On the Role of Theoretical Work in the Sciences

Harald Atmanspacher

Summary

In several sciences, theoretical work is distinguished from other procedures that are oriented numerically, empirically, or applied. The success and the productivity of scientific activity depends ultimately on a balanced interrelation among these aspects. This balance is easily perturbed if digressive considerations (such as an excessive dependence on media attention or economic interests) become overly influential. As recent developments in the sciences demonstrate, this often goes hand in hand with a misrepresentation of the role of theoretical work. As a consequence, its impact and significance are in danger of becoming severely underestimated.

Here we argue against such tendencies. First we propose a rough characterization of theoretical work with a few essential elements. Then we outline and analyze some pertinent historical case studies of theoretical work in physics and astronomy as well as in biology and psychology. Subsequently we criticize the current situation in cognitive science and neuroscience and indicate erroneous and misleading trends. Finally we sketch some examples from our own work that suggest how timely theoretical work in these areas of research could look like.

This essay is not intended to be a scientific paper but rather a basis for (controversial) discussion, a pamphlet as it were. It is based on an article published in a German collection of essays entitled “Wege zur Wissenschaft” (“Roads to Science”), edited by H.R. Yousefi (Nordhausen 2007).

What Are Theoreticians Doing ?

As a theoretically working scientist, one is occasionally confronted with the question of what one actually does when one pursues theoretical work. At first glance the question sounds somewhat impertinent, but if one takes it seriously it is not so easy to answer indeed. What does it mean to work theoretically? What is referred to if something is called a theoretically achieved result? What are essential aspects of theoretical work? – Presumably one would be most inclined to expect answers to these questions in physics, with its traditional division (at least in academic institutions) of theoretical physics and experimental physics. As a first guess one might think that it is the task of the theoretician to explain the results of experimentalists by means of theories. But – apart from the need to clarify the notion of explanation itself – this is not so simple. In many cases experimental physicists entertain and use their own ideas of how to build models for their observations. But then, what are theoreticians doing?

The following sketch of a characterization of theoretical work results from about two decades of practice with such work in various areas of science (and largely ignores the existing literature on this topic in the philosophy of science). *In my opinion it is a necessary condition for theoretical work that it attends to a conceptual* *question.* Very short-hand and sloppily speaking, theoretical work is thinking. This can be done in numerous ways. In the natural sciences, one will primarily aim for mathematical formulations, in other areas theoretical work can also be based on the precise use of natural language.

Of course, theoretical work in the natural sciences (as opposed to mathematics or many branches of philosophy) has always to be seen in conjunction with the empirical results to which theories refer. However, the goal is, *to begin with*, an answer to the question that stands at the outset, and the gain in knowledge that is related to it. *Theory in this understanding is not primarily motivated by or even derived from empirical results* – although such results are inevitable to study the validity of a theory. If it can be confirmed this way, the explanation of corresponding empirical results is a consequence (or a by-product) of the answer to the conceptual question posed initially.

Today it is a truism that theory-free empirical facts do not exist, not even in the so-called empirical sciences – each result of an observation or a measurement relies on tacitly accepted non-empirical assumptions. Without thoroughly reflecting these presumptions there is always a risk that, at the end of the day, the scientist understands about science as much as a fish about hydrodynamics. *For a sound critique of implicit assumptions it is necessary that theoretical work gets by with as few as possible theory-external inputs.* The empirical correctness of a theory is obviously important, but it is decisive is that it is conceptually compulsory. A good theoretician is not a bad experimentalist – he is a scientist whose enthusiasm for insight into conceptual relationships is even greater than his enthusiasm for interesting experiments.

Before I go on to illustrate these issues with some examples, I want to point out that the picture so far is somewhat overemphasized. There are many results in the domain of theory which could hardly have been achieved (not only confirmed) without an essential impact of experiments and observations. My concern in this context, however, is not a characterization of the body of knowledge that can be subsumed as theory, but a characterization of a *modus operandi* that is genuinely theoretical. In the most extreme case (for which string theory is an example): How does a theoretician proceed who simply has no empirical results at his disposal that would fall into the range of his work?

Historical Examples

Physics and Astronomy

An example that expresses very clearly how theoretical work can be conceived is Einstein's road to the theory of relativity, one of the two central cornerstones of 20th century physics. The decisive conceptual question for Einstein resulted from the inconsistency of two then fundamental theories, mechanics and electrodynamics. This explains in a natural way the title of the publication with which he introduced special relativity in 1905: "On the electrodynamics of moving bodies". In this publication he answers the question of how the two theories can be combined in such a way that they become compatible with one another. He achieved this considering fundamental symmetries, expressed by so-called invariance principles, in this case Galilei invariance and Lorentz invariance. The extension of Galilei-invariant classical mechanics to a Lorentz-invariant mechanics leads to the special theory of relativity.

The occasional claim that the experimental results by Michelson and Morley, testing the hypothesis of an ether, were Einstein's starting point is outright incorrect. It is true that the constancy of the speed of light that these experiments demonstrated became one of the postulates of Einstein's theory. The context in which this is the case, however, is only intelligible by the requirement that the laws of motion remain invariant in rest frames that move uniformly relative to each other.

A basically similar question of consistency marked the beginning of the other fundamental physical theory of the 20th century: quantum theory. The radiation law that Planck published in 1901 ("On the Law of Distribution of Energy in the Normal Spectrum") resulted, among other things, from the attempt to relate two incompatible radiation laws to each other. Planck's equation comprises the laws by Wien and by Rayleigh-Jeans for the limits of high and low frequencies, respectively. The description of the entire spectrum by *one* law forced Planck to introduce a quantity today known as the quantum of action, carrying his name. It took a quarter of a century until its fundamental significance for quantum theory as a whole was recognized in Heisenberg's uncertainty relations and the underlying commutation relations of incompatible (or complementary, respectively) observables.

In both these exampels the initial conceptual question can be specified by the intention to unify distinct and inconsistent theoretical frameworks, each established *per se*, in one joint framework. The applied procedure is, thus, in both cases characterized by aspiring a higher degree of generality of the theory. Ultimately this means that more empirical results can be described and understood by the same theory – the factual domain of validity of the theory is extended. However, the motif of theoretical work capable of achieving this is doubtless located outside the empirical.

Another useful historical case study is the road from the astronomical tables of the Babylonians to Kepler's insights concerning planetary motion. The Babylonian tables are obviously of empirical nature, and for centuries they served various purposes from the projection of mythological traditions to astrological predictions and to the design of calendars. Some few celestial bodies, the planets, were soon recognized as moving relative to the majority of the others, and it became a subject of study how to describe their motion. The epicycles postulated by Ptolemy in the 2nd century are an example for how lawful regularities can be constructed from empirical material. Large heaps of existing data are put together in such a way that the amount of data is (as one would say today) algorithmically compressed.

This procedure is clearly different from theoretical work in the sense mentioned above. It took until the end of the 16th century that Kepler, on the basis of the heliocentric worldview of Copernicus, suceeded with the decisive theoretical breakthrough for understanding planetary motion. His guiding idea was, different from Einstein and Planck, not so much the generalization of existing knowledge. Kepler's starting point was a mixture of basic theological and neo-Platonic convictions (it is hardly possible to be less empirical!), which – as Pauli showed later – let him asume that ideal geometric forms such as circles are reserved for the sphere of the divine and do not appear in reality. There, deviations from Plato's ideal bodies will be the rule. For this reason, planets will not move on circular trajectories but one has to expect ellipses as distorted circles. This argument – absurd for us today – led Kepler to his three laws which later, in the 17th century, were essential for the development of Newtons' theory of gravitation.

Biology and Psychology

Physics and astronomy are areas of science in which mathematical methods had and have both an established tradition and overwhelming success as instruments of thinking. Thus, the question suggests itself how theoretical work appears in less mathematically formalized sciences. Let us first have two looks into biology. In which sense can Darwin's theory of evolution be understood as a product of theoretical work? And what impact did theoretical insights have at the beginnings of molecular biology?

The core of Darwin's approach was the postulate that biological species, opposite to then dominating opinions, were not jointly created but developed successively: species variation rather than species constancy. However, as Lovejoy showed later, all empirical results that were known at the time could be used to argue both in favor of and against these two assumptions. Moreover, there were no indications of mechanisms for variation with which an evolution of biological species could have been reasonably supported when Darwin published his theory in 1859 ("On the Origin of Species"). Darwin encouraged experimental investigations of this lacking evidence, but he himself did not even speculate about such possibilities. So what could possibly have convinced him to decide categorially pro species variation?

An important role in this context may be due to the increasing rejection of teleological arguments in the natural sciences, which were still strongly represented in mid 18th century biology. Against this background, Darwin looked for a mechanism allowing him to reconstruct the variety of species as a result of cause-and-effect relations. He postulated a reason, unknown to himself, for variation as the precondition for the operation of a principle based on selection by adaptation that is capable of avoiding teleological arguments.

Later on, random (as opposed to directed) mutations of hereditary material were discovered and identified as such a variation mechanism. That these mutations are pure chance entails that the theory of evolution cannot make predictions of individual events but only permits a *post hoc* reconstruction of an evolution that already took place. In recent decades, new insights in the field of epigenetics have shown that the originally assumed blind chance must be complemented by specific molecular mechanisms in inheritance. The anti-teleological motif of the theory of evolution remains unchanged by these new developments.

The pioneer of molecular biology, Delbrück, was strongly influenced by the extremely fast accumulation of novel insights in quantum theory, in which he participated early on. He was convinced that elementary biology, occupied with the structure of genes and their mutations, could only be understood by applying elementary physics. In this spirit, Delbrück's "atomic physics model of gene mutation" of 1935, marking the birth of molecular biology, was nothing else than a successful translation of basic ideas of quantum theory to genetic problems. In this publication, mutations are described as quantum jumps between stable configurations of genes. Although Delbrück worked increasingly experimentally in later years, theoretical questions (such as the role of the idea of complementarity in biology) remained his primary driving force. He considered experiments as instruments of thinking, not the other way around.

At the end of the historical part of this essay, let me add some remarks concerning theoretical work in psychology. This theme is particularly interesting because a generally accepted consistent theoretical framework for psychology or one of its areas has not been established yet. Various possible reasons for this state of affairs have been discussed over and over. Nevertheless, a number of outstanding scholars in the more than 200 years of the history of psychology were distinctly inclined toward theoretical work.

This can be most clearly seen in the way in which early psychology defined its major problems: at first the mind-body problem (or psychophysical problem), then the problem of the impact of inheritance versus environment (in parallel to corresponding discussions in the theory of evolution). Although Fechner and later Wundt are currently known mostly for their empirical contributions to psychology, they were both intensely concerned with decidedly conceptual questions. If today, for instance, relations between brain and consciousness are discussed in terms of "dual aspects", then developed precursors of this concept can be found in the work of Fechner and Wundt.

Wundt in particular was very influential in the international institutionalization of academic psychology. In addition, his comprehensive work became worldwide standard reading together with the "Principles of Psychology" of James, also with respect to the two problem areas mentioned above, until to the present time. Subsequently, 20th century psychology was methodologically refined (and constricted) by the behaviorist emphasis on behavioral data rather than qualia and introspection. Innovative theoretical ideas as, for instance, in depth psychology (Freud, Jung), Gestalt psychologie (Köhler, Wertheimer) and developmental psychology (Piaget) at the beginning of the 20th century became more and more the exception rather than the rule.

One of those exceptions is the "Organization of Behavior" (1949) by Hebb. Still under the slowly decreasing influence of behaviorism (cf. Skinners "Behavior of Organisms", 1938), Hebb developed a series of ideas which (opposed to his own prognosis) belong to the conceptual core of cognitive neuroscience today: the Hebbian synapse, Hebbian assemblies, Hebbian learning. Basically these are now essential concepts for how to understand mental states and their dynamics (learning) neurophysiologically. Hebb's speculative ideas turned out to be an inspiration and a basis for large parts of current experimental and computational neurocience.

Also in the 1940s, the Macy conferences with von Neumann, Wiener, McCulloch and others established an influential "cybernetic" tradition for contemporary research. Using elements of (syntactic) information theory and logic, brain and cognition were conceived as feedback systems, to some extent analogous to computers. Although von Neumann as one of the pioneers of this idea ("The Computer and the Brain", 1958) was quite reluctant concerning its viability in general, the "computer metaphor" became a central dogma of various domains in cognitive science and neuroscience.

The Contemporary Situation in Cognitive Science and Neuroscience

This dogma is still held in high esteem at present. The tradition of cybernetics has produced a situation in which theoretical work is often confused with applications of computational science or numerical simulations. This is particularly obvious in groups, departments and centers under the heading of "computational neuroscience". That this represents a stretch on experiment rather than theory is revealed by the fact that mainly data analysis and modeling are the focus of these activities. Innovative theoretical ideas are still scarce goods. If one tries to simulate the functioning of human brains on networks of supercomputers (as is done at EPF Lausanne since 2005), then this has as much to do with theoretical neuroscience as a computer simulation of a planetarium has to do with astronomy.

The contemporary situation on cognitive science and neuroscience, as also in psychology, is probably correctly assessed by a substantial deficit of theoretical work and corresponding ideas. But there is more to it – often, scientists in these directions of research try to reinterpret this misery as a virtue by claiming that theoretical work is basically superfluous. For scientific progress, so it is pleaded, more experiments, more measurements, more data are all that is needed. The background assumption seems to be that the understanding, which should be the goal of any scientific effort, will arise more or less by itself if sufficiently many empirical observations are accumulated. It is true, as Poincaré said long ago, that science is built up of facts as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house.

There may be different reasons for the deficient development that has led to claims as preposterous as indicated above: the low success of theoretical work in the history of psychology, lacking education in and, thus, missing experience of precise formal thinking, a generally changed scientific culture which relies on (allegedly) illustrative images rather than actual understanding, evaluation systems for which raised funds rank higher than actually relevant research results, and much more. By and large, from the viewpoint of the development of science as a whole, the blend of theory deficit and theory phobia that has been accrued in cognitive science and neuroscience is untimely and prereflexive.

A slowly growing number of scholars in the mentioned disciplines realize this situation as being severely in need for correction since some time. What the absence of theory can entail is, among other things, visible in the populist arguments that recently keep being entertained by experts for brains in order to argue in favor of a naive neurobiological determinism (that is already questionable in itself) in discussions about free agency. These arguments have reached exorbitant extents of embarrassment and should by ignored as "not even wrong" right away – if not at the same time the leap of faith were at stake that scientific experts is still granted in large parts of society. One cannot live up to the resulting responsibility in the blissful state of prereflexive naivety.

The locus of theory in psychology has more or less shifted into meta-theoretical topics such as historical, social, cultural, and epistemological analyses (see for instance the proceedings of the "International Society for Theoretical Psychology" since 1987 and its journal "Theory & Psychology"). Theoretical approaches in cognitive science and neuroscience are predominantly found in branches of philosophy, mainly in the (analytic) philosophy of mind. As an example, the philosopher David Chalmers and the Tucson conferences that he co-founded have contributed a lot to a renaissance of the topic of consciousness, including the long and systematically dismissed discussion of qualia, in the 1990s. His name is connected with the "hard problem of consciousness", referring to relationships between phenomenal conscious states and brain states and belonging to the key issues of the "Journal of Consciousness Studies" since 1994.

Chalmers (then Tucson, now Canberra) took up the old idea of consciousness and brain as dual aspects of one psychophysically neutral reality and presented a critical discussion of the option to describe this reality information theoretically. As a further important author of (non-formal) conceptual work in this context, Kim (Brown University) is to be mentioned with his thorough accounts of supervenient mental states that are multiply realized at the neural level. Finally, Metzinger's (Mainz) innovative concept of transparent and opaque, respectively, mental representations and the related intentional relations is a significant achievement.

Central points in the renaissance of the problem of consciousness are again particular facets of the psychophysical problem. In the research practice of cognitive science and neuroscience this remains, as a rule, unaddressed; one might call it an "interior anomaly" (in the sense of Kuhn) within many empirical results about particular "neural correlates of consciousness". Numerous such correlates were and are investigated in many laboratories week for week. The theoretical question behind them is why the underlying correlations exist at all, how they can be more precisely characterized, and ultimately maybe even better understood.

Exactly this is a topic downright predestined for theoretical work, with respect to both conceptual questions and mathematical formalization. A research institution addressing this in detail is the profoundly interdisciplinary division for theory and data analysis at the Institute for Frontier Areas of Psychology at Freiburg, Germany. The way in which the question for neural correlates of consciousness is approached there portrays interesting examples for modern theoretical research combining results, hypotheses, and methods of the philosophy of mind, cognitive science, neuroscience, and theoretical physics.

A first example aims at a formalized representation of structural relations between the levels of description at which neurodynamics on the one hand and mental states on the other play a role. The corresponding relations constitute a kind of emergence (thus, we are talking about a non-reductive program) known and well illustrated by physical examples. A basic principle of the construction of emergent mental states leads to the problem of how the space of neural states can be partitioned in such a way that robust mental states emerge. This problem was solved using tools of the ergodic theory of dynamical systems, and the obtained partitions can indeed be mapped onto mental states. First applications to experimental data underscore the promising perspectives of this powerful approach.

The applied procedure was developed independent of empirical material to begin with, and it led to the development of a "structural framework theory" that is neutral with respect to possible applications. The specific details needed for a particular desired application must be successively implemented within the general framework of the theory. This requires knowledge – as a rule, also empirically gained knowledge – from the relevant disciplines.

A second example can be traced back to inspirations by Bohr 80 years ago and refers to the concept of complementarity. (Descriptions that contradict each other but are together necessary for the understanding of a phenomenon are called complementary.) After this concept had turned out to produce important insights in quantum theory, Bohr suggested to explore its significance for areas outside physics as well. He himself proposed a series of possibilities none of which, however, was worked out in detail over all the years.

As above, it was suggestive to look for a structural framework including the conditions for complementarity without specific details for applications. Such a theory provides an algebraic structure, a monoid, whose core is the existence of non-commuting observables. (An alternative approach uses algebras with Boolean subalgebras pasted together in a non-Boolean way.) If those details that are required for a discussion of quantum systems are implemented in this structure, one retains ordinary quantum theory. Other specifications result in examples outside physics. Predictions of a model for the bistable perception of ambiguous stimuli that was worked out this way are in impressive agreement with a number of old and new experimental results and suggest further investigations.

A final example takes up a proposal by James. In the much cited chapter on the stream of consciousness in his "Principles of Psychology" he calls for focusing not only on stable but also on unstable mental states *between* stable categories. Following the Swiss philosopher Gebser, corresponding acategorial states can be distinguished from both categorial states and virtually regressive non-categorial states. From future work in this direction we expect interesting contributions to a better understanding of meditative states and insight processes. Moreover, a refined understanding of instabilities will become more and more important to deal with the social and societal problems in the future ahead of us.

The road from a framework theory to its concrete application can be long and tedious, and usually this it what it is. In the indicated examples it took about five years until first ideas were worked out far enough to allow us successful comparisons with experimental results. This fairly short period can be explained by two reasons: Firstly, the required elements of the theory were to a large extent known and had just to be adapted and put together. Secondly, already established cooperations with experimenters provided easy access to data for testing the theory. In more difficult problems and a less developed theoretical state of the art, the historical examples show that the way to success can take longer or much longer. The far most significant and inevitable precondition for such undertakings is good ideas – something that no money can buy.