

Specific character of citations in historiography (using the example of Polish history)

Władysław Marek Kolasa

Received: 5 July 2011 / Published online: 13 November 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract The first part of the paper deals with the assessment of international databases in relation to the number of historical publications (representation and relevance in comparison with the model database). The second part is focused on providing answer to the question whether historiography is governed by similar bibliometric rules as exact sciences or whether it has its own specific character. Empirical database for this part of the research constituted the database prepared ad hoc: The Citation Index of the History of Polish Media (CIHPM). Among numerous typically historical features the main focus was put on: linguistic localism, specific character of publishing forms, differences in citing of various sources (contributions and syntheses) and specific character of the authorship (the Lorenz Curve and the Lotka's Law). Slightly more attention was devoted to the half-life indicator and its role in a diachronic study of a scientific field; also, a new indicator (HL14), depicting distribution of citations younger than half-life was introduced. Additionally, the comparison and correlation of selected parameters for the body of historical science (citations, HL14, the Hirsch Index, number of publications, volume and other) were also conducted.

Keywords Citation analysis · Historiography · Historical science · Poland

Introduction

Bibliometric research has its own long scientific tradition which has been shaped by substantial achievements. Over last decades numerous theoretical works have been written and professional tools have been successfully improved and implemented. For many scholars the empirical evaluation of the scientific output is no longer controversial and has already become an integral part of a standard research procedure. Regrettably, these remarks are not universal because they are mainly connected with formal and natural

W. M. Kolasa (✉)
Institute of Information and Library Science, Pedagogical University of Cracow,
ul. Podchorążych 2, 30-084 Kraków, Poland
e-mail: wmkolasa@up.krakow.pl

science. Even the most general analysis of bibliometric attainments leads to the conclusion that only the body of studies classified as *science* constitutes the basis for the lion's share of existing theoretical works which are the point of reference for the study itself. Taking into account the number of records registered in ISI databases and achieved levels of indicators, there is a bottomless gap between the body of *science*, *social science* and *the humanities* (Marshakova-Shaikevich 2001, 2009).¹ Three factors determine the worst situation of the last group: (1) the lack of representative citation indexes causes that the results of their application in the humanities are inadequate; (2) there is a serious underdevelopment of theoretical works in that field (Schubert 1999, 2001), which makes it difficult to define the real reasons of this situation; (3) finally, all of the obstacles cause open aversion of humanists towards bibliometrics and stimulate the popularization of false beliefs regarding this field. Eventually, this leads to a complicated situation because neither theory nor practice generates impulses for a more thorough analysis of the problem.

The only solution seems to be the commencement of basic research. Particularly, three vital questions require answers: (1) are international services really not representative for the humanities (to what extent)?; (2) are humanistic works governed by similar bibliometric rules as exact sciences or do they have their own specific character? (3) can and should the humanities be studied in the international context or are they local by nature? Providing answers to these questions is going to be the main aim of this paper. Due to understandable reasons only the most characteristic part of the humanities was selected to be the basis of this research, namely – historical science, whose fragment (history of the media) has been thoroughly examined.

To serve the purpose of this work the citation index (Citation Index of the History of Polish Media [CIHPM]) was created by the author and constituted a point of reference during the assessment of international services. It also provided essential data for the description of the specific character of historical science. ICHMP database has been created in Institute of Information and Library Science (Pedagogical University of Cracow). The development has been managed by the author of this paper. It contains 24,627 publications interlinked by 63,811 citations. Majority of the publications in the database regard media history: 15,920 documents cited 52,254 times (without counting self-citations); other documents regard media science. ICHMP index is retrospective in nature and includes publications created in years 1812–2009 and most of the publications have been created after 1945 (22,975 publications, or 96% of database content). The index includes scientific papers and documentation-type works only (the documentation-type works, have been selected for arguably long-lasting value, and regardless of publication language and printing form, like books, articles from journals or books, brochures, etc.). ICHMP contains records from different sources, but most of those were based on reviewed abstract bibliographies.² The index was built author's methodology

¹ According to SCI (*ISI National Science Indicators Deluxe 1981–2002*) the number of works divided into main bodies is as follows: natural science [P]—90%; social science [S]—8%; the humanities [H]—2.5% (Marshakova-Shaikevich 2009: 194). Also, selected indicators differ substantially, e.g.: the number of cited publications (P—38.32–78.48%; S—32.7–59.1%; H—7.2–20.5%); citation number per publication (P—1.36–12.98; S—0.92–3.83; H—0.10–1.80) and the number of citations in a cited publication (P—5.02–18.59; S—2.80–6.48; H—1.35–1.80) – (Marshakova-Shaikevich 2001, p.157; 2009: 20).

² E.g. *Polska Bibliografia Adnotowana Wiedzy o Środkach Masowego Komunikowania = Bibliographie Commentée Polonaise de la Science sur les Moyens de l' Information...* (1965–1987) published by OBP (Press Research Centre of Jagiellonian University), a member of International Association for Media and Communication Research (IAMCR/AIERI) and other, more focused works, e.g. *Polska bibliografia prasoznawcza = The Polish Press Research Bibliography*, ed. Commission of Press Research. Polish Academy of Sciences (Kolasa and Jarowiecki 2005).

(Kolasa 2011b)³ oriented towards achieving highly focused (clustered) results. The results achieved are comparable to ISI indexes, e.g. *Citation Impact*' (number of citations per document cited) for history was 6.66, the publication with the highest number of citations – 415, while the most cited author has achieved 1,075 citations. The database is processed in the aforementioned institute, and in August of 2011 the author has made available a strongly reduced version of the database (ICHMP lite) online, adapted for browse purpose.⁴

In order to answer the aforementioned questions the subject literature may be useful, but only to some extent. Apart from numerous works (McCrank 2001, pp.85–139), traces of deeper interest in that subject are only symbolic (Archambault and Gagné 2004; Finckenstaedt 1990; Nederhof et al. 1989; Boonstra et al. 2004). Due to the lack of adequate data some authors either use publication statistics (e.g. Hérubel 2006, 2007) or indicators different than citations. However, rarely published works based on citations present the humanities as a case study (e.g. Koniczna 2002) or treat it contextually (e.g. Marshakova-Shaikovich 2009). An inspiring work turned out to be the study by Piotr Nowak (2004) in which the author defines three areas of the specific character of the humanities: (1) differences of the paradigm of exact sciences in opposition to the humanities; (2) the influence of the publication language on the humanities and (3) different role of basic publishing forms in the humanities and natural science. Several of the theses presented in this paper have been successfully verified empirically and supplemented with new aspects.

History in international databases

The main aspect which has to be brought up at the very beginning is the empirical assessment of the content of the largest world sources. Thus, the key factor would be the answer to the question: *Do international indexes contain suitable representation of data in the field of historical science, taking into account both their quantity and quality?* It is not a simple task because of the lack of adequate comparative data concerning the number of publications and the potential of the personnel (e.g. the number of historians employed in the education sector). However, UNESCO data relating to the production of books might shed some light on this topic. Such attitude seems to be justified because in practice, in historical science (as opposed to exact sciences) the major and the most valued form of presentation are books (Dejnarowicz 1980; Nederhof Leeuwen and Raan 2010). In the light of the aforementioned data it can be estimated that among 237,332 books published in 1995 in the world about 49,891 are connected with history.⁵ Poland, with its annual number of publications being 820 works, takes the 17th place comprising 1.64% of the world output of texts of this kind. The peculiar character of this rating is worth underlining because main leaders in the field of historiography are: Great Britain (17.37% titles) and Germany (14.27%), not the USA (8.55%) which usually top the ISI lists. American historiography is closely followed by renowned European centres of historical thought (France, Italy and Spain—each 6–7%); then other highly-developed or still developing

³ Preprint - <http://hdl.handle.net/10760/16052> [2011.09.01].

⁴ CIHPM, ver. 7.72 light - <http://bazy.wbp.krakow.pl/cgi-bin/makwww2/makwww.exe?> [2011.08.29].

⁵ In the source database there was only one common History/Geography section [UKD 9]. History was interpolated as 75.4% of the source section (assessed on the basis of: *Ruch Wydawniczy w Liczbach* 1996). UNESCO data are not complete – only 72 countries were included in Table 1; missing countries are: Australia, China, Belgium, Ireland, the Netherlands, Pakistan and the majority of African countries.

Table 1 History books in the world (1995)

Country	History [N]	History (%)	Total books	Country	History (N)	History (%)	Total books
1. Great Britain	8,666	17.37	10,1764	12. Canada	913	1.83	17,931
2. Germany	7,119	14.27	74,174	13. Iran	897	1.80	13,031
3. United States	4,265	8.55	62,039	14. India	881	1.77	11,643
4. Italy	3,323	6.66	34,470	15. Finland	855	1.71	13,494
5. France	3,290	6.59	34,766	16. Norway	844	1.69	72,65
6. Spain	3,266	6.55	48,467	17. Poland	820	1.64	11,925
7. Japan	1,917	3.84	56,221	18. Sweden	790	1.58	12,700
8. Russia	1,579	3.16	33,623	19. Switzerland	616	1.23	15,771
9. Korea	1,139	2.28	35,864	20. Hungary	547	1.10	9,314
10. Brazil	1,004	2.01	25,540	Other countries	6233	12.49	13,5568
11. Denmark	929	1.86	12,478	Total	49891	100.00	23,7332

Author's own calculation (see footnote 5 for methodology) based on the basis of: *Book production: Number of titles by UPC classes* [1995] – UNESCO Institute for Statistics [Data Centre] – report from the database: <http://stats.uis.unesco.org/> [20.10.2010]

countries (each 1–2%). Although imperfect, the following rating constitutes an ideal starting point for further research Table 1

An experiment was conducted in order to answer the question to what extent Polish historiography is represented in world databases. Search results of selected databases were compared to the model reference source – the *Citation Index of the History of Polish Media* [CIHPM] and the rank of each hit was verified (Table 2). The experiment confirmed initial hypotheses because the results, in each case, were lower than 1.64% calculated above (0.38% for *Web of Science* [WoS], 1.60% for *Historical Abstract* [HA] and 1.11% for *International Bibliography of Historical Sciences* [IBHS]).

Significant differences between sources have been revealed by the study of relevance (Table 2). It was analyzed with the use of the formula of SPERES database for four sections defining the quality of a publication measured by the number of citations: A—*famous papers* (150–415 citations); B—*well-known papers* (52–149 citations); C—*known papers* (11–51 citations) and D—*other papers* [less known and unknown papers] (less than 10 or without citations). The results of the experiment revealed that IBHS is the only source including acceptable number of highly-cited records. For the A section (*famous papers*) the level of compliance was as high as 60%; for B—20%, C—5.2% and D—0.6%. For other sources the relevance was significantly lower, particularly, they did not register a single work from the sections of *famous* and *well-known* papers [sic!]. For HA the highest relevance concerned only records from C (1.5%) and D (1.3%) sections; the worst results were those of WoS – the relevance of only 0.13% for the D section (the last one). It is worth mentioning that the second international index *Scopus* was excluded from the experiment because it deals with historiography only marginally.⁶

⁶ Even the top of the SCOPUS index list – American works constitute only 11,441 documents, British – 7,250 and French – 1,990. Poland occupies the 35th place with only 95 works cited 256 times. Data based on the SCIMago database – <http://www.scimagojr.com/index.php> [2011.03.10].

Table 2 Representation of the history of Polish media in selected international sources

Citation index of the history of Polish media			Web of science		Historical Abstract		International bibliography of historical sciences	
Rank of publication	Rank	N	N	Relevance	N	Relevance	N	Relevance (%)
A-famous	1–10	10	0	0	0	0	6	60
B-well-known	11–100	90	0	0	0	0	18	20
C-known	101–1000	900	0	0	14	1.5%	47	5.2
D-other	1001–15,920	14,920	20	0.13%	207	1.3%	97	0.6
Incorrect	–	–	41	–	40	–	10	–
Total	1–15,920	15,920	61	0.38%	261	1.6%	178	1.11

The research was performed in years 2009–2010 on exhaustive number of publications registered in CIHPM (15,920 items)

Bold values show the dominant

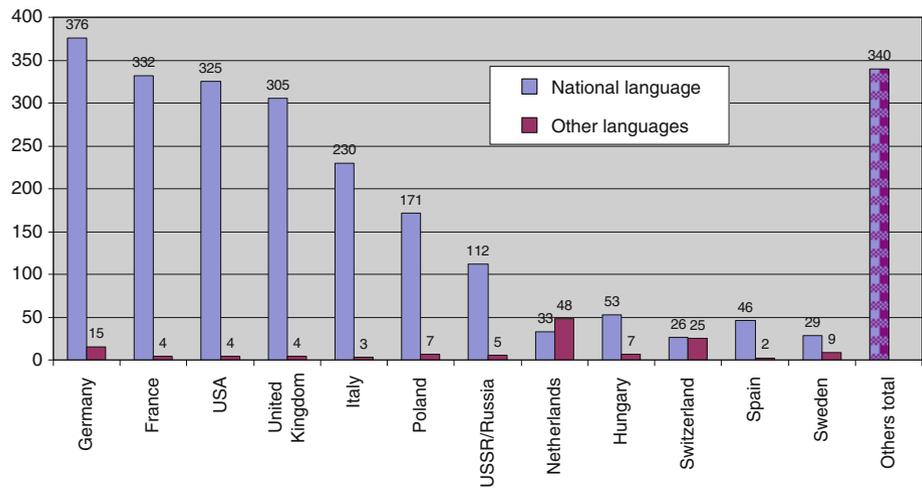


Fig. 1 Publication Language in the Field of History in the Global Scientific Circulation [n]. N = 2511 publications (1,531 books and 980 articles). Source: *International Bibliography of Historical Science* Vol. 1–72 (Paris-Berlin-London-New York 1926–2003) [Part 4. Press]; Netherlands - 33 publications in English; Switzerland - 21 publications in French

The only source which turned out to be useful in further research was the *International Bibliography of Historical Sciences* [IBHS] – the oldest international bibliography in the field of history, issued under the aegis of UNESCO, unremittingly since 1926. Up to the year 2004, 73 volumes have been published, comprising approximately 500,000 documents. Closer analysis shows that during the given period certain national historiographies were included in it in a way similar to a representative one, particularly, Polish historiography (on average 2.27% of total publication number). Therefore this source can be treated as a reliable tool for further study.

Two aspects seem to be especially important: the language of the publication and dominant publishing forms. As far as language is concerned, the analysis of a representative sample comprising 2,511 publications (Fig. 1) depicts that in historiography national language is a prevailing one.

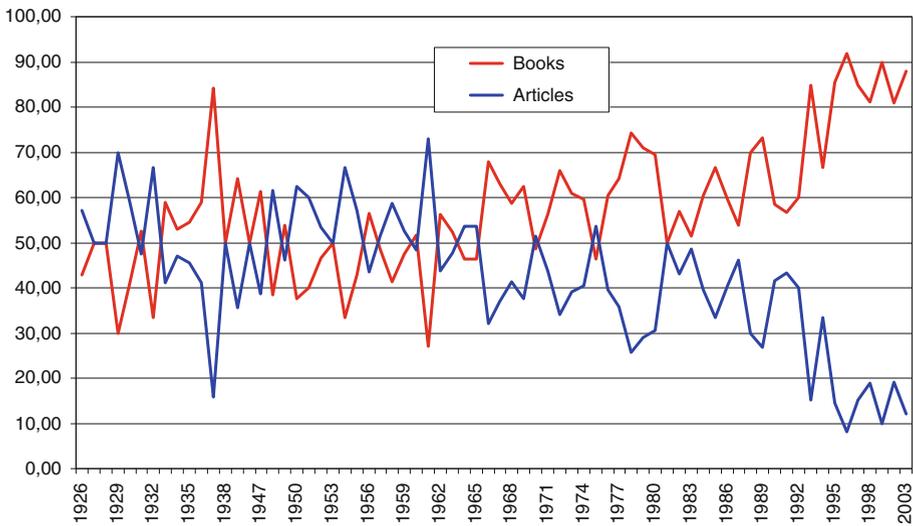


Fig. 2 Main publishing forms of historical works in the global scientific circulation [%]. $N = 2171$. Source: *International Bibliography of Historical Science* Vol. 1–72 (Paris-Berlin-London-New York 1926–2003) [Part 4. Press]

Participation rate of a national language usually fluctuates within the range of 95.6–98.8% (France—98.8%; USA—98.8%; Great Britain—98.7%; Italy—98.7%; Germany—96.1%; Poland—96.0%; Spain—95.8%; Russia—95.6%); rarely is the rate lower (Hungary—88.3%; Sweden—76.3%). Very similar dominance rate can even be found in countries using several languages (Switzerland—2%: 49.1—German; 41.1—French). The only exception are countries where seats of editorial offices of international journals are located: for the examined field there is only one such a magazine – “Gazette”, which has been published since 1955 in the Dutch city of Leiden. A great number of works in “Gazette” have been published in English (40.7%) but still the participation of national languages has been significant (Dutch—40.6%; French—12.2%; German—6.1%). Presented feature is the first discriminant of historical science and it confirms its localism.

Another source of interesting data is the examination of publishing forms (Fig. 2). The analysis of a representative sample comprising 2,171 history works from IBHS shows that the participation of books in historical science has always been high but now it is dominant. This indicator in the period of 1926–1965 equalled 50–60%, since then (1966–1980) it has been rising (60–70%) only to stabilize at the level of 80–90% at the moment. This trend can be evaluated in a number of ways, but from author’s viewpoint it seems like a peculiar reaction to a global growth in publication numbers, i.e. about 25% a year (Cybulski 1990, p. 23), while the annual limits of data registered in IBHS remained constant (6–10,000). In order not to exceed predefined limits, research and publication boards have applied systematically higher selection criteria. This process also shows that according to national editorial offices, books contain the basic stream of the most valuable knowledge in the field of historiography. As later citation analyses (Fig. 3) revealed that editorial offices’ opinion was fully justified.

On the other hand, it seems that an important feature influencing the high position of books is their volume, which is inseparably connected with their information capacity. In IBHS the total volume of analyzed books was 13 times greater than of articles (books –

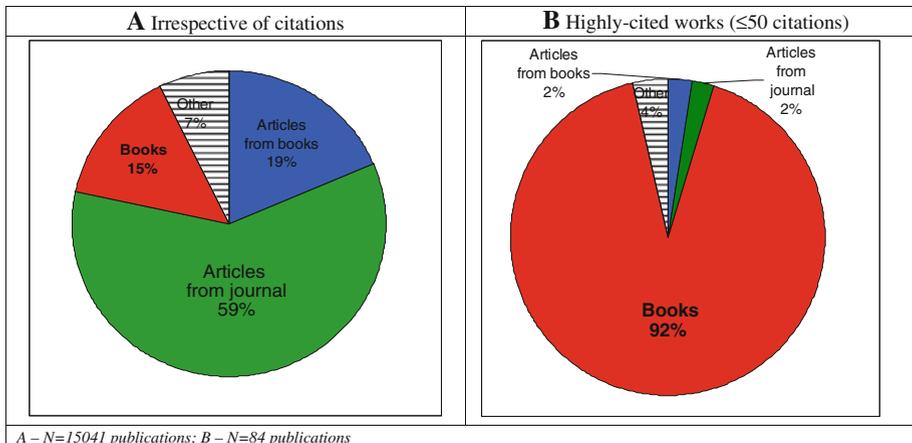


Fig. 3 Dominant publishing forms in the field of historiography in the CIHPM [%%]

509,793 pages which is 92.7%; articles – 38,533 pages which is 7%). Presented feature is the second discriminant of historical science.

The evaluation of the presence of the humanities in ISI databases has been the bone of contention for ages (Nowak 2000, pp. 146–149). As the practice shows, the humanities, as opposed to other fields, are underdeveloped; registered resources are definitely smaller (approx. 2.5-times smaller annual intake than for exact sciences) and have lower citation indicators when compared to *science & technology* (Marshakova–Shaikevich 2001, pp. 155–160).⁷ But the question which still needs to be answered is: *What is the exact number defining geographic non-representativeness and what does it depend on?* Archambault and Gagné (2004, pp. 19–21) estimate that for Polish publications in the fields of social science and the humanities, these indicators are: 87% of omitted titles and 99% of omitted works and probably they are right. In order to verify that thesis an experiment with the use of several testing phrases in the field of history was carried out (Table 3). Judging by the distribution of the data a conclusion can easily be drawn that geographical rating of hits is directly proportional to the number of indexed journals. Currently, the ISI comprises 260 titles in the field of history among which 89 titles (34% of total number) are for the USA; 57 (22%) for Great Britain; 17 (7%) for Germany; 14 (5%) for France and 83 (32%)⁸ for other; given data (except Canada) are surprisingly similar to previously obtained results. It means that irrespective of the search query only publications from the aforementioned countries are relatively well-represented and have a chance to exist on the market. Still, one interesting relationship was revealed: national themes are strongly correlated with countries of their publication, which is another confirmation of the fact that history is by nature local. For countries highly represented these indicators are two or three times higher than the average (Germany—214%; France—288%); and for other countries – even five times higher (Poland—529%).

As far as ISI results for history are concerned, there is another serious imperfection. It turns out that for history the absolute majority of publications are book reviews (on average

⁷ Indicators of ISI databases are listed in annotation 1.

⁸ *ISI Master Journal List* – <http://science.thomsonreuters.com/mjl/> [2010.09.24]. Including two Polish titles: “Acta Poloniae Historica” (since the 80s) and “Kwartalnik Historii Żydów” since 2009 (among 43 of Polish texts indexed in the system of *Web of Knowledge*).

Table 3 Selected history themes in the web of science [%%]

		Testing phrases					
		Second world war	Germany	France	Poland or Polish	Jew or Jewish	The average
Place of publication	USA	19.7	17.2	24.6	15.9	40	24.13
	England	16.9	6.5	10.5	3.2	4.4	8.35
	Canada	7.8	1.5	5.0	2.3	3.2	3.88
	Germany	6.5	11.6	1.2	2.4	3.2	5.42
	France	5.1	1.1	9.9	0.8	1.7	3.43
	Poland	0.2	0.1	0.08	8.0	0.6	1.51
	Israel	0.4	0.3	0.2	0.5	6.4	1.38
	Other	43.4	61.7	48.52	66.9	40.5	51.90
Form	Book review	46.2	80.6	65.8	73.6	62.1	65.66
	Article	43.8	15.2	27.5	21.0	31.2	27.74
	Other	10.0	4.2	6.7	5.4	6.7	6.6
Language	English	67.5	35.5	49.7	59.1	74.2	57.2
	German	14.8	52.3	6.0	18.0	11.5	20.5
	French	11.0	6.6	40.0	14.5	7.5	15.9
	Other	6.7	5.6	4.3	8.4	6.8	6.4

Source Web of Knowledge – databases: *SCI-Expanded*, *SSCI*, *A&HCI*; index: *Topic*, range: *All Years*, Subject areas: *History* – measurement 24.10.2010. In order to verify some hypotheses, some particular geographically-oriented phrases were searched for (Germany, France, Poland or Polish, Jew or Jewish) and one neutral phrase were used (Second World War) and the relation of the phrase to the publication location, language and printing form was subjected to research

Bold values show the dominant

65.6% of the works); it should be underlined that they are important publications in terms of scientific criticism but worthless as far as their informative function is concerned. This situation is clearly depicted by data concerning Poland: among 5,287 publication relating to this topic only 21% of texts are articles, the majority of the remaining works (73.6%) are reviews abundantly published in “Acta Polonicae Historica” (34%), “Jahrbücher für Geschichte Osteuropas” (8%) or “American Historical Review” (6%). So it seems that Archambault and Gagné’s (2004, pp. 19–21) assessments are close to truth. Non-representativeness is also deepened by a very strong linguistic polarization. Quite astonishing is the fact that in case of history the dominance of English in WoS is not absolute (as opposed to other fields) and in case of the national theme it is usually balanced (German—52.3%, French—40%). As a result, even WoS confirms the already formulated thesis about linguistic polarization of national historiographies, even though, because of the overrepresentation of works in English, it underestimates this number at least by half. These facts, combined with the assertion that ISI does not register books (their role has been described above) make that Philadelphian services for historical research are completely useless. Such a conclusion brings into question both the relevance of the use of WoS for the study of the humanities as well as WoS-based ratings.

The majority of aforementioned critical remarks also concern the *Historical Abstract* [HA]. However, closer examination reveals some interesting details (Table 4): firstly—HA, when compared with ISI, is five times more complete (on average 468%); secondly—similarly to ISI, there is a strong correlation between the theme and the national language

Table 4 Selected history themes in the *Historical Abstract*

		Testing phrases						Total	
		Germany		France		Polish or Poland		N	%%
		N	%%	N	%%	N	%%		
Form	Books	6081	9.13	5941	8.70	989	3.94	13011	8.13
	Articles	51808	77.8	54219	79.41	21488	85.62	127515	79.71
	Other	8704	13.07	8116	11.89	2619	10.44	19439	12.15
Language	English	30389	42.49	34647	46.29	7430	28.23	72466	41.96
	German	27247	38.10	2598	3.47	2056	7.81	31901	18.47
	French	3780	5.29	30033	40.13	1124	4.27	34937	20.23
	Polish	2190	3.06	585	0.78	12867	48.89	15642	9.06
	Other	7917	11.07	6985	9.33	2841	10.79	17743	10.27
HA/WoS [%%]		468.34		462.54		474.67		468.51	

Historical Abstracts with Full Text – measurement (su: Germany; su: France; su: Polish or Poland) – access EbscoHost – <http://www.ebscohost.com> [24.10.2010 and 10.03.2011]

Bold values show the dominant

(German—38.1%; French—40.1%; Polish—48.89%) which is yet another confirmation of the localism of historiography. The most important difference is methodical because the database includes all historically important publishing forms, particularly books and the overall picture is not blurred by book reviews as it happened when dealing with WoS. But the question arises: *Why did the measured relevance for the history of Poland turn out to be so low?* (Table 2). It is the analysis of the list of included titles that provides the answer: among over 2,300 indexed journals there were as many as 41 Polish titles⁹ (it is about 16–25% of the real body), all of them renowned; however, they were not systematically or exhaustively indexed. The magnitude of the problem is well illuminated by data for leading Polish publication named “Kwartalnik Historyczny”, which is included in a large scale (2,150 publications), but those are mainly the works published before 1989 (1,780 publications); while the number of latest articles (published since 2000) is only 132 publications; there’s a similar situation regarding “Przegląd Historyczny” (1,155); “Przegląd Zachodni” (831) and many others. These facts lead to the conclusion that the HA database is useless for Polish historiography.

Among analyzed sources the WoS database was the only one with a complete mechanism of citations; the rest either began adding citations quite recently (HA) or did not have them at all and it should not be expected that they will ever be created because they are on a printed media (IBHS). As a consequence, the way to further research in this area is closed. The only solution is the creation of local citation indexes.

History in the local citation index

Analyses which have been conducted so far, empirically confirmed the local character of historical science and its low representation in international sources. But one question

⁹ List of indexed titles in HA (January/February 2011) – <http://www.ebscohost.com/titleLists/hah-coverage.pdf> [2011.03.10].

Table 5 Comparison of selected parameters of the citation index of the history of Polish media and ISI databases

	CIHPM		SCI & A&HCI					
	History		Total		Materials science		History	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Number of works	15041		5655186		191128		No data	
Number of citations	46152		40516820		575725		No data	
Number of cited works	6924	46.0	3768822	66.6	97807	51.1	No data	19.6
Number of uncited works	8117	53.9	1886364	33.3	93321	48.8	No data	80.4
Citation Impact ^a		3.06		7.16	3.01		0.37	
Citation Impact ^b		6.66		10.75	5.89		1.91	

Data: CIHPM – numbers concern only history works from the years 1945–2009; SCI 1981–1992 – source: (Kozłowski 1994, pp. 14–15); history in A&HCI 1981–2002 (Marshakova-Shaikevich 2001, pp. 157)

^a *Citation Impact* number of citations per publication

^b *Citation Impact'* number of citations per cited publications

Bold values show the dominant

remains unanswered: *Are humanistic works governed by similar bibliometric rules as exact sciences or do they have their own specific character?* Because of the fact that WoS could not provide the answer, further research was conducted on the basis of the citation index (the *Citation Index of the History of Polish Media* [CIHPM]). The final version of the database included 24,627 documents linked with each other by a net of 63,811 citations and descriptions concerning the history of the media (15,920 documents cited 52,254 times), which constituted the core. The index was compiled according to the rules of the author's method similar to the SCI methodology (Garfield 1979), set in such a way that would enable to obtain the highest possible concentration of results allowing to draw reliable statistical conclusions. Eventually, the results were similar to Philadelphia indexes (Table 5).

Indicators presented above constitute a convincing proof of the fact that history is governed by the same bibliometric rules as natural science (Table 5). Similarities and differences are best presented by the indicator called *Citation Impact'* (number of citations per cited publications). Although this parameter for history in the CIHPM is 6.66 and it is slightly lower than the SCI average [10.75] (particularly in leading disciplines i.e. molecular biology [18.61]), at the same time it remains close to such disciplines as agricultural science [5.71] or materials science [5.89]. On the other hand, the *Citation Impact'* for history, taken from the CIHPM is almost twice higher than for social science in SSCI [2.80–4.36] and four times higher than for the humanities in A&HCI [1.35–1.80] (Marshakova-Shaikevich 2001, p. 157). Glaring disproportions have been revealed by the comparison of indicators for history between the CIHPM and A&HCI, which – as the table shows – have been lowered at least four times (*Citation Impact'* 6.66–1.99), and for some other indicators even twelve times (e.g. number of citations of one work – 3.06 to 0.37).

Among countless peculiarities of history works there are these which are worth defining – dominant publishing forms and their publication language. The IBHS analysis has proven the dominant role of books in historical science (60–90% of publications total) but the indicator was deformed by too elevated selection threshold. Slightly different conclusions can be drawn from the CIHPM analysis: books constitute 14.5% of the database, articles

from journal – 59.7% and articles from books – 18.5%. But if we focus on highly-cited works the compliance with IBHS is astonishing: among *famous* and *well-known* works (50 and more citations) as many as 95.2% are books (Fig. 3b). This indicator dwindles proportionately to the number of citations: 78.2% for 30 citations, 53.4% for 10 citations.¹⁰ Undoubtedly, the feature which influences the high position of books is their volume which is inseparably connected with their information capacity (in disciplines which do not use abstractions it seems to be undisputed). In the analysis of the volume of the main body (801,410 pages of text) books constituted 75.6% of the total volume, articles from journal – 17.5% and articles from books – only 5.4%.

Strong linguistic polarization for historical science has already been depicted during the analysis of all international sources, especially IBHS (Fig. 1). The CIHPM confirms those data: Polish historiography is dominated by works in Polish—96.2% of works; tiny representation have: English—1.3% and German—1.0% and other languages are barely represented – each less than 0.3%. Justification of this situation should mainly be connected with the “localism” of the object of our study (Nowak 2008, p. 28). Historical science requires wide contextual knowledge (culture, politics, language and other), impossible, as it may seem, to obtain without active participation in a given culture and difficult to express in a language other than the national one. The only exception to this rule are works devoted to foreign historiography and history of international relations (together 9.26%) as well as popularization texts and educational studies. This thesis has been confirmed by the CIHPM, in which, cases of the use of a language other than a national one concerned exclusively foreign history, intercultural comparative literature and issues relating to geographical borderlines or minorities – 19% of these works were published in languages other than Polish (in this: 4.67%—English, 7.61%—German, 2.31%—Russian, 4.41%—other). Not so often, works in congress languages were connected with their popularization on the international arena—2.13% of works (usually publications in special magazines – e.g. “Acta Poloniae Historica”). Still, it should not be overlooked that the main stream of history is national historiography (90.74% of publications and 97.10% of citations), in which the use of Polish is 97.87%.

Another characteristic feature of history works is the high number of texts by one author; it is totally opposite to tendencies present in exact sciences (Price 1967, p. 85; Dobrov 1969, pp. 164–176). In the CIHPM, among 15,041 of history works as many as 94.8% had one author and only 5.2% of works had many authors or were anonymous. This specific character of the historiography had already been confirmed by numerous studies, e.g. British researchers, referring to the humanities, set this indicator at the level of 69% (Bottle and Efthimiadis 1984) and one Polish scholar analyzing bibliographers’ publications – at the level of 89.2% (Skalska-Zlat 2002, pp. 311–312). The dominance of history works by one author seems to be connected with the individualistic methodology of traditional historiography, in which the key element is heuristic. The aforementioned low indicator of history works by many authors (5.2%) makes it possible to limit further analyses only to the first author.

Interesting information about historiography are provided by the results of the examination of authorship and productivity and their connections with other parameters. Analysis of these aspects requires two separate methodologies: one for the most prolific and the most frequently cited (the top of the list) authors and another for less prolific writers. In the

¹⁰ This regularity has been confirmed by other studies in the field of the humanities. D. Konieczna (2002:139) measured that in history of literature books have 73.73% of citations, articles – 24.17% and other – 1.56%.

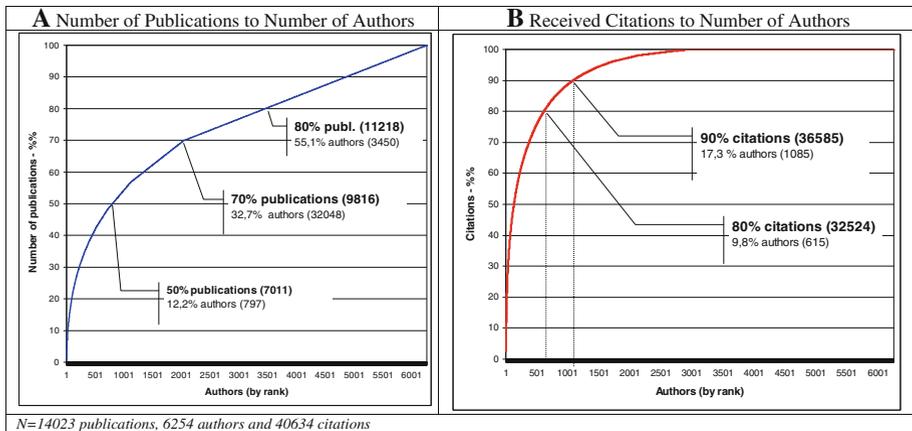


Fig. 4 Author's rank to the number of publications and to citations in the CIHPM [%]

first case the most convenient way of presentation is the Lorenz Curve and in the second one – the approximation of constants for the Lotka's Law. Irrespective of the method, in the body of authors, at the top of the list, there occurred characteristic concentrations of authors confirming the Pareto Principle, according to which a small number of reasons generate a large percentage of results (usually 20:80). This regularity, but in a weaker form, proved to be valid for the most prolific authors, because the top of the list was occupied by 12% of authors who had published the total of 50% of works (Fig. 4a), then respectively 32:70 and 55:80. It is worth noting that in the analyzed distribution at the point of 70% the logistic relationship turns into a linear one. This effect (linearization) has its source in a fact that for a large number of authors there's a tendency of an author to produce one publication only; i.e. 4,222 authors (67.5% of entire corpus) have produced one publication per author only. This assertion is very similar to W. G. Potter's conclusions in which he pointed out that ratio of authors producing one publication only (per author) in a set equals 2/3 (Potter 1980).

Far stronger polarization has been revealed by the relationship of the number of received citations to the author's rank (Fig. 4b). In this case, the Pareto Principle has been confirmed even in excess. 9.8% of authors from the top of the list generate as many as 80% of citations and to reach 90% it has to be taken into consideration (17.3%). This relationship brings to light several things: firstly—the importance of citation measurement in the evaluation of scientific knowledge; secondly—existence of, as it is called, “author's basic group”, which is a group of respectable and frequently cited specialists; thirdly—it lets us assess the number of peripheral publications – i.e. uncited ones. The last indicator for history equals 53.9% and is slightly higher than the SCI average (33.3%) but close to e.g. materials science (48.8%) – Table 5.

An analysis with the use of Alfred J. Lotka's Law provides quite detailed picture of the presence of less prolific authors in the body of disciplines. According to Lotka the product of the number of authors – y , among whom each published x number of works is a constant value – c and is described by the relationship: $x^n y = c$ where n and c are factors depending on the discipline.

As the result of the approximation it has been established that these indicators for history are $n = 1.2$ and $c = 0.67$ and that the real distribution agrees surprisingly well with

Table 6 Historiography and the Lotka's Law (analysis of the CIHPM)

Number of publications	Number of authors	x^n (for $n = 1, 2$)	Number of publications in a group/ total number of publications	Constant (c)
1	4222	1	0.675088	0.675088
2	916	2.297397	0.292933	0.672982
3	404	3.737193	0.193796	0.724253
4	192	5.278032	0.122801	0.648150
5	120	6.898648	0.095939	0.661847
6	88	8.585814	0.084426	0.724866
7	69	10.33041	0.077231	0.797824
8	43	12.12573	0.055005	0.666973
9 ... 150	200	–	–	~0.1–3.6

$N = 6,254$ authors and $14,023$ publications

the principle [$\pm 5\%$]. Presented data (Table 6) demonstrate that distribution limited to the number of less prolific authors from the end of the list (up to eight publications), who have 71% of publications. For the rest of scholars (200 most prolific, who have published from 9 to 150 works) the law is true within the slightly narrower range (the constant oscillates within the range 0.1–3.6). It should be emphasized that received indicators underline the distinctiveness of history in comparison with natural science; for example, exponent n for chemistry and physics is almost twice higher (respectively: 1.88 and 2.02). This suggests that historical sciences include much higher number of researchers that work on history occasionally or temporarily.

Extremely important role in historical studies is played by the topicality of publications measured by the frequency of citing (Fig. 5b). As far as historical science is concerned, works published 3–14 years before (3.0–4.0% of works) are cited very often but the most frequently cited are 6-year-old works (4.0%). However, these are not indicators typical only for history because similar numbers appear in the entire humanities, e.g. psychology (Zafrunnisha and Reddy 2010). Still, in history, there is a distinctive difference for older works: high ratio of citations can be found in texts published 30 years ago (over 1.29%), 40 years ago (0.66%) and 50 years ago (0.27%). To serve the practical purpose it is easier to use the standard half-life indicator – which is the time necessary to receive 50% of citations; for history in the CIHPM it is 14 years (Fig. 5a). Also in this case it is not really different from other fields of the humanities i.e. history of literature (Koniczna 2002, p. 140) but it is in opposition to exact and technical sciences where it is twice or even three times lower (physics—4.6 years, technological chemistry—4.6 years, physiology—8.1 years) (Nalimov and Mulchenko 1971).

The role of the *half-life* indicator should not be limited only to cognitive or comparative goals; it shapes the picture of a discipline seen from the diachronic perspective – when we analyze its development with the use of the time axis. Histograms above (Fig. 6a–d) depict selected parameters in a chronological perspective in order to present the deformations of the picture of the discipline over a longer period. It is worth emphasizing that basic parameters of the description of the body of historiography depend heavily on time, in particular: the number of publications usually increases in a linear way and the number of citations is definitely larger for older publications (Fig. 6a). This state of things strongly impacts and distorts the indicators. This is well illustrated on Fig. 6c that demonstrates that

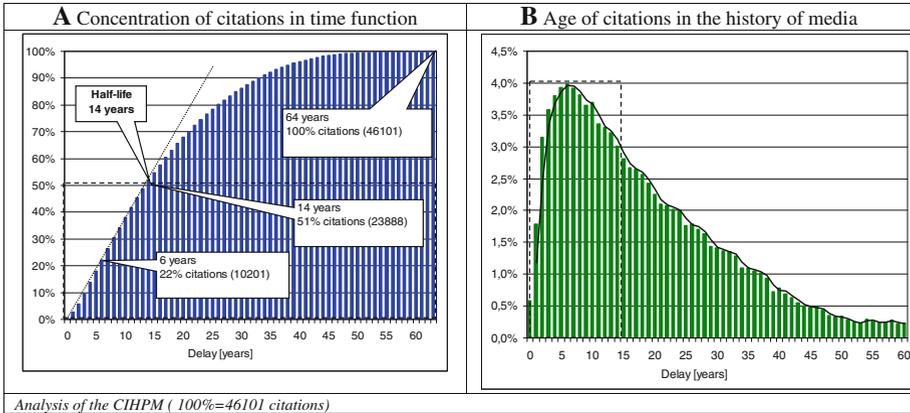


Fig. 5 Age of citations and half-life in historiography

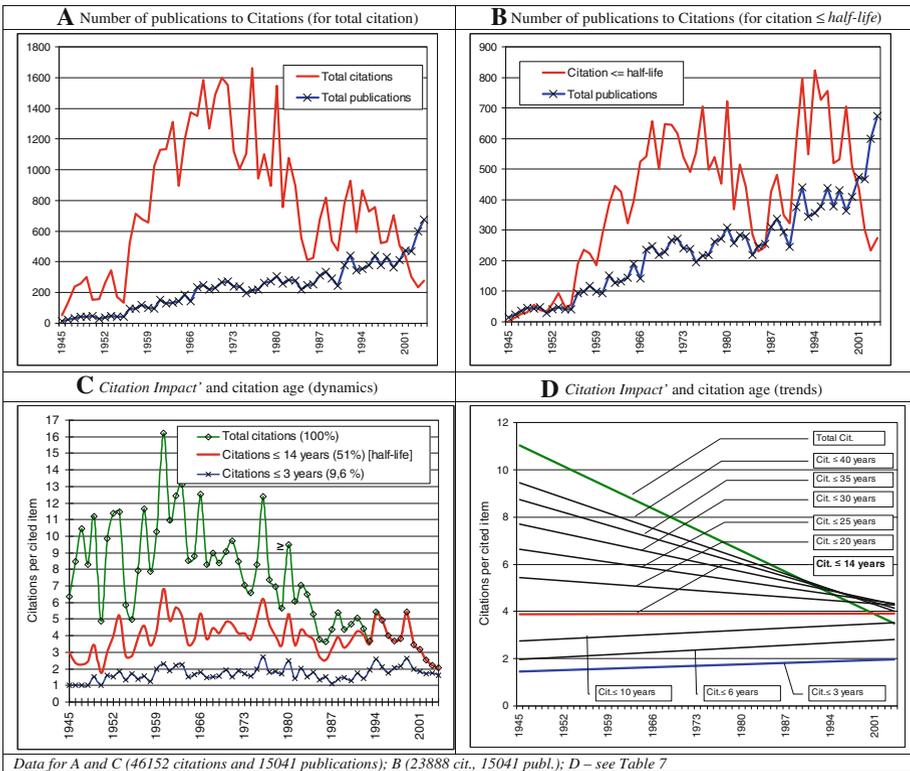


Fig. 6 Selected dynamic relationships in historiography (analysis of the CIHPM)

Citation Impact' [citation per cited item] at full citation number (total citation) is significantly higher for older publications (9.0 on average) and has strong tendency to decrease. It is easy to conclude that simple adding up citation numbers to totals leads to overestimating impact of the older publications.

Table 7 Citation impact^a for different age values of citations (analysis of the CIHPM)

Citations			Citation impact ^a		
Citations age	<i>N</i>	%%	Linear trend ^a	Arithmetic mean	Variance
Total citations	46.152	100.00	$y = -0.1298x + 11.97$	6.84	11.32
Citations ≤ 40 years	44.396	96.20	$y = -0.0950x + 9.61$	6.37	8.55
Citations ≤ 35 years	42.539	92.17	$y = -0.0812x + 8.89$	6.09	7.07
Citations ≤ 30 years	39.814	86.27	$y = -0.0610x + 7.83$	5.69	5.52
Citations ≤ 25 years	36.138	78.30	$y = -0.0425x + 6.75$	5.21	3.99
Citations ≤ 20 years	31.394	68.02	$y = -0.0228x + 5.50$	4.62	2.61
Citations ≤ 15 years	25.312	54.84	$y = -0.0029x + 4.14$	4.06	1.54
Citations ≤ 14 years	23.888	51.76	$y = -0.0013x + 3.94$	3.91	1.35
Citations ≤ 13 years	22,470	48.69	$y = + 0.0028x + 3.66$	3.60	1.29
Citations ≤ 10 years	17.586	38.10	$y = + 0.0108x + 2.78$	3.04	0.91
Citations ≤ 6 years	10.201	22.10	$y = + 0.0128x + 1.99$	2.36	0.42
Citations ≤ 3 years	4.456	9.66	$y = + 0.0087x + 1.43$	1.71	0.17

^a values illustrated on Fig. 6d

Bold values show the dominant

In order to eliminate this negative trend in diachronic analysis one can use publications published later than selected year. This, however, bears the question: how to select this value? A series of experiments included in Table 7 and illustrated on Fig. 6d demonstrates, that *half-life* can be used for the purpose, because for citations younger than 14 years the linear trend coefficient (y in Fig. 6d) is lowest ($y = -0.0013x$) and simultaneously it is close to average (3,9). One could risk postulating a hypothesis that using a number of citations selected this way [HL14 in further text] produces correct image of development of particular discipline. Figure 6b contains example use of HL14, which is HL14-adjusted version of data presented in Fig. 6a. The limitation of HL14 is incomplete representation of citations at the end of the axis (in example at year 1996) and using only half of all citations. However, despite the limitations the indicator works well in practice, which is demonstrated in first historiographical works that make use of it (e.g. Kolasa 2011a) that have been well-received by the reviewers. Of course the boundary can be shortened by the use of even younger citations (Fig. 6c – a series of 3-year-old citations), but in this case we lose a significant number of citations (90.36%) which can make our analysis highly incomplete. This example shows that to examine a general trend (e.g. research front) the use of short periods is possible and justified.

Historiography has one more interesting feature. There is a significant rank difference between publications dealing with individual subjects (contributions) and general studies (syntheses) – Fig. 7. Although less numerous, synthetic publications generate twice or three times more citations (in the example: 2.75 citations per publication) than contributions (1.22) and the number increases proportionally to the quantity of works. So it may seem that in historiography the main indicator of its development is the number of syntheses, which proves the advancement level of a discipline. What should not be overlooked is the fact that differences between types of works mentioned above are deeper and regard methodology; it is a common practice to use archives as the basis of argumentation in contributions but in syntheses the basis is formed by published study results (hence the

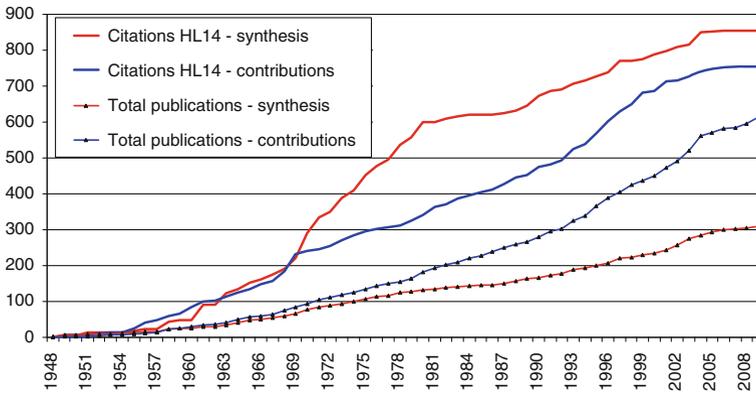


Fig. 7 Syntheses and contributions concerning the history of Cracow’s Press (analysis of the CIHPM)

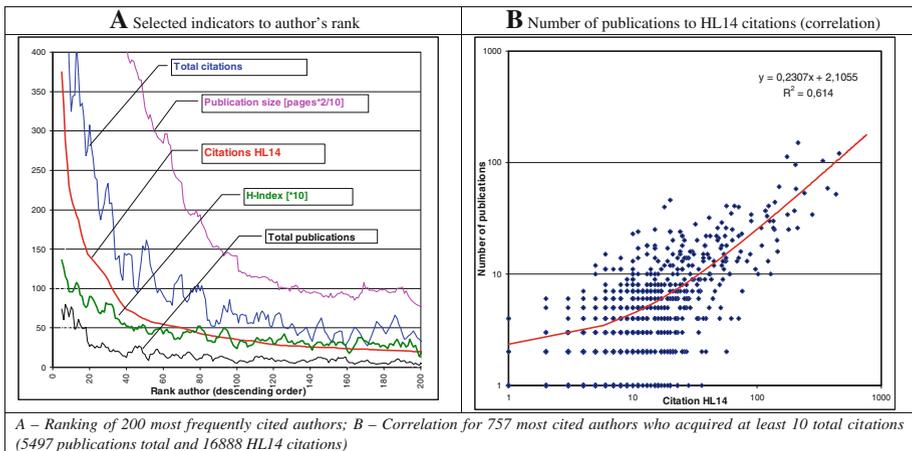


Fig. 8 Selected relationships and correlations in historical science

greater number of references); also their reception differs – because of their wide range, syntheses arise more interest.

In this paper, so far, the HL14 citations have been used. They, as has been proven, are the most just indicator in long-term analyses but still, it does not mean that they are the only useful measure in quantitative analyses of history works. In countless situations there are other, more useful methods, e.g. total number of publications, complete citations or the Hirsch Index (Hirsch 2005). Detailed analysis confirms that in history, for frequently-cited authors, all measured indicators are closely correlated (Fig. 8a).

Especially interesting is the correlation between the number of publications and received citations (Fig. 8b). This relationship is clearly visible when it is analyzed on a logarithmic scale. From a certain point (about 30 citations) there is a complete positive correlation ($R = 0.78$). It means that the body of frequently-cited authors overlaps with the most prolific ones. This relationship corresponds with the study results of American sociologists: J. Cole and S. Cole from 1960s. These authors analyzed a group of physicists

and established that: frequently-cited and at the same time the most prolific authors (33% of the group) were also frequently-awarded ones (90% of the group), held high positions (58% of the group) and were widely respected (69% of the group) (Nalimov and Mulchenko 1971, p. 102).

Closer analysis of top authors from the examined group (25 the most frequently-cited authors) shows several other relationships. There is, among others, strong correlation between author's rank and his degree – 92% of them are independent scholars (e.g. doctors), usually with a professor title (76%). On the other hand, in order to be a renowned persona in the field of historiography one has to work for many years. The average time of research activity for the selected group of authors is slightly over 37 years (which is almost entire active professional life); only in some rare cases the career is faster (19 years). Among many characteristic features of history works there is one more worth mentioning, namely, extensive publications. The average number of pages of a history work is 49 but there are some which have over 100. In this context, impressive literary achievements of some historians should not be surprising – for the analyzed 25 names from the top of the list the average number of pages was 2,245 and in case of the most prolific author it was 6,246 pages. However, it does not seem that historians' publication abilities differ from those of other scholars. In the analyzed set the average number of publications of one author was 52 works, which means that historians have published on average 1.3 publications per year (3 per year for the most prolific author). But if we include scholars strictly connected with the analyzed discipline and active during the entire analyzed period the indicator oscillates between 2.19 and 2.88. Provided data fully correspond with findings of sociologists of science: according to a representative survey sent over by 5,306 scholars, it is feasible to write between 2.1 and 3.4 of articles per year (Dejnarowicz 1980, p. 153).

Further development of these measured relationships leads to the conclusion that the majority of discussed indicators used in the field of historiography (and probably in other areas of the humanities) is in fact emanation of authority and it correlates strongly to long and industrious scientific career.

Conclusions

The analysis of international sources has proven that they barely or to a very small extent represent Polish historical science (Table 2). This defect does not concern only Polish historiography but all national historiographies, which – as further studies showed – are by nature local (Table 3). Localism is present in two aspects: linguistic and geographical. In the first case all sources confirmed strong correlation of a language and national themes (Fig. 1; Tables 3 and 4) – in sources with high relevance this indicator oscillates within the range 95.6–98.8%; and in the second case – the strong dependence of national themes on the country of publication, to which they refer was depicted (214–529%) (Table 3). It should be underlined that ISI services are totally useless as far as historical research is concerned because of their glaring geographical non-representativeness; an additional defect is, on one hand, the lack of book registration and on the other, disproportionately high representation of book reviews (65.6% of publications total) distorts the picture of the discipline.

Far more important results were provided by the analysis of the local (model) citation index – CIHPM. Comparison of main citation indicators of the model database with ISI services confirmed that from the bibliometric perspective history does not differ strongly from other disciplines, even these from the *science* group i.e. materials science (Table 5)

and seems to be governed by similar rules. Among specific features of history the role of books was highlighted; although these publications constitute less than 15% of the general set, they include as much as 92% of the most frequently cited works (Fig. 3). Also, strong linguistic polarization was confirmed: Polish—96.2%, English—1.3% and German—1.0%; which is the consequence of the very nature of historiography, which requires wide range of contextual knowledge impossible to acquire in a different way than by active participation in the given culture and difficult to express in a language other than a national one. Another specific feature of history works is a great number of publications by one author (94.8% of publications total) – this seems to be connected with the individualistic methodology of traditional historiography in which the key role is played by heuristics. The examination of authorship presents strong polarization in a group of prolific authors (50% of works by 12% of authors) and even stronger in a group of highly-cited authors (80% of works by 9.8% of authors) – Fig. 4. On the other hand, the analysis of the Lotka's Law (established indicators: $n = 1.2$ and $c = 0.67$) proves that historiography brings together far more scholars, who deal with history occasionally (Table 6). Historiography (as other disciplines from the field of the humanities) is characterised by quite a long period of topicality measured by the frequency of citing (the *half-life* is 14 years) and ability to retain its topicality even for publications 50 years old (0.27% of citations). Extremely useful for the diachronic study of the discipline turned out to be the HL14 indicator (i.e. citations younger than *half-life*) introduced *ad hoc*. The use of HL14 for longer periods prevents the picture of the discipline from being deformed by favouring older publications (Fig. 6b). From other described regularities one also deserves to be mentioned, namely, the significant difference concerning the frequency of citing between contributions and syntheses; the latter is cited on average three times more often (Fig. 7). Many conclusions can be drawn from the comparison of selected bibliometric indicators – for the most frequently cited authors all selected indicators (complete citations, HL14, number of publications, the H-Index and even the volume) are strongly correlated (Fig. 8).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Archambault, É., & Gagné, É. V. (2004). *The use of bibliometrics in the social sciences and humanities*. Montreal: Science-Metrix.
- Boonstra, O., Breure, L., & Doorn, P. (2004). *Past, present and future of historical information science*. Amsterdam: Netherlands Institute for Scientific Information.
- Bottle, R. T., & Efthimiadis, E. N. (1984). Library and information science literature: Authorship and growth patterns. *Journal of Information Science*, 9(3), 107–116.
- Cybulski, R. (1990). *Książka na świecie*. Wrocław: Ossolineum.
- Dejnarowicz, Cz. (1980). *Literatura naukowa – uczeni – wydawcy*. Warszawa: PWN.
- Dobrov, G. M. (1969). *Wstęp do naukoznawstwa*. Warszawa: PWN.
- Finkenstaedt, T. (1990). Measuring research performance in the humanities. *Scientometrics*, 19, 409–417.
- Garfield, E. (1979). *Citation indexing – its theory and application in science, technology, and humanities*. New York: Wiley.
- Hérubel, J.-P. V. M. (2006). Clio's view of the history of science: A preliminary bibliometric appreciation. *Behavioral & Social Sciences Librarian*, 24(2), 69–91.
- Hérubel, J.-P. V. M. (2007). Disciplinary and research subjects in french doctoral dissertations in press and media Studies: A bibliometric treatment. *Behavioral & Social Sciences Librarian*, 25(2), 23–46.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.

- Kolasa, W. M. (2011a). Kierunki badań nad historią prasy polskiej 1918–1939. Cz. 1, Tendencje rozwojowe, typologia, *Rocznik Historii Prasy Polskiej* (Vol. 27/28), in print.
- Kolasa, W. M. (2011b). Retrospektywny indeks cytowań w humanistyce: koncepcja, metoda, zastosowania = Retrospective citation index in humanities: the idea, methodology, applications [preprint], *E-Lis* - <http://hdl.handle.net/10760/16052>. Accessed 9 Jan 2011.
- Kolasa, W. M., & Jarowiecki, J. (2005). *Polska bibliografia prasoznawcza 1996–2001 = The Polish Press Research Bibliography 1996–2001*. Kraków: Polska Akademia Nauk.
- Konieczna, D. (2002). Bibliometryczna analiza publikacji cytowanych w czasopiśmie “Litteraria” w latach 1969–1999, *Zagadnienia Naukoznawstwa*, z. 1/2, s. 137–145.
- Kozłowski, J. (1994). *Miejsce nauki polskiej w świecie*. Warszawa: Komitet Badań Naukowych.
- Marshakova-Shaikovich, I. (2001). Nauka polska u progu XXI wieku: analiza bibliometryczna dorobku Polski i państw Unii Europejskiej, *Roczniki Biblioteczne*, R. 45, s. 147–165.
- Marshakova-Shaikovich, I. (2009). *Badania ilościowe nauki: podejście bibliometryczne i webometryczne = Quantitative analysis of science: Bibliometric and webmetric approach*. Poznań: Uniwersytet im. A. Mickiewicza.
- McCrank, L. J. (2001). *Historical information science : An emerging unidiscipline*. Medford: Information Today.
- Nalimov, V. V., & Mulchenko, Z. M. (1971). *Naukometria*. Warszawa: WNT.
- Nederhof, A. J., Zwaan, R. A., Debruin, R. E., & Dekker, P. J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioral sciences. *Scientometrics*, 15(5-6), 423–435.
- Nederhof, A. J., van Leeuwen, T., & van Raan, A. (2010). Highly cited non-journal publications in political science, economics and psychology: A first exploration. *Scientometrics*, 83(2), 363–374.
- Nowak, P. (2000). *Wybrane problemy efektywności polskich czasopism naukowych z dziedziny humanistyki*. Poznań: MotiVex.
- Nowak, P. (2004). Piśmiennictwo z zakresu nauk społecznych i humanistycznych przedmiotem oceny i analiz metodami bibliometrycznymi. Możliwości i ograniczenia. *Roczniki Naukowe PWSZ im. Komeńskiego w Lesznie*, nr 2, ser. A, *Miscellanea*, t. 2, s. 5–18.
- Nowak, P. (2008). *Bibliometria, webometria. Podstawy, wybrane zastosowania*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Potter, W. G. (1980). When names collide: Conflict in the catalog and AACR 2. *Library Resources and Technical Services*, 24(1), 3–16.
- Price, D. J. de Sola (1967). *Mała nauka – wielka Nauka*. Warszawa: PWN. Orig.: *Little science, big science*. New York: Columbia University Press, 1963.
- Schubert, A. (1999). Scientometrics: A citation based bibliography 1994–1996. *Scientometrics*, 44(2), 267–315.
- Schubert, A. (2001). Scientometrics: A citation based bibliography 1997–2000. *Scientometrics*, 50(1), 99–198.
- Skalska-Zlat, M. (2002). *Bibliografia w Polsce 1945–1996: naukoznawcza analiza dyscypliny*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Zafrunnisha, N. V., & Reddy, P. (2010, July). Citations in Psychology PhD Theses: An obsolescence study, *library philosophy and practice* (Vol. 2010). <http://digitalcommons.unl.edu/libphilprac/400> 21 Mar 2011.