Brain, mind, and limitations of a scientific theory of human consciousness [1]

Alfred Gierer
Max-Planck-Institute for Developmental Biology, Tübingen

Key words: brain / mind / consciousness / human / decodability / decision theory / quantum physics / Gödel / Heisenberg / Frayn

Abstract

Modern brain research related to consciousness has resulted in many interesting insights, for example into the neurobiological basis of attention and of language. In biological terms, human consciousness appears as a system's feature of our brain, with neural processes strictly following the laws of physics. This does not necessarily imply, however, that there can be a general and comprehensive scientific theory of consciousness. Predictions of the extent to which such a theory may become possible vary widely in the scientific community. There are reasons - not only practical but also epistemological - why the brain-mind relation may not be fully decodable by finite procedures. In particular, analogies with mathematical theorems of undecidability suggest that self-referential features of consciousness, such as multiple self-representations like those involved in strategic thought, may not be fully resolvable by brain analysis. Assuming such limitations exist, this implies that objective analysis cannot exhaust subjective experience in principle. A person's consciousness and will are accessible to external observation only within limits. In some respects, we do not even learn to know ourselves except by our actions. It thus appears that a scientific look at consciousness and the human mind, combining universal physicalism with epistemological scepticism, is not inconsistent with certain concepts of subjectivity that are current in the humanities, despite all the differences in the style and terminology of discourse.

[1] Preprint of a contribution to the symposium: Proteus im Spiegel - Kritische Theorie des Subjekts im 20. Jahrhundert (P. Geyer, M Schmitz-Emans, eds), 191-203, Königshausen und Neumann, Würzburg, 2003

The relationship between brain and mind is a genuine scientific problem

"It is is a very difficult task, in any respect, to form a solid opinion on the mind (psyche, soul)... It seems as though everything the soul experiences occurs only in association with a body... joy, as well as love and hate; in all of these cases something happens in the body as well... If this is the case, then these properties possess something material even in their very being... And that is, after all, the reason why the physicist is responsible for the investigation of the soul."

These sentences are 2300 years old, to be found in the introduction to the first systematic discussion of the mind-body problem, in Aristotle's 'De anima'. At their core they are still valid today: The mind-body problem is fascinating; it is difficult to solve; and revolves mainly around the relationship between corporeal and mental processes which are perceived in very different ways and expressed in terms of different concepts. While we can objectively observe and measure bodily processes such as speaking and blushing, and modern techniques allow us to detect brain activities as well, within our conscious mind we are directly exposed to our own state which manifests itself in the form of feelings, thoughts, intentions, memories, wishes, fears, and hopes, often independent of the senses and almost always without knowing the simultaneous physical processes in the brain. We cannot even feel that there are processes taking place in the brain at all - the ancient Greeks and other peoples often believed that the soul was in the heart. Conscious experience is not primarily given in spatio-temporal terms and therefore not directly subject to physical laws. Yet it can be expressed and communicated through speech and thus be subject to scientific discussion. Some expressions of mental states, such as laughter and crying are innate while others, particularly those associated with speech, are transmitted through culture. Using a thesaurus one is impressed by the sheer number of expressions involving the 'mental' domain. About one half of the words in 'Roget's Pocket Thesaurus' can be found under such headings as 'intellectual faculties', 'voluntary powers', 'sentient and moral powers'. Even if we reduce this number by applying very strict criteria, there remain thousands of expressions belonging to the mental domain. Combining them allows intersubjective communication of an enormous number of different mental states and processes.

Now what is the relationship between mental and physical states? Sensory perception, language and cognitive abilities are all correlated with increased activity in specific areas of the brain. Some functions, such as those involving long-term memory and thought, seem to be distributed across wide areas, but also modularly. The latest neuroscientific technology allows the non-invasive monitoring of brain activity. The spatial and temporal resolution of the next generation of this technology may increase dramatically. The noninvasive techniques allow us to discover more and more relations between brain activities and their locations. Activities in different brain areas involved in a given function may be linked by waves of activity across

the brain.

A particularly interesting topic relating consciousness to brain states is attention. For instance, whether or not you are listening attentively to my presentation determines the activity or inactivity of the hippocampus, which is involved, among other things, in the transmission of information from short-term to long-term memory. If we focus our attention on the color of an object, certain parts of the brain are active; if we focus our attention on the shape of an object, a different set of areas of the brain is active; if we focus our attention on both color and shape, the two sets of activated areas do not simply overlap; rather, a third pattern emerges. Activity patterns in the brain depend on a person's intention to focus her or his attention on a particular aspect, but the patterns are rather complex even for simple intentions, and decoding the patterns for more complex ones will certainly be difficult. Attention focusing is only one example for neurobiological processes that are connected to consciousness. By now, there are thousands of studies on patterns of activity of the human brain, employing the new non-invasive techniques, PET (positron emission tomography) and of fNMR (functional nuclear magnetic resonance) - studies on cognition and language, even on self-representation, representation of others (often called "theory of mind"), and empathy.

On the whole one should not view the nervous system simply as an intermediary, as a sort of switch in the relationship between stimulus and behavioral response. Even when there is no outward stimulus, something does happen in the brain. For example, when we imagine something or have a hallucination a similar pattern of brain areas is activated as when we have an external perception. The difference is that when we imagine things, additional areas are activated that have to do with recalling something from memory, whereas when we perceive something, primary sensory areas are activated in addition.

In the chronological organization of processes and, in particular, in the planning of future behavior, the prefrontal cortex plays a special role. The prefrontal cortex, in turn, has strong interconnections with centers involved in memory and feelings in areas of what is called the limbic system: with the aforementioned hippocampus, mediating access to memory, and with the amygdala, particularly associated with the emotive judgment of desirable and undesirable situations in the future.

Such monitoring of brain activities in relation to conscious processes represents only one approach, albeit an impressive one, in the investigation into the neurobiological principles behind the conscious mind. Due to their limited spatial and temporal resolution, this neuro-imaging techniques cannot, per se, track the rapid processing of data in neural networks, not to mention in the individual nerve cells. They are supplemented by other methods addressing neural activity, connectivity and function.

In addition, there is an entire spectrum of research fields - such as psychophysical methods, comparative investigations on other primates, the study of models of neural networks and the corresponding computer-based theoretical research - that

may contribute to the understanding of higher brain functions, such as pattern recognition and language, learning and memory, voluntary movement, chronological organization of actions and many other abilities. It would be impossible to describe or explain all of these activities here. All in all we are increasingly able to understand many relationships between processes in the brain and processes in the mind. Consciousness is mediated by the cerebral cortex, it is activated particularly by situations that are novel or that involve difficult planning or decisions. However, most of our activities are unconscious and routine; conscious processes are heavily influenced by unconscious pre-conditions, such as past experience and emotions.

Even sensual experience is not necessarily conscious experience; this is demonstrated, in particular, by the remarkable results of studies on 'blindsight'. There are patients with damages in the visual area of the cortex who are unable of conscious visual experience. However, their eyes and other brain areas are still intact. It turned out that they are able to see something without knowing it. Being asked, for instance, whether a visual stimulus was horizontal or vertical, they often give the right answer without being able to explain why. Unconscious vision.

The progress made in brain research involving consciousness has been so impressive that many neurobiologists tend to assume that there will be an asymptotic approach to a complete understanding of the mind: according to them, how far we get depends solely on the effort made; there are no limits in epistemological terms. In the following, I am mainly concerned with the question of whether there can be, in epistemological contrast to this belief, fundamental limits to a scientific theory of consciousness. But in order to do this, I would like to shift the focus and not discuss individual, neurobiological discoveries concerning the relationship between the brain and the mind, however impressive they may be, but to show, first, that there are general theoretical reasons that we can go a long way along this path. Then, I would like to explicate further reasons for the supposition that complete decoding the brain-mind relationship with respect to all interesting features is impossible in principle and not just in practice.

Neurobiology contributes to the understanding of consciousness

The neural network of the human brain consists of well over ten billion nerve cells, connected by hundreds of thousands of kilometers of neural fibers. They make up thousands of billions of synaptic connections between nerve cells. In some sense the brain is a system for storing and processing information, analogous to a computer. For computers a general rule holds: what can be formalized, can be mechanized. Every function that one can model in formal, mathematical terms can, in principle, be executed by a computer. The analogy between the brain and the computer has its limits, but the nerve cell's capacity as a building block for information processing is greater, not smaller than that of the digital yes/no switch of computers. This is

the reason why one expects all the formally representable functions of the human brain to be based on physical chemical processes in the nervous system. This argument supports that a scientific explanation is possible in principle; but it is not the explanation in itself, which can only be achieved by neurobiology. And then there is the important question of whether all features and functions of the brain can be described formally in scientific terms. What can be formally represented can be seen in the research into artificial intelligence, and the list is impressive: object recognition, conceptional abstraction, memory, planning and the comparison of different strategies for future behavior are all on this list - in other words, much of what one considers to be the higher capacities of the brain.

How far will such investigations take us? Does it depend solely on our efforts, or are there principle limits - limits not just concerning complex details, but also limits to intrinsic, interesting, and central features of consciousness? When mental processes are unambigously linked to physical states of the nervous system, and these follow the laws of physics, it is tempting to assume that a complete decoding of the brain-mind relationship should be possible. However, this assumption is subject to criticism upon closer analysis.

Consider, for comparison, physics and mathematics. In these fields there are many questions that become more answerable, the more effort we put in. Energy states of stable material systems can be determined very exactly, depending on the effort involved; in this way, for instance, we understand chemical bonds very well. But there are also questions for which there are no definitive answers no matter how much effort you put in. Predicting individual events on the atomic level is in principle not precisely possible regardless of the efforts involved in measurement and calculation. That is the content of the famous "uncertainty principle" of quantum physics that Heisenberg discovered, and this principle has a profound reason: namely the unavoidable interference of measurements or observations with the states or processes to be observed. Quantum physics has included these limits in its basic equations from the outset: uncertainty, i.e. the limitation of possible knowledge, is itself a law of nature. Uncertainty is not just limited to the atomic level, but can also apply on a large scale whenever individual events on the atomic level are amplified to lead to macroscopic changes. Sexual reproduction, for instance, is dependent on processes in individual molecules of the genetic material DNA, on the random breaking and combining of chromosomal nucleic acid, and these processes are subject to quantum uncertainty. It is for this reason that the features of future living beings resulting from sexual reproduction, including humans, cannot be predicted in principle - not even in all essential aspects.

There is another epistemological limit to our knowledge that has been uncovered by mathematical decision theory. In the 1920s it was still considered a goal of mathematics to arrive at a formal system of mathematics and logic, allowing for the proof of its internal consistency - the proof that contradicting statements cannot possibly arise within the system. It was Kurt Gödel's great discovery in the 1930s

that this is impossible for stringent mathematical reasons: logical systems, except very primitive ones, do not allow proving their consistency by using their own means. Within every such system of formal thought and calculation there are questions that cannot be decided in principle.

It is remarkable that physical uncertainty and mathematical undecidability, though applying to different domains of knowledge, are logically related: in both cases, problems of self-reference are involved, the effects of measurements on the states to be measured in physics, and the logic of logic in mathematics. I would like to name reasons why there could be limits for the decoding of the relationship between brain and mind, reasons also linked to the intricacies of self-reference - after all, the relationship involves consciousness of consciousness.

The brain-mind relation may not be fully decodable in principle

Let us select, as an example for states of consciousness, general behavioral dispositions for the future, that is, intentions of an individual for various patterns of behavior depending on various scenarios of the future. Such dispositions are stored in our brain and are at least partially accessible to consciousness. Let us perform a thought experiment: let us suppose that we can simulate states and processes of the brain by a correspondingly constructed and programmed computer. In principle, we could calculate, over time, what would happen to a given initial brain state when exposed to certain exterior conditions, and which behavioral responses would result. One could now argue that we could thus test all possible exterior conditions of the future, one after the other, with the final goal of determining the general behavioral dispositions corresponding to the initial state of the brain, but valid for different scenarios in an open future - and in this way to decode the present brain state with respect to conscious states objectively and exhaustively, at least as far as they are related to behavioral dispositions.

But, on second thought, we realize that this would not work; a procedure of this sort proves to be impossible to perform if we consider the finiteness of the world and take it seriously in epistemological terms: The intrinsic finiteness of the world also limits the decidability of problems. Even a computer made up of the mass of the entire universe, running 15 billion years - the age of the universe -, would still only be able to execute a finite number of operations - a very liberal upper limit would be 10^{120}, a one with 120 zeroes after it. But numbers of this huge magnitude do occur even in everyday problems as the number of possibilities. The number of possible letters with various contents, even when only a few pages long, is much larger. The same holds true for the number of possible future physical states that a particular behavioral disposition may apply to; and the number of possible different behavioral dispositions is also so large that they certainly could not be checked by processing them one after the other in a finite decision-making process to find out which of

these dispositions actually correspond to a given physical state of the brain. It follows that according to finitistic epistemology, mental states are not automatically included in a physical description of the state of the brain; there is no algorithm for their comprehensive deduction.

Of course, it is possible in any field of science to discover many general relations, rules and laws applying to widest domains not limited by numbers such as 10^{120}, through clever observations, experiments, and theoretical thoughts. This also holds true for the mind-brain-relationship; but there will presumably be no general process for discovering every relationship of general validity. It is more likely that some essential aspects of the body-mind-relationship are not automatically 'decodable' in a finite number of steps, such as, for example, a secret code that has been so cleverly devised that it cannot be reliably decoded using finite means.

As for the aspects of consciousness that a scientific theory may not be able to fully encompass, only more or less educated guesses are possible. We can find hints by drawing analogies to the aforementioned mathematical decision theory: the standard example for principally undecidable questions, the proof of the internal consistency of mathematical-logical systems by their own means, possesses a pronouncedly self-referential character. Analogous to this, the characteristic properties of consciousness, like the generation of behavioral dispositions, are also self-referential. We appear in our own memories, fears and hopes, desires and plans - as we are, or as we believe ourselves to be, or as we wish to be seen by others, as we want or do not want ourselves to become and as we see our past, and our future possibilities. Behavioral dispositions are influenced by these 'self images', which of course do not represent concrete spatial conceptions, but are rather abstract representations of features of the individual in his or her own brain. 'Self images' are often contradictory and can never be complete because no physically existing entity can contain a complete duplicate of itself. Self images change in the course of time and alternate within conscious experience. They interact with one another and feed back on themselves. Perhaps these multiple 'self images' belong to the aspects of consciousness that cannot be determined fully by analysis of the physical state of the brain.

To sum these considerations up, it is not a stringent consequence of the applicability of physics to the brain and the unique correspondence of mental states to physical states of the brain that all behavioral dispositions are deductable from the physical state of the brain in a finitistic process. We have more reason to believe that there are limits to the decodability of brain states with respect to mental states. According to everything that we know, the brain follows the same physical laws as do machines; but a machine that we were capable of understanding could not do everything like a human, and a machine that could do everything like a human would be impossible for us to fully understand. If we know the mental state of a human, expressed by means of language and gestures, we may know far more than would be possible to know through a purely physical analysis of her or his brain, however

elaborated it may be.

The spectrum of concepts and theories on the mind-brain-relationship is wide

As part of this gathering, in which I am the only or one of the few coming from the field of natural science, I would like to say a few words, apart from my own position and intuitive suppositions about the solution to and the solvability of the brain-mind relationship, about the wider spectrum of opinions about this question that are represented in the scientific community. Most scientists insist, much like I do, that mental states are clearly linked to physical states of the brain, and that physical laws apply fully in the brain. This view is in contrast to Descartes' postulate of influences of the psyche on the brain by means of the pituitary gland. In recent times, it has been in particular Eccles who has represented the position that psychic states cause or influence behavior through their effects on the nervous system. For physicists one conclusion appears certain: if psychic states have an effect on the brain, then physics cannot apply fully to the brain because physics does not allow influences from outside physics; and it looks much more like there are no such influences.

Another position which one finds to be represented in almost every symposium on the psycho-physical relationship is the thesis that we are dealing with an artificial, illusory problem. All one needs is to explain the terminology properly, to clearly define what one means - which are for example the semantic and epistemological problems in grasping the minds of others, and in describing mental states of oneself - and the mind-body problem will disappear. Actually it is just an artificial product of philosophy since Platon, if not only since Descartes. Popper argued rather convincingly against this: the differentiation between body and soul is older than philosophy, appearing implicitly as early as Homer. In the tenth canto of the Odyssey, Circe transforms men into swine. After this unfortunate action, the condition of Odysseus' companions is described as follows: 'They had the heads, voices and bodies of swine. Only their minds remained as before.' It seems that Homer had no particular trouble in conceiving body and mind as different entities. The relationship between our subjective experiences on one hand and physical processes, including those in the brain, on the other- this relationship exists, and no matter what rules apply, it cannot be dismissed as something not belonging to the realm of scientific inquiry.

Another line of thought in the spectrum of opinion holds that understanding the mind-body relationship is not possible in the current state of physics, but a future expanded physics could fully resolve the problem. The world of atoms was not accessible to the physics of the beginning of the 20th century, but the new discipline of quantum mechanics - and the underlying conceptional changes and expansions of the basic laws of physics - rendered atoms and molecules understandable in physical

terms. Why should an expanded form of current physics not be able to explain consciousness? This position is represented, for example, by Penrose. Although it is very much a minority opinion, it must be taken seriously; but the likelihood of this happening is something I would probably bet against.

Most practicing neurobiologists adhere to the 'asymptotic' position on the resolution of the mind-brain problem I have already mentioned, that progress in their field will be able to explain the relationship between brain and mind more and more, and that there are no limits in principle, even if some questions are too complicated and some calculations are too exhausting for concrete solutions. Consciousness is for them a property of systems of nerve cells in the brain, much like supraconductivity is the property of systems of certain metal atoms at low temperatures; after all we have learned how to understand supraconduction in terms of physics - why should this not be possible for consciousness as a system's property of neural networks in the brain? But such comparisons are not completely on track. Supraconduction is objectively defined - the electrical resistance is zero - consciousness is not. On can try to define consciousness. There are options with respect to levels of consciousness. If human consciousness is to be encompassed, one quickly agrees the self-reference is a central feature. Integration in time - past and future - and symbolic thought are also essential. But this is not yet enough for a comprehensive definition of the mental aspects; one could build self-referential properties, when they have been formally defined, into a pocket computer, but we would still not consider it conscious. Consciousness is primarily accessible through self-awareness and through the communication of the awareness of others; it is doubtful whether in principle a complete formal or objective definition can be made. For these reasons, it is also doubtful whether the evolution of human consciousness can be fully and objectively explained in scientific terms.

Finally I would like to come back to my own position within this sketch of the spectrum of opinions: there are principle reasons standing in the way of a complete decoding of the relationship between the brain and mind - a position that is supported by more than a few, but which is probably not a majority opinion in the scientific community. It is based on two premises: stringent physicalism - physics is valid everywhere, including the human brain - and epistemological scepticism. Neurobiology will take us further in our understanding, as many believed a short time ago. But there will be unsurpassable limits, particularly where self-referential processes in the brain are concerned.

A comparison of epistemological aspects of quantum physics and the brain-mind-problem

My view on the implications of quantum physics for the mind-brain-problem is that the former will not resolve the latter per se; and yet, the level and type of discussions on the foundations of quantum physics may be a model for other fields not excluding

the brain-mind relationship. I would like to include a few, more or less hypothetical, comments in this context.

An essential consequence of quantum indeterminacy as a law of physics is that it cannot be overcome by real or thought experiments, however sophisticated. In particular, it is not possible to first produce complete replicas of the quantum state of a system, followed by independent measurements on the replicas concerning all aspects of the original system, such as position and momentum of a particle. It is a law of quantum physics that single quantum states cannot be "cloned". Information on a quantum system is obtainable by a sequence of measurements, but this sequence alters the original states, and causes "decoherence" of "entangled" states in multiparticle systems, such that only limited information can be abstracted, the limitations being governed by quantum indeterminacy.

Brain states, in contrast to individual quantum states, are macroscopic features to which presumably the no-cloning-theorem does not apply. In the context of our thought experiment, a large number of replicas of the physical state of a brain could be analyzed by computer separately and subsequently. Nevertheless, it is not possible to perform the many analytical operations that would be required for testing all scenarios relevant to behaviour; their number of operations would by far exceed what can be realized within intracosmic limits. An aforementioned, upper limit of some 10^{120} elementary steps of information processing is not an arbitrary figure without epistemological status; it is given by the number of relatively stable particles (some 10^{80} nucleons) and the maximal number of operations per particle (about 10^{40}) consistent with the size and age of the universe. The figure 10^{120} can be estimated, in terms of cosmology, on the basis of elementary constants of physics, Planck's constant, the gravitation constant, the mass of the nucleon, and the speed of light. It is for this reason that, in my view, the corresponding limitations are fundamental, supporting a finitistic epistemology.

Nevertheless, we are able to gain much and most interesting information on the brain and its functions by subtle experimentation and sophisticated analysis. One may, of course, supplement the analysis of a human being and his or her brain by asking questions. The latter approach, however, requires interference with internal information processing in the corresponding brain by external stimuli, generating attention and inducing actions (including verbal expressions) in response to one given scenario among all possible scenarios. This, in turn, will alter the actual state of the brain, which then differs from the brain state without such interference, with consequences on possible future responses and actions. Therefore, the richness of internal states of the brain (in terms of information theory) could be much higher than what can be sequentially determined by neurophysiological as well as by psychological means. Analogously, the richness of quantum states is higher before than after decoherence induced by external measurements and observations. It is emphasized that this analogy is meant to apply at the epistemological, not at the physical level; the brain presumably is not a quantum computer. I suggest that in conscious

experience the richness of internal states of the brain (which is not fully accessible by external analysis) is integrated into self-images, emotional states and subtle strategic dispositions. It is not claimed, though, that this notion would exhaust what we consider as human consciousness.

At the beginning of the 20th century a most commonly held belief was that mathematical mechanics at least in principle would be capable of calculating and predicting all physical processes and states - this becomes asymptotically more and more attainable, the more effort we put into it. Since around 1927 - since quantum physics - this is no longer true. Correspondingly most mathematicians at around 1900 believed in the asymptotic solvability of all logical questions that could be reasonably formulated - including the logical validation of logic - and since about 1930 we have known that this is not true. Nowadays, around 2000, many neurobiologists and researchers into consciousness believe in the asymptotic solvability of the brain-mind relationship; they hold that our knowledge depends on our efforts, which corresponds to the mainstream position in mathematics and physics at around 1900. Will this situation look the same in 2030 or in 2130? I am one of those who think that is unlikely - one of those who believes that there are inherent questions in this field that are principally irresolvable.

Limits to the decodability of the brain-mind relation have implications for the philosophy of nature, the humanities and the arts

In the final part of my presentation I would like to choose aspects of the problem that I assume to be interesting to a gathering of researchers in the humanities who are mainly involved in art and literature. The relationship between science and art is closer than one often sees it. Both paths to understanding of self and of the world focus on the hidden shared features of processes and phenomena; science imparts general laws, and art conveys often multiple and thus common meanings and thereby fitting forms of representation.

If only for this reason it is worth the effort to build bridges of understanding between the sciences and humanities. The defensive positions which we sometimes encounter do not necessarily contribute to the expansion of our knowledge; with this I do not mean just passive disinterest, but rather active efforts using radical constructivist and historically relativistic theses to distance oneself from natural sciences. Let us instead concentrate on a few relationships between a scientific view on the brain-mind relationship and aspects of consciousness that are of interest to the humanities - particularly regarding self-reference and limits of knowledge of one's own mental states, but also that of others. Where this last aspect is concerned, perhaps better than any theoretical analysis is Hofmannsthal's story of the sandman:

A physicist constructs a lifelike doll with rather lifelike behavior. The physicist's assistant falls in love with the doll, but falls out of love just as soon as he recognizes

that he was in love with a machine. In order to recognize another consciousness, to feel real empathy, we must - as this story has just shown us - know or accept that the other entity involved is a human being like us. For example, no one would be charged with murder if he or she threw the doll out the window.

I would like to particularly emphasize one epistemologically significant consequence of principle limits to the decodability of the brain-mind relationship, since it has so much to do with understanding literature and art: subjective statements about one's own thoughts and feelings do not just practically, but also in their essence contain information that may go beyond what one could derive from a neurobiological analysis of the brain, no matter how complete it may be. With this in mind, any art that has to do with expressions of human consciousness can lay claim to its own contribution and value that need not be completely reducible to objective facts.

Of particular interest are the self-reflexive aspects of conscious processes that play such a great role in artistic expression as well as in the scientific debates on consciousness. As mentioned before, this feature has close ties to the self-limits of mathematical and scientific thought discovered in the last century; to physical uncertainty and mathematical undecidability, to Heisenberg's law and Gödel's theorem. Both have something to do with the fact that we can only know the world from the inside and are therefore in our physical selves and in our thoughts an inseparable part of the world that we would like to gain knowledge of. We cannot completely disregard the impact of the process of observations and measurements on the results of observations and measurements. The logical analysis of a logical system is also a process from inside, from within the system; there are logical prerequisites involved in the analysis of logic. Understanding consciousness has a similar logical structure, involving consciousness of consciousness. Self-references are sensitive to self-contradiction and are incomplete - but we are increasingly able to know what we can know and what not, and why.

It is precisely this knowledge about the limits of knowledge that shows us why scientific knowledge, despite of its unambiguous contents with respect to spatio-temporal processes and laws, is and remains ambiguous on the metatheoretical level. The basic limitations of knowledge are concerned with the relationship between the order of nature and human cognition, and they are linked, in this way, with fundamental questions of man's image of himself and the universe. It is for this reason that, in contrast to many ideas that existed in the nineteenth century, modern science is capable of being, and needs to be, interpreted on the philosophical, cultural and religious levels, and is consistent with different, though of course not all, such interpretations. The ancient Greek philosophers have given us building blocks for possible interpretations - logos, number, spirit, elements; interpretation itself is the task of the present, a task for science, art and the humanities, even if the languages spoken in different sections of culture are often quite different from one another.

Self-reflexive human thought is often subject to more or less hidden contradictions

and ambiguities. It is not uninteresting to see how we come to terms with this. There is a category of jokes that one could name "metatheoretical". Example: Nasredin Hodscha becomes the village judge without knowing much about the job. A participant in a dispute presents his case and Hodscha declares, "You are right". The opponent presents his point of view, and Nasredin Hodscha says, "You are right", which provokes an angry outburst from the complainant. "You said I was right, and now you're telling him that he's right. We can't both be right." And Nasredin Hodschas final judgment before he leaves the courtroom: "You are right".

The joke is clearly based on the last statement, the one about being right about being right; it is self-referential and therefore subject to contradiction. We are predisposed to handle such contradictions well. If it is serious, we become alert, otherwise we laugh about it. But the contradiction remains - why not?

Finally, the problem of consciousness is tied closely to one of the most difficult questions surrounding our understanding of ourselves: the question of free will. Naturally, it will not be solved by thoughts on possible limits of a theory of the brain-mind relationship; but they make a small contribution. They say namely that there may be principal limits to grasping the consciousness of others. The will of an other, despite being tied closely to processes in his or her brain, cannot be completely decoded by an outsider, and therefore not objectively understandable. Outsiders cannot claim to make certain statements about our own motives if we do not voluntarily share them. Luckily there are limits to intruding into the consciousness of others, and there is unfortunately often too little modesty and reservation when judging the motivations of others. In fact, complete mind reading is beyond human capabilities.

Limits to human understanding of self - Example: Heisenberg and Bohr in "Copenhagen"

Do we at least know enough about ourselves, do our own conscious experiences, emotions and thoughts lead to actual self-understanding? This is an old topic about which you from the fields of psychology, the humanities, literature and art can contribute much more than any scientist. I just want to confirm that even from a point of view of natural sciences, the answer can be 'no' - that we gain knowledge of ourselves only partially, and often only through our own actions. Through our own acts we become different from our previous selves. Often, we learn about our motives only in hindsight, and then only in a limited way.

At this point I would like to mention a play that first appeared in London, then in Essen and New York. The author is Michael Frayn and the name of the play is "Copenhagen". There, around 1927, Heisenberg and Bohr discovered and explained quantum indeterminacy, the principle that is fundamental to modern physics. This centennial achievement is often called the "Copenhagen interpretation" of quantum theory. Fourteen years later, in 1941, Heisenberg, by then one of the heads of the

German uranium project during the war, visited his elder teacher and friend Bohr in Nazi-occupied Denmark, in order to discuss in one way or another the topic of 'atom bombs'. The visit was a disaster. Bohr never wanted to talk about it again and Heisenberg suffered the rest of his life from various contradictory, mostly negative versions of what he said, thought and planned at the time. Now they are both dead and meet in Heaven along with Margareth, Bohr's wife, who plays the most critical role, with respect to Heisenberg's behavior. They want to find out what really happened in 1941, what was thought, done, said and wanted. It is still not a complete success, but the reciprocal understanding and mood both improve. That is due to the insight into the logical relationship between the limits of self-knowledge and the limits of any knowledge about reality, as shown by the uncertainty of quantum physics that the protagonists Bohr and Heisenberg themselves discovered. In the appendix to his play the author describes briefly and quite well what one really knows about the events of that time. Then, as a writer, he summarizes his ideas on consciousness which are very similar to my position as a scientist, and I would like to conclude with the corresponding statements by Frayn:

"What the uncertainty of thoughts does have in common with the uncertainty of particles is that the difficulty is not just a practical one, but a systematic limitation which cannot even in theory be circumvented. It is patently not resolved by the efforts of psychologists and psychoanalysts, and it will not be resolved by neurologists either, even when everything is known about the structure and workings of our brain..."

Notes:

This article is based on my books on 'Science and the Image of Man' (Gierer, 1998) and on 'Physics, Life and Mind' (Gierer 1985, 1991); the reference sections of these books are concerned with various aspects of my present article. It would be beyond its scope to explicitly include here more than very few references to some key issues.

An introduction to brain research with particular emphasis on aspects related to consciousness, such as attention, cognition and emotion, is given by Roth (1996) 'Das Gehirn und seine Wirklichkeit'. Roth essentially supports an 'asymptotic' view expecting that the relation between consciousness and the brain is resolvable in principle, as do Crick and Koch (1990, 1992). As an example for arguments that the mind-brain-dichotomy is an artificial construct of certain lines of philosophy to be dissolved by stringent conceptual and philosophical analysis, I would like to refer to Ryle (1949) in 'The Concept of Mind'. Mind-brain interaction beyond normal physical laws and processes is postulated by Eccles in the book: Popper and Eccles (1977) 'The Self and its Brain'. This book also contains (on pp. 128-209) a very instructive introduction by Popper into the history of the mind-body problem including thoughts from the times of Homer, Aristotle, Descartes, Spinoza, Leibniz and Kant up to the present. Popper discusses the inconsistency of mind-body interaction as

proposed by Descartes with the principles of physics and considers the mind-body relationship as a genuine problem of science. The book by Penrose (1996) 'Shadows of the Mind: A Search for the Missing Science of Consciousness', is an example for the view that present days physics is insufficient for solving the mind-body problem, but an extended physics of the future may be capable of achieving a solution. My view that there may be principle epistemological limitations for decoding the mind-body relationship is explicated in my articles: 'The physical foundations of Biology and the psychophysic problem' and 'Relation between neurophysiological and mental states: possible limits of decodability' (Gierer 1970, 1983). Basic limitations for a solution are also postulated and explained by Chalmers (1995) in his article on "The puzzle of conscious experience". For the physical basis of large numbers (such as 10^{120} in cosmological contexts) see Barrow (1990). The understanding of quantum uncertainty in relation to information transfer and the no-cloning-theorem of individual quantum states are the subjects of papers by Wootters and Zurek (1982) and Zurek (2000). Michael Frayn's play 'Copenhagen' with the postscript quoted at the end of my article is published by Methuen (Frayn 1998).

Bibliography

Barrow, J.D. (1990): The mysterious lore of large numbers. In: Modern cosmology in retrospect (Bertotti, B., Balbinot, R., Bergia, S. and Messina, A. Eds.). Cambridge University Press, Cambridge, pp. 67-93

Chalmers, D.J. (1995): The Puzzle of Conscious Experience. Scientific American, pp 62-68

Crick, F. and Koch, C. (1990): Towards a neurobiological theory of consciousness, Seminars in Neuroscience 2, pp. 263-275.

Crick, F. and Koch, C. (1992): The problem of consciousness, Scientific American, pp. 111-117.

Frayn, M. (1998): Copenhagen. Methuen Drama, London.

Gierer, A. (1970): The physical Foundations of Biology and the psychophysic Problem, Ratio 12, 40-54.

Gierer, A. (1983): Relation between neurophysiological and mental states: possible limits of decodability, Naturwissenschaften 70, 282-287

Gierer , A. (1985, 1991): Die Physik, das Leben und die Seele ('Physics, Life and Mind'). Piper München.

Gierer, A. (1998): Im Spiegel der Natur erkennen wir uns selbst - Wissenschaft und Menschenbild ('Science and the Image of Man'). Rowohlt Reinbek.

Penrose, R. (1996) Shadows of the Mind: A Search for the Missing Science of Consciousness. Oxford University Press

Popper, K.R., Eccles, J.C. (1977) The Self and its Brain. Springer Berlin

Roth, G. (1996) Das Gehirn und seine Wirklichkeit. Suhrkamp.

Ryle, G. (1949) The Concept of Mind. Hutchinson, London.

Wootters, W.K. and Zurek, W.H. (1982): A single quantum cannot be cloned. Nature 299, 802

Zurek, W.H. (2000) Einselection and decoherence from an information theory perspective. Ann. Phys. (Leipzig) 9, 855-864