Perceiving Function and Category

- Visually perceive
- Functionality of objects that we see
- Categorization: Classify into known types
Agenda

- Part 1: Perception of function
- Part 2: Categorization: Various phenomena
- Part 3: Theories of Categorization
- Part 4: Recognizing letters and words

Perception of function

- Direct Perception
- Indirect Perception

Affordances

- Flat surface
- Horizontal
- Strong

Sittable

J. J. Gibson

Categorization

All others

Sittable
Direct Perception of Function

- Traditional approach was *categorization*
- Gestalt psychologists argued direct perception
- **Affordances**: properties that prompt user interaction
- J. J. Gibson (1979) claimed that:
  - Objects can be *grasped upon, sat upon*
  - No standard categories for such affordances

Affordances

- Two important considerations:
  1. **Functional form**: Function must follow from form
     - Round wheels: rolling
     - Triangular wheel?
  2. **Observer relativity**: Affordances perceived depends upon the observer.
Affordances (continued)

- Neisser (1989)
- Functional properties that conform to both conditions are called **physical affordances**
- These are *necessary*, but *not sufficient* for direct perception (Gibson)

Affordances (continued)

- Neisser suggested:
  - Affordances and categorization are fundamentally different *modes* of perception
  - Accomplished by *different* neural systems
  - Evident in patient with damaged ventral system
Affordances (shortcomings)

- Cannot account for all functional information that we perceive
  - Example: CDs
- Hence categorization approach is important
Categorization: Four Components

1. Object representation
2. Category representation
3. Comparison processes
4. Decision processes

PART 1
Perception Phenomena
- Direct Perception
- Indirect perception
  - Object representation
  - Category representation
  - Comparison processes
  - Decision processes

PART 2
Categorization

PART 3
Theories of Categorization

PART 4
Recognizing letter & word

Categorization: 1. Object representation

- Shape is the most important
  - Templates
  - Fourier spectra
  - Feature lists
  - Structural descriptions
- Other information
  - Texture
  - Color
  - Size
  - Orientation
Categorization: 2. Category representation

- Shape is the most important
  - Templates
  - Fourier spectra
  - Feature lists
  - Structural descriptions
- Other information
  - Texture
  - Color
  - Size
  - Orientation

Categorization: 3. Comparison processes

- Object and category *representations* should be of the same *type*
- Comparison: Serial or parallel?
  - Comparing *across categories*: parallel
    - Very large number of known categories
  - Comparing elements *within representation*: Not obvious.
Categorization:
4. Decision processes

- Which category does the object belong to?
- Should support:
  1. Novelty
  2. Uniqueness
     - For mutually exclusive classes
     - A thing cannot simultaneously be cat and dog
- Get fit value for each category
- Three approaches of decision making:
  - Threshold rule
  - Maximum (best-fit) rule
  - Maximum-over-threshold rule
  - Most appropriate

PART 1
- Perception Phenomena
  - Direct Perception
  - Indirect perception
    - Object representation
    - Category representation
    - Comparison processes
  - Decision processes

PART 2
- Categorization

PART 3
- Theories of Categorization

PART 4
- Recognizing letter & word

Categorization:
4. Decision processes (cont.)

Problem with threshold: Not unique
4. Decision processes (cont.)

Problem with best fit: No novelty

Maximum over threshold: Preserves uniqueness
4. Decision processes (cont.)

Maximum over threshold: Preserves novelty

---

Perceptual Categorization: Phenomena

1. Defining categories and their structure
2. Effects of perspective viewing conditions on categorization
3. Does *part structure* help in categorization
4. Contextual effects on categorization
5. Visual agnosia
Perceptual Categorization: Defining categories

- Categorical structure is largely hierarchical
  - Dog < Mammal < Animal < Living thing …
- Two ways of representing:
  1. Hierarchical trees
  2. Venn diagrams

Perceptual Categorization: Defining categories (cont.)

- Variations within each category
  - Not all dogs look alike, nor all birds, nor cars
  - What is the basis of categorizing objects in a category?
- Classical approach: Aristotle
  - Category was designated by a set of rules
    - **Necessary and sufficient conditions** for membership
  - Conditions: List of properties that object must have
  - Example: Triangle (closed polygon, three lines)
Perceptual Categorization: Aristotelian view

- Binary category membership
  - Either in category or not
- Is it good at explaining natural perceptual categories?
  - Ludwig Wittgenstein (1953) said “No”.
  - Name features common to all games
  - Family members resemblance, but no necessary or sufficient condition definition

Perceptual Categorization: Prototype

- Eleanor Rosch, UC Berkeley (1970s)
- All natural categories might be structured in a similar way in terms of a central or ideal example
- This is called Prototype
- Prototype is an average member
  - ‘Doggiest’ possible dog
Perceptual Categorization: Aristotelian vs Prototype

- Rule-based vs instance-based representation
- Binary versus graded membership
  - How *doggy* a dog is?
- Prototypes are used naturally
  - Chihuahua rated poorly as dogs than beagles

---

Perceptual Categorization: Levels of categories

- At which hierarchical level do we categorize an object?
  - Lassie < Dog < Mammal < Animal < … ?
- Most people identify object at an *intermediary* level
- Rosch defined it as *basic-level category*
- Superordinate categories: above basic
- Subordinate categories: below basic
Perceptual categorization: Basic-level categories

- Highest level category such that:
  1. Similar shape
  2. Similar motor interactions
     - Piano, guitar
  3. Common attributes

Perceptual Categorization: Entry-level categories

- Example:
  - Category: Bird
    - Robin, sparrow: identified as ‘birds’
    - Ostrich: identified as ‘ostrich’

- For some basic-level categories with *wide* variety:
  - *Typical* objects are classified at basic level
  - *Atypical* objects are classified at *subordinate* level

- Jolicoeur (1984) called them entry-level
Categorization: Perspective Viewing

- Perspective views influence speed and accuracy of recognition
- Some views of objects are easier to recognize than others

Experiment: Palmer, Rosch and Chase (1981)
- Subjects rated many views of the same object
  - Canonical Perspective
- Other subjects named entry-level category of many objects
  - Latency was noted
Categorization: Perspective Viewing

PART 1 √ Perception Phenomena

PART 2
- Categorization
  - Categories
- Perspective viewing
  - Canonical view
  - Priming effect
  - Orientation effect
- Part structure
- Context
- Visual agnosia

PART 3 Theories of Categorization

PART 4 Recognizing letter & word
### Perspective Viewing: Canonical view hypotheses

- **Hypotheses:**
  1. **Frequency hypothesis**
     - Frequently seen views are more canonical
     - But cups seen from above are not
  2. **Maximal information hypothesis**
     - Views that provide more information about shape and use of object are more canonical
     - Best views tend to show multiple sides
- Both hypotheses are true to some extent

### Perspective Viewing: Relation to Priming effect

- **Priming effect (Bartram, 1974)**
  - Two sets of images shown. Latency noted.
  - Categorizing is faster and more accurate if the object is presented a second time
  - Heightened state of readiness
  - Repetitions need not be exact replica
  - Different perspective view may be presented
- Irving Biederman used this to study the effect on perspective viewing on categorization
Perspective Viewing: Priming effect

- Modification in position, reflection or size does not affect priming effect
- Changes in perspective does
- However, if same parts are visible in different perspective, then no effect
Categorization: Orientation effect

- Rotation of object along line of sight
- Pierre Jolicoeur (1985)
  - Faster categorization of objects in their normal, upright orientation
- Orientation effects diminish with practice
  - People may store multiple representations of the same object at different orientations

Categorization: Effects of Part structure

- Biederman and Cooper (1991)
- 2 experiments
  - Based on priming effect
  - Used line drawings of objects
- Experiment 1
  - In first image, half contours were deleted
  - Compliment image had only those lines
- Experiment 2
  - First image, some *parts* deleted
  - Compliment image has only those *parts*
Part Structure:
Experiments

- Identity priming
  - Same set was used in two trials
  - Just for baseline
- Compliment priming
  - Complement sets were used in two trials
- Different exemplar priming
  - A totally different perspective view was used in the second trial

Part Structures:
Results of experiments
Categorization: Contextual effects

- Depends upon prior knowledge
- Depends on surroundings in the view

**TAE**  **CAT**

\[ B \text{ out of context} \]

\[ C \text{ in context} \]

- TNAE
- Eye ear
- Nose mouth

Categorization: Visual Agnosia phenomenon

- Brain damaged patients
- Inability to categorize previously known objects
- Apperceptive agnosia
  - Sensory processing damaged
- Associative agnosia
  - Perceptual part intact, but association lost
- Prosopagnosia
  - Cannot recognize faces visually
Visual agnosia

Recap

- Perception of function
  1. Direct (affordances)
  2. Indirect (categorization)
- Categorization
  - Categories (basic-level, entry-level)
  - Effects of perspective on categorization
  - Effects of part structure
  - Effects of surrounding context
  - Visual agnosia is related to categorization
Theories for Object Categorization

- How objects might be perceived in the visual human system
- Most Prominent:
  - Recognition by Components (RBC) Theory
  - Irving Biederman (1985, 1987)
  - Also called Geon theory

Recognition By Components Theory

- Objects can be specified as spatial arrangements of primitive volumetric components called geons.
- Geons
  - geometric ions
  - A set of generalized cylinders which are easily distinguishable from each other.
  - Letters: Words :: Geons: Objects
PART 1: Perception Phenomena

PART 2: Categorization

PART 3: Theory of categorization
- Recog by components
- Explaining phenomena
- Other theories

PART 4: Identifying letter & words

Geons

Cross-sectional curvature
- Straight
- Curved

Symmetry
- Reflectional
- Rotational
- None

Axis curvature
- Straight
- Curved

Aspect ratio
- Equal
- Axis greater
- Cross section greater

Cross-sectional Size variation
- Expanding
- Expanding & Contracting
- Fixed

# of geons (qualitatively) = 2 * 3 * 2 * 3 = 36
# of geons (quantitatively) = 36 * 3 = 108
Nonaccidental features

- Properties to identify geons.
- Not dependent on ‘accidents’ of viewpoint.

**Figure 9.3.3** Nonaccidental properties of two geons. A brick and a cylinder can be distinguished by many properties that are present from all but a few specific viewpoints.

Geon relations

- Spatial relation of geons:: Order of alphabets in words
  - e.g. SIDE-CONNECTED, LARGER-THAN
- 108 different relations
- # of 2 geon objects > 1,000,000
A Neural Network implementation

[Image of a neural network diagram]

[Source: Hummel and Biederman, 1992]

Explaining empirical Phenomena

- Prototypes /typicality
- Basic level/entry level categories
- Perceptual viewing conditions
- Part structure
- Contextual effects
- Visual Agnosia

[Diagram with labeled parts: Features, Geon attributes, Relations, Assembly, Edges, Objects]

[Source: Hummel and Biederman, 1992]
Prototypes/typicality

- Categorization is function of geon matching
  - ‘Rough’ i.e. qualitative descriptions
  - Subordinate category (fine grained changes) can not be explained.

Basic level/entry level categories

- Typical members closely match the geons for basic level descriptions
- Atypical members do not. Hence not normally classified in Basic level.
- But, how are they classified?
  - Not clear
**Viewing conditions**

- Canonical perspective
  - Some geons get occluded
- Priming effect
  - Not studied

**Part structure**

- Component geons trigger activation, not lines and edges.
- Half the lines are enough for geon activation.
Contextual effects

- Cannot be explained as RBC looks only at parts of the object.
- But the idea can be extended to ‘scenes’ whose components are ‘objects’.

Visual Agnosia

- All views are ‘unusual’ to patients.
- Not much known!
Weaknesses

- Lack of representation power
  - 108 cylinders, 108 relations
- Finer discrimination required
  - Dog Vs Cat
  - Face recognition
- Implementation?

Viewpoint specific theories

- Multiple Views are required
  - 1 view cannot capture the 3D model
  - Multiple ‘good’ views with low latency
- Hence:
  - Aspect Graphs
  - Alignment with 3-D models
  - Alignment with 2-D view combinations
Aspect graphs

- Many views of the same object are actually very similar
- **Aspect:** *Multiple views* matched to a *common abstract representation*

![Diagram of three views](image1)

**PART 1**
Perception Phenomena

**PART 2**
Categorization

**PART 3**
- Theory of categorization
- Recog by components
- Explaining phenomena
- Other theories

**PART 4**
Identifying letter & words

Aspect graphs

- **Aspect:** *Multiple views* matched to a *common abstract representation*

![Diagram of three top views](image2)

1 common representation i.e. *aspect*
Aspect Graphs

- Different aspects connected if continuous change takes viewer from one aspect to the other.

Aspect Graphs: Issues

- Scalability
- Innate 3-D ability of the brain
- Discrimination (capture topology, not geometry)
Alignment with 3-D Models

- 3-D Models stored in brain. Mapped to 2-D images.
- Still top-down

1. Image
2. Edge features
3. Matching with 3-D model
4. Object recognition

Alignment with 2-D View combinations

- Use a ‘few’ 2-D views in brain rather than 3-D model.
- A method that can derive new 2-D views from a few stored ones.
Weaknesses

- View point theories don’t explain:
  - Innate 3-D ability
  - Novel objects
  - Non rigid objects
  - Part structure
  - Exemplar variation

Identifying Letters and words

- Perceiving letters as well as understanding words.
- Easier than object categorization:
  - Two-Dimensionality
  - Combinatorial structure
- Study:
  - Identifying Letters
  - Identifying within words
  - Interactive activation model
Identifying Letters

- Can be identified using:
  - Templates
  - Structural descriptions
  - Features

[McCelland & Rumelhart, 1981]

Identifying Letters

- Fuzzy Logic Model of Perception (FLMP)

[Massaro & Hary, 1986]
Identifying letters within words

- Letters are **not** detected independently of words.

  TAE CAT

- HOW MANY LETTERS CAN YOU REPORT NOW? 

Effects

- Word superiority effect
- Word-nonword effect
- Word-letter effect

**Word superiority effect**

- 83%

**Word-nonword effect**

- 70%

**Word-letter effect**

- 70%
Interactive Activation Model

- Proposes a multilayer neural network like model.
  - Feature level
  - Letter level
  - Word Level

PART 1  Perceptual Phenomena
PART 2  Categorization
PART 3  Theory of categorization
PART 4  Identifying letter & words
          Identifying letters
          Identifying letters in words
          Interactive activation model
          Alternative explanations

![Interactive Activation Model Diagram]

Figure 9.10: The architecture of the interactive activation model. Each level of the network represents a group of neurons corresponding to features, letters, and words, respectively. Levels are interconnected, allowing for the activation of one level to influence the next. (From McClelland & Rumelhart, 1981; From Rumelhart & McClelland, 1986.)
Alternative Explanations

- Word shape
- Serial Letter recognition
- Parallel recognition
  - Moving window effect
  - Boundary effect

Word Shape

- We recognize a word is the pattern of ascending, descending, and neutral characters [James Cattell, 1886]

<table>
<thead>
<tr>
<th>test</th>
<th>Error rates</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>testf</td>
<td>13%</td>
<td>Consistent word shape</td>
</tr>
<tr>
<td>tesc</td>
<td>7%</td>
<td>Inconsistent word shape</td>
</tr>
</tbody>
</table>
Word Shape Vs. Letter Shape

- Letter shape more important than Word shape [Monk & Hulme, 1983]

<table>
<thead>
<tr>
<th></th>
<th>Same Word shape</th>
<th>Different Word shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>than</td>
<td>nan</td>
</tr>
<tr>
<td>letter shape</td>
<td>15% missed</td>
<td>19% missed</td>
</tr>
<tr>
<td>Different</td>
<td>jan</td>
<td>ian</td>
</tr>
<tr>
<td>letter shape</td>
<td>8% missed</td>
<td>10% missed</td>
</tr>
</tbody>
</table>

PART 1 Perception Phenomena
PART 2 Categorization
PART 3 Theory of categorization
PART 4
- Identifying letter & words
- Identifying letters
- Identifying letters in words
- Interactive activation model
- Alternative explanations

Serial letter recognition

- Analogy to dictionary [Gough, 1972].
  - Start with 1st letter, then 2nd and so on.
  - Search for a letter in random strings.
    - 3rd letter 30 ms, 4th letter 40 ms.
  - Bigger words take longer to recognize.
- Effects
  - NOTUH
  - SORTH
- Cannot explain word superiority effect

PART 1 Perception Phenomena
PART 2 Categorization
PART 3 Theory of categorization
PART 4
- Identifying letter & words
- Identifying letters
- Identifying letters in words
- Interactive activation model
- Alternative explanations
Parallel letter recognition

- Moving window effect [McConkie & Rayner 1975]

- Fixate on words (200-300ms), then saccadic movement (20-35ms).
  - Fovea (3 or 4 letters)
  - Neighboring (8 or 9 letters)
  - Parafovea (15 to 20 characters)

---

Moving window effect

<table>
<thead>
<tr>
<th>Window size</th>
<th>Sentence</th>
<th>Reading rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 letters</td>
<td>An experimxxx xxx xxxxxxxxx xx</td>
<td>207 wpm</td>
</tr>
<tr>
<td>9 letters</td>
<td>An experiment wax xxxxxxxxx xx</td>
<td>308 wpm</td>
</tr>
<tr>
<td>15 letters</td>
<td>An experiment was conxxxxxx xx</td>
<td>340 wpm</td>
</tr>
</tbody>
</table>
Invisible boundary effect

<table>
<thead>
<tr>
<th>Word shown</th>
<th>Properties</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>chart</td>
<td>Identical word (control)</td>
<td>210ms</td>
</tr>
<tr>
<td>chovt</td>
<td>Similar word shape</td>
<td>240ms</td>
</tr>
<tr>
<td></td>
<td>Some letters in common</td>
<td></td>
</tr>
<tr>
<td>chyft</td>
<td>Dissimilar word shape</td>
<td>280ms</td>
</tr>
<tr>
<td></td>
<td>Some letters in common</td>
<td></td>
</tr>
<tr>
<td>ebovf</td>
<td>Similar word shape</td>
<td>300ms</td>
</tr>
<tr>
<td></td>
<td>No letters in common</td>
<td></td>
</tr>
</tbody>
</table>

[Rayner, 1975]

Summary

- Part 1: Perception of function phenomena
  - Direct Vs. Indirect
- Part 2: Categorization phenomena
  - Parts, categories, viewpoints, agnosia
- Part 3: Theory of categorization
  - Recognition By Components
- Part 4: Recognizing letters and words
  - Interactive Activation model