

Representation of Mental Images

[weekly_readings](#) [#week-1](#) [#ch-2](#) [#my-notes-755](#)

How do mental images represent?

- One way to understand a *complex cognitive ability* is to try to build a **machine** that has that ability (or at least some primitive form of it).
- The **program** that the machine runs is a *model of the ability*.
- Often the ability being modeled is a very primitive and simplified form of the ability that we are trying to understand.
- This is the case with SHRDLU, which was intended to model only a very basic form of linguistic understanding.
- But even in cases like that, we can still learn much about the *basic principles of cognitive information processing* by looking to see how well the model works.
- This is why the history of *cognitive science* has been closely bound up with the history of *artificial intelligence*.

As we saw in the previous chapter, the emergence of cognitive psychology as a serious alternative to behaviorism in psychology was one of the key elements in the emergence of cognitive science. A good example of how cognitive psychology can serve both as an inspiration and as a tool for cognitive science came with what has come to be known as the imagery debate.

- The **imagery debate** began in the early 1970s
- inspired by a thought-provoking set of **experiments on mental rotation**
- carried by the psychologist **Roger Shepard** (in collaboration with Metzler and Lynn Cooper)

Mental Rotation of 3D objects (1971) - Shepard and Metzler

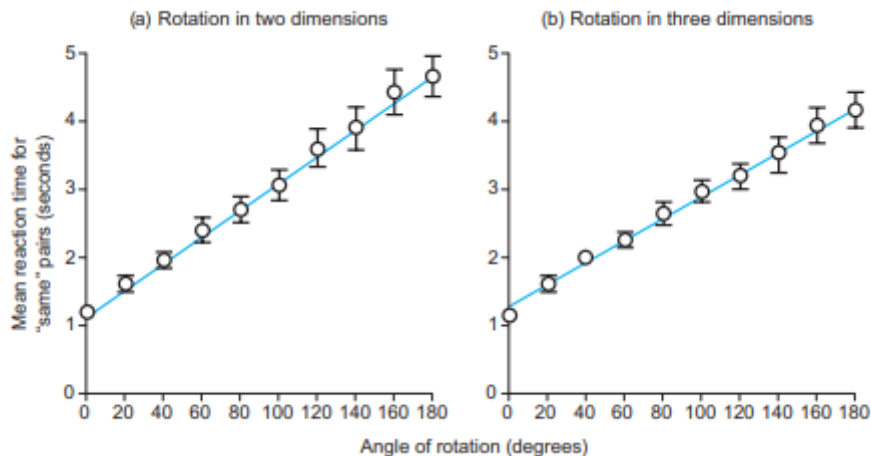
EXPERIMENT

- Subjects were presented with drawings of 3 pairs of 3D figures.

- Pair-1 and Pair-2 have same figure but rotated to different angles. Pair-3 has different figure.
- The subjects were asked to identify as quickly as possible pairs of drawings where the second figure is the same as the first

OBSERVATION

- There is a *direct, linear relationship* between the *length of time* that subjects took to solve the problem and the *degree of rotation* between the two figures.
- The length of time increased in direct proportion to the degree of rotation.



INTERPRETATION by Shepard, Metzler, Lynn Cooper (and others)

- Subjects solved the problem by "mentally rotating" one figure to see whether or not it could be mapped onto the other.

This explanation raised 2 QUESTIONS

1. How is the information represented
2. How is it transformed

Information processing in Mental Imagery

The standard way of thinking about the mind as an information processor takes the digital computer as a model.

But there is a Slight Difference

- When information is digitally encoded, the length of time it takes to process a piece of information is typically:
 - a function of ONLY the **quantity of information** (type does not matter)

- However, mental rotation experiments show that the time taken by some information processing tasks are different even if:
 - **quantity of information is the same**

RESULT

- Many cognitive scientists have suggested that mental rotation tasks tap into ways of encoding information very different from how information is encoded in a digital computer.

The basic characteristic of an imagistic representation is that representation is secured through resemblance

EXPERIMENT by Stephen Kosslyn in 1973

- subjects were asked to memorize a set of drawings
- then, gave them the name of one of the objects (e.g., "airplane") and asked them to focus on one end of the memorized drawing (eg., "tail").
- then, gave the names of possible parts of the object (e.g., "propeller") and asking them to examine their images to see whether the object drawn did indeed have the relevant part (which it did on 50 percent of the trials)

OBSERVATIONS

- Kosslyn found an effect rather **similar to that in the mental rotation** studies, namely,
- The length of time it took the subjects to answer varied according to the distance of the parts from the point of focus. (*Example: If the subjects were asked to focus on the "tail" of the plane, it would take longer for them to confirm that the plane had a "propeller" than that there was not a pilot in the cockpit*)

INTERPRETATION by Kosslyn

- The type of information processing involved in answering the test questions involves **scanning imagistic representations**. Instead of searching for the answer within a digitally encoded database of information about the figures, the subjects scan an imagistically encoded **mental image of the airplane**.

An interdisciplinary model of vision

The mind can be studied at many different levels:

1. **Bottom-up approach:** beginning with individual neurons and populations of neurons, or perhaps even lower down, with molecular pathways whose activities generate action potentials in individual neurons, and then trying to build up from that to higher cognitive functions
2. **Top-down approach:** starting out with general theories about the nature of thought and the nature of cognition and working downward to investigate how corresponding mechanisms might be instantiated in the brain

- On either approach one will proceed via **distinct levels of explanation** that often have separate disciplines corresponding to them.
- A fundamental problem for cognitive science is working out how to combine and **integrate different levels of explanation**.

Levels of Explanation

A Computational Investigation into the Human Representation and Processing of Visual Information - 1982 Book explaining *David Marr's* model of "Human Visual system"

Marr distinguishes three different levels for analyzing cognitive systems:

1. Computational Level:

The tasks of an analysis at the computational level are:

- to translate a general description of the cognitive system into a specific account of the particular information-processing problem
- to identify the constraints that hold upon any solution to that information processing task.
- A computational analysis identifies the information with which the cognitive system has to begin (the input to that system) and the information with which it needs to end up (the output from that system)

2. Algorithmic Level:

- how the cognitive system actually solves the specific information-processing task
- how the input information is transformed into the output information

3. Implementational Level:

- to find a physical realization for the algorithm, and
- to find mechanisms at the neural level that can properly be described as computing the algorithm in question

Applying Top-Down Analysis to the Visual System

Marr's model is very interdisciplinary

Marr's in his book refers to Elizabeth Warrington's work

On patients with damage to **left and/or right parietal cortex** - regions related to **perceptual recognition**

Patients with damage to:

1. right parietal cortex: are able to recognize and verbally identify familiar objects provided that they can see them from familiar or "conventional" perspectives, but difficult from "unconventional" perspectives
2. left parietal cortex: showed a diametrically opposed pattern. Generally have language problem, but can identify an object from both perspectives.

Marr's drew two conclusions from this work:

1. information about the "*shape of an object*" must be processed separately from information about "*what the object is for*" and "*what it is called*"
2. the visual system can "*deliver a specification of the shape of an object*" even when the perceiver is "*unable to recognize the object*"

Marr's explanation at various levels:

Computational Level

the visual system's basic task is to construct a representation of the three-dimensional shape and spatial arrangement of an object in a form that will allow that object to be recognized

Since ease of recognition is correlated with the ability to extrapolate from the particular vantage point from which an object is viewed

Marr concluded that this representation of object shape should be on an object-centered rather than an egocentric frame of reference

Algorithmic Level

raises a range of new questions:

- how exactly is the i/p and o/p info encoded?
- what are the system's representation primitives (basic units)?
- what operations is the system performing to do the info-processing-task

In Marr's theory the information - about surfaces in the field of view - is derived from a series of increasingly complex representations:

1. **primal-sketch:**

Distributions of light intensity and basic geom. of field of view

2. **2.5D sketch**

depth and orientation of visible surface from viewer's perspective

3. **3D sketch**

viewer-independent representation

At the algorithmic level the job is to specify these different sketches and explain how the visual system gets from one to the next, starting with the basic information arriving at the retina.

The algorithmic description of the visual system takes a representation formulated in terms of these representational primitives as the input, and spells out a series of computational steps that will transform this input into the desired output, which is a representation of the three-dimensional perceived environment

Implementation Level

- a further set of disciplines come into play
- basic physiology of the visual system
- Neural level implementation - little is discussed in Marr's own work though

Contribution to Cognitive Psychology

- Marr's analysis of the visual system clearly illustrates both:
 1. how a single cognitive phenomenon **can be studied at different levels** of explanation,
 2. how the different **levels of explanation can come together** to provide a unified analysis.
- Marr's analysis of the visual system has been taken as ***a paradigm of how cognitive science ought to proceed.***

