Myths, Challenges, Risks and Opportunities in Evaluating and Supporting Scientific Research

Martin Quack

ETH Zurich, Physical Chemistry
CH-8093 Zurich, Switzerland

Martin@Quack.CH

Citation:
Martin Quack, in “Incentives and performance: Governance of research organizations”,
I. M. Welpe, J. Wollersheim, S. Ringelhan, and M. Osterloh (Eds.),
Springer International Publishing, Cham (2014),
DOI: 10.1007/978-3-31909785-5_14

Myths, Challenges, Risks and Opportunities in Evaluating and Supporting Scientific Research

Martin Quack[1]

Abstract We summarize and discuss critically the various procedures used for decisions in evaluating and funding scientific research in view of the goal of obtaining the best research with a given funding volume. Merits, difficulties, and problems in some procedures are outlined. We identify a number of myths appearing frequently in the process of research evaluation. We indicate some similarities in deciding on funding research projects, making appointments to academic and leading research positions and selecting prize winners for high prizes for scientists. Challenges, risks and opportunities in research and its funding are identified.

1. Introduction

The present short essay on some challenges, risks and opportunities in evaluating and funding of scientific research is based on the "president's speech" given in Leipzig on the occasion of 111th meeting of the Bunsen Society in Leipzig, of which printed records exist (in German, (Quack M, 2012a, Quack M, 2012b). Our goal is to summarize and discuss critically the various procedures that are currently being used and identify their strengths and weaknesses. We identify some myths that are circulating in the community involved in evaluating scientific research and the resulting dangers from an uncritical belief in these myths. We conclude with a general discussion of the role of scientific research in terms of its value to the individual and to the society. The topic of our essay should be of obvious importance for all those being actively involved in scientific research, but also to mankind as a whole. Besides the enormous opportunities there are also hidden risks and dangers which we want to discuss.

2. Some Basics in Evaluating and Funding Scientific Research

We shall discuss the following main questions.

1. Who decides upon funding research? (Institutions, committees, bureaucracies, individuals as sponsor or as Maecenas)
2. How does one decide, what should be funded? (Procedures and criteria used).
3. What is good practice in evaluating and funding scientific research?
4. What is the goal of research funding? (Discoveries and new knowledge, future income and profit)
5. What is the goal of scientific research in a broader context?

Let us first consider the funding of research by institutions. Well known examples are in the USA the "National Science Foundation" (NSF), in Great Britain the "Engineering and Physical Sciences Research Council" (EPSRC), in Germany the "Deutsche Forschungsgemeinschaft" (DFG), in Switzerland the "Schweizerischer Nationalfonds zur Förderung der

[1] Physical Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland, Martin@Quack.ch

Wissenschaftlichen Forschung" (SNF, also "Swiss National Science Foundation" or "Fonds National Suisse", FNS, and SNFNS as Acronym), in Austria the "Fonds zur Förderung der wissenschaftlichen Forschung" (FWF) and in France with a somewhat different structure the "Centre National de Recherche Scientifique" (CNRS) or in the European Union the relatively new and quite remarkable "European Research Council" (ERC) besides many more world-wide, of course.

My closest personal contact is obviously to the Swiss National Science Foundation, where I like in particular the mentioning of the *goal* "zur Förderung der wissenschaftlichen Forschung" ("for the support of scientific research") in the German name of the institution, because all too often such institutions tend to forget their goals.

It would be of obvious interest to discuss the quality of the various institutions mentioned, but we shall refrain from this in order to avoid unnecessary embarrassments. Rather we shall turn now more generally to the procedures used by such institutions in decisions on funding research projects. Similar procedures are also used in decisions on appointments for higher academic and research positions or in attributing higher prizes for scientific research. Indeed, such decisions can be considered as providing support for scientific research in a broader sense as well. We can distinguish here the following main types of procedures:

A. The decisions are made by a committee ("panel") of competent persons, who are themselves actively involved in scientific research of the general field considered, covering at the same time a rather broad range of expertise. The decisions of the committee are based on the careful study of written research proposals, often obtaining more detailed confidential assessments from external experts in the particular topic of the research proposed.
B. The decisions are made by a group of bureaucrats, which uses various combinations of indices and other measures as well as sometimes expert assessments, or simply "gut feeling" ("Bauchgefühl").
C. The decisions are based on pure chance ("Lottery").

Of course, there exist various mixed and intermediate forms of the three basic types of decision making in research funding. Also, if there is sponsoring by an individual sponsor and not an institution, the sponsor might simply base the decisions on "personal taste ("gut feeling"), but more frequently some form of decision following the three main types mentioned is chosen also by individual sponsors.

Let us turn first to point C, the "pure chance choice" which may appear to some as a joke. However, there are serious proposals to base decisions on research funding on a random choice. We can cite here Les Allen following Buchstein (Buchstein H, 2011), where one can also find further discussion of this topic: "I suggest that the Engineering and Physical Science Research Council throw out the panels, throw out the referees and have a lottery for all the available funds. Such a system would be fairer then the present one and would also be better at supporting truly original research. Pure chance must give more hope than the opinions of a subset of my peers" (Les Allen). We might mention also the discussion of a random choice selection in the publication of journal articles by Osterloh and Frey (2012) in this general context (Osterloh M and BS Frey, 2012), see also (Osterloh M, 2013).

An argument which is frequently advanced in favour of the "random choice selection" procedure is its alleged "fairness" (or doing proper "justice").

It is certainly true that with an honest procedure, the random selection is in a certain sense "fair" (nobody is preferred). However, by those being subjected to such a procedure luck is perceived as "fair and good" only if it is "good luck" and rather as bad and unfair if it turns out to be "bad luck". We shall not waste too much time on this procedure, as it appears to us as obvious nonsense in the present context, and even immoral in a certain sense. We can make the following comparison: What would our students say if we simply made a choice of their marks by a lottery instead of basing the marks (including the decision of failure or success) on a careful expert evaluation of examination results? They would certainly complain justly about what would appear as an arbitrary, truly immoral, if very simple and time saving procedure.

Another argument sometimes given in favour of the random choice selection is the well-known bias in the alternative "expert selection", due to experts favouring "mainstream" as opposed to "revolutionary" (Kuhn TS, 1962) projects or results. We think that this argument misses an essential point in the evaluation by a good committee of experts: A "good expert" is fully aware of such a bias and takes it into account giving room in the decisions on funding to what one might call "high risk" projects. Of course, even this does not exclude expert errors, of which there are many examples. Still, a well informed and conscientiously made expert selection appears to be a far better option than the random choice selection. We can give another example for this from daily life. With a serious illness we would consult an expert medical doctor, in fact, the one with the best possible medical education and medical record and rather trust his diagnosis and therapy (perhaps after getting a second opinion from a further expert). We do so in perfect knowledge of the existence of errors of expert medical doctors. We still prefer this approach simply because of the alternative of a "random choice" selection of some arbitrary person (or even a randomly chosen medical doctor) for a diagnosis and therapy is completely unreasonable (and also definitely less hopeful and successful).

Indeed, the random choice selection is rarely used purposefully although we shall discuss below that it plays sometimes involuntarily an important role in mixed forms of "combination procedures". Most frequently the methods A. and B. of evaluation and decision taken by some committees and persons of various origins are used in practice.

Thus now we address the further question of the *criteria* used in making decisions on funding particular scientific research projects or also in selecting persons for tenured academic positions or as prize winners of high academic prizes. Here we can distinguish the following main criteria

a. The quality of the research project proposal, most frequently as assessed by detailed expert evaluations.
b. The personality of the researcher as assessed by the past achievements, research record and publications, and further information such as personal (oral) presentation of the research, interview with a committee of experts etc. The president of the Humboldt foundation Helmut Schwarz has expressed this in a concise fashion as "Fund people, not projects" (Kneißl D and H Schwarz, 2011). Evidently, this approach is mostly used in academic appointment decisions.
c. Bureaucratic indices such as citation indices, total numbers of citations, h-index (Hirsch JE, 2005), past funding record, (such as sum of previously acquired funds and number of previously funded projects), number of participating researchers or research groups in larger collaborations, number of previous publications, perhaps weighted by the impact factor of the journal, the position as "first author" or "last author", sometimes only counting publications in "Science" or the like.

According to my experience, these three main criteria are used with widely varying weights by the various institutions and committees. We state this fact at this point without any further judgement. In the section II (mathematical and physical-chemical sciences) of the Swiss National Science Foundation, where I was active as member of the research council for about 10 years, one uses mainly the criteria a. and b. (perhaps fortunately so), whereas in more recent years arguments resulting from the criteria c. are mentioned in the discussion but without much weight for the final decisions. One might mention here that the members of the

Swiss National Research Council are active research scientists, carefully selected from the members of the Swiss research community. They act in their function in the research council only for a limited number of years and contributing only a limited fraction of their time to this task. This is to be seen within the Swiss political culture of the "Milizsystem" ("militia system"), where, for instance, in principle, every Swiss citizen is a "part-time soldier" for a limited time in the year and period of their life, officers being selected from this pool as well. At the same time they remain active in their ordinary life in whatever profession they may have. Similarly, the members of the Swiss National Science Foundation remain regular active scientists in universities, research institutes etc. during their period of service on the council. Such a committee is by its very nature more influenced by scientific arguments and much less so by bureaucratic indices. However, we know also of other institutions and committees where the reverse is true and the criteria under c. can become dominant, indeed. Between these two limits one can find many intermediate situations, in practice.

After having introduced now already some judgement into our discussion, we shall reinforce these by mentioning some serious and increasing problems. We shall do so by referring to what we might call some "myths" propagating in the evaluation and funding of scientific research as well as in selecting researchers for academic and other positions or for prizes.

3. Some Myths in Evaluating and Funding Scientific Research

Myth 1: High rejection rates R (in funding schemes, journal publications etc.) are a measure of high quality of the procedure ("competitiveness"). The nonsense of this rather widespread belief can be easily established by considering the limit $R \rightarrow 1$, where everything is rejected, nothing accepted (for funding, publication etc). This would be then the ideal procedure, where people write proposals and papers, expert reviewers assess them and committees decide in the light of these expert reports to accept in the end … nothing. This would be of course in reality the complete idling and waste of time and effort. While the nonsense of such procedures is obvious, also in the cases of very large rejection rates R close to 1 (but $R < 1$), the procedures are of poor quality, by general experience. What happens in such situations is that in the end among a large number of, in principle, worthy projects only a few are selected by pure chance, a random choice selection, because other valid criteria are not available. While this is inefficient (and perhaps even immoral, see above) as one might as well use a lottery instead of an assessment in these cases, it frequently happens involuntarily in procedures with very high rejection rates. Of course, also the other limit of no rejections ($R = 0$) will not in general lead to an adequate and responsible use of research funds.

It is not possible to define a "correct" rejection rate R for an optimal funding scheme. In principle, it would be correct to fund all good projects and reject all poor projects. The corresponding rejection rates would depend on the research field, the research culture and tradition in a given field, but also the political context, the country and so forth. From my experience in physical-chemical research in a European context values of $R = 0.4 \pm 0.2$ would fall in a reasonable range and with $R > 0.8$, where less than 20% of the projects are funded, the quality of the selection process degrades rapidly and drastically.

Myth 2: Citation numbers are an adequate measure for the importance of a scientific publication. Experienced scientists know from numerous examples that this statement is pure nonsense. We shall illustrate this here with just one prominent example selected by R. N. Zare (Zare RN, 2012a, Zare RN, 2012b). The paper "A Model of Leptons" by Steven Weinberg (Weinberg S, 1967) has contributed most importantly to the current "Standard Model" of Particle Physics (SMPP) and has also contributed to the Nobel Prize awarded to Weinberg. It has even importance in relation to our physical-chemical understanding of molecular chirality (Quack M, 2011a). However, according to Zare (Zare RN, 2012a, Zare RN, 2012b), the paper by Weinberg (1967) was not cited at all in 1967 and 1968 and just once in 1969 and 1970 each (1971 it had 4 citations including one self-citation). This implies that this truly important paper has not contributed to the "impact factor" of the corresponding journal (Phys. Rev. Letters), and if the editor had taken this "measure" as a criterion for selecting papers, he should not have accepted the publication. Also, this paper would not have contributed to funding or tenure decisions for Weinberg, if the relevant people in these decisions had looked for this measure (they fortunately did not). There are numerous similar examples from many different fields, even if not all of these end up in a Nobel Prize. Molinié and Bodenhausen have generated a graphics for some "classics" of NMR spectroscopy (Molinié A and G Bodenhausen, 2011). While today (long after the Nobel prize) Weinberg's paper is highly cited (more than 5000 citations) this fact is irrelevant with respect to the period for further funding decisions (and in many cases also further tenure decisions) which would have been typically the period of 1967 to 1970 in the case of the author of this 1967 paper. Today the citation numbers for the paper are irrelevant for further funding of this research. Many fallacies of the citation analysis have been discussed over the years including "folk citations" for non-existing authors such as S. B. Preuss (Straumann N, 2008, Einstein A, 1931, Einstein A and SB Preuss, 1931). Another aspect of showing the nonsense in just counting citations is the neglect of the "sign" of the citation. A publication can generate high numbers of citations because it is criticized by many as erroneous (perhaps even with forged results). This "lack of sign" problem in the "impact" has been discussed in a humorous way by Petsko (Petsko GA, 2008), but it is a serious matter and we know numerous "real" examples from experience which leads us now to a closely related myth.

Myth 3: The "impact factor" of a journal as derived from the citation statistics of the first years after publication is a measure of the quality of the journal (*Science*, for example with its enormous impact factor, would then be an outstandingly good journal). The example cited under myth 2 already demonstrated that the very basis of such a statement as derived from the citations of an individual paper in the journal would be erroneous. However, one can find statements that the cumulative use and assembly from many different individual papers in a journal in the end leads to the "impact factor" being a meaningful measure of its quality. Again, experienced scientists know that this is wrong. In the author's own field physical chemistry and chemical physics, for instance, there are good journals with very modest "impact factors" (J. Chem. Phys., PCCP, J. Phys. Chem, Mol. Phys., etc) compared to *Science*, which is a journal of very questionable quality. Even if one does not subscribe to the nasty comment that "the condition for a paper to be published in *Science* is that it is either stolen or wrong", there would be many experts, at least in some of our fields of research, who would agree with the milder statement that the relation of Science to the four journals mentioned for this particular field (and related cases in other fields) is similar to the relation of quality of newspapers like "Blick" or "Bildzeitung" to the "Neue Zürcher Zeitung" or "Frankfurter Allgemeine Zeitung" to give examples from the German speaking areas (there are analogues for English, French etc.)

As examples for wrong results published in *Science* we might mention the interesting critical discussion of such an example by Volkmar Trommsdorff (Trommsdorff V, 2002) or the wrong paper on ortho- and para-water by Tikhonov and Volkov (Tikhonov VI and AA Volkov, 2002), commented upon by Albert et al (Albert S, BH Meier, M Quack, G Seyfang and A Trabesinger, 2006) and Manca Tanner et al (Manca Tanner C, M Quack and D Schmidiger, 2013), besides many more examples. Well known and particularly serious examples are the many wrong papers published by H. Schön et al in *Science* (a case of "wrong" even by forgery).

Of course, there are also journals of good quality with relatively high impact factors, although not the very highest, such as "Angewandte Chemie". Thus a high impact factor does not necessarily imply low, "Boulevard Journal" type quality. There is in fact no simple relation between quality and the impact factor of a journal. There exist good and bad journals with low and high impact factors. The next myth to be discussed is even more serious in that it deals with individual scientists.

Myth 4: The so-called h-index (Hirsch-index) is a suitable measure for the importance or quality of a scientist. Hirsch (Hirsch JE, 2005) has introduced this bibliometric measure and has made such a claim and has, indeed, seriously proposed to use the h-index in professional appointment and tenure decisions. The nonsense in such a statement is again well known to experienced scientists involved in such decisions and we shall return to the question of academic appointments below. Here, we shall cite the critical and very competent discussions by Molinié and Bodenhausen (Molinié A and G Bodenhausen, 2010a, b, 2011), as well as by Richard Ernst (Ernst RR, 2010a, b), who provide ample evidence rejecting Hirsch's proposal. Thus, without going into more detail of this absurd "quality measure for scientists", we shall turn to a further myth widely promoted in the science bureaucracy.

Myth 5: The amount of research funding acquired by a scientist (or a group of scientists) is a good measure for the corresponding importance (or quality) of the researchers. One might express this "Funding importance" F_{I} by Eq. 1.

$$F_{\mathrm{I}} = \frac{\text{Sum of aquired research finances}}{\text{Number of participating scientists}} \tag{1}$$

Of course, such a number can be easily derived for every researcher or research group, thus its popularity, and "the higher the F_{I} the better the research group". However, after giving some thought to this matter, one quickly comes to the conclusion that for an optimal use of research funds one should rather use a measure (if any), where the sum of research funds used would appear in the denominator, and one might call this the *research efficiency* R_{E}.

$$R_E = \frac{\text{Scientific knowledge generated}}{\text{Sum of finances (acquired and used)}} \tag{2}$$

We may quote here Martin Suhm (Suhm MA, 2010) (freely translated): "It would not be misleading, if the sum of acquired funds appeared in the denominator, the place, where it should appear in the name of efficiency and sustainability, instead of appearing in numerator of such measures". Of course, the bureaucratic use of Eq. (2) is hampered by the "Scientific knowledge generated" not being measurable by some number, unless one replaces it by "numbers of publications generated" or "numbers of citations generated", which is, indeed, sometimes done, but would be nonsensical as discussed above for the myths, 2, 3, and 4. Thus, if Eq. (2) is to be used in a sensible way, one has to interpret it as a symbolic, not as a numerical equation and thus no simple number can be generated from it for some rankings.

We shall conclude this section on the various myths propagated in decisions on supporting scientific research by some comments on the particular dangers arising from the use or rather abuse of bibliometric data in this context. Indeed, in recent years we are increasingly confronted with this abuse, be it by the science bureaucracy or by scientists themselves believing in bibliometry. I can quote here from personal experience from an expert report in an appointment procedure for a professorship (in anonymized and somewhat altered form for reasons of confidentiality) "… in our country bibliometric counts are most heavily weighted". The "expert" was from a country in Northern Europe and in the end drew his conclusions based on bibliometric data for the candidates to be evaluated. Fortunately, the appointment committee considered this particular expert report as irrelevant (for several good reasons) and did not take it into account in the final decisions. One must be afraid, however, that some poor appointment committees would follow such poor advice. Indeed, uncritical belief in bibliometry can be occasionally found in serious publications supported by major academies (see, for instance, (Gerhards J, 2013)) and also with some active scientists who contributed important work in their own specialty. From one such scientist, I heard the comment that there is "no objective alternative" to bibliometry. Again this is pure nonsense. An obvious alternative has been formulated by Richard Ernst (2010) in his most relevant article. "And there is, indeed, an alternative: Very simply start reading papers instead of merely rating them by counting citations". Of course, following such an advice requires time and knowledge of the subject, and the bureaucracy lacks the knowledge and does not want to invest time.

We shall discuss now further, related alternatives, in relation to the very important question of procedures in appointments for professorships or other higher academic and research positions.

4. Criteria Used in Academic Appointments

Good appointments of academic positions (professorships or research positions) at universities and research institutions are among the most important ways of supporting scientific research. They have a long term effect and are truly efficient investments for the institution as for science overall, if they are carried out successfully. We can cite here almost literally from Zare (Zare RN, 2012b, Zare RN, 2012a) for what we might call a summary of criteria used in appointment procedures with good academic practice (here for the example of the Chemistry Department at Stanford University):

1. First of all they must be good departmental citizens.
2. Second they must become good teachers.
3. The Department wants them to become to become great researchers (This last criterion is the most difficult).
   We ask experts, whether the research of the candidate has changed the view of the nature of chemistry in a positive way.
   … it is not based on the number of papers, with an algorithm on impact factor, etc.
   … do not discuss h-index metrics
   … do not count publications or rank them as to who is first author.
   We just ask: has the candidate really changed significantly how we understand chemistry?

From my experience in presiding appointment committees at ETH in many departments as a delegate of the president (for more than 15 years), I would add that similar considerations prevail at ETH Zürich, even though every once in a while attempts are made to introduce bibliometry and the like into the discussion, but without much effect. Particularly the younger people looking for an appointment tell me, however, that they in fact know universities, where bibliometric and similar data are used importantly, even dominantly, in professional appointments. My reply to this is, sure, there exist also poor universities, and if I am then asked how to tell apart the good from the poor universities, I would answer certainly not by using bibliometry and “rankings”, but for instance, by looking at the procedures they use in professional appointments, among other criteria. Of course, even good procedures do not exclude occasional erroneous decisions. Prominent examples include a Nobel prize given for a wrong discovery (see (Quack M, 2013)), although the person selected deserved the prize, but obtained it for the wrong reason.

That the criteria 2. and 3. concerning teaching and research are important for appointing professors at universities can appear as self-evident. The first criterion, requesting a “good citizen” might come as a surprise to some, and might support a suspicion that only “well adjusted” candidates are selected. This, however, is not the meaning of this requirement for a “good citizen”. It arises from the often painful experience that “bad citizens can damage good science”. This problem is frequently covered with silence in the scientific community, or given only minimal weight. It is, however, a very serious problem because the damage caused to scientific research by some “bad citizens” can be huge, directly and indirectly. Bad behavior can appear as a straight-forward, criminal fraud and forgery of research results, and then the damage caused is much larger than any potential advantage hoped for by the criminal (see, for example, (Pfaltz A, WF Van Gunsteren, M Quack, W Thiel and DA Wiersma, 2009)). However, cases of fraud and deception towards a partner in cooperations also exist. One such case, the quarrel between O. Piccioni and E. Segré in the discovery of the antiproton, was brought to the courts and reached a wider public, with enormous indirect damage for the reputation of scientific research ((Heilbron JL, 1989)). That Segré did a severe wrong to his colleague Piccioni can hardly be subject of doubt. This wrong was not punished, however, perhaps it was even rewarded, which puts this area of science in this period in a somewhat dubious light. That the problem reached the public is a rare exception. Most of the time, such events in scientific research are covered with silence. I shall not mention such a case from physical chemistry, as the aim of this section of our essay is not to sing a “dies irae”. In principle, a request for “good citizenship” in the republic of science is not specific to research or science, it is rather a generally valid principle in human relations.

ETH Zürich has a motto, which fits in this context “Prima di essere ingegneri, voi siete uomini”. It was formulated by one of its first professors, Francesco de Sanctis (1817-1883) in his inaugural lecture. Correct human behaviour supports scientific research by preventing damage, among other things. However, also reliability of universities in establishing contracts with their professors and keeping long term promises and escaping the temptations of later breaches of contracts are important elements in supporting research, which in recent times have been increasingly eroded even in the best institutions. Fundamental research, though, needs adequate freedom (Kneißl D and H Schwarz, 2011) provided by the generous and reliable appointment contracts at the top universities. Freedom of teaching and research is the most important pillar of science, innovation and creativity, which are also strengthened by reducing bureaucracy. Indeed, the incessant and unchecked growth of bureaucracy is one of the greatest risks in the current support of scientific research, a risk which can cause great damage (Szilárd L, 1961, Quack M, 2013). Again, this phenomenon is not restricted to science, but it is a general phenomenon in modern society. Much has been said and written about this and Fig 1 shows results from a classic in this field, here with the growth of bureaucracy in the British colonial administration following Parkinson’s law (Parkinson CN, 1957). Parkinson’s Law has given rise to many joking comments; it is, however, a serious matter. Also the growth of cancer cells follows such a law of growth until it is ended by a catastrophe. An analysis of the growth of the staff in science and university bureaucracy shows close analogies upon which I will not further comment, here. I do not intend to enter here into a general bashing of university administration. Indeed, there exists also the truly “good administration”, which serves and supports science. However, the staff in this part of administration does not grow and its relative importance rather decreases. We have elsewhere identified the general growth of bureaucracy as one of the great risks of mankind, besides nuclear war and climate change (Quack M, 2013, 2011d), and shall conclude this discussion here with this brief remark, turning finally to some aspects of opportunities of scientific research, rather than risks, after a brief summary of what might be considered good practice.

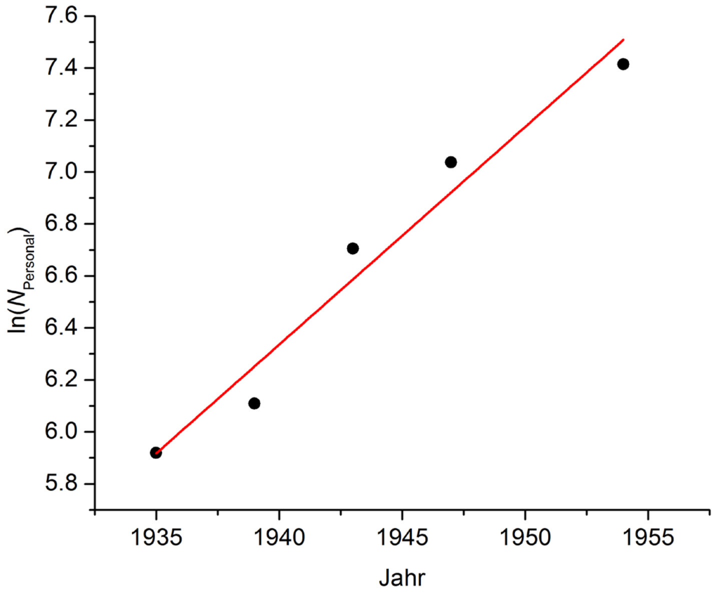

Fig. 1 Logarithm of the number of persons in the staff of the British Colonial Administration $\ln(N_{Personal})$ as a function of time (Following an exercise in the Kinetics course at ETH with data from Parkinson (Parkinson CN, 1957)

5. A Brief Summary of Good Practice

While implicitly contained in the text of the previous sections, it may be useful to summarize here what may be called the "good practice in evaluating and funding scientific research and in making academic appointments". We mention advantages of this "good practice", but also the associated difficulties (Diederich F, 2013). A good procedure for distributing available research funds and making academic appointments can be summarized by the following steps:

1. Bring together a group of competent and trustworthy experts covering a sufficiently broad range of the field under consideration in order to avoid too narrow a view and to neutralize possible conflicts of interest.
2. Get every research proposal and person looked at in detail by the relevant experts and the whole group, obtaining more detailed specialized outside reviews, if necessary.
3. Have the group discuss every individual case in depth before coming to a conclusion by the group as a whole, by consensus or by vote, if necessary.

This procedure is not new and it is used by good funding institutions such as the SNF or appointment committees at good universities, academies as well as prize committees for major prizes. It has many advantages and minimizes the chances for severe errors, although even the best procedure cannot exclude occasional errors completely. It has one major disadvantage: It is time consuming and costly (Diederich F, 2013). It also requires the cooperation of experts, sometimes difficult to obtain. These disadvantages have led some institutions to use one or another shortcut as mentioned under the 5 myths. However, using such shortcuts should be considered foolish or even fraudulent, as it replaces expert knowledge by bureaucratic superstition and forgery. It would be as immoral as the "shortcut of the bad citizen" replacing serious experimentation and carefully analysed data in scientific research by some invented results, when proposing or testing a hypothesis.

Sometimes it is argued that using statistical data, say from bibliometry, can be justified by some perhaps existing correlation with "real data". For instance, it is claimed that really good scientists (as judged by a careful evaluation as outlined above) statistically have a higher h-index, than scientists of lower quality. While one may have some doubts, whether this correlation really exists, even if so, it would be at most very rough, with numerous easily proven "outliers". On the other hand, the decisions to be taken (in deciding on a proposal or an appointment) are often very important decisions on individual cases. A rough statistical correlation (even if it existed), is of no use in this case. We may use here the analogy of evaluating written examination papers: From long experience we definitely know that there is, indeed, a very rough (but not tight) correlation between the number of pages in a written solution of an examination paper and the final examination result, as carefully assessed by an expert examiner. The more pages, the better the results, statistically. However, we also know that there are many outliers, some brilliant solutions are very short, and sometimes also many written pages of an exam paper contain just erroneous solutions, thus a poor final result. Using the "time saving short cut" of simply counting written pages would correspond to a fraudulent and immoral procedure of the examiner. There are many analogies of this example with the use of bibliometric or other indices (for instance neither a page count nor a citation count or h-index needs an expert, it could be done by some administrative staff). No more needs to be added on such fallacious "short cut" procedures.

6. Why Scientific Research? The Opportunities for Creativity and Innovation

We should address also the general question as to why one should consider supporting scientific research at all. In a commencement speech of 2004 which in the meantime has been printed in several versions (see (Quack M, 2011b) and references cited therein), I have summarized some important reasons for fundamental scientific research:

1. For the personal satisfaction of discovery and knowledge.
2. As contribution to the edifice of knowledge of mankind, towards understanding the world as well as the human condition.
3. Directly and indirectly to contribute to improving the conditions for human life and of mankind – and for its survival.

The first reason is an important, intrinsic, subjective, personal motif of the researcher. The second and third reasons provide objective grounds, why society should support science financially and otherwise as an investment in the future of mankind and society.

The first, intrinsic motif was formulated already 2400 years ago by Demokritos of Abdera (ca. 460–370 B.C.) in an inimitable way.

βούλεσθαι μᾶλλον μίαν εὑρεῖν αἰτιολογίαν ἢ τὴν
Περσῶν οἱ βασιλείαν γενέσθαι

He talks about the scientist-philosopher, freely translated here "He would rather make a single fundamental discovery than become the king of the Persians." We can also cite Rose Ausländer (Ausländer R, 2002) with one of her poems (imperfectly translated here)

You are
irresistible
Truth
I see you
and name you
Bliss

One can also translate these texts in a somewhat extended, completely free way as applied to the scientist and researcher (Quack M, 2012a, Quack M, 2011b, Quack M, 2011c):

He would rather make and teach a single fundamental discovery than:
… become president of the United States.
… obtain the wealth and power of Bill Gates.
… build a large bomb.
… have 10 publications in Science – the magazine.
… to reach the top position in citations of scientists.
… have 100 presentations on TV.

One can also phrase this in the form of a "non-motivation" (Quack M, 2012a, Quack M, 2011b, Quack M, 2011c):

Fundamental research: Why not?
1. Not to damage others.
2. Not to beat somebody in competition.
3. Not to have power.
4. Not to become rich.

These positive as negative points concern the personal motivation. There are, however, also the objective reasons concerning the service to mankind. This aspect is obvious with all forms of applied research, but is frequently forgotten, when fundamental research is considered. Nevertheless, fundamental research, innovation and creativity can be considered to be among the most important driving forces in improving the human condition (Perutz MF, 1982). Indeed, the support of fundamental research can be considered to be the greatest opportunity of all investments of society and mankind in their future.

I shall illustrate this here with the Schrödinger equation, which is one of the fundamental equations for physics and chemistry (Schrödinger E, 1926a, b, c, d, e)

$$\mathrm{i}\frac{h}{2\pi}\frac{\partial\Psi(q,t)}{\partial t} = \hat{H}\,\Psi(q,t) \tag{3}$$

This equation was introduced by Schrödinger in 1926 in order to provide a deeper formulation and understanding of quantum theory (see also (Merkt F and M Quack, 2011)). It was "pure research" in theoretical physics in its purest form, far removed from any practical goals or technical applications. Today there are estimates that about 20% of the gross national product of modern industrial countries are based in a general way on applications of quantum mechanics. Eq. 3 prevails in all applications of optical spectroscopy which range from the study of the Earth's atmosphere to combustion in car engines and industrial processes (Merkt F and M Quack, 2011, Quack M and F Merkt, 2011). We can also mention here nuclear magnetic resonance spectroscopy (NMR) with applications for example in MRI (Magnetic Resonance Imaging) available today in hospitals all over the world to name just these selected examples (Ernst RR, G Bodenhausen and A Wokaun, 1987, Ernst RR, 1992) among many more.

Another example, comparable to the development of quantum mechanics in the 20th century can be named with the development of electricity and magnetism in the 19th century. Technical applications of these are visible everywhere in our daily life today. However, the original developments were made long before these uses were obvious, although such future uses were predicted by Faraday with great foresight. He is reported to have replied to a question concerning the allegedly non-existing "profit" from his research "Lord Gladstone, one day you will tax it" (see (Kneißl D and H Schwarz, 2011)). We know, how true this prediction was, although this became obvious only many decades later than the government under Lord Gladstone existed. Many further examples could also serve to illustrate the different time scales of political governments, on the order of a decade in modern democracies, and the time delay of often many decades from a fundamental discovery to make it into textbooks and finally some practical, technical use (Quack M, 2014).

We shall conclude with an anecdote on the Schrödinger equation, which can serve as another illustration of what support to fundamental research implies. If one reads in Moore's Schrödinger biography on the history of the Schrödinger equation (Moore WJ, 1989) one finds that the first success in the search for this equation occurred during a stay of Schrödinger over the Christmas and New Year's holidays in Arosa 1925/1926, thus quasi as a private person in his free time without "financial support for this research". Obviously, a corresponding research proposal for "holidays, with the goal of discovering an equation" would hardly be funded by SNF today (it did not exist then), but the research was nevertheless indirectly publicly supported by giving Schrödinger the freedom of a professor with his salary at the University of Zürich to do research when and wherever he wanted to. Just how private this holiday was, can be guessed from some further gossip (Moore WJ, 1989), on which I shall not expand (Popper K, 1989). Moore states that Schrödinger stayed in "Villa Herwig" in Arosa, and provides a correct photograph of the house where holiday guests stayed (in contrast to the tuberculosis patients of Dr. Herwig in Arosa). However, if one looks in Arosa one finds that the house is actually called "Frisia", it still exists today. There is ample evidence that Schrödinger made his breakthrough during this holiday, and the first paper of the famous series was written and submitted immediately after return from this holiday in January 1926. However, we do not know, of course, whether the illumination came in this house, or perhaps some Café in Arosa. My preferred (rather unfounded) historical hypothesis is a visit of Schrödinger's to the Bergkirchli, a little church built by the Walser in 1492 and located in a most beautiful spot at 1900 m altitude only a short walk away from the Villa Frisia (Fig. 2).

Fig. 2 Bergkirchli in Arosa (built 1492, Photograph by R. Quack)

Such a heavenly inspiration of Schrödinger to find his equation describing the "music of atoms and molecules" can at best claim circumstantial evidence, although it is known that, for instance, Einstein's attitude towards science as a means of understanding the world had a strongly religious component (Quack M, 2004), not so different from the inspiration drawn by writers of "real music" such as Bach (Quack M, 2004, 2012a). This is, however, only one of several attitudes to be found with scientists, motivating their research as a route towards discovering fundamental underlying truths of the world. Another, more modest attitude aims at just providing fruitful models of the world, but in the end these two approaches may be closer relatives than obvious at first sight (Quack M, 2014). Independent of the particular philosophy of research, it deserves support as a basic part of our culture to understand and shape the world in which we live.

Acknowledgment: Our own scientific research receives support from the ETH Zurich, Schweizerischer Nationalfonds and by an Advanced Grant from the European Research Council. I am grateful to the hospitality of the Schmid family in their house in Arosa, where the present manuscript was written during a holiday and I am greatly indebted to Ruth Schüpbach for transferring my handwritten notes into a clean manuscript. The paper is dedicated to Richard Ernst on the occasion of his 80th anniversary.

Albert S, Meier BH, Quack M, Seyfang G, and Trabesinger A (2006) On the Possibility of Stable Nuclear Spin Symmetry Isomers in H_2O. Chimia 60:476

Ausländer R (2002) Und nenne dich Glück. Gedichte. Fischer Verlag, Frankfurt, The original text of the poem in German: Du bist - unwiderstehlich - Wahrheit - ich erkenne dich - und nenne dich - Glück (- indicates a new line each time in print)

Buchstein H (2011) Der Zufall in der Forschungsförderungspolitik. Swifts "Akademie von Lagado". Forschung und Lehre 8:596-597 (cites Les Allen after Neil Duxbury, Random Justice, Oxford 1999, page 89)

Diederich F (2013) Are We Refereeing Ourselves to Death? The Peer-Review System at Its Limit. Angew Chem Int Ed 52:13828-13829

Einstein A (1931) Zum kosmologischen Problem der allgemeinen Relativitätstheorie. Sitzber Preuss Akad Wiss Phys-Math Kl:235-237

Einstein A, and Preuss SB (1931) Akad Wiss:235 (This is the very common false version of the citation)

Ernst RR (1992) Nuclear Magnetic Resonance Fourier Transform Spectroscopy (Nobel Lecture). Angew Chem Int Ed Engl 31:805-823 Angew. Chem., 1992, 104, 817-836.

Ernst RR (2010a) The Follies of Citation Indices and Academic Ranking Lists, A Brief Commentary to 'Bibliometrics as Weapons of Mass Citation'. Chimia 64:90-90

Ernst RR (2010b) The Follies of Citation Indices and Academic Ranking Lists, A Brief Commentary to 'Bibliometrics as Weapons of Mass Citation'. Bunsen-Magazin 5:199-200

Ernst RR, Bodenhausen G, and Wokaun A (1987) Principles of nuclear magnetic resonance in one and two dimensions. Clarendon Press, Oxford

Gerhards J (2013) Der Deutsche Sonderweg in der Messung von Forschungsleistungen. In: Wissenschaftspolitik im Dialog, Schriftenreihe der Berlin-Brandenburgischen Akademie der Wissenschaften, Berlin

Heilbron JL (1989) The detection of the antiproton In: De Maria M, Grilli M, and Sebastiani F (eds) Proceedings of the International Conference on the Restructuring of Physical Sciences in Europe and the United States 1945-1960 (Rome, 1988). World Scientific, Singapore, p 161-209

Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 102:16569-16572

Kneißl D, and Schwarz H (2011) Editorial: Grundlagenforschung braucht exzellente Wissenschaftler -- und Freiräume. Angew Chem 123:12578-12579

Kuhn TS (1962) The structure of scientific revolutions. University of Chicago Press, Chicago

Manca Tanner C, Quack M, and Schmidiger D (2013) Nuclear spin symmetry conservation and relaxation in water ($^1H_2{}^{16}O$) studied by Cavity Ring-Down (CRD) spectroscopy of supersonic jets J Phys Chem A 117:10105-10118

Merkt F, and Quack M (2011) Molecular Quantum Mechanics and Molecular Spectra, Molecular Symmetry, and Interaction of Matter with Radiation. In: Quack M, and Merkt F (eds) Handbook of High-Resolution Spectroscopy. John Wiley & Sons, Ltd, Chichester, p 1-55

Molinié A, and Bodenhausen G (2010a) Bibliometrics as Weapons of Mass Citation. Chimia 64:78-89

Molinié A, and Bodenhausen G (2010b) Bibliometrics as Weapons of Mass Citation. Bunsen-Magazin 5:188-198

Molinié A, and Bodenhausen G (2011) The Kinship or k-Index as an Antidote Against the Toxic Effects of h-Indices. Chimia 65:433-436

Moore WJ (1989) Schrödinger, life and thought. Cambridge University Press, Cambridge

Osterloh M (2013) Das Paradox der Leistungsmessung und die Nachhaltigkeit der Forschung. Nova Acta Leopoldina NF 117:103-113

Osterloh M, and Frey BS (2012) Input control and random choice improving the selection process for journal articles (Manuscript preprint); Ranking Games, to appear in Evaluation Review (2014).

Parkinson CN (1957) Parkinson's Law and other Studies in Administration. The Riverside Press, Cambridge - Massachusetts
Perutz MF (1982) Ging's ohne Forschung besser? - Der Einfluss der Naturwissenschaften auf die Gesellschaft. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart
Petsko GA (2008) Having an impact (factor). Genome Biol 9 Article Number: 107
Pfaltz A, Van Gunsteren WF, Quack M, Thiel W, and Wiersma DA. 2009. An investigation with respect to the Possible Fabrication of Research Data reported in the Thesis ETH No 13629 and in the Papers Journal of Chemical Physics 112 (2000) 2575 and 113 (2000) 561, July 2009, ETH. Available: http://www.ethlife.ethz.ch/archive_articles/120123_Expertenbericht_tl/120123_Expertenbericht
Popper K (1989) Popper on Schrödinger. Nature 342:337-337
Quack M (2004) Time and Time Reversal Symmetry in Quantum Chemical Kinetics. In: Brändas E J, and Kryachko E S (eds) Fundamental World of Quantum Chemistry A Tribute to the Memory of Per-Olov Löwdin. Kluwer Adacemic Publishers, Dordrecht, p 423-474
Quack M (2011a) Fundamental Symmetries and Symmetry Violations from High Resolution Spectroscopy. In: Quack M, and Merkt F (eds) Handbook of High Resolution Spectroscopy. John Wiley & Sons, Ltd, Chichester, New York, p 659-722
Quack M (2011b) Naturwissenschaften! Warum überhaupt? Warum nicht? In: Gegenworte, Heft 26, "Zweckfreie Forschung". Berlin-Brandenburgische Akademie der Wissenschaften, p 28–35
Quack M (2011c) Naturwissenschaften! Warum überhaupt? Warum nicht? Bulletin der Vereinigung der Schweizerischen Hochschuldozierenden VSH/AEU 37:7-14
Quack M (2011d) Rede des Ersten Vorsitzenden anlässlich der 110. Hauptversammlung der Deutschen Bunsen-Gesellschaft für Physikalische Chemie in Berlin, 2. Juni 2011 (printed version of the speech). Bunsen-Magazin 13:138-143
Quack M (2012a) Rede des Ersten Vorsitzenden anlässlich der 111. Hauptversammlung der Deutschen Bunsen-Gesellschaft für Physikalische Chemie in Leipzig, 17. Mai 2012 (im Druck publizierte Fassung). Bunsen-Magazin 14:181-189
Quack M (2012b) Risiken und Chancen in der Forschungsförderung (Auszug aus der Rede des 1. Vorsitzenden der Deutschen Bunsen-Gesellschaft für Physikalische Chemie, Leipzig Mai 2012). Deutsche Zahnärztliche Zeitung 67:726-730
Quack M (2013) Error and Discovery: Why Repeating Can Be New. Angew Chem Int Ed 52:9362-9370
Quack M (2014) The Concept of Law and Models in Chemistry. European Review 22:S50-S86
Quack M, and Merkt F (eds) (2011) Handbook of High Resolution Spectroscopy. Wiley, Chichester, New York:
Schrödinger E (1926a) Der stetige Übergang von der Mikro- zur Makromechanik. Naturwissenschaften 14:664-666
Schrödinger E (1926b) Quantisierung als Eigenwertproblem I. Ann d Phys 79:361-376
Schrödinger E (1926c) Quantisierung als Eigenwertproblem II. Ann d Phys 79:489-527
Schrödinger E (1926d) Quantisierung als Eigenwertproblem III. Ann d Phys 80:437-490
Schrödinger E (1926e) Quantisierung als Eigenwertproblem IV. Ann d Phys 81:109-139
Straumann N (2008) Materie, Antimaterie, und Dunkle Energie. In: Walde P, and Kraus F (eds) An den Grenzen des Wissens. Vdf Publishers, Zürich, p 103-126. Here it is shown that a paper of A. Einstein (1931) (see citation) has been transformed by citations (obviously from people who never read or even saw the paper) until it reached its

"canonical form" with a second coauthor S. B. Preuss (A. Einstein and S. B. Preuss (1931)).
Suhm MA (2010) Fahren mit angezogener Handbremse. Bunsen-Magazin 12:200
Szilárd L (1961) The Mark Gable Foundation. In: The voice of the dolphins, and other stories. Simon & Schuster, New York
Tikhonov VI, and Volkov AA (2002) Separation of water into its ortho and para isomers. Science 296:2363-2363
Trommsdorff V (2002). Gesteinslaboratorium Alpen (Abschiedsvorlesung ETH Zürich). Bildarchiv der ETH-Bibliothek [Prod.],
Weinberg S (1967) A model of leptons. Phys Rev Lett 19:1264-1266
Zare RN (2012a) Editorial: Assessing Academic Researchers. Angew Chem Int Ed 51:7338-7339
Zare RN (2012b) Research impact. Curr Sci 102:9