover, invariance to some transformations, such as rotations in depth, appears to depend strongly on familiarity, as well as on the nature of the object and task.**

**Size and position**
- Performance is mostly insensitive to size/position changes (with small effects in same-different judgements and highly trained conditions)

**Rotation in picture-plane (2D)**
- Relative small costs of misorientation for simple, familiar stimuli (letters, digits, line drawings)
- Trained novel stimuli show sensitivity to specific trained orientations (also shown in monkeys)

**Inversion**
- **Face inversion effect:** much harder to recognize upside-down faces (compared to effects size for other objects)
- Effect also observed for other highly trained, homogenous categories (dogs, cars, birds, “Greebles”)
- Effect is absent in **monkeys** and mild in **young children**

**Action-based representations**

**Question:** Are similar representations used when the perceived object elicits a visually guided motor action?

When normal human subjects reach for an object, they move their fingers into a certain spatial configuration appropriate for grasping the object. If, however, the perceived size of an object is different from its actual size, as may occur with some form illusions, then a dissociation is observed between the perceived size and the size of the object represented in the systems mediating the grasping behavior.
Rotation in depth (3D)

- Object-centered vs. viewer-centered representations
- Canonical view(s): one or two views of common objects are judged as subjectively better (more informative?) than all others
- Little effect of depth rotation for common objects
- Trained novel objects are strongly viewpoint-dependent

Accounts
- Decomposition into viewpoint-invariant parts/features
- Normalization to object-centered "structural description" ("prototype")
- Storage of multiple viewpoint-specific representations ("exemplars")

Neuropsychology: Category-specific breakdowns

Dissociable clinical deficits in the visual processing of objects, for purposes of categorization, identification, and goal-directed action, strongly suggest a multiple systems architecture for recognition.

Face recognition: Prosopagnosia
- Relatively selective impairment in recognizing faces
- Also have problems distinguishing individuals within other objects classes, such as fruits, playing cards, housefronts, and automobiles...birds (for birdwatcher); cows (for farmer)
- Reduced/reversed face inversion effect

Facial expressions
- Amygdala lesions (but confounded with representations of emotions per se)

Visuo-motor representations: Balint’s syndrome, optic ataxia

- Impairments in visual guided reaching

Disturbance of [hand] preshaping can be dissociated from visuospatial perception, suggesting a double dissociation between the representations used for action and those that may underlie the perception of an object. Agnosic patient [DL] was unable to perceive the orientation or size of objects, although she could still accurately use orientation and size information for visuomotor actions.

Anatomical framework underlying visual processing in the primate brain

Two cortical visual systems (Ungerleider & Mishkin, 1982)
- dorsal ("what") to posterior parietal cortex (7a/areas 7/39)
- ventral ("where") to inferotemporal cortex (AIT/areas 20/21)
  - area TEO = PIT
  - area TE = CIT & AIT
**Electrophysiological studies**

- Single-cell response properties: changes in **topography, receptive field size, stimulus selectivity** in moving posterior to anterior
  - TEO: coarse topography, complete rep. of contralateral visual field
  - ant TE: no topography, large RFs including fovea

*IT appears to have all the machinery requisite for the formation of object descriptions. Cells respond selectively to stimulus attributes such as color and texture, to simple and complex patterns, and to complex natural objects such as faces. They also show a certain degree of translation and scale invariance.*

**Neural representation of visual objects**

- **Questions**
  - Are they represented explicitly by the firing of a few gnostic units? (so-called “grandmother cells”)
  - Are they represented by the firing of a small population of neurons, each encoding some features, aspects, or single views?
  - Are they represented by a large population of cells, each acting as a specialized pattern filter that combines certain shapes with different surface properties of objects, such as their texture, color, or lightness?

- Electrophysiological findings suggest the existence of at least **two possible neural mechanisms** for object representation.
  - One system may code the **prototypes** of objects that can be **decomposed into parts** and recognized by indexing these parts and their metric or spatial relationships.
  - A second, separate, system may be used when **holistic configuration** rather than individual features is important and may rely primarily on small populations of neurons with **strong configurational selectivity**.

**Conclusion**

*[I]*t is becoming increasingly clear that there are no universal principles of [object recognition] and that facts discovered about one form of [object recognition] need not hold for other forms. This is why systematic classification of [object recognition] systems, both psychological and physiological, is an essential prerequisite for the successful pursuit of the empirical and theoretical understanding of [object recognition] processes and mechanisms. The systems approach combined with appropriate processing theories seems to provide the most direct route to the future.

adapted from Tulving & Schacter (1990)