Reference Standards and Reference Multipliers for the Comparison of the Citation Impact of Papers Published in Different Time Periods

Werner Marx
Max Planck Institute for Solid State Research, Heisenbergstraße 1, D-70569 Stuttgart, Germany. E-mail: W.Marx@fkf.mpg.de

Lutz Bornmann
ETH Zurich, Professorship for Social Psychology and Research on Higher Education, Zähringerstr. 24, CH-8092 Zurich, Switzerland. E-mail: Lutz.Bornmann@gess.ethz.ch

Manuel Cardona
Max Planck Institute for Solid State Research, Heisenbergstraße 1, D-70569 Stuttgart, Germany. E-mail: M.Cardona@fkf.mpg.de

In this study, reference standards and reference multipliers are suggested as a means to compare the citation impact of earlier research publications in physics (from the period of "Little Science" in the early 20th century) with that of contemporary papers (from the period of "Big Science," beginning around 1960). For the development of time-specific reference standards, the authors determined (a) the mean citation rates of papers in selected physics journals as well as (b) the mean citation rates of all papers in physics published in 1900 (Little Science) and in 2000 (Big Science); this was accomplished by relying on the processes of field-specific standardization in bibliometry. For the sake of developing reference multipliers with which the citation impact of earlier papers can be adjusted to the citation impact of contemporary papers, they combined the reference standards calculated for 1900 and 2000 into their ratio. The use of reference multipliers is demonstrated by means of two examples involving the time adjusted *h* index values for Max Planck and Albert Einstein.

Introduction

It is possible to determine which papers are highly cited (and which ones are seldom cited) only on the basis of field-specific reference standards, because of differences in the expected citation frequencies within different fields. Radicchi, Fortunato, and Castellano, (2008) found, for example, that papers in "engineering and aerospace" are cited on the average 5.65 times, whereas papers in "astronomy and astrophysics" are cited on average 23.77 times. To make the citation performance of papers comparable over field and subfield boundaries, the different field-specific expected citation frequencies must be related to one another through the use of predetermined reference standards.

Two paths have emerged for the construction of field-specific reference standards: (a) A mean value for the citation impact is calculated for those papers to which a literature database attributed a given (sub)field. For Chemical Abstracts (CA), for example, the Chemical Abstracts Service (CAS) categorizes chemistry publications into 80 different subject areas (chemistry subfields, called "sections"). Every publication becomes associated with a single principal entry that makes clearly apparent the most important aspect of the work. The calculated mean value represents the average impact of the publications in a subfield. If, for example, a reference standard for a chemistry paper from 2000 is required, which was attributed by CAS to the section "surface chemistry and colloids," the average citation rate is calculated across all publications that appeared throughout the year 2000 and were attributed to this section by CAS.

(b) Instead of calculating the average citation rate for publications that are attributed to a field category in a database, the possibility also exists to construct a reference standard that reflects the average citation rate of the publications that were published in a specific journal. A relevant paper, for which a reference standard is needed, may have been published, for instance, in the year 2000 in *Tetrahedron Letters*,

Received March 25, 2010; revised April 30, 2010; accepted April 30, 2010

 • Published online 18 June 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21377

in which case, the average citation rate is calculated on the basis of those papers that have appeared in that journal during that year.

At the moment, the conventional way to standardize the citation impact of one individual paper, and, thereby, make the impact comparable for other fields, is to divide its citation count either (a) by the field-specific reference standard or (b) by the journal-based reference standard (see Opthof & Leydesdorff, in press; Radicchi, et al., 2008). The reference standards correspond to the expected citation rates, which underlie the (journal-based or field-specific) relative citation rates (van Raan, 2004). If that division yields a value of approximately 1, then the paper in question was cited as often as an average paper (a) from its field or (b) from that journal (depending on the procedure used). Values higher or less than 1 correspond to the above or below average influence of a paper respectively—according to citation rates.

Because the citation indexes among the Thomson Reuter's Web of Science (WoS) extend back to 1900 (Century of Science Project; see http://science.thomsonreuters.com/news/2004-11/8254441/), the papers from the first half of the 20th century are generally available and their citations are also available for bibliometric investigations. The question emerges as to how the citation impact of these papers may be evaluated in comparison with the impact of contemporary papers. Just as different scientific fields and subfields have different citation practices, we must also compare the widely diverse publication and citation cultures of different periods leading up to the development of modern science. A comparison of the citation impact of papers by the pioneers of science (such as Albert Einstein) with the impact of papers by contemporary researchers (such as the physicist Edward Witten) requires that one either standardizes citation rates with the help of reference standards or readjusts them to suit the present practice within the field. The citation counts of papers written by the pioneers of physics, given under heading "times cited" in WoS, can be compared only with the citation counts for papers by contemporary scientists, if the relationship between citation counts of earlier and contemporary papers can be ascertained. (This kind of comparison affects not only the greatest scientists but also the lower tiers.) The following specific question guides this investigation: How many citations made in the present time correspond to one citation made at the beginning of the 20th century? The central motivation of our inquiry is to discover by what factors are contemporary papers more often cited than earlier papers.

The availability of bibliometric data from the first half of the 20th century in the WoS leads to a situation where these data are also drawn on for performance evaluation studies. De Visscher (2010), for instance, calculated a series of bibliometric indicators for Albert Einstein and contemporary physicists and arranged them in a comparative overview. Without the consideration of reference standards or without adjusting citation impact for time period, however, the indicators for physicists could hardly be interpreted as significant, because they worked in such widely differing time periods. (The author hints indirectly at this shortcoming when he refers to the results for Einstein's citation rates as "surprising results"; see p. 324.)

In the following pages, we envision approaches—on the basis of the approaches to field standardization in evaluative bibliometry—with which we can undertake (1) a *time standardization* of citation impact from the first half of the 20th century or (2) a *time adjustment* of their impact to reflect contemporary circumstances. Next, after the Methods section, the necessity of time standardization and time adjustment will be demonstrated in our discussion of certain publication and citation rates that are used to measure scientific output and citation impact in the eras of "Little Science" and "Big Science." After that, we introduce our approaches to the procedure for time standardization and adjustment. These approaches are illustrated by using bibliometric data from physics; they are, however, fully transferrable to other natural science fields. Finally, we apply the time adjustment of citation impact to two specific cases: We calculate the adjusted *h* index (Hirsch, 2005) values for Max Planck and Albert Einstein that would correspond to contemporary standards.

Methodology

We investigated the number of publications in physics (and related fields) since 1900 up until today in the Inspec database. The Inspec database of the Institution of Engineering and Technology (IET) is a comprehensive index of the global journal and conference proceedings literature in physics, electrical/electronic engineering, computing, control engineering, and information technology. For the present analysis of the data from Inspec, papers of every document type were considered; these results are given at the beginning of the Results section.

The data, on which we base our citation analyses, presented in the Results section, were accessed from the Science Citation Index (SCI) within WoS. We drew exclusively from the SCI and not from the other WoS citation indices (such as the Social Science Citation Index), because they do not extend as far back as 1900 (and are also nearly irrelevant concerning physics). For the calculation of the average number of references in papers related to physics or from a particular physics journal, the SCI was accessed through the database provider STN International, because of the system limitations of the WoS to only 100,000 records of citing papers. The citations of papers by Einstein on special and general relativity were captured including the erroneously written references. In particular, among the papers from the first half of the 20th century, the "times cited" results in the WoS become less reliable.

Three physics journals could be singled out for the determination of *journal-based* reference standards. These journals, *Physical Review, Philosophical Magazine*, and *Astrophysical Journal*, have continually published papers from 1900 up through today. (Other important physics journals were founded later or discontinued by then.) *Physical Review* is one of the oldest and, nowadays, most esteemed journals in the field of physics. It has existed since 1893 and has been published since 1913 by the American Physical Society (http://www.aps.org/). *Philosophical Magazine* was founded in 1798 and is now the oldest commercial scientific journal in the world. At the time of its foundation, this journal published papers from the entire spectrum of natural sciences, whereas today it solely covers condensed matter physics. In 1978, the journal was split into two different parts, which have been reunited since 2003. Since 1987, *Philosophical Magazine Letters* has been published regularly. *Astrophysical Journal* was founded in 1895 and is still issued today by the American Astronomical Society. Supplementary material has been published under the name *Supplement Series* since 1953.

On the basis of these three journals, the citation rates of papers from the ages of Little Science and Big Science may be directly compared, because the relevant journal can be held constant in the process of calculating reference standards. We accessed all papers from the years 1900 and 2000 from these three journals and ascertained the number of citations within the first decade after their publication (1900–1909 and 2000–2009, respectively). Counting the citations within the first decade after publication and not across many decades is inevitable here but clearly omits the (few) papers with delayed recognition, sometimes called sleeping beauties (van Raan, 2004). From the number of citations within the first decade, the average number of citing papers per paper was calculated for the two publication years (rather than the average number of citations per paper, see also the Discussion section). One citing paper can, in principle, contain more than one citation of a specific volume of a given journal. This happens, for example, whenever a citing paper cites several papers from the same journal and the same year. Because of the relatively low-mean number of references in early works, this probability remains slight for those cases.

To determine the *field-specific* reference standards adequately, we selected the entire literature on physics (all types of documents) from the years 1900 and 2000 from all subject areas relevant to physics in WoS (such as "physics, condensed matter" or "physics, nuclear"). Then, the citation rates were established for papers for the period within the first 10 years after those papers were published. With these numbers, the average number of citing papers per paper was calculated for papers from each publication year.

For the calculation of the *h* index (Hirsch, 2005) of Planck, his papers since 1900 were collected in the source journals of the SCI, as a first step of our searching iteration. Papers by namesakes (the same first given names initials and the same surname) were sought and excluded manually. Work that emerged after 1950 was excluded from this study, because Planck died in 1947 and the papers that appeared later were classified as irrelevant (only two reprinted articles; both with zero citations). In this way, 88 papers by Planck could be compiled into a publication list. These papers were sorted by the number of times they were cited (from the date of publication until the present). In a second iteration step, we employed the Cited Reference Search of the WoS to determine the citation rates of the papers by Planck that appeared before 1900 and, thus, were not stored in the WoS database. Our research does not include all papers by Planck that appeared in the Cited Reference Search, but rather only those with at least two citations. Most of the entries with one citation are erroneous citation variants, which could be attributed only to a publication by Planck through an interpretive leap. Over the course of the century, references become irrevocably corrupted by errors (similar to genetic mutations) that authors make while writing up bibliographical information. We could identify 12 papers by Planck, which appeared before 1900 and were cited after 1900 (the citations from the year of publication up until 1900 are currently still not included in the WoS). These papers have been supplemented with their citations in the publication list established for Planck during the first step of searching.

An analogous procedure to the one used for Planck allowed us to determine the *h* index for Einstein as a second example. Einstein had not published any (noteworthy) papers before 1900, a fact which made our procedure significantly simpler. Einstein died in 1955, but the 13 posthumous reprints of papers and speeches compiled by SCI were not cited except for two papers (with seven and one citations respectively). We found a total of 165 papers by Einstein and arranged them into a publication list along with their respective numbers of citations.

Results

Little Science Versus Big Science—The Necessity of Time Standardization and Adjustment for Citation Impact

In the first half of the 20th century, the number of researchers active in the natural sciences was relatively low compared with the number today. Accordingly, much less was published. One measure for the output of the research in a disciplinary field derives from the number of publications compiled yearly in field-specific databases. The Inspec database is very well-suited for the investigation of the temporal development of the output in physics. It has compiled a more thorough collection of literature in physics than WoS, for example.

As shown in Figure 1, the temporal development of the physics publications, which this database compiles every year, shows a marked rise in the publication rate between 1955 and 1960. The development of the literature in chemistry shows a similar course, albeit with a less formidable rise (Marx & Schier, 2005). Following the well-known book by Derek de Solla Price (1965), one may refer to the time period before 1955 and after 1960 as Little Science (individual science) and Big Science (team science), respectively. The rapid ascent of the number of publications between 1955 and 1960 marks the transition from Little Science to Big Science, evidently set off by the so-called "Sputnik Shock" (Dickson, 2007). After the former USSR had put the first satellite into orbit around the earth in 1957, funds were hurriedly allocated in previously unheard of quantities to the funding of research and development in western industrial

DOI: 10.1002/asi

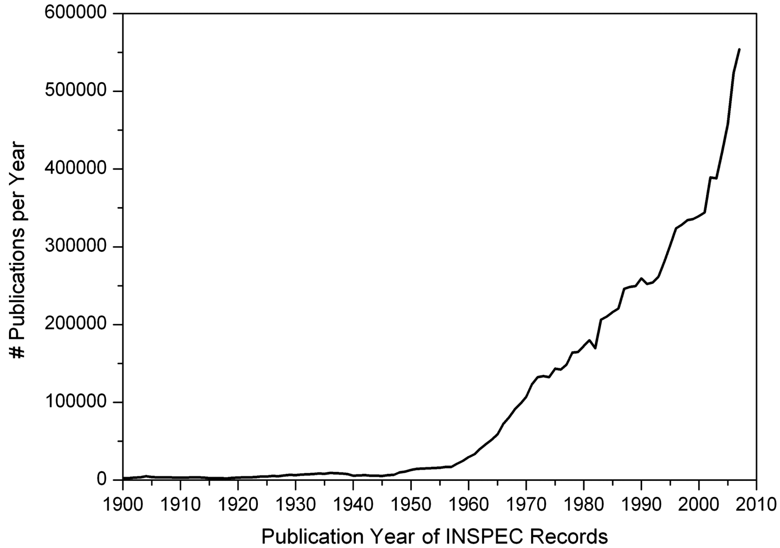

FIG. 1. Number of physics publications per year compiled by the Inspec database.

nations, and especially the United States. As a result of this expansion process, for example, the annual budget for the National Science Foundation was suddenly quadrupled. Then, space exploration and military technology became the predominant fields, and the education system was comprehensively reformed as increased funds were channeled into it (see the National Defense Education Act, NDEA, passed by the U.S. Congress in 1958). These measures promptly enhanced the number of scientists, so research activity and scientific output dramatically increased.

Papers from the period of Little Science are not only less numerous but also significantly less frequently cited by each other than papers produced in the period of Big Science (for a given paper in the early 1900s, there were much less potential readers than there are now). This becomes apparent when considering the low mean number of references (citations) in the appendices of earlier papers as compared with the higher mean number of references in the papers today. The probability that a certain paper will be cited—its "citation potential," as Garfield (1979) defines it—depends above all on the mean number of references in the citing papers. The (arithmetic) mean number of references in physics journals from the year 1900 (810 papers in total) is 6.0 (median: 2). The (arithmetic) mean number of references in the papers that appeared in physics journals in 2000 amounts to 22.2 (here, unfortunately, the median cannot be determined because of the high costs of the required database access).

Most papers published at the beginning of the 20th century were hardly cited after 1960. Of the 810 physics papers retrieved from 1900, the WoS shows only 12 citations in physics papers during the year 1960 (none corresponding to Planck, one corresponding to Einstein). Papers of fundamental significance, such as those of Planck and Einstein, do, however, show lasting effect. Figure 2 plots the cumulative citations of the papers by Einstein on special relativity theory (Einstein, 1905) and on general relativity theory (Einstein 1915, 1916). It is significantly apparent that these papers are also frequently cited after 1960—even significantly more often than before 1960. The sharp rise in publication rates in physics around 1960 is also clearly reflected in the varied distribution of citations for these papers. At that point in time, one observes a left-skewed distribution (see Figure 2), as opposed to the typical right-skewed distribution of citations found in contemporary papers, a pattern thoroughly discussed by Seglen (1992).

Calculation of Reference Standards for Time Standardization

The once renowned German physics journals (such as the *Physikalische Zeitschrift*) were discontinued after World War II, ceased publication in later years (as in the case of the *Zeitschrift für Physik*), or lost their esteem among Anglo-Saxon scholars (as the *Annalen der Physik* was replaced in reputation by the *Physical Review*). Therefore, these journals cannot be drawn upon for the calculation of reference standards in the execution of time standardization. Of the three journals employed for this study, the *Astrophysical Journal* has been investigated over the period of a whole century with bibliometric methods (Abt, 1995). In addition, Redner (2005) has conducted a comprehensive study on the *Physical Review* (however, only citations of papers within that journal were used as data pool).

In Table 1, for the years 1900 and 2000, the total number of papers in the three physics journals—*Physical Review*, *Philosophical Magazine,* and *Astrophysical Journal*—is presented alongside the total number of physics papers published within each of those 2 years. Additionally, the number of citing papers within the first 10 years after publication is displayed. Using the published papers and their citations, the average citation rates are calculated. For papers from 1900, the values vary between 0.26 and 0.48 citing papers per paper.

DOI: 10.1002/asi

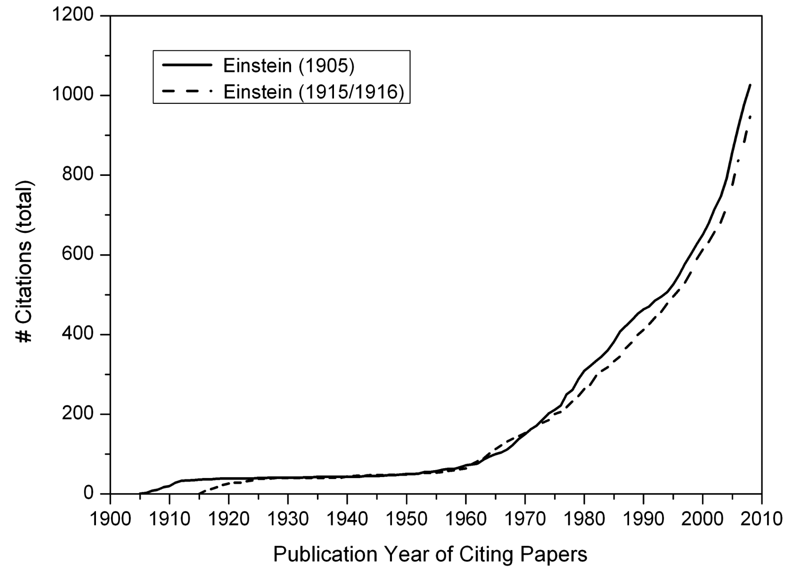

FIG. 2. Cumulative number of citations of two papers by Albert Einstein on special relativity theory (Einstein, 1905) and on general relativity theory (Einstein, 1915, 1916) from publication until 2009.

TABLE 1. Number of papers, number of citing papers, number of citing papers per paper (reference standards) for three physics journals and for the physics literature on the whole.

Journal title	# Papers (1900)	# Citing papers (1900–1909)	Citing papers per paper (1900)	# Papers (2000)	# Citing papers (2000–2009)	Citing papers per paper (2000)
Astrophysical Journal (including supplements)	104	27	0.26	2600	53479	20.57
Physical Review (sections A-E, including letters)	52	22	0.42	14343	237681	16.57
Philosophical Magazine (sections A-B, including letters)	70	21	0.30	452	4185	9.26
Physics total	810	390	0.48	86977	632050	7.27

The rates are, accordingly, at a similar very low level. Papers from the year 2000, in contrast, yield significantly higher values: between 7.27 and 20.57 citing papers per paper. These average values, which we calculated for the publications from the years 1900 and 2000, constitute the reference standards, with which the time-standardized citation impact values can be calculated for individual papers from the period of Little Science and Big Science. The lowest reference standard for a year (1900 or 2000) designates the lower limit (e.g., 7.27 for 2000), and the highest value designates the upper limit of the expected citation rate (e.g., 20.57 for 2000) for a journal or field.

Accordingly, time-standardized citation impact values were established for certain sample papers: the most cited paper in *Philosophical Magazine* from the year 1900 was written by Ernest Rutherford, one of the leading forerunners of atomic physics (Rutherford, 1900). This paper garnered three citations by the end of 1909. If we divide this number by the journal-based reference standard (0.30) and, thereby, time standardize it, the value of 10.0 emerges. The most cited paper from the same journal in 2000 (Legros, Elliott, Rittner, Weertman, & Hemker, 2000) was cited 159 times by the end of 2009. The time-standardized value (the journal-based reference standard is 9.26) yields 17.2 for this case. Compared with the paper from 1900, the paper from 2000 still has a higher citation impact. However, an analysis of the impact of another highly cited paper by Rutherford (1911), which he also published in *Philosophical Magazine*, yields a wildly different impact relative to the later paper of Legros et al. This paper treats Rutherford's famous gold foil scattering experiment, which led to the modern atomic model and had been cited 29 times by the end of 1920. If this number is divided by the journal-based reference standard for the year 1911 (0.67), a value of 43.3 emerges. Only half of the 1911 volume of *Philosophical Magazine* is available on WoS, and, thus, it is an inadequate basis for the construction of a reference standard. When we divide the number of citations (29) by the journal-based reference standard for the year

DOI: 10.1002/asi

TABLE 2. Four reference multipliers calculated in order to adjust the citation impact of early (astro)physics papers to the impact of contemporary papers.

Journal title	Citing papers per paper: 2000/1900
Astrophysical Journal	79.12
Physical Review	39.45
Philosophical Magazine	30.87
Physics (total)	15.15

TABLE 3. Example for the application of reference multipliers to a specific paper: Einstein's (1905) paper on special relativity theory.

Basis for reference multiplier	Reference multiplier	Calculation	(Adjusted) Citation count
Actual cites from 1905 to 1960	0		69
Physics (total)	15	15 × 69	1,035
Physical Review	40	40 × 69	2,760
Astrophysical Journal	80	80 × 69	5,520

1910 (0.41), we obtain the value 70.7. The time-standardized citation impact is, thus, significantly higher than that of the above-mentioned papers from 2000.

Calculation of Reference Multipliers for Time Adjustment

When we use the numbers in Table 2 to determine the relationships between citing papers per paper (reference standards) for papers cited from the two different years in question (1900 and 2000), the collated reference multipliers in Table 2 result in time adjustment. The average citation rates for the entire literature in physics in the year 1900, for instance, equals 0.48 citing papers per paper and that of the year 2000 equals 7.27. From these two numbers, the relationship of the citation rates from 2000 to the one from 1900 is 15.15 (7.27/0.48 = 15.15). This ratio suggests which factor a single citation of an early paper should be multiplied by to compare its impact with that of a contemporary paper. Therefore, it constitutes the field-specific reference multiplier for time adjustment. For both journals, *Philosophical Magazine* and *Physical Review*, reference multipliers turn out to be around 30 or 40; the multiplier for *Astrophysical Journal* is almost 80, significantly higher than for the other two journals.

The number of references (reference count) of a given paper depends on many parameters (Bornmann & Daniel, 2008). Abt (1987) showed that the average number of references per paper is linearly proportional to their length, and that the linear relation is the same for astronomy, chemistry, geophysics, and physics. The reason why astronomers have more references in their papers is obviously because their papers are systematically longer. Hence, the currently higher citation rates and the resulting comparatively high reference multiplier for *Astrophysical Journal* could be because of the trend in physics to shorter publication units.

The reference multiplier for physics literature as a whole must be regarded as a lower limit, because a larger difference between the number of citing papers and the number of citations must be expected especially for the year 2000. It is significantly more probable that a citing paper would cite several different papers from any physics journal published in 2000 than it would cite several papers from only one journal from that year. This fact raises the reference standard for 2000 and with it, the value of the reference multiplier.

We explain the application of reference multipliers here by using sample papers. Whereas the citation frequency of a paper is divided by the reference standard for field standardization and time standardization, a reverse route is taken in this case: the citation frequency of the paper is multiplied by the reference multiplier. The paper by Einstein on special relativity theory (Einstein, 1905) received 69 citations by the end of 1960. If we multiply this number by the field-specific reference multiplier of 15 and, thereby, adjust it for the context of all work in physics from the period of Big Science, the result becomes 1,035 citations. The use of journal-based reference multipliers on the basis of the *Physical Review* (a factor of 40) produces a total of 2,760 citations. If one considers that precisely this paper took on special meaning for astrophysics, the inclusion of reference multipliers of 80 is justified on the basis of *Astrophysical Journal*, from which a time-standardized value of 5,520 citations results (see Table 3). Drawn from the conditions of Big Science, this value correlates neatly to the order of magnitude for the impact of contemporary papers. The most cited papers (theme: string theory) by E. Witten (a contemporary physicist with one of the highest h index values) have received around 2,000 citations (such as Witten, 1981), and the most cited papers in *Astrophysical Journal* (theme: cosmological radiation) receive about 5,000 citations (such as Schlegel, Finkbeiner, & Davis, 1998).

Application Cases for Time adjustment: The h Index of Max Planck and Albert Einstein

In the year 2005, a bibliometric indicator was introduced for evaluating individual scientists. It is the so-called h index (also called the h factor or the h number): "A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each" (Hirsch, 2005, p. 16569). An h index of 40, for instance, means that the relevant researcher has published 40 papers, which were cited at least 40 times each. The h index combines output (the number of papers) with impact (the number of citations per paper) into a single measure. Bornmann and Daniel (2007, 2009a) have published an overview of the research on the h index.

Hirsch (2005) has calculated the h index values for some of today's prominent physicists. The citation time interval extended from the period after the publication of their papers until the point when Hirsch published his analysis. E. Witten had reached the highest value with $h = 110$; further physicists with very high values were A. J. Heeger ($h = 107$), M. L. Cohen ($h = 94$), and A. C. Gossard ($h = 94$). One of the authors of this study, M. Cardona, was given an h value of 86. If one undertakes a citation analysis for the early pioneers of the science, such as Planck, one may be disappointed by the low citation rates and the low bibliometric indicators based on them, such as the h index. Planck has an h index of 13, if one takes the temporally and disciplinarily refined list of his 88 papers, compiled on WoS (thus, in accordance with the procedure for the calculation of this indicator for contemporary scientists). Today, an industrious doctoral student could achieve an h index value at that level solely on the basis of work published before his doctoral defense. Note that the h index increases roughly linearly with the scientific age of a researcher. The manual integration of the papers by Planck that appeared before 1900 from the publication list hardly changes the h index. If, however, the reference multipliers calculated in this study are utilized, then a completely different picture emerges.

For the time adjustment of the h index of Planck, we used the extended publication list and ordered it by citation rates. We then multiplied the citation counts alternately by factors of either 15 or 40. A reference multiplier of 15 had emerged as the ratio of citation rates for physics overall, whereas the reference multiplier of 40 had emerged as the ratio of citation rates for the journal *Physical Review* (see Table 2). On the basis of the time-adjusted citation rate, h index values have been calculated which are comparable with any other author's h index value by contemporary standards. Because reference multipliers of 15 and 40 construct the lowest and highest values in Table 2, respectively, we did not consider the reference multiplier that we calculated for *Philosophical Magazine*. (We decided that the reference multiplier for *Astrophysical Journal* is hardly relevant for Planck's publication list, because this journal is not seen as representative for physics in general.)

Because of the adjustment of citation rates with the help of reference multipliers of 15 (that is, on the basis of physics as a whole), one obtains the time-adjusted h index of 53 for Planck (as opposed to 13, see above). If one applies the reference multiplier of 40, based on the *Physical Review*, an h index of 68 emerges. Next, if the papers written before 1900 are taken into consideration in the calculation of the h index, the h index reaches values between 85 and 90. The time-adjusted h index values that we calculated correspond adequately to the impact (and scientific significance) of Planck's work in physics, and they are, thus, more in line with the aforementioned h index values given by Hirsch for prominent contemporary physicists. In fact, Planck is disadvantaged in his h index, because his accomplishments are concentrated in relatively few papers, in contrast to today's publication practice, and were highly cited by the standards of the time. The 100 papers by Planck that we compiled correspond to a several times higher publication rate than that of prominent contemporary researchers. One should always consider this fact in the calculation of h index values for scientists from the period of Little Science.

The total of 165 papers that Einstein published yielded a time-unadjusted h index of 50. If one multiplies the times cited of the 165 papers compiled in WoS with a reference multiplier of 15 (thereby adjusted to the present standard for physics as a whole), then an h index of 120 emerges. If one uses a reference multiplier of 40 based on *Physical Review*, one obtains an h index of 139 for Einstein. (For Einstein's publication list, the reference multiplier for the *Astrophysical Journal* is not relevant either.) Compared with contemporary physicists, Einstein would have an extraordinarily high h index value after adjustment; this value would surely correspond to his unique significance in the field of physics.

When we compare the h index of Planck with that of Einstein, we should keep in mind the following two considerations: (a) While Planck and Einstein founded quantum physics, Einstein is the founder of relativity theory as well. Unlike Planck, who alternated from research work to scientific administration over the course of his career, Einstein was consistently involved with scientific research during his whole life. (b) Planck's life and activity were severely affected by the two World Wars and serious strokes of fate in his family (Cardona & Marx, 2008).

Discussion

In this study, reference standards and reference multipliers are suggested, which enable the comparison of the citation impact of early papers in physics (from the period of Little Science at the beginning of the 20th century) with the impact of contemporary papers in physics (from the period of Big Science beginning around 1960). For the development of time specific reference standards, we determine (a) the mean citation rates of papers in selected physics journals as well as (b) the mean citation rates of papers in physics as a whole published in 1900 (Little Science) and in 2000 (Big Science); this is accomplished by relying on the processes of field-specific standardization in bibliometry. For the construction of reference multipliers, with which the citation impact of early papers can be related to the citation impact of contemporary papers, the reference standards from 1900 and 2000 are brought into relation with one another. The effect of reference multipliers is illustrated on the basis of two of the most prominent early physicists: time-adjusted h index values are calculated for Max Planck and for Albert Einstein.

Time standardization and time adjustment are not meant to represent an *exact* calculation of the impact of early papers, such as those of Planck and Einstein, which would correspond to contemporary standards. Rather, our efforts aim at constructing a method to gauge the effect of a paper from the early 20th century as compared with a paper from the present. Reference multipliers are not used to achieve precise figures, but rather to establish an order of magnitude adjusted for contemporary standards, in other words, to find an approximate confidence level for impact adjusted to the present. Exact calculations are impossible because of the limitations inherent

DOI: 10.1002/asi

to our approach. We would like to articulate some of these limitations below.

A first limitation to the approach of time adjustment results from the fact that a suitable reference multiplier could not be constructed for every individual paper, but rather it was only plausible to construct one reference multiplier for papers in physics generally and one for each of three particular journals. Ideally, the (sub)field or journal, in which the paper to be adjusted was published, should be identified for every *single* paper in a publication list and a corresponding reference multiplier should be constructed. For example, Planck had published his papers over a period of nearly 50 years within several physics subfields and journals. Because, however, only a few journals in which the early papers from Little Science were published still exist today, this process was impossible in physics. The extent to which more such journals exist in other fields should be tested in future studies.

As a second limitation, we would like to point out that we limited ourselves in the construction of reference standards and reference multipliers to only 2 years (1900 and 2000), as we undertake time adjustment and time standardization. Both of these years represent, respectively, the eras of Little Science and Big Science, as considered in this study. However, these eras extend over a long time period, a fact which probably makes the use of reference multipliers and reference standards for more than only 2 years necessary. Within shorter time periods like one decade, average citation rates for field and journal generally vary little with time. For instance, during the period from 1900 to 1910, the number of citing papers per paper rose only from 0.42 to 0.59 for the papers that appeared in *Physical Review*. The period from 1900 to 1910 is the stretch of time in which the most important papers by Planck and Einstein appeared. In theory, a suitable reference multiplier should be determined for every paper in a publication list on the basis of papers from the *same* publication year, and that exact reference multiplier should be considered in the calculation of a time-adjusted value.

A third limitation concerns the difference between citing papers and citations. A citing paper, in principle, can contain more than one citation. If, for instance, a citing paper cites two *Physical Review* papers from 2000 in the data set for this study, then we measure one citing paper, although there are two citations. Considering the proportionally low mean number of references in the papers from 1900, the difference between the number of citing papers and the number of citations is small. In the case of the year 2000 for *Philosophical Magazine,* the 4,228 citing papers comprise 4,522 citations (difference: 305 citations). In view of the course of the year 2000 for physics literature as a whole, the difference is even greater, but it cannot be ascertained in SCI under STN International because of system limitations. The reference standard of 7.27 (see Table 1) and the reference multiplier of 15.15 (see Table 2) should, therefore, be taken as lower limits.

We examined the use of reference multipliers for time adjustment in the cases of two physicists: We calculated time-adjusted h index values for Planck and Einstein. As the basis for the calculation of h index values, we used citations of their papers from the dates of their publication up through today. Because the reference multipliers were calculated on the basis of reference standards that related to a fixed citation window of 10 years, we should have taken into consideration only the citations from the first decades into the h index calculation of papers by Planck and Einstein. We decided to forego such a procedure in this study when we calculated the h index for Planck and Einstein, because h index values are not normally calculated using the citation frequency over a fixed time window and were not even calculated for present day physicists by Hirsch (2005).

To calculate the time-standardized impact values, we suggested dividing the paper's citation counts by the reference standard, a technique that relies on field standardization. Bornmann and Daniel (2009b) tested the advantage of using z-scores for field standardization. z-scores have been used in psychological testing for many years for standardization of test scores. Instead of simply dividing by the reference standard, it is recommended to divide the difference by the test score x and the population mean μ by the population standard deviation σ. Because the results of Bornmann and Daniel (2009b) indicate that z-scores are very well-suited to producing field-standardized impact values, future studies on time standardization should also examine the advantages and disadvantages of z-scores. In such studies, one could also investigate the merits of developing a reference standard on the basis of the median instead of the arithmetic mean. Because of the basically skewed distribution of citations among publications (Seglen, 1992), the median is better suited as the measure of central trends. We avoided calculations based on the median in this study, because of the very high financial costs involved in the required database research.

Future studies concerned with time standardization and time adjustment (e.g., in other fields than physics), may focus on the comparison of the impact of papers from the eras of Little Science and Big Science, as this article does. However, the need for time standardization and time adjustment not only is required when comparing papers from the first half of the 20th century with contemporary papers, but it is even necessary whenever one compares the citation impact of papers by researchers whose publication volume peaked in the sixties, for example, with the citation impact of researchers publishing today (see also the aforementioned second limitation). Eugene Garfield (1998) supplies specific numbers concerning this topic: "Due to continuous growth of source journal coverage and increasing references cited per paper, the ratio of citations to published papers increased about 75% from 1945 to 1995—from 1.33 to 2.25 over the past 50 years. It is the inflation of the literature which increases the ratio each year" (p. 72). Future studies ought to examine which reference standards and reference multipliers should be used in physics and other fields in the natural sciences for the comparison of citation impact in various time periods.

The procedure for time standardization and time adjustment that we recommend in this study ought to be of interest especially for studies in the history of science. The object of this discipline is often the examination of early pioneers in the sciences and their lifework (usually highly recognized scientists, such as early Nobel laureates). The reaction to individual publications is often discussed in this context (see here Gingras & Wallace, 2010). Bibliometrics offers quantitative data on them in the form of time-dependent citation counts (citation history). If the citation history extends over the boundary-year of 1960, which divides the eras of Little Science and Big Science, then the schemas offered in this study should facilitate a better interpretation of the impact data.

References

Abt, H.A. (1987). Reference frequencies in astronomy and related sciences. Publications of the Astronomical Society of the Pacific, 99(622), 1329–1332.

Abt, H.A. (1995). Some statistical highlights of the Astrophysical Journal. Astrophysical Journal, 455 (part 1), 407–411.

Bornmann, L., & Daniel, H.-D. (2007). What do we know about the h index? Journal of the American Society for Information Science and Technology, 58(9), 1381–1385.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. Journal of Documentation, 64(1), 45–80.

Bornmann, L., & Daniel, H.-D. (2009a). The state of h index research. Is the h index the ideal way to measure research performance? EMBO Reports, 10(1), 2–6.

Bornmann, L., & Daniel, H.-D. (2009b). Universality of citation distributions. A validation of Radicchi et al.'s relative indicator $c_f = c/c_0$ at the micro level using data from chemistry. Journal of the American Society for Information Science and Technology, 60(8), 1664–1670.

Cardona, M., & Marx, W. (2008). Max Planck – A conservative revolutionary. Il Nuovo Saggiatore, 24(5–6), 39–54.

De Solla Price, D.J. (1965). Little science, big science. New York: Columbia University Press.

De Visscher, A. (2010). An index to measure a scientist's specific impact. Journal of the American Society for Information Science and Technology, 61(2), 319–328.

Dickson, P. (2007). Sputnik: The shock of the century. New York: Walker & Co.

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper [On the electrodynamics of moving bodies]. Annalen der Physik, 17(5), 891–921.

Einstein, A. (1915). Über die allgemeine Relativitätstheorie (plus addendum). Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie. Die Feldgleichungen der Gravitation [On the general theory of relativity (addendum). Explanation of the perihelion motion of Mercury from the general theory of relativity. The field equations of gravitation]. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften Part 2, 778–847.

Einstein, A. (1916). Die Grundlage der Allgemeinen Relativitätstheorie. [The foundation of the general theory of relativity]. Annalen der Physik, 49(7), 769–822.

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? Scientometrics, 1(4), 359–375.

Garfield, E. (1998). Random thoughts on citationology. Its theory and practice. Scientometrics, 43(1), 69–76.

Gingras, Y., & Wallace, M. (2010). Why it has become more difficult to predict Nobel Prize winners: A bibliometric analysis of nominees and winners of the chemistry and physics prizes (1901–2007). Scientometrics, 82(2), 401–412.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the USA, 102(46), 16569–16572.

Legros, M., Elliott, B.R., Rittner, M.N., Weertman, J.R., & Hemker, K.J. (2000). Microsample tensile testing of nanocrystalline metals. Philosophical Magazine A, 80(4), 1017–1026.

Marx, W., & Schier, H. (2005). CAS kontra Google [CAS versus Google]. Nachrichten aus der Chemie, 53(12), 1228–1232.

Opthof, T., & Leydesdorff, L. (in press). Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance. Journal of Informetrics.

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. Proceedings of the National Academy of Sciences, 105(45), 17268–17272.

Redner, S. (2005). Citation statistics from 110 years of Physical Review. Physics Today, 58(6), 49–54.

Rutherford, E. (1900). A radio-active substance emitted from thorium compounds. Philosophical Magazine, 49(296–01), 1–14.

Rutherford, E. (1911). The scattering of α and β particles by matter and the structure of the atom. Philosophical Magazine, 6(21), 669–688.

Schlegel, D.J., Finkbeiner, D.P., & Davis, M. (1998). Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. Astrophysical Journal, 500(2), 525–553.

Seglen, P.O. (1992). The skewness of science. Journal of the American Society for Information Science, 43(9), 628–638.

van Raan, A.F.J. (2004). Sleeping beauties in science. Scientometrics, 59(3), 467–472.

van Raan, A.F.J. (2004). Measuring science. Capita selecta of current main issues. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems (pp. 19–50). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Witten, E. (1981). Dynamical breaking of supersymmetry. Nuclear Physics B, 188(3), 513–554.

DOI: 10.1002/asi