

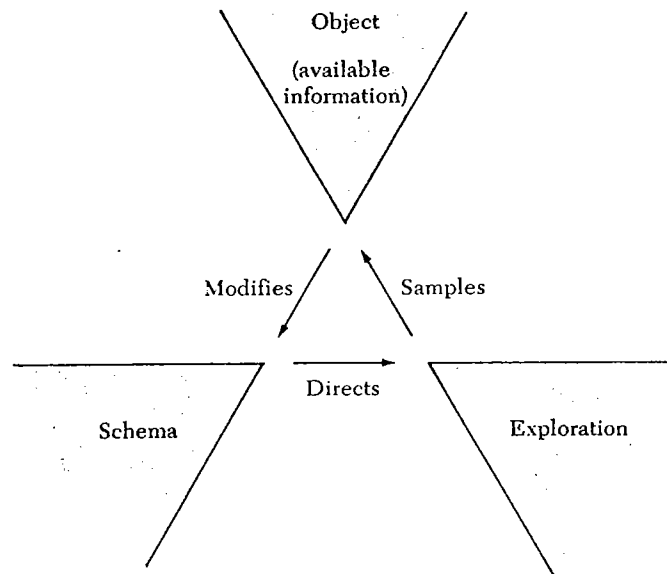
imagery which is, I believe, the most promising version of the 'perceptual activity' theory which has yet appeared. This is because Neisser is able to deploy a more complete notion of perceptual activity than we find in either Gregory or Gibson. The theory still has its problems and remains in many ways very sketchy. However, I believe that I can elaborate the theory in such a way as to make it apparent that its defects can, in principle at least, be overcome.

§II.D.5. Imagining and Cycling.

So far we have looked at three types of theory of perception and imagery which all stress the active nature of the processes. But the stress of each of these three types of theory falls on a different aspect of this activity. Eye movement theories stress the actual physical movements involved in seeing **where** things are, and the "hypothetical instantiation" theory stresses the determination of **what** things are, and concentrates on the activity which goes on inside the head. But surely no conception of activity can be complete which does not pay due attention to both its internal, mental aspect and its external physical one. Gibson's theory is concerned neither with **what** nor **where** but **what for**, and he is reluctant to say anything about inner processes. In fact he has little specific to say about the outer processes of exploratory

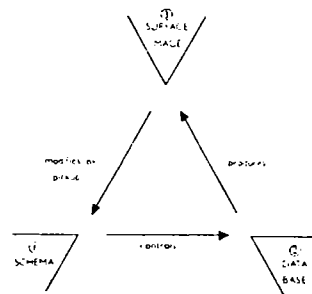
movement either. Most of his effort was devoted to redescribing the environment in such a way as to make apparent the potentially perceptible structures in the stimulus energies which could specify 'affordances'. However, although 'affordance' perception may be biologically primary, object perception certainly also occurs, in humans at least.

In his book *Cognition and Reality*, Neisser {1*} has tried to formulate a theory which lays a proper stress on both aspects of activity. Also, Neisser accepts Gibson's direct realism but tries to give a theory which will account for object as well as affordance perception. The fact that he allows himself to postulate mental structures (which are **not**, however, 'encodings') means that he can account both for object perception and imagery much more comfortably than can Gibson. As Gibson himself insists, the information in the environment must be actively looked for, and we surely cannot look for everything at once. Also, if Gibson is at all on the right track, there must be an enormous number of possible 'invariants'. To look for them at random would be horrendously inefficient. It thus seems reasonable to propose that there must be some sort of control structure, presumably located in our brains, which makes sure that our perceptual activity is arranged efficiently. It will tell us what features of the world (an 'invariant', it seems to me, can be treated as 'higher-order' feature {2*}) it will be worth looking for next, on the basis, presumably, of those which we have



The perceptual cycle.

Figure II.D.5_1
(Reproduced from Neisser [1976a].)



The imagery cycle illustrating 1, the image schema, which controls 2, the abstract data base, which produces 3, the surface image, which is picked up by 1, the image schema.

Figure II.D.5_2
(Reproduced from Hampson & Morris [1979].)

already picked up and our past experience of what tends to go with them. Such regular clusters of features in effect constitute objects. Neisser calls the control structure appropriate to a particular sort of object the "schema", and he regards perception as a cyclical interaction between three elements: the relevant schema; the exploratory behaviour of the perceptual systems; and the actual things in the external world. The schema directs the exploratory behaviour in its examination of the environmental object, and the resulting findings react back on the schema, causing it to be modified, so that it 'learns' from experience and becomes progressively more efficient at directing the exploration. Neisser calls this tripartite process the "perceptual cycle" (see figure II.D.5_1).

The exploratory behaviour is probably best conceived of as the making of specific tests for the presence of particular invariants, with the schema determining which test it will be most useful to make next. This is the sort of interpretation put on Gibson's theory by McArthur {3}, which, if he is right, opens up the possibility of a fruitful interaction between Gibsonian and computational theories, where they have more usually been seen as strongly opposed to one another. The idea of actively testing for features is fundamental to much of the work on 'computer vision', and McArthur argues that such research can thus be construed as an (unwitting) contribution to the Gibson's project of constructing an "ecological optics", a description of the environment in

terms of visually perceptible invariants {4*}. Note that Gibson himself does occasionally speak of perceptual "tests" in something like this sense {5} - although, of course, he does not have the notion of a schema controlling them. There must be many possible tests, however, and it is very hard to see how such a multiplicity could be effectively deployed unless under the control of some such structure as the schema. 'Computer vision' programs to deal with all but the simplest and most stereotyped of scenes are generally thought to almost certainly call for an element of such so called "top down" control. One of the most influential conceptions amongst AI researchers of how such control should be structured is Minsky's theory of "frames" {6*}, and Neisser {7} acknowledges a significant parallelism between Minsky's "frames" and his own "schemata" {8*}.

It is not stressed by Neisser, but it is of particular significance for our overall project, that there is no reason to think that a schema can call up exploration by only one perceptual system. If something appears to sight in a certain way then the most appropriate perceptual activity to perform next may well be to feel it, or to sniff it, testing for a certain texture or smell. On the other hand, spatial information, whether obtained on the basis of sight or of touch, may very well give rise to just the same response or continuation of exploration in either case. In this way the schema functions very much like the "common sense" of Aristotle, dealing with the spatio-

temporal and integrating the deliverances of the various "special senses" (9*). In fact, I believe that Neisser's theory is the nearest thing we yet have to a description of a mechanism adequate to the functions which Aristotle assigned to the *sensus communis/phantasia*. As we are about to see, Neisser argues that the schema not only controls perceptual activity, but is also responsible for imagery.

Because of the way the schema directs perceptual activity, it can be considered as a structured set of perceptual anticipations or expectations. Given what we have perceived up to now, it tells us what we (or rather, it tells our perceptual systems what they) are likely to be able to perceive next, or what will be most significant if it can now be perceived - in general, what it will now be best to look for. In this way schemata bear some resemblance to the "perceptual hypotheses" of Gregory, but by stressing the schema's function as a control structure we avoid the temptation, inherent in the term "hypothesis", to think that what is involved is a set of descriptive propositions, to be accepted or rejected as a whole, and which constitutes the perceptual experience (10). True, the schema does embody a form of mental representation, but it is an "interactive" representation (in the sense of Bickhard & Richie (11)), a representation inasmuch as it is a structure adequate to appropriately direct our interaction with the environment. It is not an "encoding" or set of "encodings" which has to be read by some inner interpretive agent. Obviously, perceptual expectations will

not always be fulfilled, but the fact that we do not find some feature which we are looking for is perceptual information in itself, and normally we will absorb this, re-order our expectations accordingly, and carry on looking for what is now the next most likely or most significant feature. If we consider that this re-ordering becomes built into the structure of the schema then we see how perceptual learning can take place on such a theory, how our ability to quickly recognise things improves as we grow up.

But there is no reason to believe that it is not possible to call into action a schema which embodies perceptual anticipations entirely inappropriate to our current environment. We could, say, make ready to pick up the invariants or features specifying a dog, when there is no dog present. According to Neisser, it is in just such a situation that we are imagining the dog, experiencing an image of it:

I believe that the experience of having an image is just the inner aspect of a readiness to perceive the imaged object.

{12}.

Provided one has acquired the necessary skill, he believes, one can call up such schemata, such inappropriate perceptual expectations, more or less at will:

To imagine something that you know to be unreal, it is only necessary to detach your visual readinnesses from your general notions of what will really happen and embed them in schemata of a different sort. When you have an image of a unicorn at your elbow - while quite certain that unicorns are purely mythical animals - you are making ready to pick up the visual information that the unicorn would provide, despite your being fully aware that your

preparations are in vain. I do not know whether animals are capable of these schematic rearrangements; certainly children take a good deal of time to master them. In any case, they seem to be easy enough for adults. According to the hypothesis proposed, however, even counterfactual images are still potentially functional anticipations. If the right kind of unicorn were suddenly to appear at the elbow of a person who was imagining one, he would see it more quickly and easily than if he had been imagining anything else.

{13}.

I must say that the last assertion made in the above quote sounds rather dubious to me. I feel that I would be caught just as unprepared, perhaps more so, if a unicorn suddenly appeared to me whilst I was imagining one, as I would be if one appeared when I had been thinking of something quite different. Nevertheless, the assertion does have the advantage that it provides some leverage for experimental testing. Unfortunately the experimental support for it seems to be quite weak. Neisser {14} cites the work of Posner and his colleagues {15}. In these experiments the subjects were shown, say, a capital A, and they could more quickly recognise another capital A flashed before them as being the same letter than they could recognise a lower case a as 'the same'. That is to say, letters with the same visual form were more easily identified as the same than were letters with merely the same 'meaning'. This is taken to imply that the original letter has set up some specifically visual expectations, which facilitate recognition beyond the mere knowledge that certain shapes represent certain sounds. However, this effect is very short lived. With a slightly longer delay

between the two letters to be compared, the advantage of similar visual form disappears. In fact, it is not at all clear that the processes uncovered by these experiments involve imagery in the usual sense at all. Indeed, the work of Perky {16} and Segal {17}, which we discussed in §I.C.1, seem to demonstrate the reverse of Neisser's suggestion; imaging seemed to interfere with visual detection in these experiments, and all the more so if the image and the actual (faint) projection were of the same thing. However, in these experiments the subjects were simply told a **type** of thing to image, there was no guarantee that the images formed would be identical to the actual stimulus presented. Neisser {18} dismisses Perky's results as being the product of "demand characteristics".

In a more recent experiment by Peterson & Graham {19} subjects were shown a line drawing and were told to try and maintain an image of it whilst trying to pick out from a series of slides containing "visual noise" (i.e. a random tangle of lines) the one also containing the original picture. The subjects who thus saw and imaged the picture were able to do this better than control subjects who had not been given any such previous cueing. However, subjects who had been merely **verbally** cued as to the nature of the picture, and who were not instructed to form an image, did just as well as the imagers. Therefore, this does not really seem to be specifically an imagery effect. There **was** an imagery effect in this experiment {20}, in that the subjects who formed images performed **worse** than

uncued control subjects when the picture they imaged was different from that in the slide. In these case, the verbally cued subjects did not show such an interference effect. However, the significance of this result is not very clear.

Some more recent work by Finke {21} also provides rather mixed evidence. He had his subjects imagine simple bar patterns in various orientations, and gave them the task of indicating whether similar bar patterns flashed before them were either vertical or horizontal. These subjects proved to be faster than non-imaging control subjects when the presented bar pattern was at a similar orientation to the imagined one, and slower when the angular difference was large. However, in a further experiment where the subjects merely had to indicate when some pattern or other was flashed before them (i.e. it did not matter which pattern it was) imagery seemed to interfere, increasing the response times and error rates, and this interference was greatest when the pattern presented exactly matched that imagined {22}. Such experimental data cannot really be said to support Neisser's hypothesis, although they do not refute it either. More work is surely needed in this area.

In any case, for Neisser, it seems that to have a mental image of something is to have an anticipatory schema as it were 'activated' in the absence of the relevant object. Indeed, he says that "images are anticipatory

schemata" {23}. This suggestion has been criticized by Hampson & Morris {24*}, and their criticism has particular force because they seem essentially to accept the perceptual theory, the notions of schema, perceptual exploration and the perceptual cycle, which goes with it {25*}. They note the similarity between Neisser's account of imagery and that of Ryle {26*}, and draw on the criticisms which Hannay {27} makes of Ryle. In particular, they point out that unfulfilled expectations or anticipations normally give rise to disappointment or surprise. Neisser seems to identify imagery with unfulfilled perceptual expectations, but images are not standardly accompanied by these emotions. Neisser's reply to this objection is that, in fact, not all unfulfilled do give rise to disappointment or surprise. For example:

The American military defense system is (supposedly) always ready for a Soviet sneak attack - it is deployed in anticipation of such an attack, one might say - but it experiences no surprise when the attack fails to materialize day after day. In short, my use of "anticipation" does not intend as much surplus meaning as Hampson and Morris suppose. What I have in mind is a specific state of readiness for specifically **perceptual** information. Images are anticipations, but not all anticipations are images. {28}.

However, it is not immediately apparent that unfulfilled **perceptual** expectations are of the type which would not occasion surprise. If I expect to see a tree round the next corner on my familiar route home, but I do not find it there, then I will be surprised. In truth, the anticipations of something like the American defence system are of a rather special and complex kind. The American

military's expectation of a Russian attack in the near future is, surely, (one hopes) normally very low. However, such an attack would be so momentous if it did occur that they deem it prudent to constantly behave as if their expectation of its imminence were very high. Of course, in our discussion of Ryle {29} we found that he linked imagery with as if behaviour, or, in other words, pretending. The problem which we ran up against there was that it did not seem, on his account, that perceiving was a form of activity or behaviour at all - and if it were not behaviour then one could hardly pretend to do it. For Neisser seeing is certainly an active process, but it is not made at all clear that imagery is likewise active. In a later paper, Hampson & Morris point this out:

In fact, when dealing with imagery, he [Neisser] fell into the same trap against which he warned when discussing perception. (...) This allowed him to consider imaging as a state of the person rather than as a process. The result is an essentially inert set of perceptual anticipations devoid of any information input, rather than a system of information uptake, schema modification and hence knowledge production. We suggest that if the relation between imaging and perceiving is taken seriously, and Neisser's general perceptual theory is accepted, then it becomes both useful and necessary to speak of imaginal as well as perceptual cycles. {30}.

This, I think, is a fair point. However, the model of an imaginal cycle which Hampson & Morris {31} produce seems to me to be most unsatisfactory. They claim to provide an account which combines the best aspects of analogue, 'propositional' and 'activity' theories, but to my mind they succeed in capturing all the worst rather than the best aspects. Their model of the "imagery cycle" (figure

II.D.5_2) involves a "surface" (i.e. quasi-pictorial {32*}) image, which is produced from a 'propositional' "data base". So far this theory does not differ significantly from that of Kosslyn, except that less detail is provided about the nature of the "surface image". What is added is a schema, which is to control the production of the image from the data base, and is to be modified by the information it picks up from the image. The notions of "surface image" and 'propositional' "data base" here seem to me to be subject to all the difficulties we found them to be subject to in the previous two chapters. All that is really added to Kosslyn's account is the concept of schema, but it is not obvious that the term signifies anything coherent here. It seems to function as a 'mind's eye'. Despite the name, the Hampson & Morris schema does not seem to be the same as Neisser's: **that** controlled exploratory behaviour (such as eye movements), which then picked up information from the ambient stimulus energies; **this** apparently picks up information directly itself, from an internal mental image, and the notion of perceptual exploration (which linked Neisser's theory to Gibsonian realism) seems to have been lost.

The account given by Hampson & Morris {33} of Neisser's imagery theory as it appears in **Cognition and Reality** {34} seems to be true to that text, and if it is right then many of their criticisms seem justified. Although Neisser clearly there regards perception as active, imaging still seems to be essentially passive, the

mere 'having' of some schema. However, I wonder if that is what Neisser truly intended. Some remarks from his paper which replies to Hampson & Morris seem to imply that he regards schemata as active in themselves:

Schemata that have picked up information about local details of objects (say, edges and corners) are not passive conduits for data that some homunculus will eventually interpret. They do not simply send reports of their discoveries to higher levels, but engage in perceptual cycles of their own. (...)

When an active schema (at any level) fails to find the information to which it is attuned, the character of its activity changes. The state of the perceptual system then becomes rather a peculiar one, and gives rise to a correspondingly peculiar mental experience. This peculiar state certainly is not perception itself: anticipations are going unfulfilled, and the characteristic information-driven changes of schemata are not taking place. Nevertheless it may be said to resemble perception inasmuch as the perceptual schemata themselves are active. It is under these conditions, I submit, that we have mental images. {35}.

We must, I think, reject the implication here that we have imagery whenever perceptual expectations go unfulfilled (which must be happening all the time). Surely we could only be said to be imaging if our perceptual anticipations were **systematically** going unfulfilled. But to make this objection is perhaps only to point to a slight carelessness of expression on Neisser's part. The important question to ask is: what is the force of "active" here? I am afraid that Neisser still does not make this entirely clear. However, the normal mode of activity of a schema is in controlling exploratory behaviour during the operation of the perceptual cycle. Indeed, the very definition which Neisser gives of "schema" is in terms of

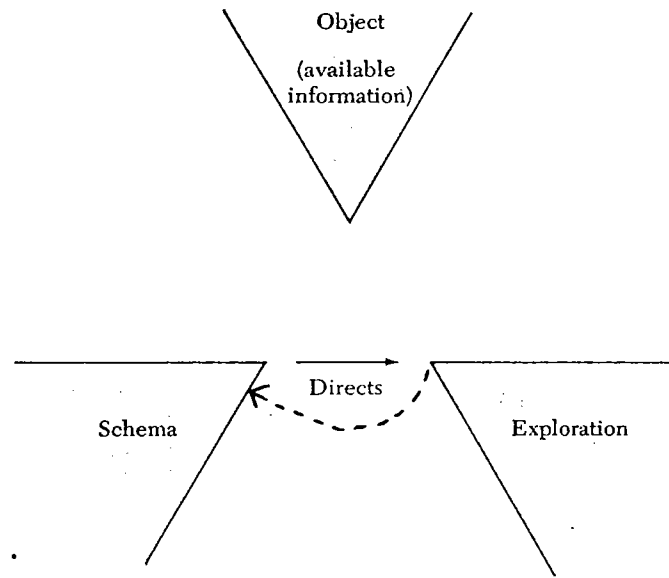


Figure II.D.5_3
 (Adapted from Neisser [1976a].)

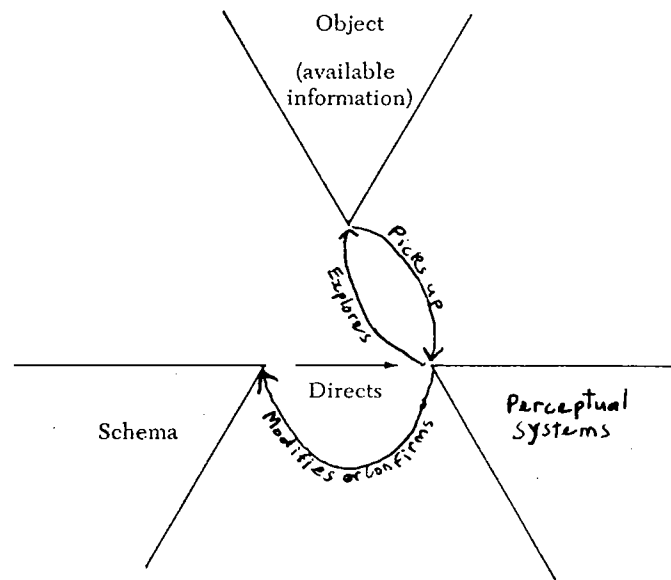


Figure II.D.5_4
 (Adapted from Neisser [1976a].)

