Navigation and Taking Action

Wu Xihong

Peking University
School of Artificial Intelligence

Introduction

- We have discussed some of these things
  - perceiving a scene and individual objects within it,
  - scanning the scene to shift attention from one place to another;
- This lecture considers the processes involved in being *physically active* and interacting with objects within a scene.
- In other words, we are taking perception out into the world.

Some Questions we will consider

- What is the connection between perceiving and moving through the environment?

- How people navigate through the environment?
- How people take action on objects, such as reaching and grasping?
- What happen when observing other people’s actions?

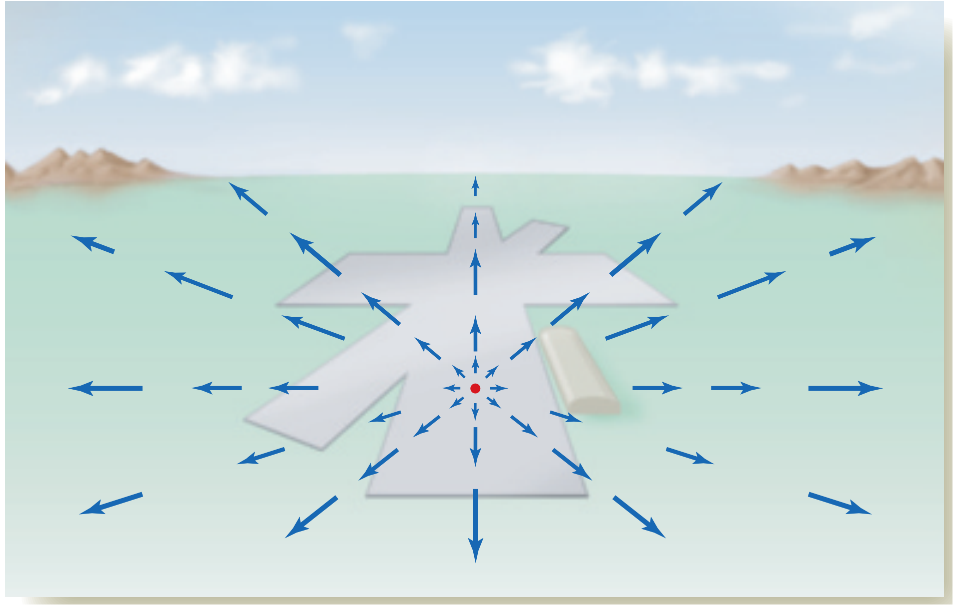
Part 1: Navigation

1 The Ecological Approach to Perception

- From the 1950s until the 1980s, the dominant way perception research was carried out was by having stationary observers look at stimuli in a laboratory situation.
- During World War II, J. J. Gibson studied the kind of perceptual information that airplane pilots use when coming in for a landing.
  - Gibson proposed that the perceived movement of the terrain provides information that helps the pilot guide the plane in for a landing.
  - This focus on observers moving through the environment was the starting point for the ecological approach to perception.

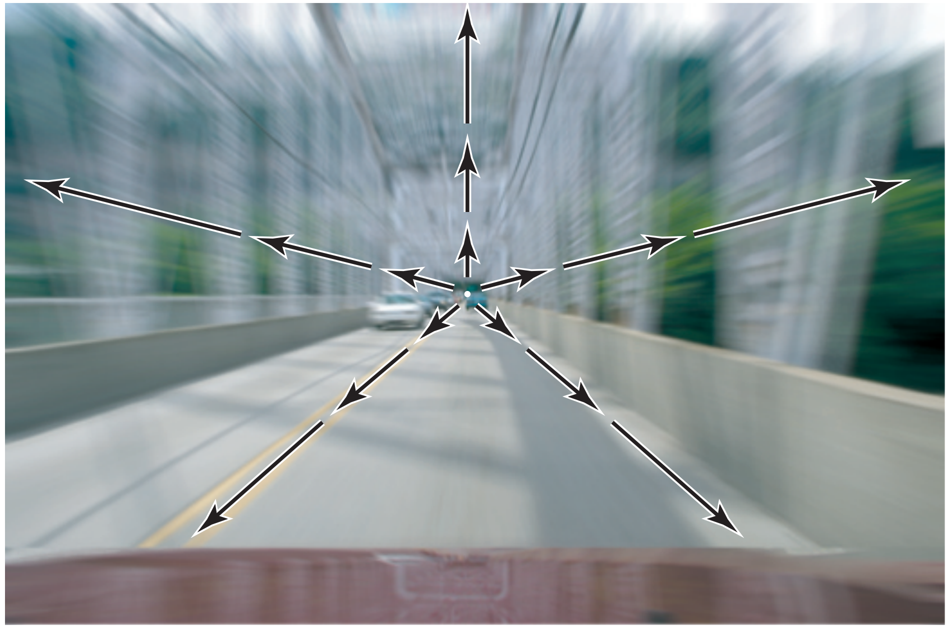
1.1 The Moving Observer Creates Information in the Environment

- Imagine that you are driving airplane for landing.
- All of the movement you are seeing is called optic flow.
- Optic flow has two characteristics:
  - The different speed of flow is called the gradient of flow. According to Gibson, the gradient of flow provides information about how fast the observer is moving.
  - The absence of flow at the destination point is called the focus of expansion (FOE). The FOE indicates the place where the plane will touch down on the runway.



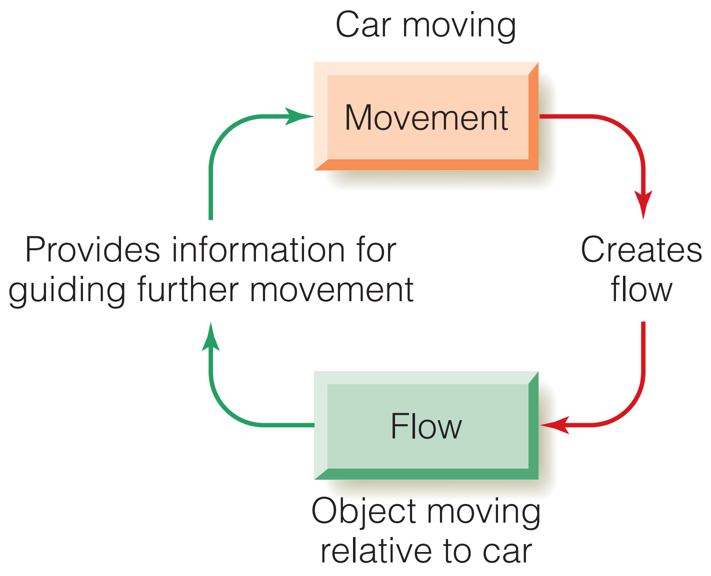
Invariant Information

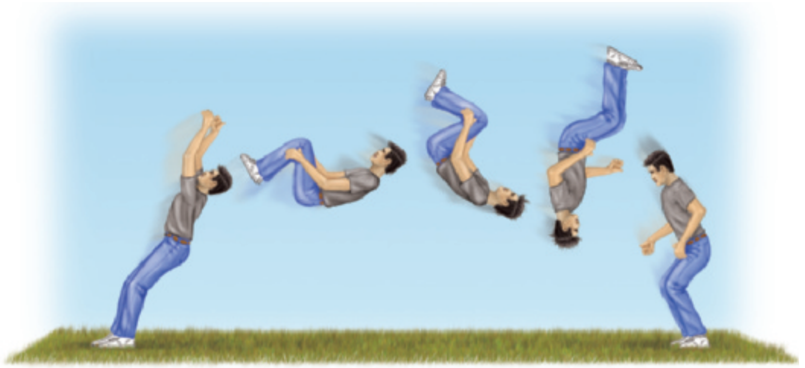
- The concept of the ecological approach is the idea of invariant information: information that remains constant even when the observer is moving.
    - How fast is moving: Optic flow provides invariant information because flow information is present as long as the observer is moving through the environment.
    - Where is heading: The FOE is another invariant property because it always occurs at the point toward which the observer is moving.



1.2 Self-Produced Information

- Another idea of the ecological approach is self-produced information: When a person makes a movement, that movement creates information, and this information is, in turn, used to guide further movement.
  - When a person is driving down the street, the observer uses the flow information to steer the car in the suitable speed and right direction.
  - Another example of movement that uses self-produced information is provided by somersaulting. Expert gymnasts performed somersaults better with their eyes open by making in-the-air corrections to their trajectory.





1.3 The Senses Do Not Work in Isolation

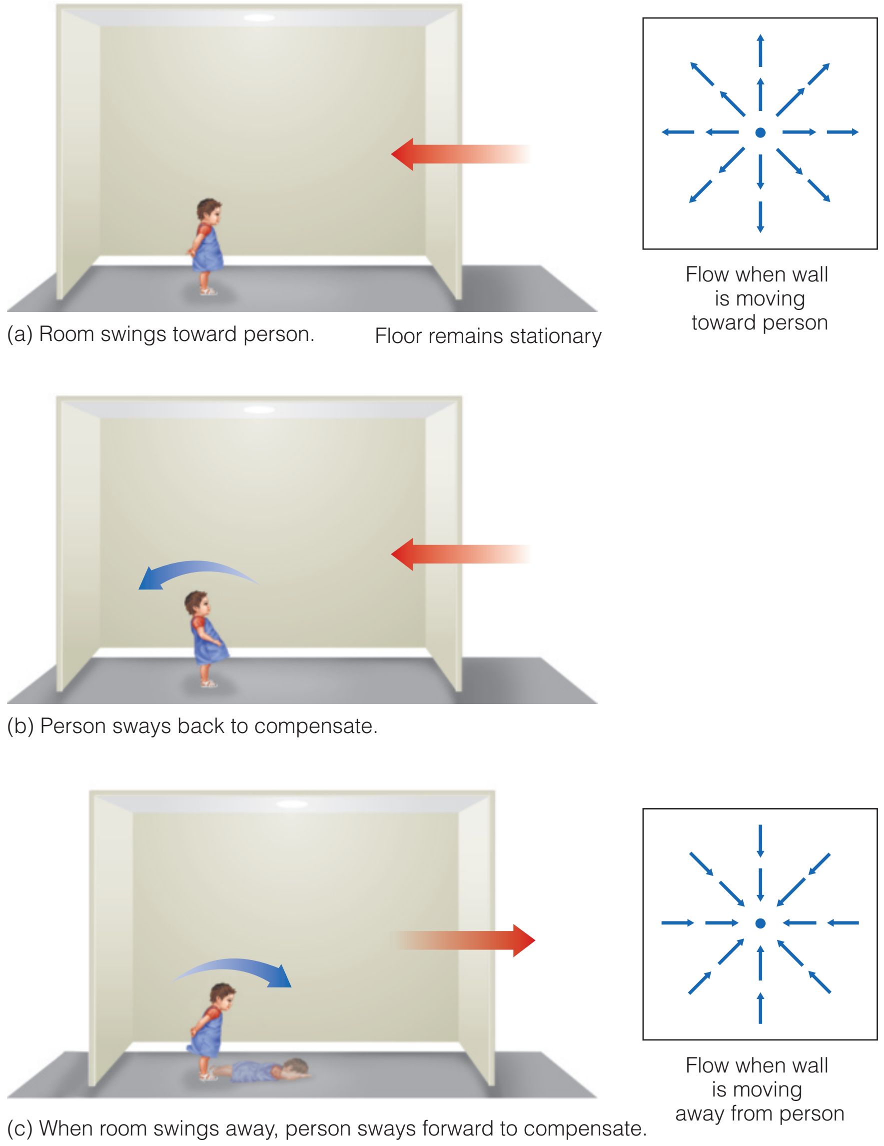
- Gibson also proposed that the senses do not work in isolation. Vision, hearing, touch, smell, and taste as separated senses, should be considered to provide information for the same behaviors.
- Your ability to stand up straight and to keep your balance while standing still or walking depends on systems that enable you to sense the position of your body.
  - These systems include the vestibular canals of your inner ear and receptors in the joints and muscles.
  - However, the information provided by vision also plays a role in keeping our balance.

DEMONSTRATION

Keeping Your Balance

Keeping your balance is something you probably take for granted. Stand up. Raise one foot from the ground and stay balanced on the other. Then close your eyes and notice what happens.

- Vision provides a frame of reference that helps the muscles constantly make adjustments to help maintain balance.
  - The swinging room experiments: Oscillating the experimental room through as little as 6 mm caused adult subjects to sway approximately in phase with this movement.
- Show that vision is such a powerful determinant of balance that it can **override** the traditional sources of balance information provided by the inner ear and the receptors in the muscles and joints.

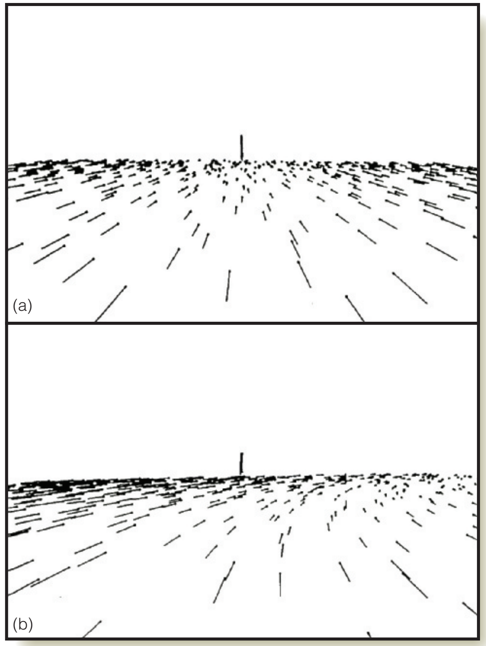


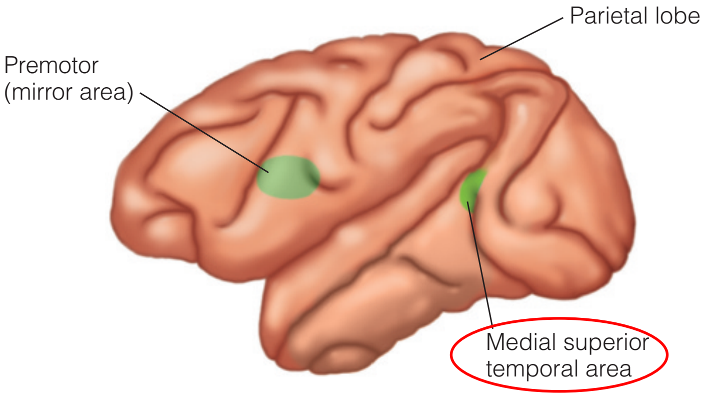
2 Navigating Through the Environment

- Gibson proposed that optic flow provides information about where a moving observer is heading.
- But can observers actually use this information?
- We consider this question next and then consider sources of information in addition to optic flow that help people navigate through the environment.

2.1 Do Observers Use Optic Flow Information?

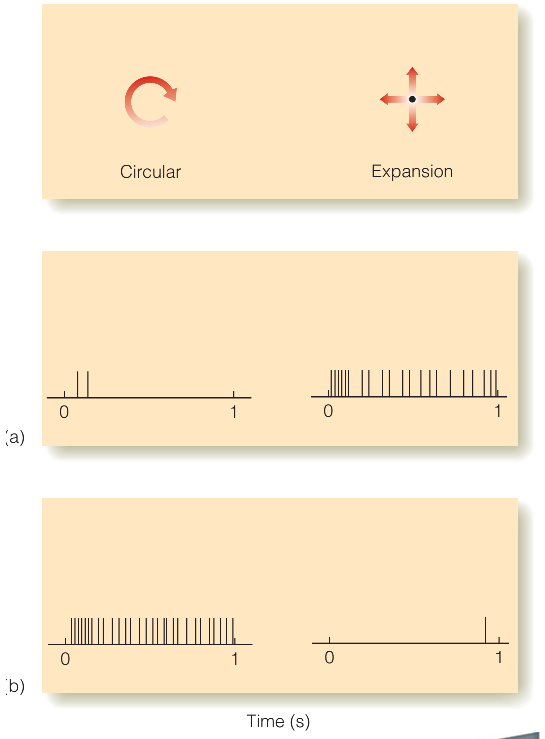
- Psychophysical results support Gibson's idea that optic flow provides information about where a person is heading.
  - The observer's task is to judge, based on optic flow stimuli, where he or she would be heading relative to a reference point such as the vertical line.
  - Observers viewing stimuli can judge where they are heading relative to the vertical line to within about 0.5 to 1 degree.
- Researchers have also identified neurons in the medial superior temporal area (MST) that respond to flow patterns, which is also important for perceiving movement.

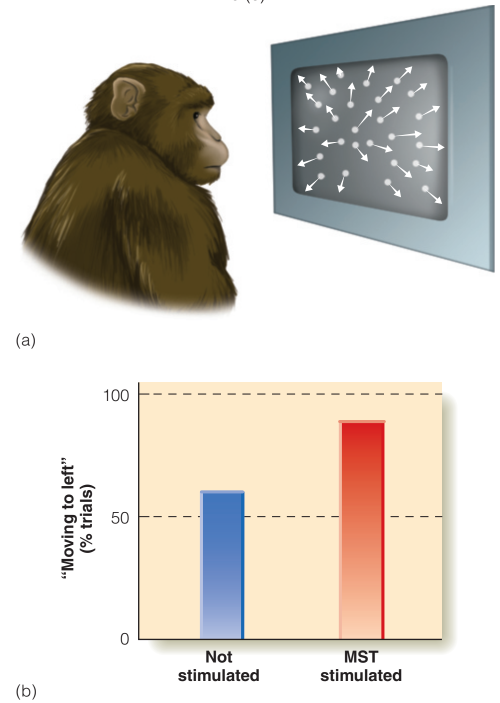




颞上内侧区

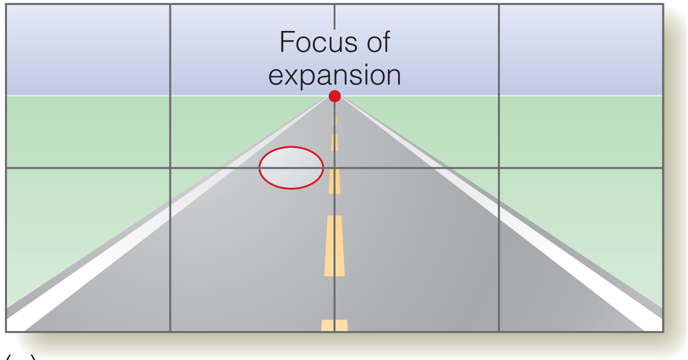
- Finding a neuron that responds to a flow stimulus:
  - The response of a neuron in a monkey’s MST that responds best as a pattern of dots are expanding outward (a)
  - Another neuron that responds best to circular motions (b).
- Demonstrate the connection between the neuron’s response and behavior:
  - First training monkeys to indicate whether the flow of dots on a computer screen indicated movement to the left or right of straight ahead.
  - The left bar (b) shows that the monkey responded to a stimulus by judging the movement as being to the left.
  - The right bar (b) shows that electrically stimulated MST neurons that were tuned to respond to flow associated with movement to the left, the monkey’s judgment was shifted even more to the left.
- Supports that flow neurons can, in fact, help determine the direction of perceived movement.



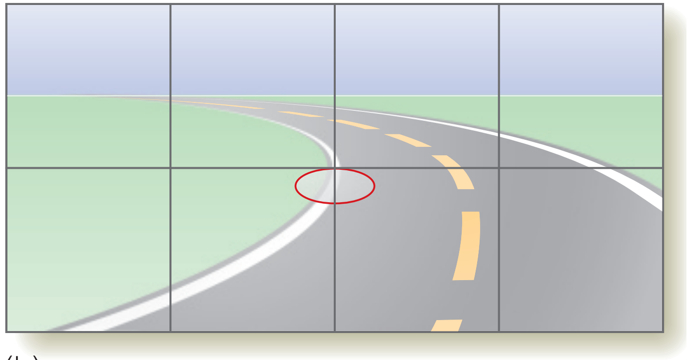


2.2 Driving a Car

- What about the flow that occurs in an actual environmental situation such as driving?
  - Although drivers look straight ahead while driving, they tend to look at a spot in front of the car rather than looking directly at the FOE (a).
  - When going around a curve, drivers don’t look directly at the road, but instead look at the tangent point of the curve on the side of the road (b).



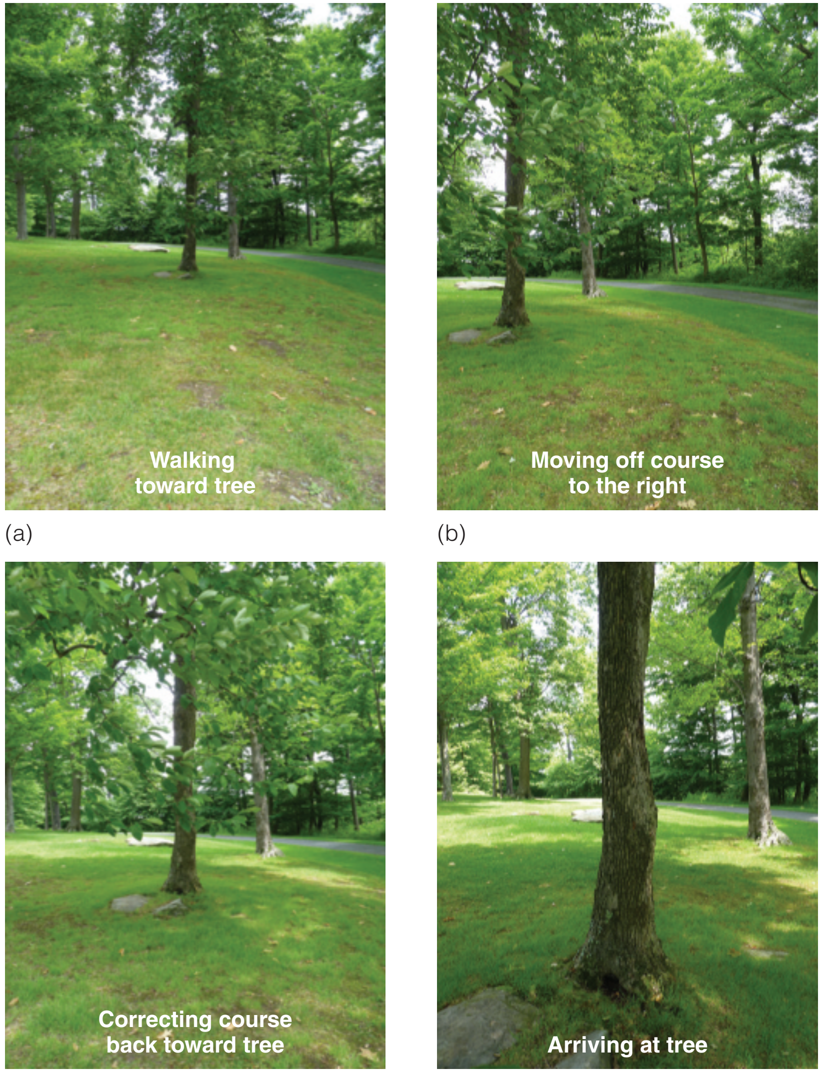
(a)


(b)

- Because drivers don’t look at the FOE, which would be in the road directly ahead, suggested that drivers probably use information in addition to optic flow to determine the direction they are heading.
  - This additional information would be noting <u>the position of the car relative to the lines in the center of the road or relative to the side of the road</u>.

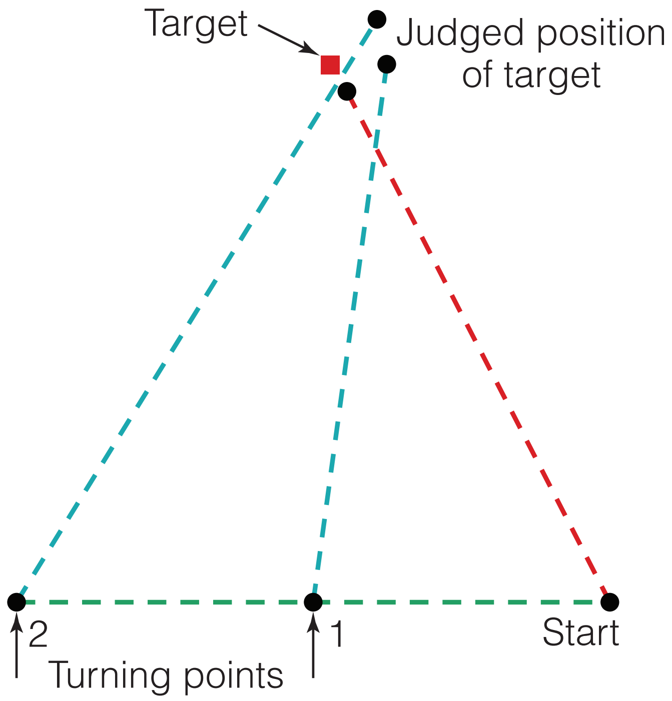
2.3 Walking

- How do people navigate on foot?
- An important strategy used by walkers that does not involve optic flow is the visual direction strategy, in which observers keep their body pointed toward a target.
- If they go off course, the target will drift to the left or right. When this happens, the walker can correct course by recentering the target.



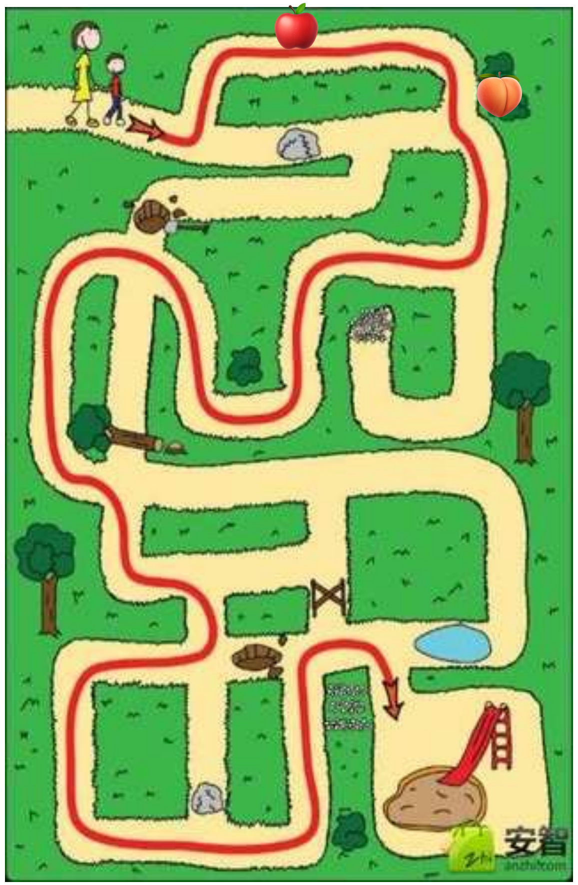
(a) (b) (c) (d)

- Another indication that flow information is not always necessary for navigation is that we can find our way even when flow information is minimal, such as at night or in a snowstorm.
- "Blind walking" procedure: observe a target object located up to 12 meters away, then walk to the target with the eyes closed.
  - These experiments show that people are able to walk directly toward the target and stop within a fraction of a meter of it (red lines).
  - Some records from these "angled" walks are shown by the blue lines, which depict the paths taken when a person first walked to the left from the "start" position and then was told to turn either at turning point 1 or 2 and walk to a target that was 6 meters away.
- The fact that the person generally stopped close to the target shows that <u>we are able to accurately navigate short distances in the absence of any visual stimulation at all</u>.

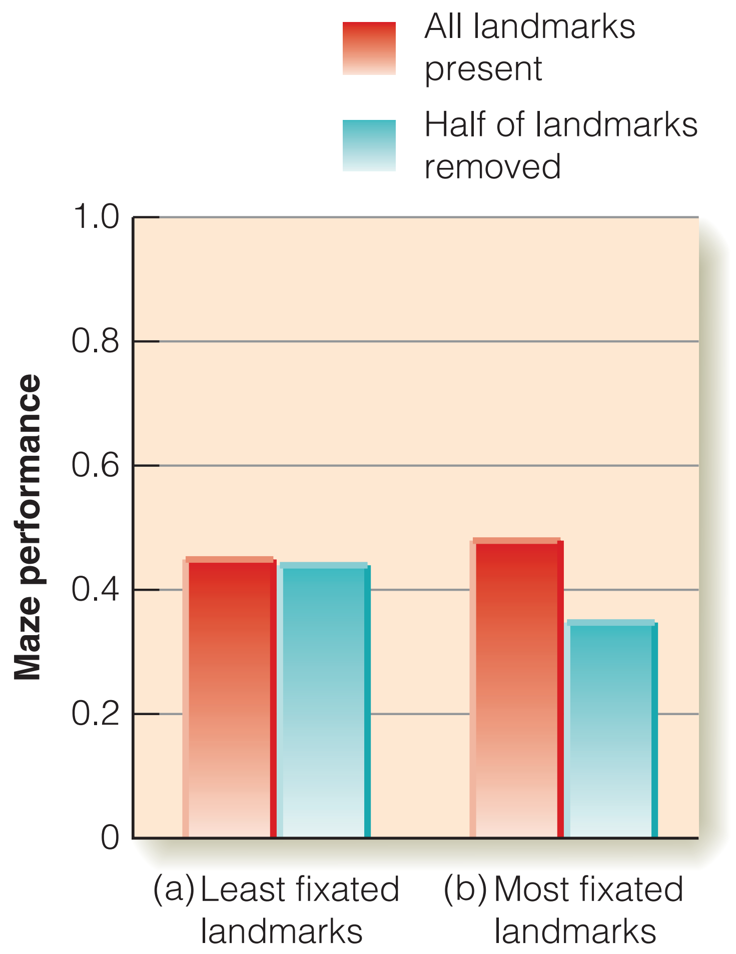


2.4 Wayfinding

- But we often travel to destinations we can’t see from the starting point. This kind of navigation is called wayfinding.
- **Landmarks**: objects on the route that serve as cues to indicate where to turn.
  - *Decision-point landmarks:* objects at corners where the subject had to decide which direction to turn
  - *Non-decision-point landmarks:* objects located in the middle of corridors that provided no information about how to navigate.
- Subjects used landmarks as they learned to navigate through a mazelike environment.
  - Training phase: Subjects first navigated through the maze until they learned its layout
  - Testing phase: told to travel from one location in the maze to another.
  - Subjects’ eye movements were measured using a head-mounted eye tracker.

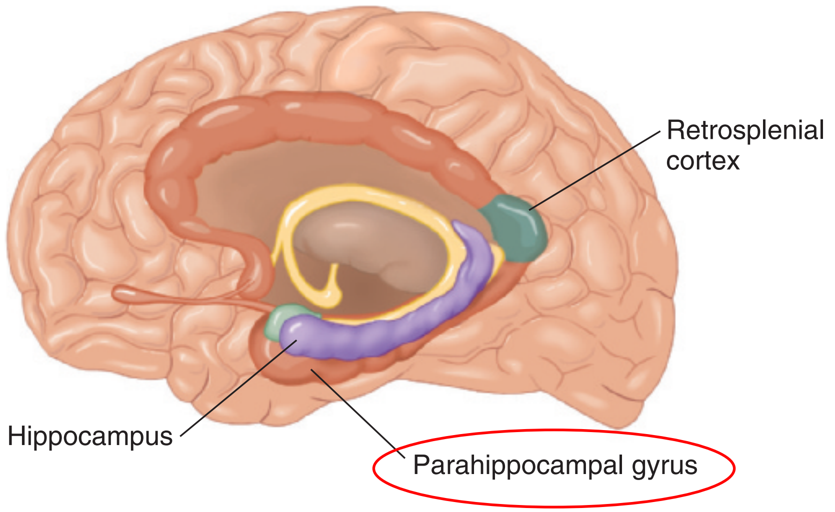


- Showed that subjects spent more time looking at decision-point landmarks than at non-decision-point landmarks.
- When maze performance was tested with half of the landmarks removed,
  - Removing landmarks that had been viewed less had little effect on performance (**a**).
  - Removing landmarks that observers had looked at longer caused a substantial drop in performance (**b**).



- It makes sense that landmarks that are looked at the most would be the ones that are used to guide navigation.
- Subjects were more likely to recognize pictures of landmark that were located at decision points than those located in the middle of the block.

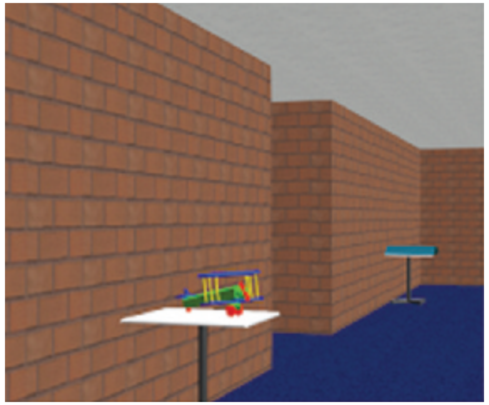
- The studies have measured eye movements, maze performance, and recognition, all of which are behaviors related to landmarks.
- But what is happening in the brain?
  - When subjects in a campus navigation study were shown pictures of buildings, the brain response in areas of the brain known to be associated with navigation, such as *the parahippocampal gyrus*, was larger than the response to non-decision-point buildings.

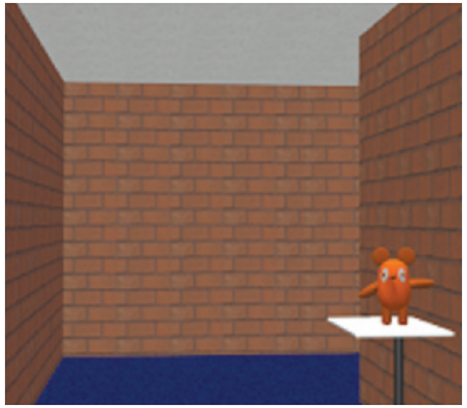


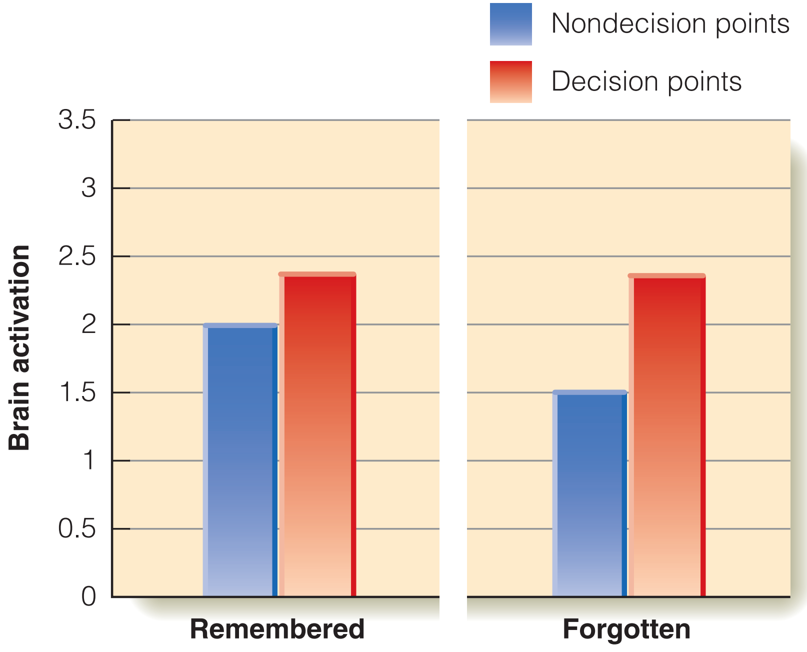
海马旁回

- Thus, decision-point landmarks are not only more likely to be recognized than non-decision-point landmarks, but they generate greater levels of brain activity.

- Observers first study a film sequence that moved through a "virtual museum", then were given a recognition test while in an fMRI scanner.
    - Activity in the right parahippocampal gyrus for remembered and forgotten.
- Show that the brain automatically distinguishes objects that are used as landmarks to guide navigation.
    - The brain responds not just to the object but also to how relevant that object is for guiding navigation.
    - Activity may automatically be "highlighting" landmarks even in cases when you may not remember having seen these landmarks before.
- Identifying landmarks is just one of the abilities needed to find one's way.
- Before you begin a trip, you need to know which direction to go, and you probably also have a mental "map" of your route.

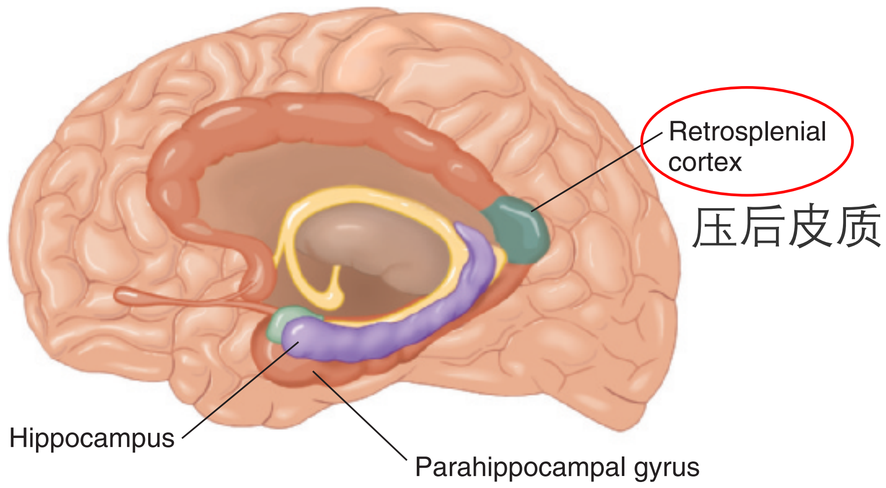

(a) Toy at decision point

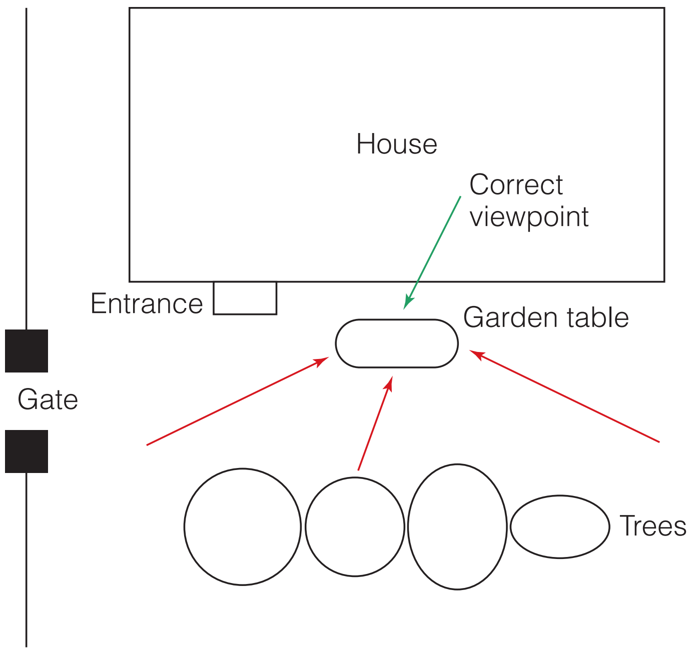

(b) Toy at nondecision point



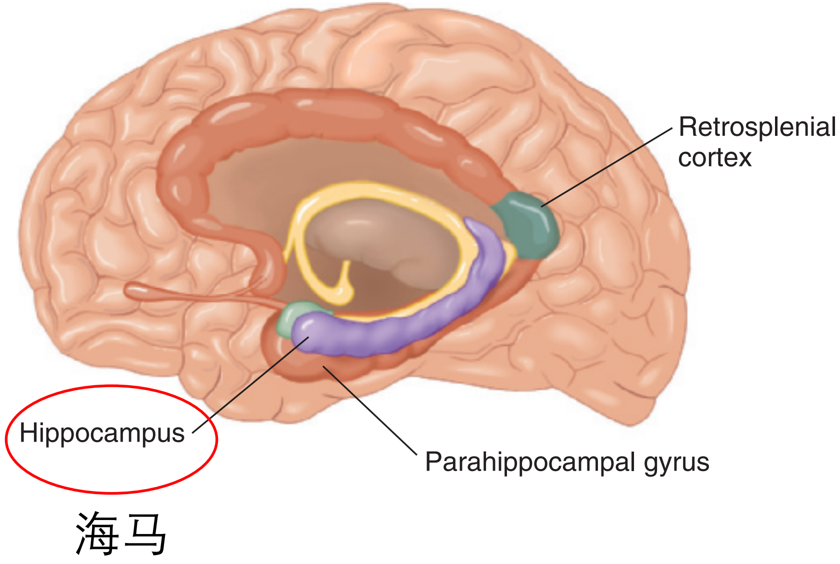
(c)

- **The Effect of Brain Damage on Wayfinding**
- **Retrosplenial cortex damage**
  - Could identify common objects and remember the positions of objects in a room.
- But
  - Lost his *directional ability*: couldn't determine the direction of any familiar destination or landmark with respect to his current position.
  - Unable to determine the viewpoints from which photographs of familiar places were taken.

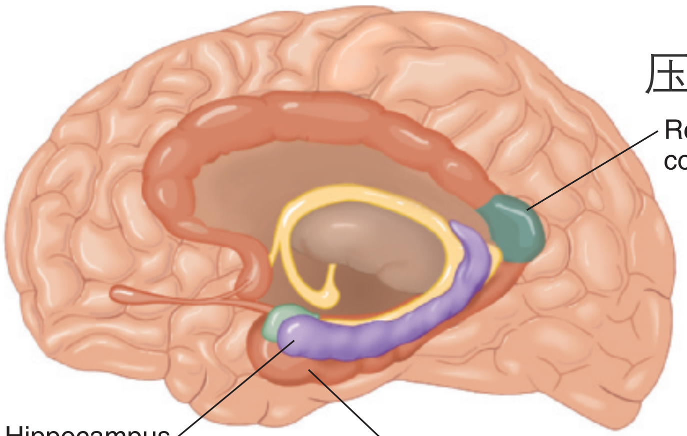




- **The Effect of Brain Damage on Wayfinding**
- **Hippocampus Damage**:
  - unable to find his way around his own neighborhood.
- Playing an interactive computer game called "The Getaway":
  - able to control taxi drivers, but only if the route involved just main roads.
  - Lost to navigate along side streets, even though he had been familiar with before.
- Show that the hippocampus is important for accessing details of routes that were learned long ago.



- **The Effect of Brain Damage on Wayfinding**
- The research on how the brain is involved in wayfinding has focused on three structures:
  - involves seeing and recognizing objects along a route (perception)
  - paying attention to specific objects (attention)
  - using information stored from past trips through the environment (memory),
  - and combining all this information to create maps that help us relate what we are perceiving to where we are now and where we need to go next.



Part 2: Taking Action

1 Acting on Objects

- So far, we have been describing how we move around in the environment. But our actions go beyond walking or driving. One of the major actions we take is reaching to pick something up.
- One of the characteristics of reaching and grasping is that it is usually directed toward a specific object, to accomplish a specific goal.
  - We reach for and grasp a doorknob to open a door.
  - We reach for a hammer to pound a nail.
- An important concept related to reaching and grasping is ***affordances***, which we describe next.

1.1 Affordances: What Objects Are Used For

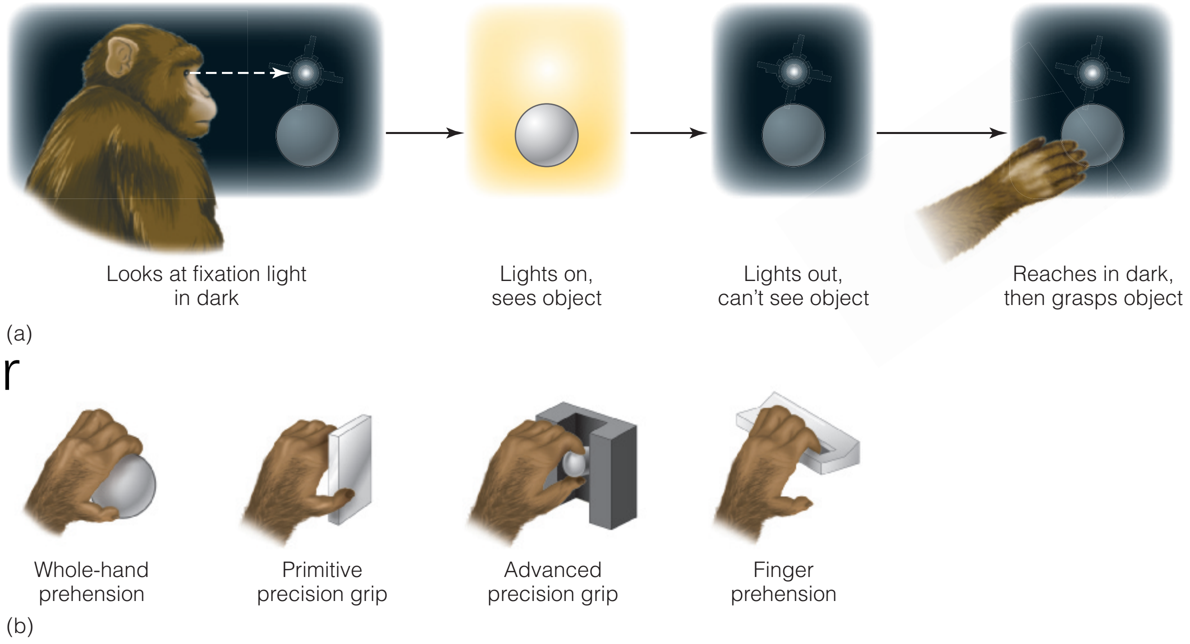
- Affordances:
  - In Gibson's (1979) words, "The affordances of the environment are what it *offers* the animal, what it *provides for* or *furnishes.*"
  - A chair, or anything that is sit-on-able, affords sitting; an object of the right size and shape to be grabbed by a person's hand affords grasping.
- Perception of an object not only includes physical properties, such as shape, size, color, and orientation, that might enable us to recognize the object; our perception also includes information about how the object is used, to guide our actions toward it.
- Temporal lobe damage: impaired the ability to name objects.
  - Given a cue, either (1) the name of an object ("cup") or (2) an indication of the object's function ("an item you could drink from").
  - He was then shown 10 different objects and was told to find the object.
  - The results showed that he identified the object more accurately and rapidly when given (2).

1.2 The Physiology of Reaching and Grasping

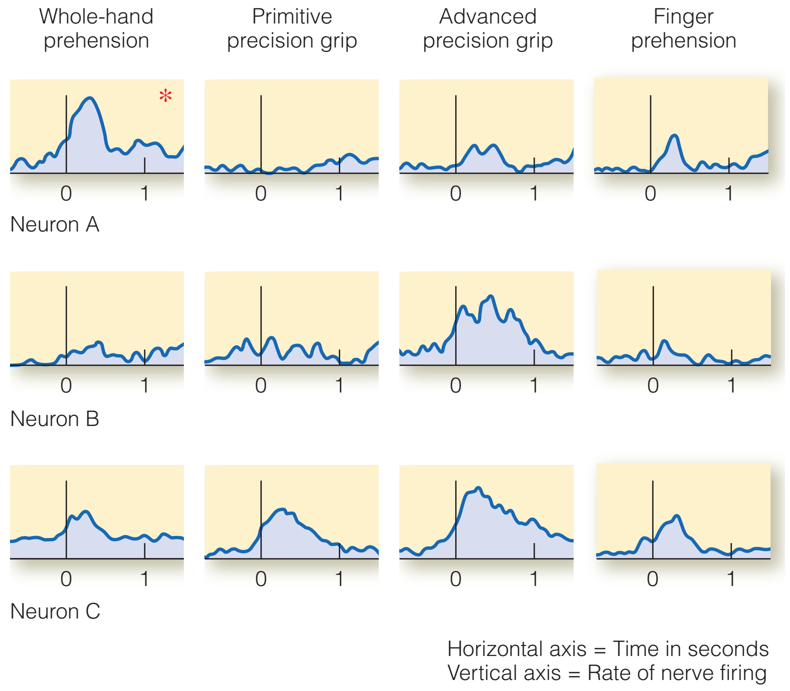
- **The Dorsal (or where/how) and Ventral (or what) Pathways**
- Even an action like picking up a coffee cup involves a number of areas of the brain, which coordinate their activity to create perceptions and behaviors.
  - First identifies the coffee cup among the flowers and other objects on the table (ventral).
  - Once the coffee cup is perceived, she reaches for it, taking into account its location on the table (dorsal).
  - As she reaches, avoiding the flowers, she positions her hand and fingers to grasp the cup (dorsal)
  - Taking into account her perception of the cup's handle (ventral)
  - Taking into account her estimate of how heavy it is based on her perception of its fullness (ventral)
  - She then lifts the cup with just the right amount of force (dorsal)

(a) Perceive cup (b) Reach for cup (c) Grasp cup

- **The Parietal Reach Region (PRR)**
- One of the most important areas of the brain for reaching and grasping is the parietal lobe at the end of the dorsal pathway.
  - There are a number of different PRRs in the human parietal lobe,
  - Neurons in an area next to the PRR that respond to specific types of hand grips.
- The monkey hand grip experiment:
  - Observes a small fixation light in the dark;
  - Lights are turned on for half a second to reveal the object to be grasped;
  - The lights go out and then, after a brief pause, the fixation light changes color, signaling that the monkey should reach for the object.
- While the monkey is reaching for the object in the dark, it adjusts its grip to match the object.

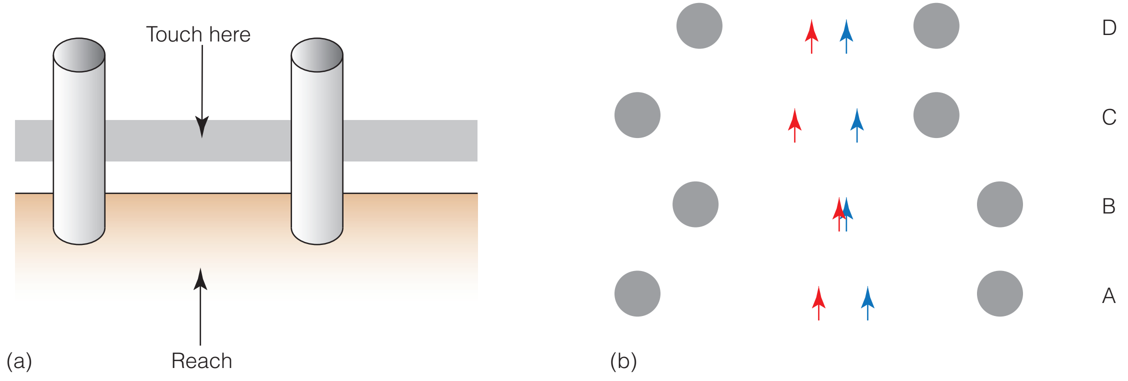


- The key result of the experiment is that there are neurons that respond best to specific grips.
    - neuron A responds best to “whole hand prehension”
    - neuron B responds best to “advanced precision grip”
    - neurons, like C, that respond to a number of different grips.
- Remember that when these neurons were firing, the monkey was reaching for the object in the dark, so the firing reflected not perception but the monkey’s actions.



- A follow-up experiment on the same monkeys discovered a type of neuron, called ***visuomotor grip cells***, initially responds when the monkey sees a specific object, and then also responds as the monkey is forming its hand to grasp the same object.
    - This type of neuron is therefore involved in both perception (identifying the object by seeing) and action (reaching for the object and gripping it with the hand).

- **Avoiding Other Objects When Reaching**
- The fact that obstacle avoidance is also controlled by the parietal regions responsible for reaching was demonstrated by
  - Two patients with parietal lobe damage who had trouble pointing to visual stimuli. These ataxia patients and a group of normal control subjects were presented with two cylinders.
  - Their task was to reach between the two cylinders and touch anywhere on a gray strip. The cylinders were moved to different positions.
  - The control subjects (red arrows) changed their reach in response to changes in the cylinders' position.
  - The reach of the ataxia patients was the same for all arrangements of the cylinder(blue arrows), didn't take account of the varying locations of the obstacles.

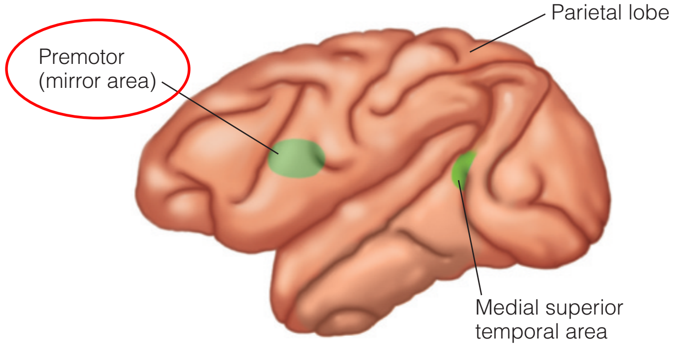


- The result shows that the *dorsal stream*, which was damaged in the ataxia patients, <u>not only provides guidance as we reach toward an object but also guides us away from potential obstacles</u>.

2 Observing Other People’s Actions

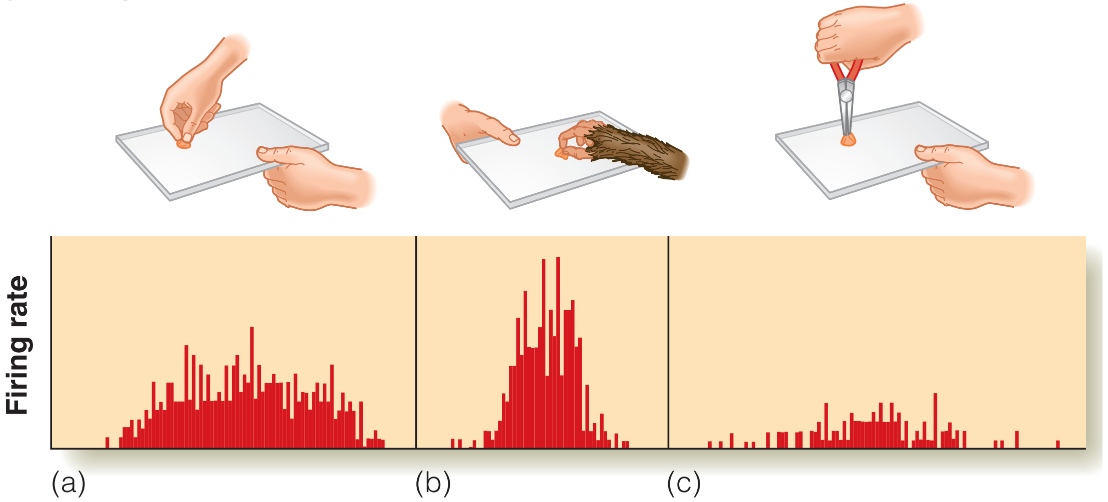
- We not only take action ourselves, but we regularly watch other people take action.
- This “watching others act” is most obvious when we watch other people’s actions on TV or in a movie, but it also occurs any time we are around someone else who is doing something.
- One of the most exciting outcomes of research studying the link between perception and action was the discovery of neurons in the premotor cortex called ***mirror neurons***.

运动前区

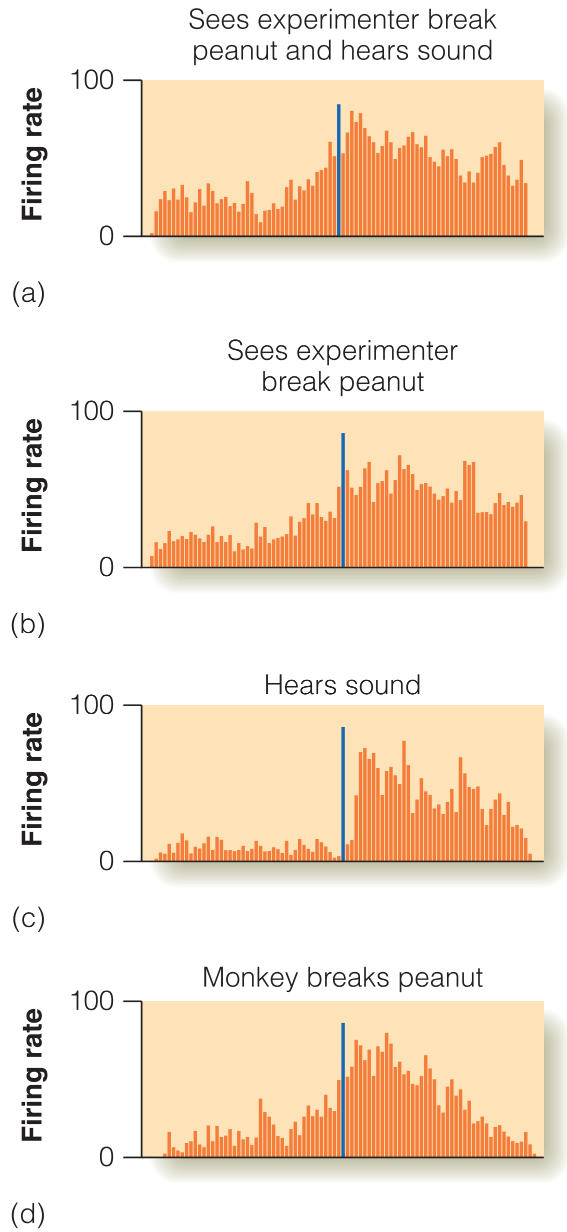


2.1 Mirroring Others' Actions in the Brain

- In the early 1990s, investigating how neurons in the monkey's premotor cortex fired as the monkey performed actions like picking up a toy or a piece of food.
  - To determine how neurons fired as the monkey carried out specific actions.
  - When one of the experimenters picked up a piece of food while the monkey was watching, neurons in the monkey's cortex fired.
- Leading to the discovery of **mirror neurons**: neurons that respond both when a monkey observes someone else grasping an object such as food on a tray and when the monkey itself grasps the food.
- Most mirror neurons are specialized to respond to only one type of action, such as grasping or placing an object somewhere.

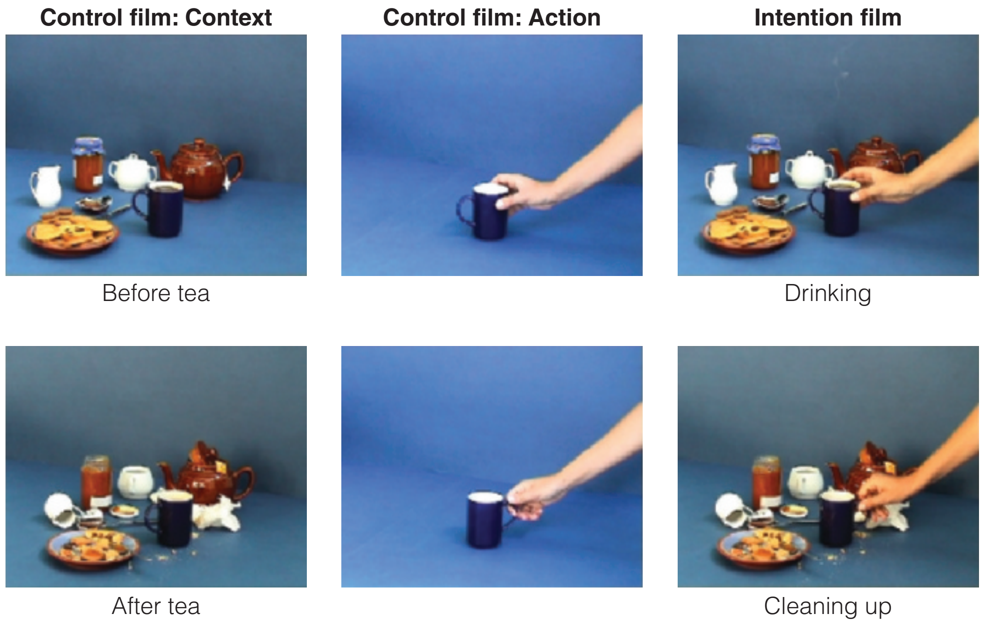


- Could the mirror neurons simply be responding to the pattern of motion?
  - The fact that the neuron does not respond when watching the experimenter pick up the food with pliers argues against this idea.
- Further the discovery of neurons that respond to sounds that are *associated with* actions, called **audiovisual mirror neurons**.
  (a) when sees and hears the experimenter break a peanut
  (b) when just sees the experimenter break the peanut
  (c) when just hears the sound of the breaking peanut
  (d) when breaks the peanut
- These neurons are responding to what is "***happening***" (breaking a peanut) rather than to a specific pattern of movement.
  - Just *hearing* a peanut breaking or just *seeing* a peanut being broken causes activity that is also associated with the perceiver's *action* of breaking a peanut.

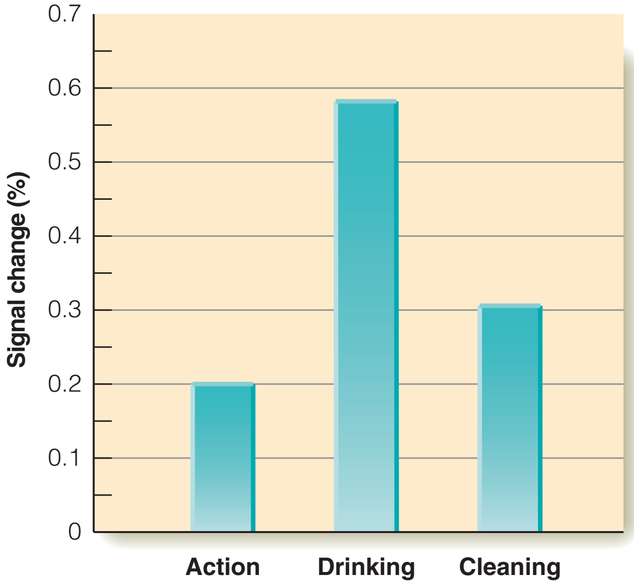


2.2 Predicting People‘s Intentions

- Some researchers have proposed that there are mirror neurons that respond not just to ***what*** is happening but to ***why*** something is happening, or more specifically, to the ***intention*** behind what is happening.
- A number of different intentions that may be associated with the same action.
  - As we see her reach for her coffee cup, one obvious answer is that she intends to drink some coffee.
  - If we notice that the cup is empty, we might instead decide that she is going to take the cup back to the counter to get a refill.
  - If we know that she never drinks more than one cup, we might decide that she is going to place the cup in the used cup bin.
- What is the evidence that the response of mirror neurons can be influenced by different intentions?
- Hypothesizes: it is likely that
  - Viewing the top film would lead the viewer to infer that the person picking up the cup intends to drink from it.
  - Viewing the bottom film would lead the viewer to infer that the person is cleaning up.



- Compared the brain activity in the Intention films to the activity in the Context and Action films, the Intention films caused greater activity than the control films.
  - The amount of activity was least in the Action condition, was higher for the Cleaning Up condition, and was highest for the Drinking condition.
  - Suggest that the mirror neuron area is involved with understanding the intentions behind the actions shown in the films.



- If mirror neurons do signal intentions, how do they do it?
  - One possibility is that the response of these neurons is determined by **the chain of motor activities** that could be *expected* to happen in a particular context.
  - According to this idea, mirror neurons that respond to different intentions are <u>responding to the action that is happening *plus* the sequence of actions that is most likely to follow, given the context</u>.

- The exact functions of mirror neurons in humans are still being actively. In addition to proposing that mirror neurons signal what is happening as well as the intentions behind various actions, researchers have also proposed that mirror neurons help us understand
  - Communications based on facial expressions; gestures used while speaking; the meanings of sentences; differences between ourselves and others.
- It has also been proposed that mirror neurons play an important role in guiding social interactions.
- Some of the proposed functions will be confirmed, but others may need to be revised.
  - Consider that when feature detectors that respond to oriented moving lines were discovered in the 1960s, some researchers proposed that these feature detectors could explain how we perceive objects. With the information available at the time, this was a reasonable proposal.
  - However, later, when neurons that respond to faces, places, and bodies were discovered, researchers revised their initial proposals to take these new findings into account.
  - In all likelihood, a similar process will occur for mirror neurons.

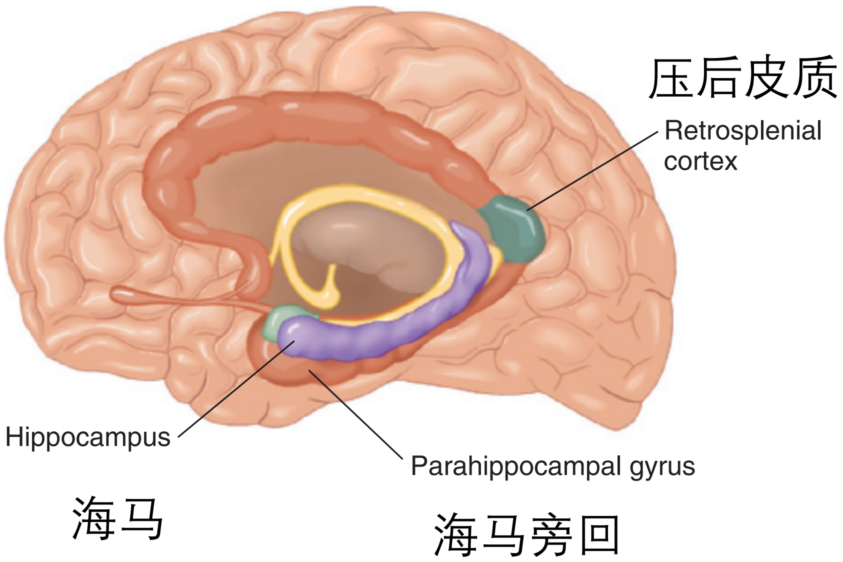
3 Action-Based Accounts of Perception

- The traditional approach to perception has focused on how the environment is *represented* in the nervous system and in the perceiver's mind.
- But many researchers believe that the purpose of vision is not to create a representation of what is out there but to guide our actions.
  - Seeing the fruit is crucial, because it makes the monkey aware that the fruit is present. But reaching for the fruit is just as important in order to survive.
  - <u>Brains evolved not to enable us to think (or perceive), but to enable us to move and interact with the world"</u>
- The vast majority of our experience involves a two-step process: first *perceiving* an object or scene and then *taking action* toward the objects or within the scene. (X)
  - From "action depends on perception" to "perception depends on action" or "people perceive their environment in terms of their ability to act on it."
  - Both ability of a person, difficulty in tasks can affect the perceptual judgements.

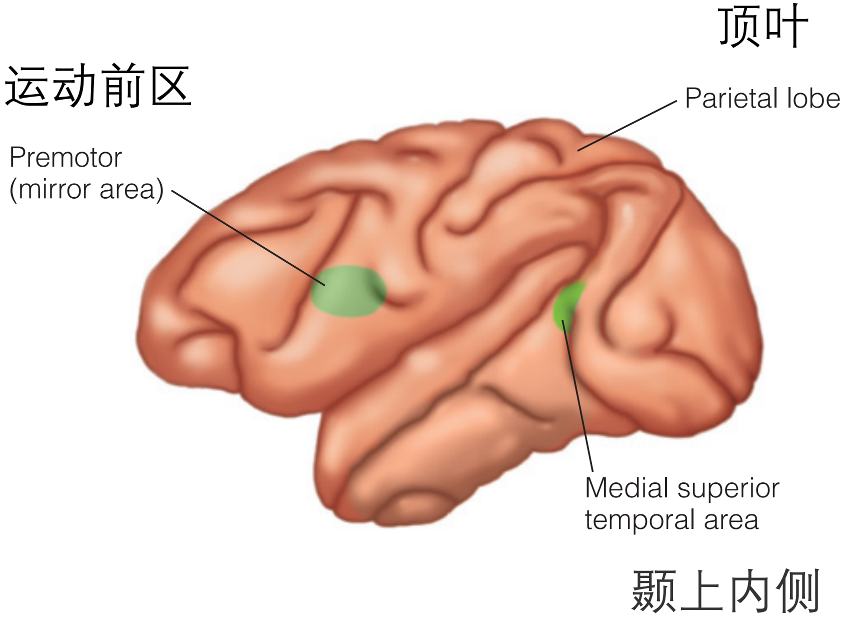
Summary

- The Ecological Approach to Perception
- Navigating Through The Environment
- Acting on Objects
- Observing Other People's Actions
- Action-Based Accounts of Perception

Summary



压后皮质
Retrosplenial cortex
Hippocampus
Parahippocampal gyrus
海马
海马旁回



顶叶
运动前区
Parietal lobe
Premotor (mirror area)
Medial superior temporal area
颞上内侧

Thanks