How does the world get into our head? Is what we perceive really true? What happens in our brain when we see, hear, smell, or touch? How do we find our way in the world? Research groups at the MAX PLANCK INSTITUTE FOR BIOLOGICAL CYBERNETICS *in Tübingen are using so-called "psychophysical investigation methods" to answer these questions. This means that sensory impressions of test subjects are manipulated and their perceptual reactions are studied. Colleagues working with* HEINRICH H. BÜLTHOFF *in the* PSYCHOPHYSICS DEPARTMENT *are attempting first and foremost to trace the principles of perception and the co-ordination of movement in space.*

Perceiving is more than Seeing

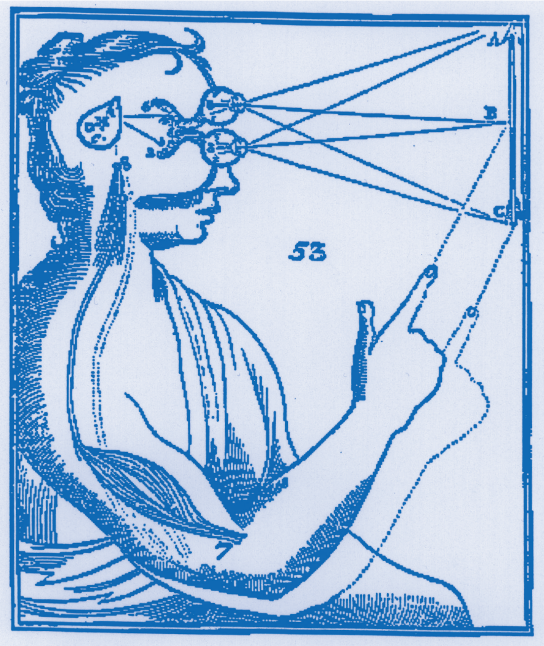

FIG.: MPI FOR BRAIN RESEARCH

Our whole nervous system is designed to interact with the environment. As a result, we must collate the various bits of information that come from our senses. In the end, this enables us to carry out motor actions successfully. For a long time, there has been speculation as to the ways in which this integration of sensory information takes place. René Descartes, for instance, assumed that the pineal body was the "seat of the soul" – it was here that all information was supposed to come together and where motor actions were initiated.

Perceptions are the basis for our understanding of the world. For the act of seeing, for instance, it is not enough to simply project the outside world onto the retina and from there onto a sort of screen in the brain. Perception is not just the passive recording of sensory stimuli, but rather an active mental reconstruction of the real world that surrounds us. As a result, our brain dismantles what appears on the retina into highly abstract bits of information that correspond in the final analysis to a sort of symbolic representation of the outside world; a self-made model of the world. What we perceive really depends essentially on unconscious cognitive decisions and conclusions. The brain usually makes these on its own without us having to bother. In doing so, it uses previously collected knowledge, experience, expectations, and prejudices.

Once the brain has learned something, it is often no longer especially bothered about the actual realities. We cannot force our way to the outside through the nerves in order to reach the true reality, to get to Kant's "thing itself". Everything that gets into our consciousness from the outside is transmitted through the clearing centres of our sensory organs – shapes, faces, and movements. But also seemingly absolute things such as matter, space, time, and even the I we experience ourselves, are, as we experience them in everyday life, something that is artificial, self-made and constructed by our brain.

It is quite clear that our brain uses an unbelievable amount of detailed information which it has at its disposal, amongst which is information we are not even aware of. How does the brain manage to select the "correct" information and pull it together? What strategies have unfolded during the course of evolutionary development in the interplay between the brain and the senses that allow us to find our way in this world? In three projects at the Max Planck Institute for Biological Cybernetics in Tübingen, light is being shed on parts of this sweeping question. They form spotlights on the way to looking over the brain's shoulder as it works.

SIGHT AND MOVEMENT – ALLIES IN THE CONTEST

We have already started to become uneasy on the way up. However, we only experience the tingling in our stomachs when the roller coaster reaches its highest point before finally plunging downwards at top speed from one loop to the next. If it gets too bad, then there is only one thing left to do: close your eyes. And if it is not too late already, then perhaps that way you will be able to leave your dinner where you put it with enjoyment when you ate it. It is the interplay between visual and

Fig. 1: Test subject on motion platform. The perception of intrinsic movement and sight are decoupled and manipulated separately. As a result, the researchers obtain feedback as to how the world is represented in the brain.

vestibular information, i.e. information originating from our vestibular organ in the ear, that causes our body trouble. The body can usually cope quite well with us not staying in one place and moving about in space. The eye (or rather the brain) has learned that a movement of the image on the retina need not necessarily mean that the surroundings are moving too. On the contrary, the central nervous system asks the muscles whether it is a spontaneous movement that is the cause of the “film” on the retina, and remaining satisfied if the data agrees as normal. However, if this is not the case, the brain sounds the alarm.

How is the complex visual-vestibular interplay organised? Which system, eye or vestibular organ, is predominant? How are these two sensory modalities offset? As yet, there are only a few investigations existing in scientific literature which relate to complex situations that are close to reality. The change in perceived intrinsic movement in space, “spatial updating”, is the research topic of information scientist Markus von der Heyde and physicist Bernhard Riecke. Three cognitive outputs are significant for this. On one hand, memory, both short-term and long-term memory, plays a role alongside immediate perception. On the other, how we behave, i.e. how we move in space, is crucial.

As a result, von der Heyde has set up a motion platform with which he intends to record and quantify the contribution of both sensory modalities to our overall perception (fig. 1). The test subjects sit on this movable apparatus and are moved linearly or rotated through space. The perceived intrinsic position in space is measured by asking the people to estimate their position in the room. In the tests which follow, the real visual stimulus is replaced by a virtual image. By way of a “head-mounted display” (HMD), a sort of “data helmet”, both eyes are presented with an image that is linked to the person’s movement. This image may correspond to the one a person with a restricted field of vision would have (in the real version of the experiment, this is achieved by using sight-restricting glasses); however, the virtual image may also be decoupled from the actual movement and offer completely different visual stimuli. The test subjects produce the same outputs under both real and virtual conditions as long as the scenes match up. This fulfils a basic condition in order to obtain meaningful results even under altered virtual conditions. In these initial experiments, it was possible to confirm the dominance of the visual system already known from literature.

The question the two researchers from Tübingen are now asking is whether there are also conditions under which we concede greater significance to the sense of balance than to our optical sensory impressions. It is precisely virtual reality methods that are so well suited to investigations of this type, since researchers can “change” the world at will by manipulating the appropriate parameters. For example, they can decouple the visual and vestibular sensory information. Whilst the eyes are offered a static image, like one would see if one looked straight ahead without moving, the test subjects are turned through 90 degrees in space. The whole thing is now supposed to be simulated in an optically more varied environment with photo-realistic versions of Tübingen’s market place, and then compared to the real experience in the market place.

It is not only the eyes that contribute to spatial perception – in cases of doubt the brain also uses other sources of information to reconstruct the three-dimensional environment. A year ago, Marc Ernst from Heinrich H. Bülthoff’s work group, together with collaborators from the University of California (Berkeley/USA), was able to demonstrate that even the hands can serve as “visual aids”. In their experiments the researchers discovered that feeling with the fingers can significantly and effectively influence visual perception.

Seeing Correctly Means Feeling Through Your Fiingers

To be able to perceive the world in three dimensions, the visual system uses different optical stimuli from which it is possible to reconstruct the spatial position and structure of objects observed. These stimuli include shadows, perspective distortions of the object, and the irregularities between the retinal images due to the different positions of the eyes known as disparity. The brain’s task is to combine the information available into an overall picture whilst at the same time weighting the various sensory impressions according to their reliability. Thus, a plausible signal contributes more to the reconstruction of the spatial image than a less plausible one. Therefore, the combination of the sensory impressions determines how we see our environment. But how does the brain know which stimulus it can trust most?

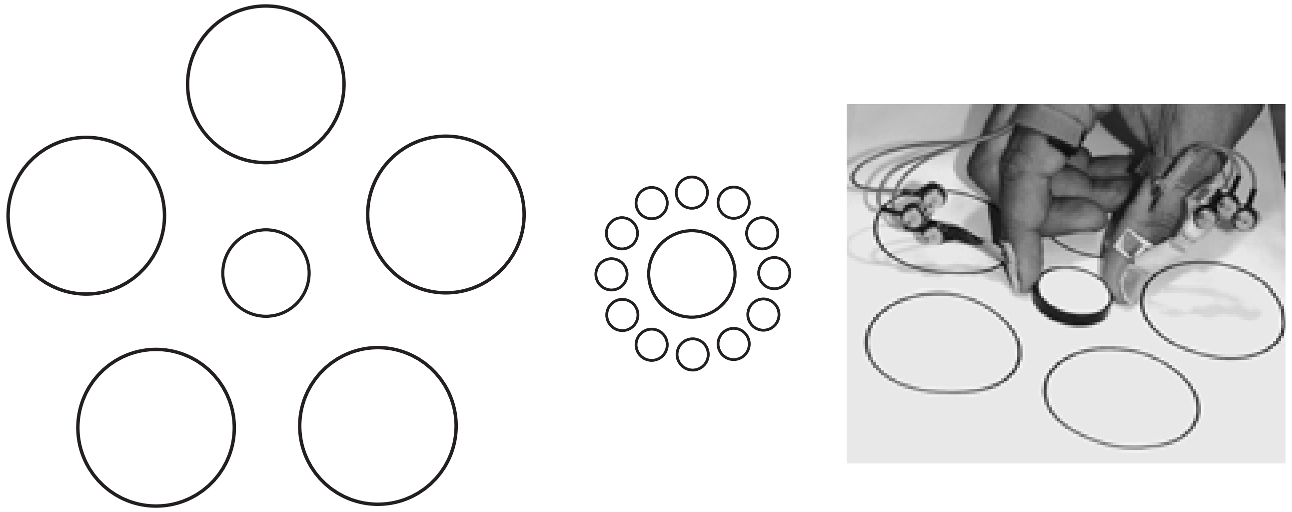

Fig. 2: The Ebbinghaus illusion. The test subjects look at a table with one circle of large and one circle of small rings. In the centre there is an aluminium disk of identical size in each case. In the grasping experiment, the test subjects are asked to pick up the disk. Infrared-emitting diodes on the thumb and first finger make it possible to measure the grasping distance between the fingertips. In the perception experiment, the test subjects are supposed to match a reference circle on a monitor to the size of the aluminium disk. Volker Franz’s results show that, quantitatively, the motor action is duped just as much as the visual system – in both experiments the tests subjects estimate the disk surrounded by the large circles to be smaller than that surrounded by small circles.

Fig.: MPI for biological Cybernetics

To clarify this question, Marc Ernst designed an apparatus that deluded the test subjects into accepting the image of an inclined surface. The inclination of this surface was transmitted by means of two optical parameters. One by way of a texture gradient, generated by means of perspectively converging vanishing lines, the other by way of a disparity difference based on the stereo effect due to binocular vision. It was possible to alter both these signals independently of each other and differently to each other so that in each case, they displayed different inclinations of the surface – thus bringing the observer’s brain into a conflict situation. The results show that the brain resolved the contradiction between the two stimuli by weighting each signal and thus came to a compromise, an interim value. The test subjects actually perceived an evenly inclined surface.

In the next stage of this experiment, the participants in the test not only saw the inclination of the surface, they also felt it. In order to achieve this, a virtual dice that they could push with one finger was projected onto the image of the inclined surface. The resistance offered to the finger when pushed in one direction or another was simulated by a computer. This finger-felt inclination was then adjusted respectively in the experiment, so that it corresponded to one of the two optical stimuli that was either the inclination that was transmitted by the disparity difference or that generated by the texture gradient. This contact really did influence the weighting of the visual signals. The signal that was finger-felt was given a higher weighting in each case, i.e., it was obviously considered more reliable by the brain. “One and the same surface is perceived differently after finger-feeling than in the previous purely optical test”, says Marc Ernst. “For the observer it appears to be inclined more strongly in the direction of the optical signal that agrees with the slope felt. It is still possible to demonstrate this effect up to a day after the touch test.”

Small or Large Circles – Everything is Simply an Illusion

The researchers were able to confirm what takes place unconsciously and continuously in everyday life when we reach for objects or climb over steps. In addition to visual information, the brain also uses tactile check-back signals from the hands and legs to obtain a spatial impression. In this case, information from the motor system is matched with that from the visual system and can affect the weighting of optical stimuli almost like a teacher.

These results rock the foundations of theoretical considerations that

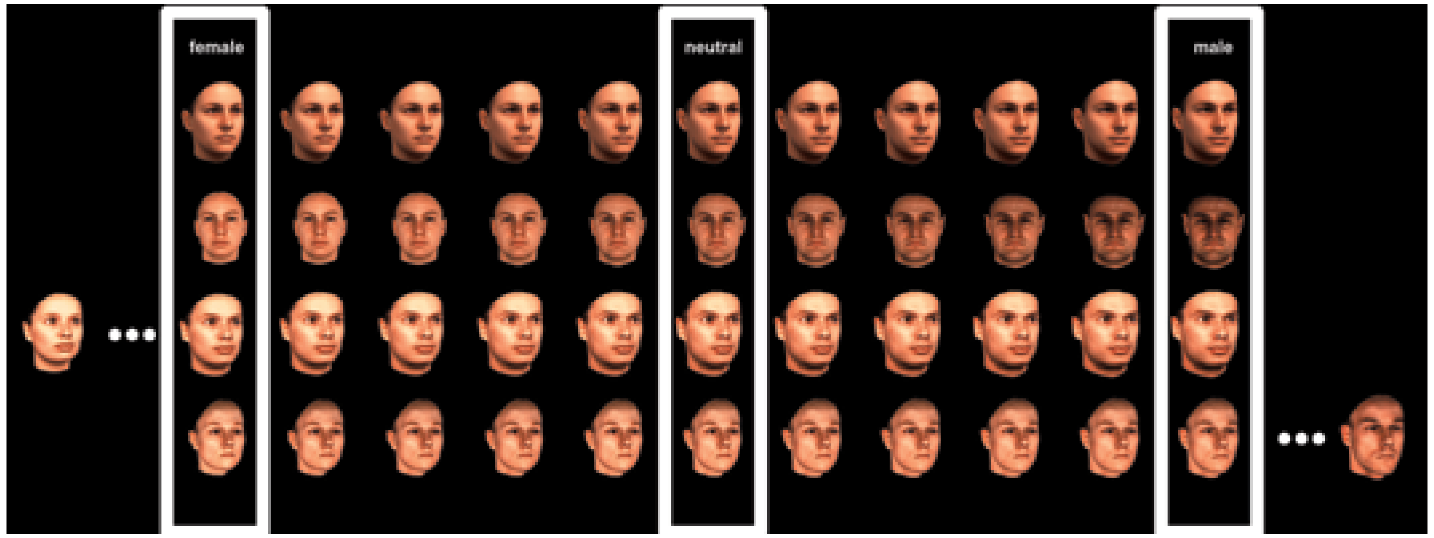

Fig. 3: A neutral average value face was calculated from 100 male and 100 female faces (top row). On the left of the picture is the female average value face, on the right the male, with the series of images showing the intermediate stages in each case. The second row shows morphs of the faces of two people of different sexes. The faces in the third row originate from a female person and were masculinised along the row of images from left to right (the original is in the "female" box). The faces in the fourth row originate from a male person and were feminised along the row of images from right to left (the original is in the "male" box). In each case there is also one image of a "super" female and one of a "super" male face (to the left and right respectively of the three points).

have emerged primarily on the basis of investigations using brain-damaged patients in whom one or other branch of perception is disturbed in its function. There are supposed to be two independent paths in the brain that play a part in the perception of objects – one path is responsible for the conscious and immediate perception of stimuli ("perception"), whilst the other checks the object's relevance of action ("action"). Scientists from the University of Verona (Italy) and the University of Western Ontario (Canada) appeared in 1995 to have been able to confirm this theory neatly in healthy test subjects. Their experiments seemed to confirm that the so-called Ebbinghaus illusion, a well-known perceptual illusion, only leads us astray when we see it and not when we reach for it. Although the two circles in the middle (fig. 2) were considered to be different sizes, the test subjects selected by Aglioti, DeSouza and Goodale each opened their hand equally wide in both situations when they were asked to touch the circular disc.

However, for Volker Franz from the MPI psychophysics work group there remained doubts not least because he and the two professors, Manfred Fahle from the University of Bremen and Karl Gegenfurtner from the University of Magdeburg, believed they had discovered methodical deficiencies in the study referred to above. They repeated the experiment on the spot under modified conditions and in actual fact, Franz obtained exactly the opposite results. On grasping the Ebbinghaus figures, the test subjects from Tübingen opened their hands in different widths, as (incorrectly) told to do by their visual system. Therefore, the motor action is misleading and from a quantitative point of view it is similar to our sight. The "rediscovered" illusion therefore led to disillusionment. The apparently wonderful evidence for the hypothesis of separate processing paths had been ruined.

Nevertheless, the psychologist from Tübingen plays the role of the spoilsport with composure. "In science one has to be prepared for surprises like these. It is for precisely this reason that the repetition of fundamental experiments is so important. The result does not necessarily refute the hypothesis, but for the time being merely strips it of one basic proof that is, admittedly, seductively simple." In the meantime Volker Franz is continuing his work in collaboration with Prof. Odmar Neumann's work group in Bielefeld. In masking experiments the researchers offer their test subjects two slightly different visual stimuli in quick succession. The two stimuli consist in each case of the inferred outlines of a square and differ in size as well as in the orientation presented (see diagram at bottom of page 48). The stimuli are shown one after another so rapidly, in a technique known as "metacontrast", that most of the test subjects say that in each case they only noticed the second stimulus. However, if one asks the same test subjects to press a button as a result of the stimulus and either right or left depending on the orientation, something strange happens. The test subjects react more quickly if the first stimulus (which they have not consciously perceived) has the same orientation as the second. Conversely, they need longer if a square with a different orientation is offered before the second, now consciously perceived stimulus. Therefore, although they cannot give any information about the first stimulus, it evidently plays a part in the execution of an action – the "action" path is affected without the "perception" path apparently taking any notice of it. If these initial results of the masking experiments were to be confirmed in further tests, then this would support the Milner-Goodale theory of the two independent processing routes, and the scientists' world would be back on course again.

Bringing images to life is not only the aim of Hollywood's dream factory, but is also on the mind of many a Max Planck scientist. Scientists are asking how it is possible to reconstruct the complete spatially characterised profile of a person from a picture or a photo of them. The problem – referred to as an underdetermined problem – is that of reconstructing a three-dimensional object from a purely two-dimensional document. To be able to achieve this, one has to fall back on empirical knowledge, the knowledge of what objects of this type (such as faces) look like if one looks at them from a different perspective.

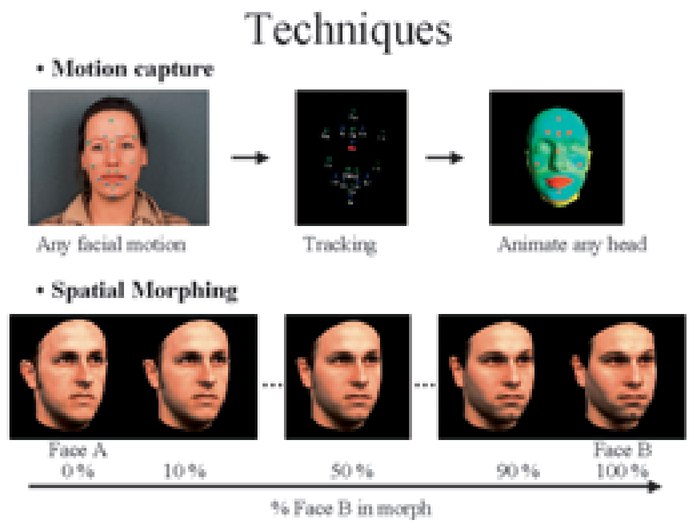

Fig. 4: The most up-to-date computer technology was used to investigate the relative significance of shape and movement information in the recognition of faces. Using the Famous Technologies' face animation system, it is possible to record facial expressions and transfer them to any number of computerised models of human heads. Using the "morphable model", designed by scientists in Tübingen, it is possible to systematically change the shape information in a face. Combining both techniques makes it possible to blend shape and movement information of faces at will.

Morphs – Artificially Manufactured Heads

Evidently, our brain has no difficulties doing this since it solves problems of this type on a daily basis. Our eyes constantly transmit nothing more than illustrations of the outside three-dimensional world to a two-dimensional retina – our brain has learnt to establish the third dimension from this data. In this case it establishes the open parameters from experience – we have already seen countless faces in our lives and know what they should look like, even if the detailed information necessary to do this is not available to us immediately.

In 1999 the physicist Volker Blanz working in the group in Tübingen succeeded in imitating the brain. He developed an algorithm which he used on the computer to reconstruct a complete head from the two-dimensional image of a single face. To do this, however, he first had to "supply" the computer with the foreknowledge of how faces in our world should look. Prior to this, Thomas Vetter and Niko Troje of the same work group had already developed a method making it possible to model and record faces in three dimensions. Using a special device, the scientists scanned as many faces as possible, viewed from all sides, and fed the data into a computer. In the meantime, the database of faces in Tübingen has increased to over 200 heads, strictly divided into half male and half female faces. Based on specific characteristics such as the position of eyes, nose, mouth, ears etc. the computer is now able – with the help of a mixed image sequence called Morph – to establish an average face from the over 200 recorded data heads (fig. 3), the so-called average value head (generic face model).

This average value head was the starting point for the algorithm that Volker Blanz created to reconstruct a spatial face profile from a model pic-

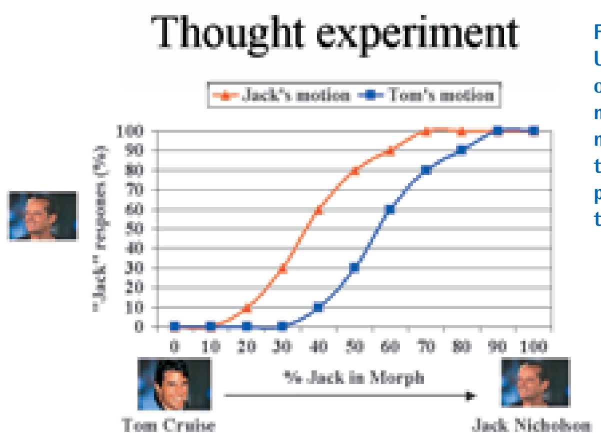

Fig. 5: Thought experiment. Do characteristic movements of faces play a part in recognition? Using a "morphing sequence", which reflects a successive change in the facial expressions of Tom into Jack, test subjects are asked which faces in this sequence look more like Tom or more like Jack respectively. If facial expression plays a part in recognition, one would expect more "Jack" answers if the face moved like Jack's and vice versa. The effect of the information by way of facial expression ought to be particularly marked in the region of the "fifty percent morph" as the shape information is ambiguous at this point. The diagram shows the anticipated curves. In actual fact the experimental findings confirm this expectation.

ture. To do this, the average value head must be photographed under the same conditions (such as position, lighting etc.) as the model picture. This photo is compared pixel by pixel with the model, and a sort of "average value" is created using a special process. The process is complicated because one has to pay attention to the appropriate correspondence – naturally one can only compare an eye with an eye and not with the mole next to it. The researchers then transfer the picture thus obtained back to the three-dimensional head. To do this, they simply reverse the process which they used in the beginning to make a photo from the three-dimensional average value head. The result even surprises professional film makers. Audrey Hepburn looks at us, turns her head and we can look at her profile and even set her face in motion. All this is possible even though we only had a photo of her as a model.

In the meantime, the database of faces and the "morphable model" based on it has become the starting point for other exciting scientific questions. Until now, only a few scientific studies have been concerned with the movement or animation of faces and the effect of this on recognition of people. The mathematician and biologist Barbara Knappmeyer would like to discover what role facial expression plays in the recognition and identification of faces. Everyone knows the inimitable smile of famous actor Jack Nicholson. Would we recognise this smile again if it were not Jack Nicholson who was snarling, but Tom Cruise? Therefore, in collaboration with the British psychologist Ian Thornton, Barbara Knappmeyer has brought movement into "her" faces. First of all she highlighted characteristic points in the faces of test subjects (fig. 4) and filmed them whilst speaking, chewing, laughing, frowning, and creating other facial expressions. She played the film sequences into the computer which carried out the rest. It was then possible to link every face from the database with every characteristic movement of a film sequence – "Tom" and "Jack", both randomly selected virtual faces from the database, can now exchange their smiles at will.

Moving Features – Recognition in the Brain

In the first phase of the experiment the test subjects were allowed to become familiar with the moving faces of Tom and Jack. In this case every face was given its appropriate expression. Once the test subjects had finally made friends with their virtual companions, the test phase began. A chewing face, either Jack or Tom or even a mixture of Jack and Tom, a so-called morph, appeared on the screen. On a continuous scale, the morph may lie nearer to Jack, nearer to Tom or somewhere in the middle (fig. 5) – that is the dividing point on the scale between Tom and Jack at which the test subjects indicate that they recognise Jack instead of Tom in the face. The researcher then asked herself the question as to whether the dividing point shifted towards Jack if the morph was simultaneously animated with Jack's characteristic chewing.

The result is clear. The characteristic movement of the faces learned during the training phase demonstrably influenced recognition of the face. On the way from Tom to Jack, Jack clearly came to light earlier if the morph moved like Jack, and conversely appeared later if it moved like Tom. "Thus it appears that the face recognition system in the human brain not only uses statistical information with regard to the face shape, but also the specific movement of faces," Barbara Knappmeyer commented on the initial results of her studies. In this case, the movement can be as varied and complex as the facial expressions of a person. Thus, the brain has learned to deal with such complex questions and in daily life can fall back at any time on an appropriate tool to identify the world around us. This all happens without our (conscious) help - and, like Alice in Wonderland, we watch in fascination when the cat disappears and only its grin is left behind.

RAINER ROSENZWEIG