Dynamic Assembly of Figures in Visuospatial Reasoning

Stefan SCHNEIDER [a,1], Ulf KRUMNACK [a] and Kai-Uwe KÜHNBERGER [a]
[a] *Institute of Cognitive Science, University of Osnabrück*

Abstract. An exploratory, qualitative experiment sheds light on the depictive theory of mental imagery. The study analyzes the very operations subjects undertake when solving visuospatial tasks. Preliminary results indicate that subjects do not make use of stable mental images: instead, they continuously assemble and reassemble different perspectives through the guidance of heuristics and prototypes. These observations allow a reinterpretation of mental imagery. We want to forward the hypotheses that a) the assembly process itself is of much higher importance than usually acknowledged; b) that an assembled perspective (or figure) is defined by one's orientation towards certain operations; and c), that heuristics and prototypes are instantiated by a heterarchical organization of mental operations.

Keywords. mental imagery, visual synthesis, model construction, spatial orientation

Introduction

What characteristics does the cognitive representation of a shape have? How do humans reason *about* and *with* shapes? Besides from subjective reports of the appearance and the usage of mental images, the existence of a faculty of imagery – however it may be conceptualized – has been proven to be functionally involved in cognitive processes. Evidence is for example presented from studies showing that humans can mentally synthesize and subsequently recognize figures. In the paradigmatic experiment of Finke et al. [1], subjects were asked to compose objects following verbal instructions like the following:

> Imagine the letter Y. Put a small circle at the bottom of it. Add a horizontal line halfway up. Now rotate the figure 180 degrees. ([1, p. 62])

Two opposed conceptualizations of mental imagery are established: the depiction theory and the description theory. These have been forwarded by Kosslyn [2] and Pylyshyn [3] respectively, who strongly opposed the other's account; the dispute is known as the *imagery debate*. Despite substantial doubts whether the depictive approach is computationally feasible and cognitively plausible [3], the intuitive apprehension that in *imagery* we operate with *images* of some sort, has become a hardened paradigm and thereby obstructs alternative conceptualizations. In the following we will present such an

[1]Corresponding Author: Stefan Schneider, Institute of Cognitive Science, University of Osnabrück, Germany. E-mail: stefschn@uni-osnabrueck.de.

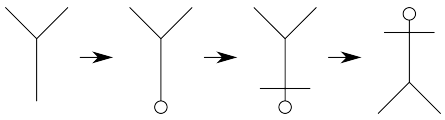

Figure 1. A figure that can be easily mentally synthesized and recognized. Re-drawn from Finke et al. [1].

attempt. First, we introduce the depictive approach, and then confront it with hypotheses derived from a reinvestigation of the very operations subjects execute when solving visuospatial tasks.

1. Stable images at your disposal?

In the experiments of [1], the idea to test for possible reinterpretation is cleverly chosen: One cannot argue that recognition of the figure rests upon semantic information in the given description (e.g. as in "Imagine two legs on a torso, add two arms and a head: what have you got?"), but has to rely on cognitive operations with visuospatial information. Figure 1 illustrates the intended assembly of a stick-figure. The reader probably has made his/her own interpretation, possibly similar, by now. It is important to note that subjects in the experiment only received the verbal instruction and did not use any external aids. The visual presence of the depiction in Figure 1 might therefore be misleading, and this, so we want to suggest, shaped the folk-psychological intuition about mental images as *pictures in the head*, and laid the foundation of the dominant depictive theory of imagery.

According to Tversky "[t]here are two basic tenets of the approach, one regarding representations and the other regarding operations on representations: that mental images resemble percepts, and that mental transformations on images resemble observable changes in things in the world, as in mental rotation, or perceptual processes performed on things in the world, as in mental scanning" [4, p. 211]. In a nutshell, the depictive account says that in solving visuospatial problems, one can distinguish two types of operations: (a) image-construction operations, and (b) subsequent inspection of the image, analogous to visual perception.

So far the theory. There is something in disorder, though. Apart from transformations of the image, Kosslyn highlights that images *fade* – "they are transient and begin to decay as soon as they are activated" [2, p. 50]. A frequent *reconstruction* is necessary, which eventually brings about transformations of the imagined objects or a new perspective. The fading of images stands in opposition to their asserted analogue nature to external images (or scenes). This seems to be a flaw in what otherwise could be a pretty useful imagery faculty. We suggest that shifting the research focus away from a supposed presence of images, and towards the process of *reconstructing* these will be of great value.

2. Operations in imagery. Current research

In an exploratory protocol study, six subjects had to solve visuospatial tasks like the above with a slightly higher level of complexity. Interviews lasted for about one hour, and were guided according to an interview protocol devised by Petitmengin [5]. In the

analysis, reports about the fleetingness of the images became overly present (which subjects often experienced as a nuisance), and accordingly the necessity to continuously assemble and re-assemble the intended figure. This becomes evident if the reader considers the following task:

> Imagine a square. Imagine its vertical lines being separated each into three parts of equal length. Now connect the resulting division marks by two horizontal lines, such that the square is divided into three identical rectangles. Now add the two diagonals to the square. How many triangles does the pattern contain? (Adapted from Wiener [6, p. 82f].)

Even though the elements of the figure are easy to imagine, and the whole figure is not more complicated than the stick-figure, it is very difficult to find all the triangles in this task (there are much more then one initially might think). Interestingly, constructing the figure seemed feasible. A typical report of the resulting image reads like this:

> *Well . . . I have to . . . juggle a bit to keep the components in focus. Everything is there. But I successively imagine how the horizontal lines join the square, and how the diagonals cross these lines, that's all a bit waggly.*

But when asked to count the triangles, *the picture changes*. A typical report is:

> *I lose the whole image. It isn't as clear as before any more. (Can you still examine it?) Yes, but I have to build it up anew. And in considering a detail of the image I . . . somehow have to retrace the constituents, for example putting in the diagonal again.*

This change of the subject's representation of the figure suggests that the task – in this case, constructing the image versus finding triangles – guides the process. In the latter case, for example, subjects rely on the given elements of the figure and assemble subsets of these in order to construct triangles. That is, instead of *looking at the image* and *recognizing* triangles, subjects actively try to construct triangles from the given material. The following transcript illustrates such an attempt; the subject has the idea that there could be another type of triangle and actively tries to assemble it:

> *There is another idea - to search for another line with which the diagonal can form a triangle - there have to be more, because the diagonals cross the lines of the two horizontal lines. I begin in the upper left corner and . . . trace the diagonal . . . and it it goes a level deeper, and forms a triangle where it crosses the second horizontal line. I didn't really search for the horizontal line, but looked how the diagonal runs, whether it could form a triangle somewhere. . . . I never have the whole image, but only - well, I often lose the whole image. I trace lines, of which I should already know how it runs. . . . Then I felt reminded of the symmetry, and the thought comes up that there have to be four of them.*

The subject has an idea – we want to call it a prototype – of a certain type of triangle which only half of the subjects found. It is being formed by one of the vertical sides, an adjacent diagonal, and the horizontal line which is further away from it; this description would already suffice to deduce a triangle. But for the subject, this is only a hypotheses so far which he tries to instantiate: he traces the diagonal and tries to coordinate it with the horizontal line. But he cannot simply see this line and the intersection. Instead, he makes the horizontal line up again from the initial figure, and coordinates the intersection with the diagonal. Another interesting aspect about his description is the intuition of symmetry, which provides a heuristic to look out for more triangles.

We want to briefly paraphrase our findings. Focusing on the assemble operations, we can see that different perspectives are being assembled over and over. This process

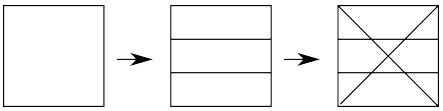

Figure 2. Another figure that can be easily synthesized. The task though consists in (mentally!) recognizing *how many triangles the figure contains*. This should be easy! Is it?. From Wiener [6].

is guided in a very fine-grained manner by a scaffolding of the task structure (triangles) and on-the-fly produced prototypes and heuristics. We consider these to be ontologically equivalent, because all guide the assembly of elements of the initial figure and thereby provide a tightly meshed, heterarchical organization of mental operations. Because all of these guides path the way to assemble components, we suggest that an assembled perspective (or figure) is defined by one's orientation towards certain operations. Eventually, an assembled perspective confirms the instantiation of a certain property or relation.

3. Summary and hypotheses

Based on our exploration of visuospatial reasoning we want to propose a renewed interpretation of mental imagery. In imagery, one actively attempts to assemble an intended structure; but *one does not simply see it*. One can indeed think of the whole process as being constituted through operations of type (a) and (b) from above, but with a reversed order. We do not make up an image to see something, but we want to see something, and in the absence of a visual stimulus we have to construct an orientation that fits.

Is the visual metaphor of depictive theory – imagery as internalized perception – therefore invalid? This only holds if we think of seeing as a one-way, passive uptake of input. But in an account of active perception, like that of Rensink, he states that one looks at scenes in a structural manner: "scene representations are no [. . .] structures built up from eye movements and attentional shifts, but rather, are structures that guide such activities [7, p. 36]".

With regard to the top-down guidance that we found in imagery we are sympathetic to such a conceptualization of vision. But do we therefore have to rephrase and say, somewhat awkward, perception is externalized imagery? Based on our observations there would be a missing piece, which are the overly present dynamic processes of assembling figures. We think it holds that imagery has its ontogenetic foundation in intentionally guided visual operations, and go with Piaget [8] who envisages that based on ones representation of actions, one eventually develops the capability to substitute external through internal processes. The dynamic assembly of figures might be just the organism's way around the lack of available sensor stimulation.

Our results are preliminary. In order to further our understanding of assembly processes, we are currently executing a study on the development of spatial orientation in familiar environments. Our approach also shows promising relations to a recent study, where [9] investigated assembly processes in the apparently unrelated faculty of mathematic cognition.

References

[1] Ronald A. Finke, Steven Pinker, and Marthaj Farah. Reinterpreting Visual Patterns in Mental Imagery. *Cognitive Science*, 13:51–78, 1989.

[2] Stephen Michael Kosslyn. The Medium and the Message in Mental Imagery: A Theory. *Psychological Review*, 88(1):46–66, 1981.

[3] Zenon W. Pylyshyn. What the Mind's Eye Tells the Mind's Brain. *Psychological Bulletin*, 80(1), 1973.

[4] Barbara Tversky. Visuospatial Reasoning. In Keith J. Holyoak and Robert G. Morrison, editors, *The Cambridge handbook of thinking and reasoning*, chapter 10, pages 209–240. Cambridge University Press, 2005.

[5] Claire Petitmengin. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5:229–269, 2006.

[6] Oswald Wiener. Materialien zu meinem Buch Vorstellungen, 2000.

[7] Ronald A Rensink. The dynamic representation of scenes. *Visual Cognition*, 7(1-3):17–42, 2000.

[8] Jean Piaget. *The Psychology of Intelligence*. The International Library of Psychology, Philosophy and Scientific Method. Littlefield, Adams & Co., Paterson, New Jersey, 1960.

[9] Stefan Schneider. *The cognitive structure of the enumeration algorithm: Making up a positional number system*. Master's thesis, University of Osnabrück, 2012.