Qualifying the performance evaluation of Big Science beyond productivity, impact and costs

Heidler, Richard; Hallonsten, Olof

Published in:
Scientometrics

DOI:
10.1007/s11192-015-1577-7

2015

This is the postprint version of an article published in *Scientometrics* 104(1): 295-312.

Qualifying the performance evaluation of Big Science beyond productivity, impact and costs

Richard Heidler
Department of Education and Social Sciences, Wuppertal University, Gaussstr. 20, 42119 Wuppertal, Germany

Olof Hallonsten
Department of Business Administration, Lund University, P.O. Box 7080, 220 07 Lund, Sweden. Corresponding author: olof.hallonsten@fek.lu.se

Abstract
The use of quantitative performance measures to evaluate the productivity, impact and quality of research has spread to almost all parts of public R&D systems, including Big Science where traditional measures of technical reliability of instruments and user oversubscription have been joined by publication counts to assess scientific productivity. But such performance assessment has been shown to lead to absurdities, as the calculated average cost of single journal publications easily may reach hundreds of millions of dollars. In this article, the issue of productivity and impact is therefore further qualified by the use of additional measures such as the immediacy index as well as network analysis to evaluate qualitative aspects of the impact of contemporary Big Science labs. Connecting to previous work within what has been called "facilitymetrics", the article continues the search for relevant bibliometric measures of the performance of Big Science labs with the use of a case study of a recently opened facility that is advertised as contributing to "breakthrough" research, by using several more measures and thus qualifying the topic of performance evaluation in contemporary Big Science beyond simple counts of publications, citations, and costs.

Keywords
Big Science; quality assessment; performance assessment; productivity; network analysis

1. Introduction

One of the most manifest consequences of the emergence of a globalized knowledge economy and intensified competition for funds and recognition in science is the seemingly unremitting flood of performance evaluation exercises. At the core of this culture of appraisal are found more or less standardized bibliometric measures used to keep track of the productivity, quality, relevance and impact of individuals, groups, institutes, and whole universities. Large science infrastructures (usually called Big Science) are not spared from this, but voluntarily advertise on websites and in annual reports their claimed excellence in the shape of numbers of publications, to prove that they contribute to scientific productivity and quality in the communities they serve, and thus are worth their cost. Publication counts therefore nowadays join two other classic performance measures for large scientific user facilities, namely operations reliability/technical performance and oversubscription rates/general user demand for experimental time, in what could be called "facilitymetrics" (Hallonsten 2013). In comparison with the former two fairly uncontroversial measures, bibliometric assessment of large scientific user facilities is highly problematic. Besides the general objections towards publication counts, e.g. that they are misrepresentative for quality, relevance and impact, calculations of return for investment show that if the output of a Big Science lab is measured solely by the amount of journal publications, the result is rather absurd, with the average cost per publication reaching several millions of dollars (Hallonsten 2014). An obvious counter-argument would be that the assessment of Big Science facilities should take into account not just numbers of publications produced but also the impact and scope of the respective publications, e.g. by the use of citation analysis. However, as this article will show, the simple summing up of citation numbers can also be misleading, and there is therefore a need for an even more balanced approach in the search for viable metrics for the quantitative performance assessment of contemporary large scientific facilities.

This article seeks to develop such an approach. It builds on the apparent disconnect between on one hand the continuing push towards using quantitative performance assessments in science, evidently reaching also onto the area of large scale scientific facilities (Hallonsten 2013), and on the other hand the clear inaptness of simplified publication and citation counts to evaluate the performance and quality of these facilities and labs (Hallonsten 2014). Extending the application of bibliometric analysis of large scientific facilities from straightforward counts of publications and citations to more complex and better nuanced measures and methods such as impact factor, immediacy index, and the construction of citation networks to make qualitative evaluations of productivity and relevance, the article qualifies the concepts performance and impact in the case of contemporary Big Science and calls for some self-reflection with regard to how the quest for excellence is translated into performance appraisals. A case study of reasonable size and with desired characteristics (cutting-edge technological setup and scientific program), for which the relevant data is also readily available, is used, namely the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory in Menlo Park, California.

The article conveys two conclusions, of which one is rather trivial and reductionist, namely that the scientific performance of large scientific facilities is difficult to quantitatively measure, and that bibliometric performance evaluation of Big Science is perhaps therefore also altogether unadvisable. The other conclusion has a constructive ambition and is somewhat more complex, and it translates to a call for a broader and more creative use of performance assessments of large scientific facilities that can also be interpreted as policy advice: Rather than using bibliometric measures in attempts to straightforwardly demonstrate productivity, impact or quality, the quantitatively oriented performance assessment of Big Science should use bibliometrics to show the distinctive contributions to science made by these facilities and labs in terms of supporting projects of inherently cutting-edge nature, and their apparent proneness to foster interdisciplinary research and recombinant science. The article takes some crucial initial steps in such an effort.

After a background and contextualization of the topic (section 2) and some basic information on the case (section 3), the basic publication and citation counts for the case are presented and compared with investments to show the outcome of such a first, rudimentary, performance assessment (section 4). Thereafter, the same data set is used to calculate impact factor and immediacy index of the publications and to construct citation networks, to show that with such slightly more sophisticated methods, other conclusions arguably more nuanced and reasonable are possible to draw regarding the productivity and impact of a contemporary cutting-edge Big Science lab (section 5). The concluding section reiterates the points made and finishes with some policy implications and suggestions for future research.

2. Facilitymetrics

Previous studies of large scientific labs and facilities, in history, sociology and research policy studies, have identified a clear shift in the political and social dimensions of Big Science. For the better part of the Cold War era, Big Science was largely motivated by a (remote) military connection and a modernist/rationalist (over)confidence in scientific and technological progress for the benefit of society, which resonated well with the bipolar global geopolitical situation where technological superiority was one key area of superpower competition (Galison and Hevly 1992; Greenberg 1999/1967; Westwick 2003). Beginning in the 1970s, other uses of large scientific instrumentation (accelerators and reactors) oriented to the study of materials (including living materials) began to grow and take over the organizations and physical infrastructure of Big Science labs, and became the rescue for many of them as they faced budget squeezes and extensive questioning of their usefulness from the political side in the 1980s and on (Hallonsten and Heinze 2012, 2013; Stevens 2003; Westfall 2008, 2012). This mounting pressure towards the end of the century was of course largely due to the broader shift in science policy frameworks in Western Europe and North America in the mid-1970s and on, originating in the economic downturn and the rise of neoliberalism and new modes for governance including a pressure for (demonstrable) productivity and efficiency as guide for public spending on science (Elzinga 2012; Greenberg 2001; Kevles 1997). For Big Science, the shift was further accentuated by the end of the Cold War, which took away much of the original rationale for heavy spending on nuclear and particle physics and put an end to its hegemonic position in publicly funded R&D. Interestingly, however, neither budget squeezes and neoliberal science policy, nor the end of the Cold War, did away with Big Science – quite the reverse, the new uses of large instrumentation were allowed to expand greatly and accelerator complexes are still built at many places in the world, although they are nowadays synchrotron radiation, neutron scattering and free-electron laser facilities that are used for experiments and measurements with applications for biology, materials science, chemistry, and condensed matter physics (Doing 2009; Hallonsten and Heinze 2013; Hallonsten 2013).

While the accelerator facilities of the old regime hosted projects protracted several years and employing hundreds or thousands of scientists, engineers and technicians, the accelerator facilities of today are essentially user-oriented and sustain wide assortments of radically different experiments by academic and industrial users across a wide variety of fields, who make occasional trips to those Big Science labs that make available the most favorable instrumentation for their work at the specific point in time. By reciprocity with globalization and internationalization of science, and technical advancements, these Big Science labs have come to be actors on a global market where they compete for the best users and most promising projects, which, by extension, make them prone to advertise their (technical) reliability, their popularity, and their productivity in the shape of a most prestigious publication records. Hallonsten (2013) has named this “facilitymetrics” and established a connection between the organization of contemporary Big Science labs as user-oriented service facilities for a wide range of sciences and the growth of an audit society and the increasing influence of managerialism on science (originating in the broad shift of science policy regimes in the 1970s and on, see above), paired with the proliferating use of quantitatively oriented evaluation of scientific excellence as discussed by several authors (e.g. Wildavsky 2010; Hazelkorn 2011; Münch 2013). The core of the argument is that Big Science facilities, just like all other institutes and organizations in publicly funded science, have come under increasing pressure to demonstrate productivity, excellence and competitiveness. In conjunction therewith, increased mobility of scientists and an expanded market of users for large scientific facilities, and the growth of the breadth of experimental opportunities offered by synchrotron radiation, free electron laser and neutron scattering labs, has made the new Big Science labs a prime feature of the multifaceted and complex R&D enterprises in fields considered strategic, such as nanotechnology, proteomics, and drug development. The labs therefore habitually advertise themselves and the competitive advantage they claim to have, by disclosing figures on technical reliability, user oversubscription, and publication/citation counts as proofs of productivity and quality (Hallonsten 2013: 504-510; 2014: 485-486).

In an attempt to analyze the consequences of Big Science labs' own use of (simplified) bibliometric measures to prove competitive advantage, Hallonsten (2014: 495) concluded that such simple publication counts lead to "rather absurd" results and therefore should be declared "irrelevant". But this conclusion is in one sense premature since Hallonsten (2014) only compared simple publication counts with expenditures, thus not at all making use of the those slightly more nuanced measures developed and nowadays routinely used for bibliometric analyses, such as impact factor and the immediacy index. This article thus takes the analysis one step further, making a more exhaustive analysis with the use of several more measures, and therefore makes another addition to the body of knowledge developed within the emerging area of "facilitymetrics".

3. The case

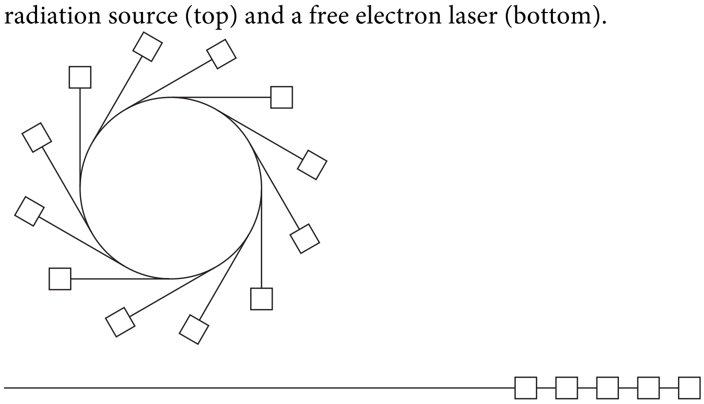
The choice of a case for the analysis was made on basis of the criteria that it should be of manageable size, reliable and complete, so that the data is searchable in the Web of Science (WoS hereafter) and comparable with other data in the same database. Moreover, in order to be representative and topical, it should be a user facility advertised as a flagship lab in its national science policy context and claiming to be technically groundbreaking and supporting cutting-edge research efforts (cf. Hallonsten 2013). The choice fell on the LCLS, which is one of two main user facilities at the SLAC National Accelerator Laboratory (formerly known as the Stanford Linear Accelerator Center, with the acronym SLAC), a dual-mission United States national laboratory for particle physics/particle astrophysics and so called "photon science".

SLAC was founded in 1962 as a single-mission, single-machine US national lab for particle physics, and started operation of its first experimental facilities (a linear accelerator, or *linac*) in 1966. Since then, it has built and run several machines for particle physics, and importantly, undergone a gradual transformation from a single-mission particle physics lab to a dual-mission and multi-purpose center nowadays dominated by its service to the scientific communities utilizing synchrotron radiation and free electron laser (Hallonsten and Heinze 2013). The LCLS is a free electron laser facility built on a recent (twenty-first century) extension of the original SLAC linac, and it was the first free electron laser to deliver x-ray laser light, which opens up a series of possibilities in materials and life sciences and at their intersection. Indeed, the original scientific justification for the LCLS as laid down in a series of reports in the late 1990s and early 2000s was that it would facilitate the exploration of completely new areas in the study of matter, within physics, chemistry and the life sciences, as well as intersections of these (Birgeneau and Shen 1997; LCLS 1998, 2000, 2002; Leone 1999).

The idea to turn the original SLAC linac into a free electron laser emerged already in the early 1990s, but there were severe technical uncertainties and risk associated with the proposal and SLAC's status as predominantly a particle physics lab with auxiliary (or, as it was called, "parasitic") synchrotron radiation activities on site also delayed the project's movement towards realization. In the late 1990s, as particle physics faced inevitable decline in importance on global level as well as on the SLAC site, synchrotron radiation and its extension free electron laser emerged as a viable alternative for securing the lab's future. In 2002, after a technical design had been worked out and the necessary scientific and political support had been won, the federal government decided to bet on the LCLS as the future infrastructural centerpiece of SLAC. Because of the leap in performance aimed for, the project was still considered a huge risk-taking, but the availability of a fully operational linac on the SLAC site plus world-leading expertise in both accelerator and instrument development worked in favor of it in policy circles. Compared to building the LCLS facility on green field, it is estimated that the use of the SLAC linac for the purpose saved the project "hundreds of millions of dollars" (Woods 2006: 12). The LCLS, states its website,[1] is "the world's most powerful X-ray laser" and "creates unique light that can see details down to the size of atoms and processes that occur in less than one tenth of a trillionth of a second" which enables the lab to facilitate "groundbreaking research in physics, structural biology, energy science, chemistry and many other diverse fields".

In the context of this article's ambitions, it is important to note the fundamental technical difference between the LCLS and its predecessors, i.e. previous facility projects that have broken new ground in synchrotron radiation technology and related research. The major investments in cutting-edge facilities for synchrotron radiation made in the late 1980s and early 1990s were put into the "storage ring" technical concept, which means a circular accelerator that permits the operation of dozens of experimental stations at once because the radiation utilized emerges evenly out in the tangential direction everywhere, at all times of operation. This means that world-leading synchrotron radiation facilities support 30-50 simultaneous experiments (Hallonsten 2013), whereas a free electron laser like the LCLS utilizes radiation from a linear accelerator which means that only one experimental station can be served at once; the LCLS has five stations in a row instead of the typical storage ring-based synchrotron radiation source's several tangentially placed around the ring (see figure 1). Planning and scheduling can of course enable some overlaps of e.g. sample preparation and calibration of instruments with actual radiation operation, so that several experiment teams can work at the same time, but the overall productivity of a free electron laser is severely limited in comparison with storage ring-based synchrotron radiation sources, although importantly the investments and operating costs are actually in the same range (Hallonsten 2014: 492).

Figure 1: Schematic sketches of the basic infrastructural difference between a storage ring-based synchrotron radiation source (top) and a free electron laser (bottom).



[1] http://portal.slac.stanford.edu/sites/lcls_public/Pages/Default.aspx

4. Publications and citations counts in relation to costs

With the same case study as in this article, and two historical cases for comparison, Hallonsten (2014) has shown that simple counting of publication numbers leads to the rather trivial conclusion that Big Science is extremely expensive: The cost of one single journal publication based on results from the LCLS, published in its third full year of operation, is calculated to $9.5 million(!). This is based on a count of total number of publications and total accumulated investment and cost of operations, which does not include all indirect use of existing resources of SLAC (see above) and is therefore an estimation that remains on the low side.

Although the absurdity of this exercise suggests that it indeed is questionable to merely count publications as a measure of the performance or productivity of Big Science labs, such simple counting is routinely done by the LCLS and other labs for promotional purposes (Hallonsten 2013: 501, 509). Since the LCLS is advertised as a breakthrough innovation that sustains and facilitates cutting edge science, it is probably only a question of time until the performance evaluation of the facility, for policy purposes or advertisement of "quality" or "excellence", incorporates bibliometric assessments and impact-oriented citation analysis. In fact, the lab itself invites such an exercise: Not only is a list of publications with results from work at the LCLS found on the facility webpage,[2] the same website also states that it is "extremely important in demonstrating the scientific impact of LCLS" and this publication list is instrumental for that. The list is compiled by the soliciting of publication data from the users of the facility, who are required to report publications. In addition to the list, a link on the website leads to a *Google Scholar* collection, where citation counts of the LCLS publications in simple yearly added up numbers can be inspected. The website publication list is divided into "accelerator science" and "x-ray science", with the former category relating to the design, construction and continuous technical refinement of the machine, and the latter to scientific experimentation performed at the LCLS. The list of publications in the accelerator science collection therefore starts with articles published already in 1995 and 1996 (see figure 2) when the first design concepts were developed, whereas the first publications in the x-ray science collection are from 2009, the year of start of scientific user operation.

The publication list was used to make a citation analysis on basis of the professional database conventionally used for such analyses, which is also more reliable than Google Scholar, namely WoS. The publications were identified using the DOI, or if not possible, the publication title. For the x-ray science publications 83 out of 93 were found and incorporated in the analysis, and for accelerator science publications 41 out of 47. The differing publications were almost all in the *Proceedings of SPIE*, a conference proceeding not indexed in the WoS (and thus for reasons of stringency also excluded from the analysis here), but aside from that, the publication output of the LCLS is well covered in the database.

In order to complement the rough analysis done by Hallonsten (2014) which pointed at an extremely high (but also every year radically diminishing) cost per publication of the LCLS, first, it could be hypothesized that the key contribution of cutting edge Big Science labs is not the quantity of output (pure number of publications), but quality and impact – i.e. the publication of "groundbreaking" results, as the facility itself states (see above). A next natural step in the analysis is therefore to compare construction and operations costs with a citation count instead of the publication count of Hallonsten (2014). Figure 3 shows graphically the yearly costs for construction and operation of the LCLS. As seen in the figure, construction

[2] https://portal.slac.stanford.edu/sites/lcls_public/Pages/Publications.aspx

costs reach their peak in 2007, while operations costs start to grow quickly from 2007 onwards and, at the start of user operation in 2009, saturate at $120-130 million annually.

Figure 4 shows raw year-wise citation numbers from 1995 and on. The numbers remain low in the period 1995-2008, when only a few articles in accelerator science were published, and from the start of user operation in 2009, citation numbers for the accelerator science publications grows linearly. In contrast, citation numbers for the x-ray science publications exhibits an exponential growth pattern at the beginning, but from 2012 onwards, the numbers start to show signs of a logistic s-curve pattern with saturation tendencies. As will be returned to later, this is already a sign of a diminishing impact, since the number of publications still grows exponentially, but it can of course also be (partly) explained by the time lag of citations.

Figure 2: Number of publications LCLS

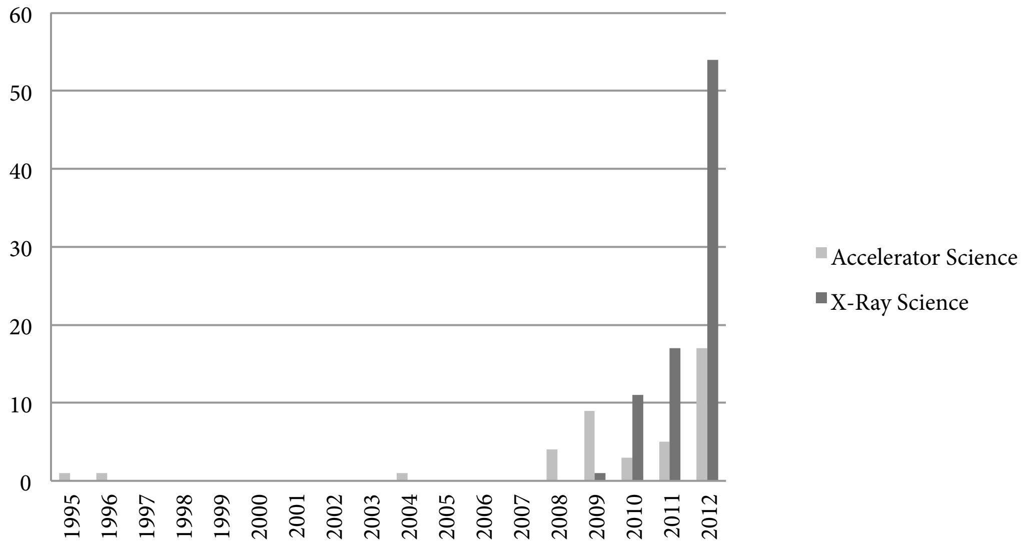
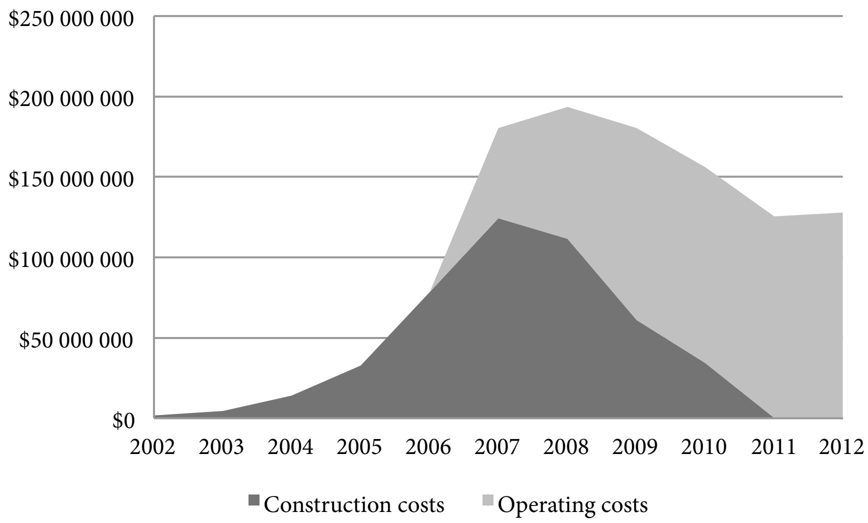


Figure 3: LCLS construction and operations costs, 2002-2012 (all in 2012 $).



There are different methodological ways to subsequently relate the citation numbers to the machine costs. One straightforward way is to divide annual accumulated construction and operation expenditures with annual number of accumulated citations, similar to the exercise in Hallonsten (2014). For 2009, as an example, the numbers would then be $686,204,410 (all construction and operations costs until the end of 2009 added together) and 182 citations (all citations recorded between 1995 and 2009 going to all articles published in the same period), which yields a cost of $3,770,354 per citation. Separating accelerator science and x-ray science, while maintaining the logic of the analysis that all costs (and gains) from the past should be added up in the year-wise values, a time series of costs per citations can be constructed and along with it, a characteristic curve, which is seen in figure 5. The curve progression shows that the total cost per citation grows strongly from 2006 onwards, as total expenditures grow dramatically but the number of citations remains low (cf. figures 3 and 4 above). The total cost per citation reaches a peak in 2009, where each citation to an LCLS publication which has been published until then is equaled to $3.77 million in accumulated costs. As user operation starts in 2009, citation numbers start to grow exponentially and expenditure is declining (construction costs) and saturating (operations costs), as shown above, the total cost per citation is also naturally declining. Relating construction costs to accelerator science and operation costs to x-ray science citations, a comparable pattern is discernible, as also shown in figure 5. All three curves converge in 2012, when each citation going to a LCLS publication has a calculated cost of $550,000.

Figure 4: Citations, 1995-2013

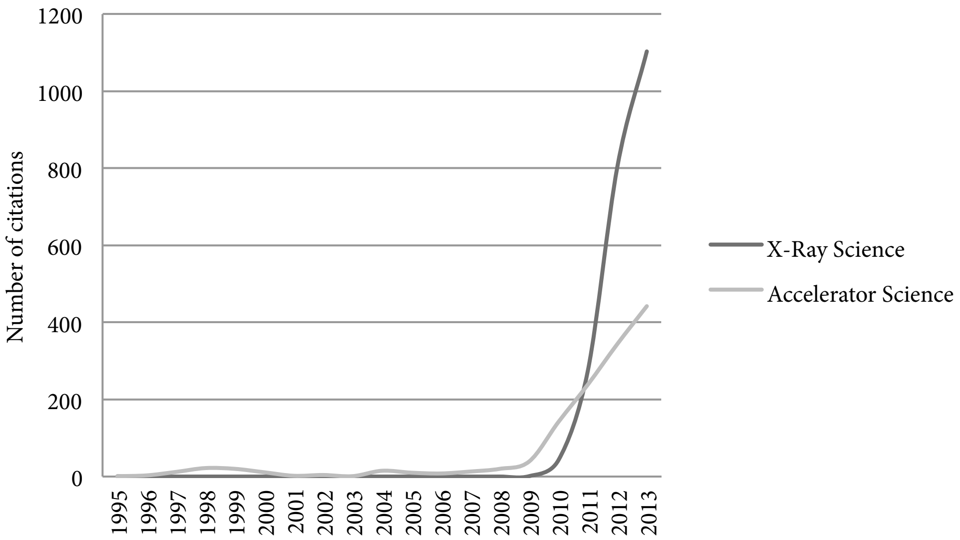
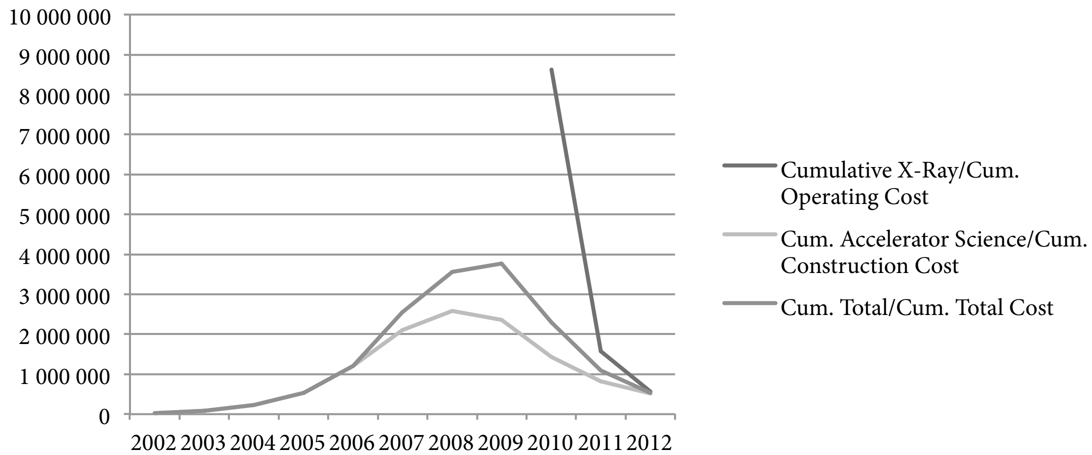


Figure 5: Expenditures (accumulated) per annual total number of citations



From an accounting perspective this computation is flawed in two respects, as the accountant's view would be how much funding is necessary in each year to produce one citation going to one article produced in the same year. From this viewpoint, first, the accumulation method needs to be discarded and replaced by a counting of only the investments in a particular year. Likewise, only citations going to publications that have been published in the same year should be counted. But since there is a time lag of citations, i.e. articles from one year require some time until they get cited, instead a reasonable time frame should be chosen, and depending on the length of that timeframe, the numbers can be computed retrospectively. Given the steep citation curves shown in figure 4, it is desirable to have a time frame as long as possible, but at the same time, the data set only allows for a time frame of three years since start of user operation (and hence start of real production of articles) was in 2009. Second, the accumulation approach to the calculations used above lacks differentiation of fixed and variable costs. There, construction costs could be treated as investments and the operation costs as variable costs. As an example, for the year 2011, the operating costs of $125,583,000 can be divided with the total number of citations (2011-2013) going to articles published in the year 2011, which is 854 (113+364+377), yielding a cost per citation of $147,052.69. In contrast to the first approach, this way of calculating could potentially lead to equilibrium of the average costs per unit, given that operation costs seem to be saturating (see figure 3 above).

With this second approach, another curve is produced, as shown in figure 6. Again, a decline in costs per citation is visible, but the curve seems to somewhat faster reach a continuous flat level. The approach leads to lower costs per citation, since future citations are incorporated in the current years and past costs are not integrated, and also a continuously declining cost per citation, since the total cost for the year in question remains intact whereas the citation numbers can be expected to grow. This results in a comparably low number of $74,610 per citation for the year 2012.[3] But although the sums of money reached in these calculations are significantly smaller than those calculated by Hallonsten (2014) for the costs of publications, they are still high, meaning in a sense that Big Science is extremely expensive also when attempting to measure impact and calculate cost per citation. A continued sharp decline in costs per citation over several coming years would perhaps lead to a situation where this tentative conclusion should be revised, but such a sharp decline will only occur if one of three following conditions is satisfied: First, declining operation expenditure, which seems improbable since operation costs are rather stable and historical comparison rather yields gradual increases in operation costs for Big Science labs (Hallonsten 2014). Second, growing number of publications with constant impact, which could happen, although as noted, the technical characteristics of the LCLS puts some fundamental limitation to such a development. Third, a growth in the average impact (number of citations) of the publications, which it is hard to estimate the plausibility of, but for which some guidance can be given by the further analysis presented in the next section, where average impact per publication is further qualified by the use of several different indicators and methods of analyzing impact.

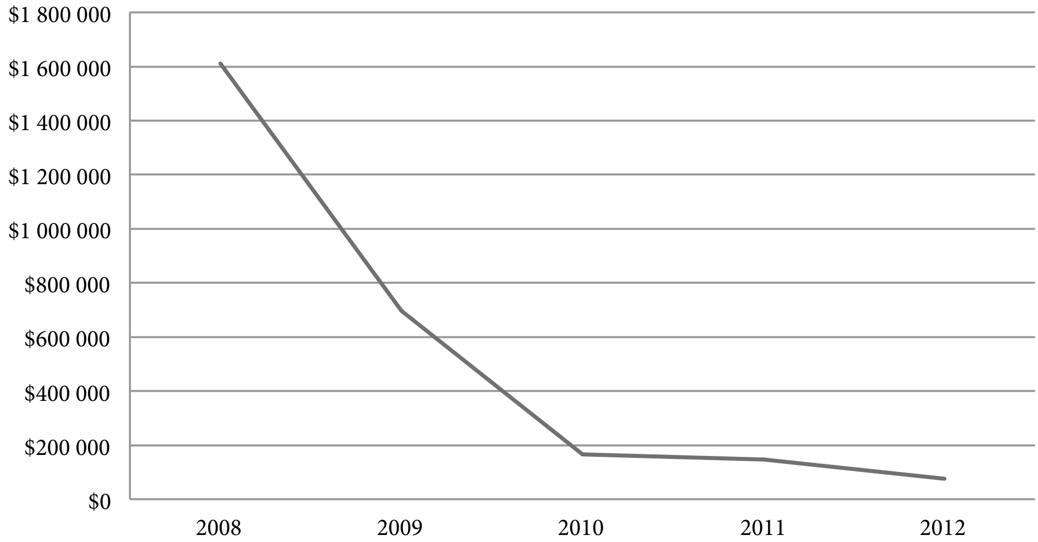
5. Qualifying impact

The calculations of costs per citation undertaken in the previous section only bolstered the previously made argument that current Big Science as represented by the LCLS, is an extremely expensive endeavor. However, it is necessary to nuance this conclusion by pointing

[3] The number of citations going to articles published in 2012 was estimated for the year 2014, since by the time of writing the article this period has not finished. The estimation was made by extrapolating on basis of the growth rate from the second to the third year from the citation curve of article published one year earlier.

at qualitative differences between different results and breakthroughs, in other words, to show that all impact is not the same. One goal of such an effort can be to try to determine whether there is a trend in the way the impact develops over the years of operation, in order to assess how the costs per impact will crystallize in the long run. Another is to compare indicators of impact scope and velocity to other well-known entities of impact assessment, such as journals. A third and significantly more complex way of qualifying impact is the use of citation network analysis, which would elucidate whether perhaps mere summing up of citations misses a potential special characteristic of the science produced with the aid of infrastructure like the LCLS, namely that it is interdisciplinary or recombinant, producing hybrid scientific results that in themselves have "breakthrough" character, such as endeavors at the intersection of e.g. physics and biology.

Figure 6: Operations costs divided on the citations (three-year time frame) to articles published in that year, 2008-2012



The first and second, development of impact over time, could be studied by the use of the well-known concept of *impact factor* as originally launched by Garfield (1955) as a means to evaluate the impact of academic journals and calculated by dividing the number of citations going to articles from the past two years in the current year, through the number of publications in the past two years. Such a calculation has a crucial advantage compared to the raw citation numbers presented above (figure 4) because it is standardized with number of publications, which means that it allows for the average impact per publication to be assessed. Figure 7 shows the yearly impact factor for LCLS publications in total, as well as divided on x-ray science and accelerator science. In the first three full years of operation, the impact factor is growing in both categories and reaches a top value of 21.167 (for all articles), only to decline in 2013. The decline is especially clear for accelerator science but seen also for x-ray science. If the trend continues, the annual costs per citation computed with the accounting method as discussed above and shown in figure 6 will at some point increase again, given that the number of publications per year (and the operation costs) remains roughly constant.

Comparing the 2012 LCLS impact factor with those entities for which this factor is usually calculated, namely journals, renders it a place as 44th among the 8,471 journals in the ranking of the Journal Citation Reports (Science Edition) of WoS. Thus although it would rank among the top one per cent in such an exercise, it is still not at the very top, and it trails e.g. the highly reputed multidisciplinary journals *Science* and *Nature* by several points: These journals had impact factors of 31.027 (*Science*) and 38.597 (*Nature*) in 2012, thus almost 1.5 and 2 times that of the LCLS, respectively. Among the disciplinary journals relevant to compare with the LCLS in terms of scientific areas concerned, namely journals in the WoS category "Applied Physics", "Optics" and "Biophysics", gives the following: LCLS would rank 3 out of 128 in Applied Physics, behind *Nature Materials* (impact factor 35.749) and *Nature Photonics* (27.254), and ahead of *Advanced Materials* (14.829). It would place itself in the second position (out of 80 journals) in the category "Optics", once again behind *Nature Photonics* (which is listed in both categories) but way before *Laser & Photonics Reviews* (7.976) and *Laser Physics Letters* (7.714). In the category "Biophysics", LCLS would be ranked number one (out of 72 journals), clearly beating *Acta Crystallographica Section D-Biological Crystallography* (14.103) and *Annual Review of Biophysics* (12.630). In this clearly asymmetrical but nonetheless highly intriguing comparison of impact factors, the LCLS at its height (2012) hence comes out quite strong in comparison with leading journals in some related scientific fields, but its publications' average impact is still clearly below the average publication in the high impact multidisciplinary journals *Science* and *Nature*. Figure 8 shows these comparisons visually.

Figure 7: LCLS publications, annual impact factor

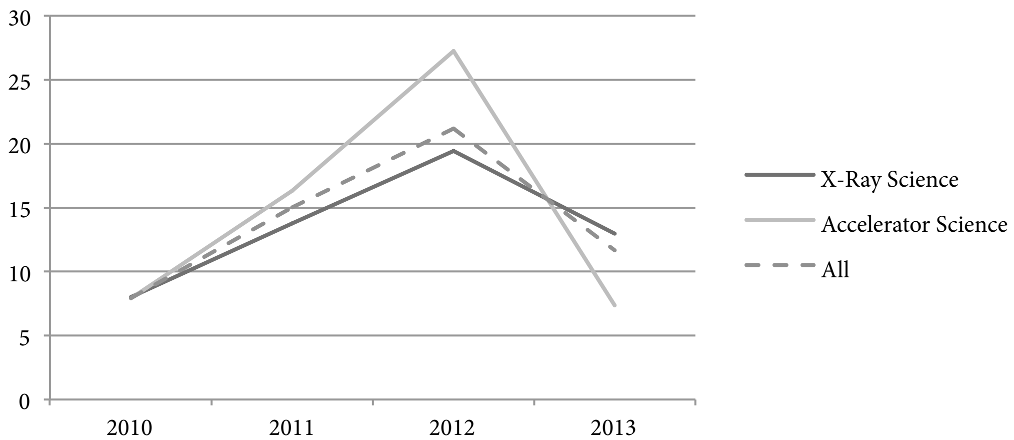
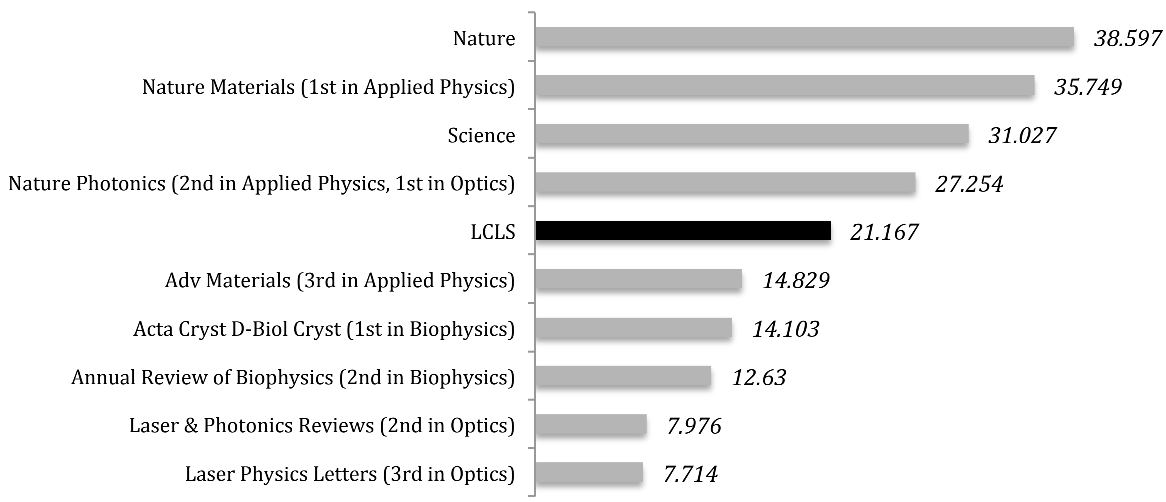


Figure 8: LCLS impact factor compared to some leading journals in relevant fields (2012)



As noted repeatedly above, a central argument for the construction of Big Science facilities in general, and also LCLS in particular, is that they produce groundbreaking research, which partly translates to research that would presumably fill urgent gaps in current scientific knowledge and thus would potentially lead to a high "immediacy index" for its publications. The "immediacy index" is, namely, a measure of how fast research is picked up and acknowledged in relevant communities, calculated by dividing the number of citations to articles published a specific year with the total number of publications in the same year. As figure 9 shows, when applying this measure, the values for LCLS are more volatile, especially for accelerator science, but the overall measure of immediacy index shows a certain pattern. A peak of 5.136 is seen in 2011, two years after start of user operation, and is followed by a slight decline, which is similar to the development of the impact factor (figure 7) although one year earlier. Again comparing the figure to those of journals, in the top year of 2011 the LCLS reaches the 37th position among all 8,471 journals (thus slightly better than general impact factor). It competes unfavorably with *Nature* (immediacy index: 9.690) but quite well with *Science* (6.075); in the category "Applied Physics" the LCLS reaches the 2nd position tightly behind *Nature Materials* (6.246), in the category "Optics" the LCLS reaches first position slightly above *Nature Photonics* (5.031), and in "Biophysics" it again reaches 2nd position between *Physics of Life Reviews* (10.917) and *Acta Crystallographica Section D-Biological Crystallography* (3.347). The comparisons are displayed graphically in figure 10, and yield a conclusion similar to that for impact factor. Although the general performance of LCLS is somewhat stronger when comparing immediacy index values than impact factor values, which indicates that emphasizing the cutting edge character of the science supported by the facility has some ground, the argument is not unambiguous.

Figure 9: LCLS publications, annual immediacy index

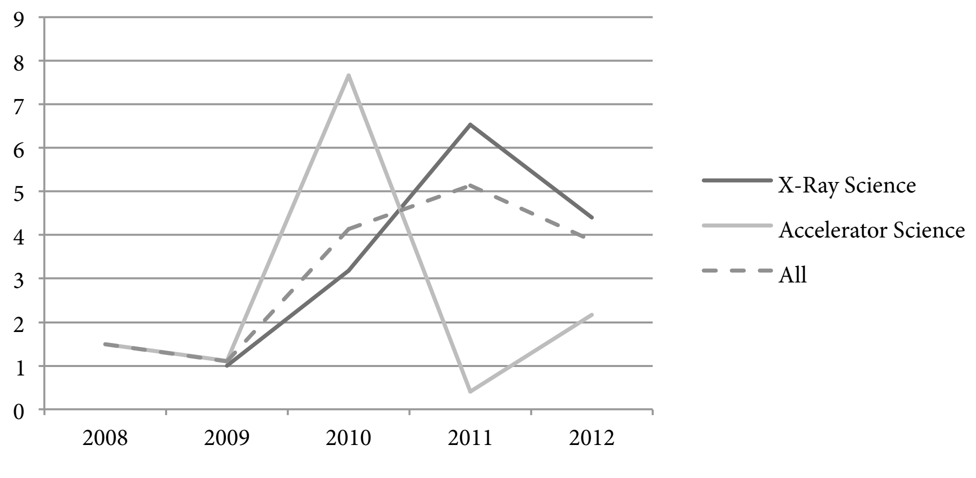


Figure 10: LCLS immediacy index compared to some leading journals in relevant fields (2011)

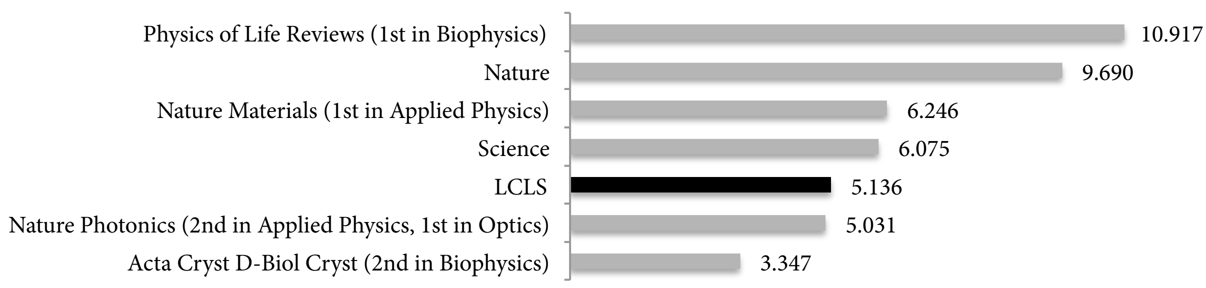
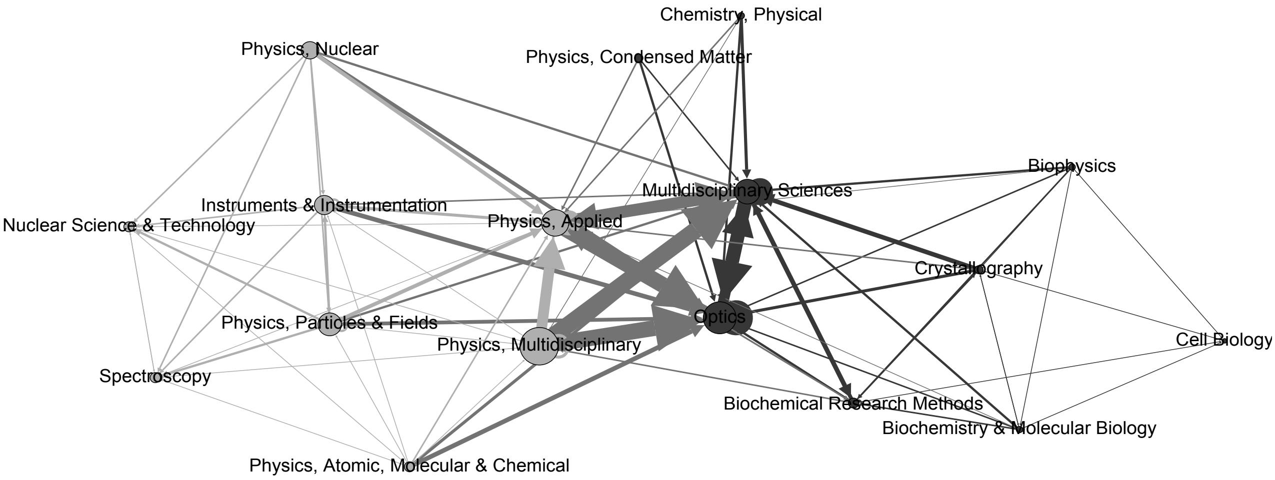


Figure 11: Citation network inside the set of LCLS articles (x-ray science and accelerator science) for the year 2012. Circles size represents the number of publications published in that year.



The third point raised above, that the LCLS facilitates interdisciplinary work and uniquely recombinant collaborations that can produce new forms of science, can be tried by the construction of annual citation networks with WoS subject categories as vertices. Those categories have been used in several previous studies of interdisciplinary constellations in various scientific contexts, where also the methodological problems inherent in using the WoS categories for classification of interdisciplinary research have been thoroughly discussed (e.g. Gowanlock and Gazan 2013; Porter and Rafols 2009; Small 2010). In spite of these problems, for reasons of simplicity and convenience, the categories are used here to create the exemplary visualization shown in figure 11, where all articles published in 2012, and all older articles cited in that year, are included. The figure shows, first, that there is a high interrelatedness of the articles published on basis of experimental work done at the LCLS. Second, it shows that this interrelatedness does cross scientific borders and brings together applied and basic physics and biology research. Having established this, the discipline-citation network was searched for community structures (Blondel et al 2008) which resulted in two main communities: A physics-oriented community (light gray) and a biology-oriented community (dark gray). This analysis clearly shows that the LCLS places itself at the intersection of biology and physics, with optics and applied physics as bridging disciplines. While the result of this analysis is rather unambiguous and quite intriguing, and perhaps therefore should be highlighted as one advisable method of proving a worth of kinds of the investments in contemporary Big Science labs like the LCLS, it carries a central problem in the context of this article. Although previous analyses have argued (and quite forcefully shown) that such hybrid fields based on new instruments are especially valuable for the growth of science, since they produce recombinant knowledge stimulating the whole system (Heinze et al 2013), it is difficult to relate this essentially qualitatively argued point to the issue of the expense and monetary worth of a Big Science lab like the LCLS, since it fails to present any hard numbers like cost per publication/citation as used above, and thus also fails to provide a counterargument to the statement that Big Science is extraordinarily (or even absurdly) expensive. This, of course, also has to do with the fact that standard performance indicators do not account for the complexities of this kind of output. Even a performance analysis going beyond the mere counting of publication and citation numbers by showing essentially qualitative values produced – like the citation network in figure 11 quite obviously does – would therefore not undo the claim made on basis of the conventional performance indicator systems. The fact that the latter thus are obviously curtailed has so far not prevented them from widespread use in many other contexts, although this might perhaps be an advisable strategy for moving towards better performance and impact evaluation of Big Science.

6. Concluding discussion

This article has used a topical, distinct and manageable case study to explore ways of using bibliometric impact assessment to go beyond mere calculations of cost per publications, which previous analyses have shown to render rather absurd results (Hallonsten 2014). Therefore, it adds an important piece to the analysis of performance, research quantity and productivity of Big Science labs as launched by Hallonsten (2013) under the name "facilitymetrics", adding the dimension of impact. Given the increased pressure on nearly all parts of public R&D systems to adhere to accountability demands by demonstrating clearly measurable performance, it is only a question of time until science policymakers and science managers in charge of funding streams will propose to use these types of analyses to determine the worth of continuous investments in Big Science.

The very rudimentary calculation of cost per unit of productivity (scientific articles) yielded the indirect conclusion that Big Science of the type represented by the case in question, the LCLS, is unsustainably expensive, and a similar calculation of cost per unit of impact (citations) confirmed it. The extreme costs of productivity and impact is partly due to the physical restrictions of the LCLS facility, which also logically leads to the search for alternative methods of measuring the worth of the investments made in the LCLS, totaling way over one billion dollars (Hallonsten 2014: 492), which is still a conservative estimate given the existing SLAC resources also mobilized around the project. As has already been argued, it is hard to justify the expenditure on the LCLS only by the use of a pure impact assessment that contents itself with the summing up of citation numbers – put bluntly, such an exercise inevitably leads to a questioning of the effectiveness of the facility as a scientific instrument, in comparison with ordinary "Small Science" or also other Big Science labs with similar costs but significantly higher output numbers, such as e.g. the European Synchrotron Radiation Facility, ESRF, whose construction and operations costs have been fully comparable with those of the LCLS but whose publication output was more than ten times higher in the comparable time period of a few years after start of user operation (Hallonsten 2014: 493).

Two counterarguments are possible to invoke. First, it can be argued that the LCLS (like the ESRF) is first and foremost a *user facility* and thus has the core purpose of providing excellent technical research *opportunities*, which does not mean that it can guarantee the quality of the research – this is up to the users to deliver on basis of their skills, competences and choice of research topics and problems. This argument is generalizable to most contemporary Big Science: Facilities for synchrotron radiation, free electron laser and neutron scattering are in operation and being built across the globe with the purpose of providing excellent research opportunities in the shape of cutting-edge instrumentation, for diverse scientific communities to utilize to the best of their ability. Hence, impact assessments focusing only on the facilities as such are essentially flawed and should be combined with some form of quality assessments of the actual users in order to give a fair picture.

Second, and related, it can be argued on basis of the final paragraphs of the previous section and the network image of figure 11 that the unique characteristics of the LCLS facility lies in its capacity to stimulate and sustain the forming of hybrid fields at the intersection of traditional scientific disciplines. This also implies that in order for "facilitymetrics" to make real sense, there is a need for (partly) new measures that can fully capture the unique capabilities of these facilities and their contributions to science. Unfortunately, this seems to mean going beyond the conventional and widespread measures of impact as used above, which this article (indirectly) has declared unfit for performance evaluation of Big Science.

Therefore, on this specific account, this article must leave the reader dissatisfied: Even this analysis, using several more metrics (cost per citation, impact factors, immediacy index, and network graph), has an ambiguous result and renders as its only clear conclusion that traditional performance indicators are not useful here. Therefore, the question of how to properly evaluate the productivity and impact of contemporary Big Science labs remains open. Unfortunately, this article does not manage to make any significant advances on the path to answering that question.

Which of course leads to the issue of research desiderata. It deserves to be reiterated here that the facilities themselves may not be so well prepared for meeting the challenges of an increased pressure to demonstrate productivity, impact and relative return for investments. Big Science labs like the LCLS should take the initiative in defining their goals and how these should be followed up in terms of (quantitative) performance assessment in order not to be unfairly judged. This would preferably include elements of the network analysis presented above, which can be used to pointing at e.g. the suitability of these facilities for fostering research that contributes to solving "grand challenges" and other inherently interdisciplinary issues, as well as varieties of applications of the immediacy index, and possibly other measurements more or less specifically tailored for assessing the benefits of cutting-edge large scale instrumentation of the type the LCLS provides to the scientific communities. Clearly, further studies are needed that attend to more cases and use larger data sets with longer time frames and further combinations of measures and comparisons, and these studies should also preferably involve a qualitative inquiry of what agendas and activities the labs themselves have in defining their productivity, quality and impact and take the initiative in proving their performance in a way that is both adequate and sufficiently convincing for policymakers and funders.

References

Birgeneau, B. & Shen, Z.-X. (1997). *Report of the Basic Energy Sciences Advisory Committee Panel on D.O.E. Synchrotron Radiation Sources and Science*. U.S. Department of Energy's Office of Science's Office of Basic Energy Sciences.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008(10): P10008.

Doing, P. (2009). *Velvet Revolution at the Synchrotron: Biology, Physics, and Change in Science*. Cambridge, MA: MIT Press.

Elzinga, A. (2012). Features of the current science policy regime: Viewed in historical perspective. *Science and Public Policy* 39: 416–428.

Galison, P. & Hevly, B. (eds) (1992). *Big Science – The Growth of Large-Scale Research*. Stanford, CA: Stanford University Press.

Garfield E. (1955). Citation indexes to science: a new dimension in documentation through association of ideas. *Science* 122: 108-11.

Gowanlock, M. G. and R. Gazan (2013) Assessing Researcher Interdisciplinarity: A Case Study of the University of Hawaii NASA Astrobiology Institute. *Scientometrics* 94(1): 133-161.

Greenberg, D. S. (1999/1967). *The Politics of Pure Science*, 2nd ed. Chicago: The University of Chicago Press.

Greenberg, D. S. (2001). *Science, Money and Politics: Political Triumph and Ethical Erosion*. Chicago: The University of Chicago Press.

Hallonsten, O. (2013). Introducing facilitymetrics: A first review and analysis of commonly used measures of scientific leadership among synchrotron radiation facilities worldwide. *Scientometrics* 96(2): 497–513.

Hallonsten, O. (2014). How expensive is Big Science? Consequences of using simple publication counts in performance assessment of large scientific facilities. *Scientometrics* 100(2): 483-496.

Hallonsten, O. and T. Heinze (2012). Institutional persistence through gradual adaptation: analysis of national laboratories in the USA and Germany. *Science and Public Policy* 39: 450- 463.

Hallonsten, O. and T. Heinze (2013). From particle physics to photon science: Multidimensional and multilevel renewal at DESY and SLAC. *Science and Public Policy* 40(5): 591-603.

Hazelkorn, E. (2011). *Rankings and the reshaping of higher education. The battle for world-class excellence*. Hampshire: Palgrave Macmillan.

Heinze, T., Heidler, R., Heiberger, H. & Riebling, R. (2013): New patterns of scientific growth. How research expanded after the invention of Scanning Tunneling Microscopy and the discovery of Buckminsterfullerenes. *Journal of the American Society for Information Science and Technology* 64: 829–843.

Kevles, D. J. (1997). Big Science and big politics in the United States: Reflections on the death of the SSC and the life of the Human Genome Project. *Historical studies in the physical and biological sciences* 27(2): 269-297.

LCLS (1998). *Linac Coherent Light Source (LCLS) Design Study Report*. SLAC-R-521, Rev. December 1998, UC-414.

LCLS (2000). *LCLS. The first Experiments*. Report, SLAC, September 2000.

LCLS (2002). *Linac Coherent Light Source (LCLS) Conceptual Design Report*. SLAC–R–593, April 2002, UC-414.

Leone, S. R. (1999) *Report of the Basic Energy Sciences Advisory Committee Panel on Novel Coherent Light Sources*. U.S. Department of Energy's Office of Science's Office of Basic Energy Sciences.

Münch, R. (2013). *Academic Capitalism*. London: Routledge.

Porter, A. and I. Rafols (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics* 81(3): 719-745.

Small, H. (2010) Maps of science as interdisciplinary discourse: Co-citation contexts and the role of analogy. *Scientometrics* 83(3): 835-849.

Stevens, H. (2003). Fundamental physics and its justifications, 1945-1993. *Historical Studies in the Physical and Biological Sciences* 34(1), 151-197.

Westfall, C. (2008). Surviving the Squeeze: National Laboratories in the 1970s and 1980s. *Historical Studies in the Natural Sciences* 38(4), 475-478.

Westfall, C. (2012). Institutional Persistence and the Material Transformation of the US National Labs: the Curious Story of the Advent of the Advanced Photon Source. *Science and Public Policy* 39(4): 439–49.

Westwick, P. J. (2003) The National Labs: Science in an American System 1947-1974. Cambridge, MA: Harvard University Press.

Wildavsky, B. (2010). The Great Brain race: How global universities are reshaping the world. Princeton: Princeton University Press.

Woods, H. R. (2006). New life for a linac. Symmetry, 3(7), 10–15.