Robert Schlögl, doctorate in chemistry (1982); post-doctoral research fellow, Department of Physical Chemistry, Cambridge, and at the Institute for Physics, Basle; habilitation at the Fritz Haber Institute of the Max Planck Society (1989); professor of inorganic chemistry at the University of Frankfurt/Main (since 1989); director and scientific member at the Fritz Haber Institute (since 1994).

Changing the Structures of Research from the Perspective of an Active Scientist

Robert Schlögl

It is one of the predictions of this conference that there are two groups of people: those who organize science and those who suffer from that. I happen to belong to the second group, the smaller community at this meeting. What I am going to talk about is dissemination and evaluation as factors affecting the creativity process. As I am an active scientist I will take the opportunity to introduce my subject a little.

1. Structure-function relations in inorganic chemistry

I work as an inorganic chemist on structure-function relationships in heterogeneous inorganic reactions. We have solid surfaces and a gas and we want to bring these two to react. It sounds very simple because we work with very small molecules that have about ten atoms. The individual atoms form arrays at solid surfaces and they have a certain structure which we can determine. They also have some defects at this structure. We think that this is a good model for reacting interfaces. However, when we do the reaction under the conditions we would like to have as chemists, then we find out that massive changes occur to these surfaces because they do single turnovers of one molecule and then everything is dead. If you bring it to conversion you see that the surface changes from a well-ordered state into a disordered array. It means that the complexity of the system and the disorder which we see is essential to the function which we want to have.

But as long as we do model studies, we have certain problems. In our system we cannot extrapolate from ideal to real states, although we know all the atoms involved. We have essentially no methods to investigate the dynamics and we have no basic understanding so far of the underlying chemistry. We do not know why it turns forms from a flat state into a disordered array. This we have to find out if we want to understand structure-function relations.

The underlying topic we are looking at is the science of heterogeneous catalysis. If you want to analyze these structure-function relations you find from a strategic point of view that you work in a rather complicated field of different sciences. This means that this is by definition an interdisciplinary subject. If you do it using only one discipline, which would mean that you look either at the macro-, or the meso- or the micro-aspect, you will just miss the whole story. You have to look at all three aspects at the same time. This is a beautiful interface between physics and chemistry. We need to understand a lot about solid-state physics but we also need to understand a lot about inorganic chemistry and chemical engineering if we want to solve this problem. But it is not so easy to devise a useful strategy because there is no single person or group who actually can do all this together.

Science in catalysis used for about thirty years a kind of approach with two dimensions: a complexity domain and a pressure domain. In this approach model surfaces were devised and then investigated by surface science. Finally a linear extrapolation was done. After about thirty years of effort, it became clear that this model fell terribly short. The real structure and the result after linear extrapolation diverge enormously. As a qualitative problem you want to have a bridge between your model systems and "real" catalysts. But you cannot bridge that gap by linear extrapolation.

So we decided to try a different approach. We accept first that we don't know exactly what our system looks like. We devise and define new methods in order to find out what our system is and from there we go downwards in complexity via model structures and end up with a suitable model for the system. Then we do our surface analysis. This different approach is doable now, as we have about thirty years of method development behind us. We have now analytical tools which will allow us to understand how the real system works, whereas in 1970 or 1980, when people started to do this kind of research, they just did not have the tools to know how complex the real system actually is.

2. Methods and criteria of evaluation

My main target is that I want to look into the details and the nitty-gritty about evaluation as I feel about it, as somebody who is being evaluated. You can evaluate on different levels: you can evaluate projects, people, institutions and whole systems (national or international science policies). It is obvious that all these evaluation targets require different methods. A fundamental hypothesis from my point of view is that science is lacking an external regulative pressure like the economy does with the market. For this reason it is necessary that you have a suitable evaluation. You also have to take into account that science happens by and large outside of the social context of the financing society. So there are only a few connections between science and society and we need an external regulatory medium for evaluation.

As an active scientist I will concentrate on the evaluation of projects. If you look at the projects you will find that there are different types of projects which require again different strategies of evaluation. First of all there are the mega-projects (the Human Genome Project, space station, astronomy, etc.). Second, there are group and network activities. Some of these network activities are scientifically required or wanted, others are politically reinforced – for example, European Community projects. Third, we have long term projects, for example in astronomy, the humanities or physics, where you need many decades of research to achieve something, because it is a difficult problem and the time-frame of your observation is long. Fourth, you have an extreme diversity of many little individual projects. Chemistry is a very typical field for that: you have many thousands of small projects. Fifth, you have industry-academia relationships – chemistry or materials science are fields where they are frequent – which also should be sometimes evaluated.

What evaluation criteria can you use? I have found two classes. There are *non-accountable criteria*. These are usually qualitative arguments that rely on in-depth content information. This is where the peers come in. You require expert knowledge for that but it always remains debatable, whether they are the right criteria and the right judgements. And of course there are *accountable criteria*. They are transparent, they are so-called inter-subjective but they are usually also context-insensitive. For this reason these kinds of criteria are very often superficial (citation analysis or impact factors, for example).

3. Impact factors of journals and their relevance

If you compare the different impact factors from a selected journal collection where we publish our daily work you see strong distortions between the practical relevance of a given journal and its impact ranking. In catalysis the most important journal is the *Journal of Catalysis*. But with an impact factor of 2.939 points it ranks very low down. *Surface Science Reports* have the highest impact factor (15.522). This journal just gives you some reviews that are usually not even very good, because they are far behind the time when they are published. If you want to have many impact points for evaluation, then it is much better to write a *Surface Science Report* than to write ten papers in the *Journal of Catalysis*, although the latter would perhaps be a hundred times more in terms of effort. This is one point where I want to draw very clear conclusions from the standpoint of a practitioner. Taking these figures is very dangerous if you don't weight them properly.

The relevance of impact factors has of course also been the subject of scientific evaluation. I take as an example the journals of analytical chemistry because I know most of them. The impact factors of nearly all of them fall within a very narrow range on the score index. Arguing from a statistical point of view the distinction between their impact factors is insignificant. All the journals have the same impact when you do this analysis. For this reason some authors have developed statistical manipulations of the sequence. We also get data from the science citation index (SCI) and then you end up with an entirely different ranking. If you compare the ranking from the SCI to the ranking of the statistical analysis, you get very often large deviations. Some of the journals have no deviations but other journals deviate massively. This means that it is very dependent on how you numerically do these analyses, what the outcome will be. I don't know what the deeper meaning in these analyses is. You can do whatever you like with them. One reason for that is that the criteria of the SCI are too oversimplified. In order to determine the impact the SCI just takes a ratio between the number of citations and the number of papers in two successive years. The SCI parameter is an equilibrating ratio method precluding a lot of information on the relevance of a journal for the scientific dissemination in the field.

The relevance of impact factors in science other than evaluation is maybe even more dangerous because there are people who think that the journal in which an analytical method has been presented is a quality indicator for the method described. An example which I found in *Analytical Chemistry*, a good journal: the European Union funded a project in order to find out what the right method is to determine the content of calcium in food. The Union's standard on calcium is now based on a rating where the impact factor has fifty percent of the meaning. That is, impact factors are used as steering criteria for the selection of analytical methods in food chemistry – a very clear example to show how dangerous these things are if they come into the hands of misguided people. It is extremely risky to think that our health could depend on the journal in which an article was published.

You can do other things with all these nice science citation indices, grand accumulations of papers to evaluate national strengths in science, for example. In the Internet I found a national indicator which purports to show how good Germany is in science. This indicator lists the percentage of papers from Germany in certain fields and the relative impact of these papers worldwide. On average about nine percent of the world's scientific effort comes from Germany. It is interesting to follow up the subjects where we are strong in. Apparently we are very strong in sciences like astrophysics, physics, chemistry, geo-sciences, plant and animal science and engineering with fourteen to thirty-two percent over the world average of citations coming from Germany. Then there are a lot of disciplines where Germany's impact worldwide is plus/minus zero: mathematics, microbiology, neuroscience, immunology, etc. In several disciplines we have big deficits: clinical medicine, psychology/psychiatry, education, economics and business, social sciences. These are areas where a normal person would think that Germany is traditionally strong. If this comes into the hands of politicians I would shudder to think what the outcome might be. Luckily this is only in the Internet and you cannot find it easily. Maybe it should be left there.

4. Criteria for quality in science

Now I want to discuss a few criteria for scientific quality. The first one is *accountability*. Accountability is an important keyword from economics and it is well-known that we are in a phase where science is being economized. It is not clear what implications this will have, but it is quite clear that economics and science have, by definition, not very much in common. Accountability is essential for controlling and distributing resources. No scientist can dispute that and if we only rely on soft factors this would be very bad, because we need hard figures. But the need for accountability might very easily replace the discourse over contents by playing around with figures: everyone can interpret what he or she likes to see in them. This just obstructs the discussion about important scientific questions. The charm of quantitative measures for quality has two dangerous consequences. First. It affects the problem-choosing strategy of scientists because it creates fashions. Second. It allows politicians and decision-makers to atomize responsibility, because they can say that they have it from an "independent authority" – from the Institute for Scientific Information (ISI), for example. This authority determines how good scientists are. The politicians don't have to determine it, they just find it in the Internet. – I think as a politician or a science organizer you should take responsibility for a decision and not give it away to an "independent authority" (and pay for it!).

Speed in science is another interesting issue. First. Speed is not an implicit quality of science. In the definition of science, speed just does not occur because it is not immediately necessary. Second. Speed is different from having a "critical mass of resources" in relation to the size of the problem you are working on. It is not important that you do something in six, nine or twelve months, but it is important that you have enough resources allocated so that you can sensibly solve the problem. Third. Speed is a keyword from sports and we should ask ourselves whether sports is a good model for science. Many people will think it is a good model because scientists have all these ideas and they have to have competition and accountability. But when you think about sports where you have referees working on the issue this might be quite questionable. Fourth. Speed is an accountable quantity, it is an "objective measure". That is good, but it has dangers. It inhibits in-depth research because by definition in-depth research takes more time than "normal" research. It calls for repetition and fragmentation because you get the same things out more quickly. It creates fashions and it produces very often superficial and glamorous statements about what is necessary and what can be done if only certain scientists get a lot of money. We have seen many examples of that, for instance the fullerenes. What did not people promise for that: a second periodic table, new cancer drugs and god knows what. Nothing of that has appeared. This is because everybody wants to have things quickly and doesn't think first about what is possible and because some scientists promise everything when they go in front of the public.

Quality in science has never been clearly defined. Everyone uses the term, saying that we need high quality. But what is high quality? There is no unambiguous definition so far, only a few criteria. Of course quality means always a *contribution to our knowledge base*. This is a must. But very often it is used in the sense that the results of a scientific effort are envisaged as a "*product*". This is that economization factor I referred to. At this point we have to talk about *cost-effectiveness* and *innovation*. But what is innovative? This depends essentially on what you rate it on. When you take all the scientific information available you might find that many things which are considered innovative are maybe not as innovative as you think. Is it essential that science is *useful for application*? Surely that is also questionable in certain areas in science apart from engineering and applied science.

There are some other arguments: quality in science is the *originality of the work*. Again, a very useful criterion but again this just transforms the problem of the definition. What does the term originality actually mean? We have the parameter "novelty and generic character" of "seminal" papers. But this might also depend on how long it takes. There are prominent examples where papers have become seminal only twenty or thirty years after they were published. What would be responsible for measuring things so long? As a rule the indices measure for two years. Of course you can also have a debate with politicians about the general value of scientific research, for mankind, for social structures, for the environment, for whatever. That is a very soft criterion. But it is very persuasive because it can be used to convince lobbies from whom you might need to get funding.

Necessary are *changes in evaluation*: reduce the load of evaluation but deepen its quality. It is useless to evaluate every project every two years. It is good to evaluate when it is necessary. And do it in a qualitative manner. In the Max Planck Society we are in very good shape because the advisory board (*Fachbeirat*) system we have is adequate. I am a big supporter of this system. Transform evaluation from a threatening into a challenging procedure. Think that there is not somebody who is banging on you but somebody who will give you advice to do something better. In internal evaluations in companies there is a real threat and you can be kicked out. But in the public sector there is no real threat. So you have a lot of evaluation reports which are standing shells and have no consequences whatsoever.

Emphasize the content-oriented evaluation in science. Stay away from all these formal requirements which we have – the huge books we have to prepare about statistics, for example. These books are not really helpful. It is much more important that the people do a content-oriented evaluation. Recognize the negative effects of accountability. In general you have to adapt the evaluation procedure to the desired purpose. Do not overdo these things. Depending on projects, the size of projects and the purpose of what you want to evaluate you have to find a suitable amount of means. Twenty-five referee comments is just ridiculous. Nobody needs this. Make a decision. It will be much better. And never use (formal) evaluation as an excuse for neglecting the necessary discussion about targets and priorities. You cannot make the difficult decisions in science – where to get the priorities – by evaluation. You need a dialogue between the people who are working there. And you need the option of doing things outside the evaluated routine system if you want to hit on important targets.

With respect to *dissemination* I will only emphasize a few points. Dissemination should be of genuine concern to each scientist as it is outlet and input to any scientific production. It is strange that there are still a lot of people who don't care about this. It is not enough that the publishers are the ones who will do something about it. We face a dissemination revolution by the existence of reliable Internet functions. The Max Planck Society has taken measures to give access to the electronic journals to all its co-workers, not only to the directors, and prepares for innovative uses of the e-media. I will now go a little bit into that.

5. The dissemination revolution

What do we understand by "dissemination revolution"? We have four factors: finance, production (this is essentially for people like me, my colleagues and my fellow scientists), dissemination (publishers and secondary source publishers) and retrieval (libraries, electronic media). Traditionally the finance body gives the money twice. First to produce knowledge and then to retrieve it. This is not really necessary. In the Internet era, we don't need the dissemination structures any more, because they are an additional and time-consuming factor and an independent body (publishers, archival systems, etc.). This can be done much quicker if we have a single chain from finance over production to retrieval. Then we have to ensure that a dialogue takes place which goes backwards and forwards in both directions. We have to prepare structures in such a way that this can be done as simply as possible. This has drastic consequences for libraries because they will be transformed from archiving places into information exchange centers, which is something quite different from having shelves with books. Also librarians will in future be quite different from what they used to be. They will not just put labels on books or give the users cards to register, but they will have to have the ability to foster information exchange, which is a different view and a different purpose from what they were trained for.

We are facing *non-traditional areas of dissemination*. Electronic media exhibit new properties going beyond the simple replacement of paper media. This paradigm change would not be very useful if you just take paper media and make them electronically accessible. We live in a transition period because we don't know how quickly things will develop. We will have highly parallel access to complex information and we will have unhindered access to interdisciplinary information. Usually you only have those things in your library which are from your field. When you work in a small institute it is a real problem. My own group for example has no access to chemistry journals because we are a physical chemistry department; we have lots about physics but nothing about chemistry. We need these things from outside. Non-discriminating access of information means: it does not matter where a scientist works or how rich he or she is. Everybody doing science in the world should have immediate access to the process. As important as that is an interactive access to information which means that many people can work simultaneously on the same amount of data and are not hindered by where they are and how much money they have.

E-media is a lot of things going beyond just archival information (books, papers, reviews, etc.). We can have retrieval information, but this information has to be integrated in all the other kinds of information. We can have strategic information on electronic media which never would be published in print. For example the strategy of how a working group tackles its problems can be published in the Internet but you would never find a paper where a group outlines its strategy. We can deal with information which is so complex that it cannot be printed. We can work with multimedia information. There is the fruitful example of living reviews where many people interact to bring science further. All these interactive activities are still visions and it will take quite some time before it will work for practical scientists. But we are moving in this direction.

E-media and evaluation are closely connected to one another. They play a major role in directing evaluation towards content-oriented judgement. In the future the evaluators will have access to all the information. Then they can look behind the "curtain" before they make their judgement. They don't have to rely on publications. Usually evaluators only look at the publication list. Because of their evaluation load they don't have the time to go in depth, except for very few, selected cases. Formal- and prejudice-oriented evaluation can be overcome if they have easy access to more information. The e-media give you world-wide access to achievements and strategic considerations of individuals and institutions. Second-level thinking can be published on the Internet and everybody can look at it, which was impossible until now. Nobody has to rely on numerical oversimplifications like the impact factor. Th e-media is a flexible means to accommodate a very large variety of work habits in scientific communities. Not every scientific community looks like physics or chemistry. There are very diverse habits in science and they all have to be accounted for. And the e-media give access to tutorial information. It happens quite often that a referee is not familiar enough with the object he or she is judging. So it is very useful if him or her additional tutorial information on that can be given.

Dissemination via print media can be expected to become more important as a review and teaching medium, not as a medium for original science. Books will be needed, but it will be necessary to write them with more original thought than today. Today books are usually a collection of second quality material because high quality material goes normally to scientific journals. Generalizations are dangerous, but there are numerous books in our libraries which could be simply thrown away because there is no new information in them. Quick review collections should vanish and conference print media are totally obsolete in the future. They should be banned, in particular from inflicted subsidies by conference fees. You pay a lot of your conference fee to get a printed book although you don't need it. In all international bodies of the small catalysis community we have decided to ban to all conference proceedings.

Dissemination in the Max Planck Society is a special question. The Max Planck Society is a bit different from other institutions because it is complex. It has eighty institutes working in eighty different branches of science. So it is very broad and very diverse. But in the future we will have an information system which is a network from our institute libraries. These libraries have – and will continue to have – books, monographs and print journals. Books and paper versions will never disappear because all the old science will remain. In addition there are many branches of science where the historical part is still valuable and might become even more valuable in the future. So we should not do away with the paper versions. Then we have electronic archives which contain factual information; we have primary and secondary databases; we have e-journals. These e-journals might disappear in the long run because of the expenses of these integrated publications by which I mean all the multimedia applications. This will probably be the future in say five to ten years from now. And of course we have the external world. So within the Max Planck Society we have to make a fundamental decision whether we want to do this ourselves, or whether we acquire the content from the outside world, or whether we just buy it from a provider. We must have this discussion because it has implications for both sides. In order to do this we need the Center for Information Management (ZIM) to give us a solid reasoning for what we should do. We are a medium-sized enterprise and we cannot do everything. But all the chances can only be tackled for the Max Planck Society if we do it ourselves. We cannot wait for someone to provide it for us.

Changes in dissemination. We should reconsider the division of labor between industry and academia. These publishers, secondary publishers and database providers: are they really needed or not? I think they are not needed. – Recognize the Internet as a prime information for content, exceeding that of papers. We can put the papers *and* the facts in the Internet. This is very new and gives us much better access to scientific information than before. – Recognize the Internet as a novel medium for interactive work. This is a new dimension because the traditional media are non-interactive and when we published a paper, it was like a snapshot. That has passed now. – Archive selectively only final documents and databases. Not everything which is open and which is in the Internet really needs to be archived. So we have to find a way to judge from that. Because there is too much information out there. – Maintain the minimum regulatory level in the Internet (at the expense of perfection). I see the danger that the rules in the Internet will be formalized and will be new obstacles to interactive work. Protect the freedom of the Internet as much as you can!

6. Some conclusions

In my view science is organizing itself by the actions of individuals in diverse communities beyond traditional disciplines. Interdisciplinarity is not something which we are heading for, we already have it. In that part of science which I know more about, interdisciplinarity is more the rule than the exception, even if that is not the case yet everywhere. – Allow for a flexible organizational framework. You don't have to foster or plan flexibility but you have to allow the individuals to do that. It is the individual scientist who really brings about this new interdisciplinarity. – Do not attempt to administer this process. You cannot plan interdisciplinarity. It just happens. Refrain from standardizing procedures of dissemination and evaluation, don't put new hurdles up. We have a strong tendency to fractionalize everything into old or even new disciplines, but they are still disciplines. If you have a border between two disciplines, this border is a difficulty in itself. – Dissemination is an intrinsic affair of scientists and their organizations. Get involved. Don't let some strange organization or commercial people do this for you. Scientists have to do it themselves. – E-media reflect the diversity of scientific activities and need support from within science on the generic level. We still lack tools for doing that. We have some of them but we do not have enough and we need some support quickly. This is not expensive and it is important to do it rather quickly. – Formal standards for traditional and innovative Internet science products are needed. You need some standardization in order to be able to cope with this problem of archiving and transporting the information. If everybody puts everything in the Internet. In a few years it will be garbage because it cannot be transported any more. – Archiving is an international affair. You have to have organizational structures and it is the duty of all of us to bring this to the attention of politicians and the public. We cannot leave this to some ambiguous structure. – There are still awareness and acceptance problems within the Max Planck Society that not everybody knows what is going on and think that it might be good for physicists or for computer scientists but not for themselves as a lawyer or whatever.

Science at large is not business. There are other factors which determine the quality of science than just economic ones. The correctness and the credibility are fundamentals of science. If you evaluate something, it is not helpful to say that this is all fraud. We should not help people to do this by putting speedometers on and demanding that things should be done very quickly. This is adverse to correctness and credibility. Define targets and priorities on the basis of facts, not on the basis of formalities or preconceptions. Try to be factual in decisions and listen to what the community says. Do not think that there are a few peers who know exactly what has to be done. We have seen so often how this can fail. So put things on a broader basis. Allow for a mix in evaluation criteria and rapidly use electronic media for dissemination and evaluation. This is a consequence of what I said before. Creativity and diversity is the most important part in science. They require each other. Whatever we do in the area of evaluation and dissemination, beware of limiting any of the two. If you limit any of the two, you will immediately limit or even destroy the other one. Evaluation is absolutely necessary and dissemination should be as broad as possible. We need a minimum of regulation for that. We can overdo it on both sides. Overdoing it is the big danger.