# 3 MODELS AND COMPONENTS OF IMAGERY

## 3.1 Kosslyn's model

Research concerning how cognitive processes operate in mental imagery generation is a task for the supporters of the analog model. These authors, in fact, aim to comprehend how imaginative system and processes (generation, scanning or inspection and transformation of mental imagery) work. In this area the most relevant contribute is due to Kosslyn and Shepard.

Following an approach that we could define as computational-analogic simulation, Kosslyn and coll. (Kosslyn and Swartz, 1977; Kosslyn, Brunn, Cave and Wallach, 1984) claimed that mental imagery has a permanent spatial structure (visual buffer) in which there are two representation forms, one " superficial" and the other "deep". The first one has a temporary character and is generated by an abstract and general system of representation; the second one is stored in long term memory. In Kosslyn's theory the visual buffer is constituted by a bidimesional "medium" formed by set of cells in an ordered system, the position of which correspond to positions in the visual field. The cells, when activated, represent the surface zone of a represented shape, in the way that the activation pattern in the buffer is equal to the object's visible surface. So, this "medium" may be occupied with information coming from the visual system or coming from the long term memory. This process is what Kosslyn calls "generation of mental imagery". In Kosslyn's theory long term memory representations of objects, shapes and surface properties, used in generation of mental imagery, are used also in recognition and have the form of a structural description.

Generation processes may occupy the buffer with patterns that represent many objects coming from different long term memory files and can link objects in different positions and orientation. The computational theory of imagery generation

(Kosslyn and Swartz, 1978) says that this generation is due to three elaboration modules, which have the function to transform the input. The "picture" elaboration module is used in an iterative way, when an object's image is created on the basis of an assemblage of more parts stored in memory and generated in succession. The "put" elaboration module has the task to control the relation between couples of components of an object, in order to calibrate a part relative to another. Finally the "find" elaboration module has the function to correctly recognize, indentify and classify spatial patterns when their forms can be reported. The "find" module represent what Kosslyn and Shwartz call " the eye of mind" and it is used to "see" the imagined object and to "scan" it not only when we imagine but also when we physically perceive it.

We can summarize Kosslyn's model in this way:

1) The generation process is a recovery of visual information from the long term memory to short term memory;

2) The transformation process occurs inside the short term memory (or in the working memory): for example, in a rotation task, the images are temporally memorized and one of them is rotated to verify the perfect correspondence with the second one. In this case there is no reference to long term memory;

3) Finally, in the inspection process (scanning) an image (for example a map) is analyzed focusing on a point, able to shift on the image itself. Unlike the simple recover of images from long term memory, here we have an "eye of mind" that inspects, in working memory, the image recovered from long term memory. Considering that also perceptions generated from external stimuli are represented in visuo-spatial working memory, these relations may be represented as in Fig. 3.1:

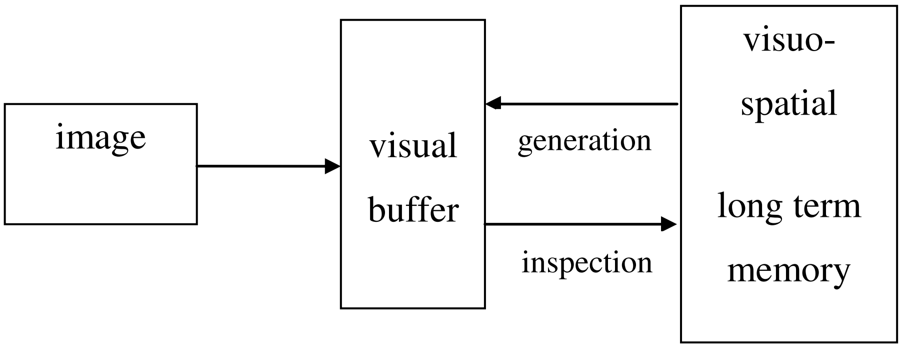


Figura 3.1: Kosslyn's model of visual memory

## 3.2 The mental scanning

According to Kosslyn (1994), image scanning is a process that allows people to shift their attention across visual mental images. Images are patterns within a visual buffer that functions as if it were a coordinate space. This is not an actual physical space, but a functional one that is defined by the way processes access representational structures. In all image representations in the visual buffer, every part of the representation corresponds to a part of the object, being represented in such a way that the relative distances between object's parts are preserved in the distances between the corresponding representation's parts. Thus, visual mental images, and the processes that operate on them, are thought to be functionally analogous to the external objects and the perceptual processes they simulate. This functional equivalence has been assessed using the paradigm of selective interference. Farah (1985) showed that a given letter is easier to detect if the participant has previously visualized the same letter rather than a different one.

The first image scanning paradigm was designed to investigate the spatial properties of images (Kosslyn, 1973). In the original experiment, participants memorized drawings of elongated objects (such a tower). They were then asked to visualize one of the objects and to focus at one of its end (for example, the bottom of the tower). Lastly participants heard the name of a feature on the object (as a flag), and they had to look for it. Participants were not instructed to scan their image, but were simply told to focus on the original location until the probe was delivered and then to focus on the named part or feature. They had to press one button when they had focused on the named part and another button if they failed to find it.

A variant of the original image scanning paradigm was later designed by Kosslyn, Ball and Reiser (1978). The key feature of this new paradigm was that there were no intervening objects between focus and probes. In a learning phase participants memorized a map of an island containing seven landmarks (such as a rock, a beach etc.). The map was designed in such a way that the distances between all pair of landmarks were different. In the test phase, at the beginning of each trial, participants heard the name of one of the landmarks. They were instructed to form an image of the intere island and to focus mentally on the named landmark. Shortly thereafter, the name of a second object was displayed. If the second object was one of the landmarks located in the island, participants had to scan it by following mentally a little spot and to press a button when the spot reached it. In half of the trials the second object was not one of the landmarks and in these trials the paticipants were instructed to press a different button. As in the original experiment, the participants took longer to scan greater distances and scanning times increased linearly with increasing distances. Kosslyn et al. (1978) interpreted this findings as evidence that mental images incorporate the metric information present in the original stimulus.

But there was a problem, because the images used in the conventional scanning experiment were constructed by reviving previous perceptual experience in some form. Images include features that are derived from previous perceptual episodes. In fact, this situation limits the investigation of image scanning to "reproductive imagery" and therefore overlooks many forms of imagery, such as the creation of novel images that are not directly traceable to immediate or recent sensory experience. Pioneer reaserch has highlighted the distinction between the so called "memory images" and "imagination images" (Holt 1964). It's necessary to create a situation in which a person builds the image of a complex object or spatial configuration from a verbal description in such a way that metric information is included in that representation. In other words, can images exhibit metric properties without the representation having been constructed from a perceptual experience directly involving metric information? In the experiment by Denis and Cocude (1989) the image of a map was constructed by processing a verbal description. The description located each landmark at a metrically defined point using the conventions of aerial navigation. The result was that the participants, who learned the map under conditions similar to those used by Kosslyn et al. (1978), displayed the typical increased distance scanned.

Some critics aroused about these experiments. According to Pylyshyn (1981), the results of the scanning experiment simply reflect the tacit knowledge that participants have about visual processes. Phylyshyn argued that imagery processes are cognitively penetrable, because they can be altered by the participants' beliefs, goals, expectations or knowledge.

For this reason Finke and Pinker (1982) designed an image scanning paradigm in which the response given by the participants were not biased by prior knowledge, no explicit imagery instructions were given, and participants were not able to infer that the experimenter was interested in the relationships between scanning time and distance. In their experiments the participants first memorized a pattern of four dots. This pattern was then replaced by an arrow in an unpredictable position in a blank field. The participants were to decide as quickly as possible whether the arrow was pointing to a location previously occupied by one of the dots. They were not instructed at any time to form or scan visual mental images. Results revealed a strong linear relationship between response time and distances separating the tip of the arrow and the target dot. These findings could not be explained by experimenter expectancy effects, or by task demands and provide further evidence that scanning does reflect the spatial structure of image reperesentations. The results of these experiments converged also to demonstrate that different scanning processes are involved in the two classical image scanning: tasks a) a scanning process involving a simulation of a movement between two points, which involves a transformational process, but distinct from that used in mental rotation b) a scanning process involving the localization of a target in which people shift their attention window along a straight line from the tip of the arrow to the target point.

## 3.3 Mental rotation

Another important mechanism related to the spatial properties of mental imagery and with its transformation is Mental Rotation. Shepard & Metzler (1971) introduced the concept of mental rotation and used it to investigate spatial properties of mental images with one of the best-known experiments in the field. The experiment is associated with a set of memorable graphical images, some of which were presented on the cover of the issue of Science where the paper was first published.

Shepard & Metzler presented subjects with pairs of drawings of three-dimensional, asymmetrical assemblages of cubes (see figure 3.2 A, B, and C). In each pair the right-hand picture either showed an assemblage identical to that shown on the left, but rotated from the original position by a certain amount, or else it showed an assemblage that was not only rotated, but was also the mirror image of the one to the left (figure C). Participants had to tell, as quickly as possible (by pressing a button), whether the two objects depicted were in fact identical (except for rotation) or were mirror images. Shepard's hypothesis was that the task would be performed by forming a three-dimensional mental image of one of the depicted objects and by rotating mentally this whole image, to see whether it could be brought into correspondence with the other picture. Results clearly supported this expectation, because it was found that, for each subject, the time taken to say that both objects of a pair were identical increased in direct proportion to the angular rotational difference between them. It was as if the subjects were rotating their mental image at a steady rate (although this might be different for each subject), so that the further they had to go to bring their image into correspondence with the reference picture, the longer it would take them. On post-experimental questioning, most of the subjects confirmed that this was indeed how they believed that they had done the task. (Interestingly, it made no difference whether the rotation was in the plane of the page or in depth). These findings seemed directly to refute the propositional hypothesis that thought processes depend entirely upon language.

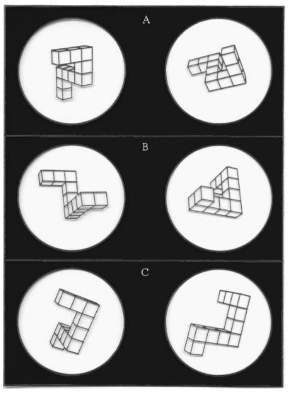


**Figure 3.2**

Some of the stimulus figure pairs used by Shepard & Metzler (1971).

A- Identical objects differing by a rotation in the plane of the page

B – Identical objects differing by a rotation in depth

C – Mirror-image objects (also rotated in depth)

Despite the elegance of this experiment and the clear-cut results (nice linear plots of reaction time against rotational angle), some researchers challenged Shepard's contention that his results show that images are rotated as a whole, rather than their parts being compared in a piecemeal fashion (Hochberg & Gellman, 1977; Pylyshyn, 1979a; Yuille, 1983). More radically, some doubted whether the rotation task really involves imagery at all (Marks, 1999). For example, Just and Carpenter (1976; Carpenter and Just, 1978), who tracked subjects' eye movements whilst they did a version of the Shepard and Metzler's task, argued that the linear increase in reaction time arose not from the inner rotation of an image, but from a need to make more eye movements between the two pictures (in order to compare their features) the more they were rotated relative to one another. In the icon-phobic culture that still prevailed in psychology in the early 1970s such an "imageless" interpretation of the results was still a current option.

However, the case for mental rotation does not rest solely on this celebrated experiment. Shepard and his students (especially Lynn Cooper) were subsequently able to demonstrate mental rotation and other related effects in quite a number of different experimental designs (see Shepard and Cooper et al., 1982) mostly designed to oppose alternative interpretations of the results that would avoid the need to postulate rotating imagery. Most of these experiments did not involve the comparison of two simultaneously visible pictures, thus leaving no scope for the sort of eye movement explanation that Just and Carpenter had suggested. For example, Cooper and Shepard (1973) presented their subjects a letter of the alphabet rotated out of its normal, upright orientation and participants had to decide whether the letter was displayed in its normal version or as a mirror image, regardless of its orientation in the picture plane (see figure 3.3). Once again the usual finding was that response time increase with how far the letter has been rotated from the standard upright position (even though the relation was not so neatly linear as in the earlier experiment). The implied explanation is that the subjects rotated the image of the non-upright letter that they showed in its canonical upright orientation, in order to compare it to their memory of how the letter would normally look.

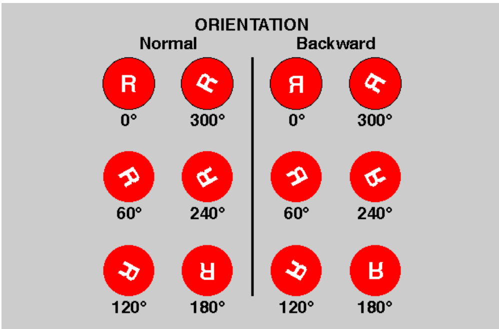


**Figure 3.3**

Rotated letter stimuli of the sort used by Cooper & Shepard (1973).

Cooper (1975, 1976) also did several experiments on the mental rotation of complex irregular polygons and some of this work depended on the subjects being explicitly told to form and rotate mental images of the polygons (whose shape they had previously memorized). Once again, a linear dependence of reaction time upon angle was found, suggesting smooth, regular rotation of an image. An advantage of the explicit imagery instructions used in these experiments is that it forges a more explicit link between the measured results and imagery as a conscious experience under voluntary control. On the other hand, it opens up the very real possibility that the results might be explained away as the result of “experimental demand” (essentially, the subjects deliberately trying to produce the results that they believe the experimenters want, regardless of the actual underlying cognitive processes involved). This is well known to psychologists to be a very real problem in many areas of experimental psychological research. However, Cooper's experiments are taken together with the other rotation experiments that did not use explicit imagery instructions, and gave little foothold for experimental demand to influence the results (Shepard and Metzler, 1971; Shepard and Cooper, 1973), they make a powerful case for the reality and robustness of the mental rotation phenomenon.

Controversy about the underlying mechanisms of mental rotation continues, but these have recently been investigated by several neuroscientific techniques, such as eye movement measurements (De Sperati, 2003), direct recording from electrodes implanted in the brain (Georgopoulos et al., 1989), functional magnetic resonance imaging (fMRI) (Cohen et al., 1996; O'Boyle et al., 2005) and transcranial magnetic stimulation (Ganis et al., 2000). Some of these studies have been focused on “motor imagery” rather than visual imagery. However, there are reasons to believe that motor processes and motor areas of the brain may be involved in mental rotation quite generally (Kosslyn, 1994; Wexler, Kosslyn & Berthoz, 1998; Tomasino et al., 2005; Eisenegger et al., 2007).

## 3.4 Mental imagery and memory

The close relationship between mental imagery and memory was a theme dear to our culture since its origins. Aristotle said that "memory objects are those that fall under its own imagination" and Hobbes suggested that "memory and imagination are the same thing with different names given to the different way of considering it." Since ancient times the link between these two cognitive functions was clear, but the experimental verification of the effectiveness of mental images for memory had to wait until the birth of experimental psychology. In 1894 Kirkpatrick noticed that, by inviting subjects to form mental representations of words, he obtained an increase in memory and that this benefit was greater for concrete words than for abstract words. Several studies showed positive effects of imagination on memory. Indeed, concrete verbal materials (words, phrases, passages), and therefore easily imaginable, has a better chance to be memorized, as well as the use of mental images increases the mnestic performance; furthermore "good imagers" usually have a better memory than "bad imagers".

A crucial point was to discover how information is represented in our permanent memory. As already seen, Paivio (1971, 1986) says that the different ways in which the verbal and imaginative systems are involved on the basis of the stimuli are responsible for different memories. The figures are easier to memorize because they immediately activate an image codification (analog) and, if the object is familiar, even the verbal codification, which gives a verbal tag to the stimulus. In this way the item is encoded twice: once by the imaginative system and second by the verbal system. Stimuli with high “value of image” are immediately coded by verbal system, but they can also have additional coding by imaginative system. Verbal stimuli with low “value of image”, words that are difficult to associate to a mental image, such as "purpose", may be more difficult to take advantage of an imaginative encoding and rely mainly on verbal one. So they are more difficult to memorize because encoded by only one system. The value of image is thus a measure of vividness and speed with which a word can evoke a mental image. The measure of the value of the image is highly correlated to words concreteness. Paivio said that mental images are not only processes, but also permanent representations. An item can be stored in memory in the form of analog-perceptual representation because it retains some perceptive properties as size, shape and colour.

### 3.4.1 Visuo-spatial working memory

Baddeley (1986) provided several proofs about the existence of a specific system of visuo-spatial working memory, different from an auditory-articulatory system (articulatory loop). The evidences come from the double task methodology and selective interference, which showed that a subject can perform more easily two tasks, concurrently or in immediate succession, that implicate different systems rather than two tasks involving the same system (Cornoldi, 1978). This highlights the relationship that exists between the concept of active processes imagination and visuo-spatial working memory. In a recent analysis of the image buffer's characteristics (Kosslyn 1980) and of the visuo-spatial store, a substantial overlap was noticed. This does not exclude that, in specific theoretical or experimental paradigms, both the one and the other concept assume particular characterizations only partially overlapping. For example, in "active" image, generated in absence of visual stimulation, the attention focus is more directly involved than in information passively collected by a visual stimulation and then preserved in the store.

Also the reference to experience and common language should generate distrust in the perfect "identification" of these two functions. "Remember" and "imagine" are not lived as the same thing: the experience of "remember" a person's face or a picture just seen, does not match the experience imagining them. A short-term visual memory is not necessarily an image, even if we can suppose that it is the same system to store and process both. A fundamental short-term memory characterization is represented by the length of the track in the system. Such a track is stored for a limited time. However the question of the duration of a track cannot be entirely solved, but can be compared to the metaphor of the impressed sand. The sign can be imprinted in different measure and then vanishes after different periods. To assess the degree with which the trace has been imprinted we cannot take account only of physical characteristics of the stimulation, but also the impact it has on the subject.

### 3.4.2 Mental imagery in thought processes

Considering the role of mental imagery in thought processes, it can be assumed that they have a function in cognitive processes 1) because they represent something (in addition or substitution of other kind of representation) 2) for what they represent 3) for how they represent it.

Shepard (1978) suggests that through mental imagery subjects have access to an alternative mental code besides the verbal one. A reason why mental images can facilitate thinking lies in what they, but not other forms of representation, allow to keep in mind during the course of reasoning. The images allow the representation of details and relevant particulars of situations that logic or verbal encoding, which is more abstract, doesn’t re-transcribe (Shepard, 1978). The schematization function that certain mental images can play makes subject's attention quickly and easily oriented towards the key aspects of the situation, avoiding unnecessary aspects that are useless or misleading. Moreover, mental imagery simplifies the situations, allowing persons to work easily (Kosslyn, 1983). Finally, the opportunity that picture representation gives to keep in mind the totality of the field, and to highlight its structural aspects, implies a lower mental workload and a consequent processing and flexibility.

So mental images are useful in thinking, not only because of what they represent, but also how they represent it. Primarily, they support a simultaneous parallel information processing (Kaufmann, 1980). The simultaneous examination of various aspects of a situation provides an opportunity to collect relationships between the different elements. In fact the ability to restructure problems is connected with the ability to perform a synthesis of mental images of objects (Barolo, Masini and Antonietti, 1990). Conversely, sequential transformations of mental imagery- such as bending and roll over - are related to logical and analytical skills but not to the ability to solve in a productive manner mathematical, geometric, instrumental and verbal problems (Antonietti, Barolo and Masini, 1988).

This facilitating contribution of mental imagery proves to be true not only in imagery flows produced after the submission of a problem solving tasks, but also when they are produced before and even if the subject doesn’t know previously the problem. According to Kaufmann (1980, 1985) mental imagery is particularly effective when the required level of operation is characterized by novelty. Phylyshyn, although critical about the cognitive role of mental imagery, recognizes its importance in the construction of new information, "while picture-like entities (mental images) are not stored in memory, they can be constructed during processing, used to generate new interpretations” (Phylyshyn, 1973).

Studies about chess players (Milojkovic, 1982) confirmed that mental image is most useful when being confronted with new tasks, which proved that beginners, but not experts, make use of visual mental representations. It seems, therefore, that picture representations operate, or prove to be particularly useful, mainly in the early stages of reasoning process (Kaufmann, 1985), at the phase where:

- people must consider the whole situation, when is not yet clear which are the relevant elements
- people have to understand the essential structure, ignoring unnecessary details or misleading information
- is more profitable to explore at the same time, through parallel strategies, various research directions rather than to use only one
- it is better to keep the cognitive field fluid, so to apply even unusual transformations and revision.

Kosslyn (1983) claimed that if we can solve the problem quickly, using only propositional information, mental imagery loses its effectiveness, but as opposite, if special circumstances prevent this, people have benefit of visual representations provided by mental images.

### 3.4.3 Conclusions

The models and experimental results mentioned before support the analog nature of mental imagery, and both have established a solid foundation for those who considered valid the hypothesis that these two activities have similar characteristics and processes. Kosslyn (1994) has tried to solve some of the debates about the visual mental image format in the perspective of cognitive neuroscience founded on the belief that "an image is a pattern of activation in some areas of the visual cortex topographically mapped" and that there is a "processing system that produces, transforms and interprets descriptive and visuo-spatial representations”.