ISSN 2707-0476 (Online)
University Library at a New Stage of Social Communications Development. Conference Proceedings, 2020, No V
UniLibNSD-2020

UDC 001.811

LAZAREV V. S.
Scientific Library, Belarusian National Technical University (Minsk, the Republic of Belarus),
e-mail: vslazarev@bntu.by, ORCID 0000-0003-0387-4515

ON NEW EXOTIC ATTEMPTS TO QUESTION THE RELATIONSHIP BETWEEN CITEDNESS FIGURES AND THE USE / VALUE OF CITED SCIENTIFIC DOCUMENTS

The objective of the present paper was to analyze a number of concepts related to the scientometric method "citation index". Some experts still interpret “negative citations” and “disproportionately large share of self-citations” as “shortcomings” of the “citation index” scientometric method. In the author’s opinion, the reason is an indelible belief in the “normative theory of citation”, which implies a conscious choice of citations, a conscientious selection of the best quality works for the reference-lists, and a “desire to return the intellectual debt” to the cited authors. But there are also new attempts to question the adequacy of the “citation index” method. For example, one of the publications of the year 2020 questions the causal relationship between the citations to the scientific documents and their use, and between their use and their value; the reason why this fundamental pattern is called into question is mainly based on the actions of some people who do not meet ethical standards and may be arbitrarily directed against its manifestation. The mentioned paper claims, in particular, that very common are the following phenomena: deliberate refusal to use the necessary scientific documents; refusal to cite the documents used (including plagiarism and restraint from citing for technical reasons); practice of evaluating little-known and inaccessible documentary sources as not valuable enough (value is confused with quality in this case); fake citing to unread works. The present article refutes the interpretation of these phenomena as indicating the absence of the considered causal relationship and tries to demonstrate that, on the contrary, some of them are involuntary confirmation of the adequacy of the “citation index” method.
Keywords: citation; use; value; quality; scientific document; scientometrics; manipulating citations

Some experts still interpret “negative citations” and “disproportionately large share of self-citations” as “shortcomings” of the “citation index” scientometric method. In the author’s opinion, the reason for this is an indelible belief in the “normative theory of citation”, which implies a conscious choice of citations, a conscientious selection of the best quality works for the reference-lists, and a “desire to return the intellectual debt” to the cited authors.

But there are also new attempts to question the adequacy of the “citation index” method. For example, one of the publications of the year 2020 (Krulev, 2020) questions the causal relationship between the citations to the scientific documents and their use, and between their use and their value.

The reasons presented in the cited paper look both new and unexpected for such kind of papers; so, no matter how obviously weak we would find these arguments; their novelty induces us to consider them (and put forward our counter-arguments).

Krulev (2000, pp. 85-87) announced that it is possible that cited documents were “**used but not valuable**”. But within the framework of the concept of value (Lazarev, 2017, 2019a, 2019b, 2019c) (which was *not* the subject to any revision attempts by *Krulev*) the only comment to this statement is possible: it just never happens. It just never happens because in information science, the concept of value is defined as “a property of information determined by its suitability for practical use in various areas of purposeful human activity to achieve a certain goal” (Dictionary…, p. 464), so the use itself *determines* value of the used documents. Moreover, “if in philosophy *value* is interpreted as a criterion of preference in the situation of alternative choice (Minkina, 1983, p. 111), then the choice itself of a specific reference for the

https://doi.org/10.15802/unilib/2020_220647

reference-list from the set of possible ones <...> already inevitably indicates the value of the cited document!“ (Lazarev, 2017, p. 6).

However, instead of discussing the issue on its merits, the section of the Krulev's paper entitled “Used but not valuable” is filled with examples of how the cited document may not be identifiable–though *Krulev* himself admits that “identification is **not** an indicator of value” (Krulev, 2020, p. 85). Also, he discusses the reasons of unsubstantiated editorial assessment of papers as being of poor *quality* because of the predominance of citations to little-known or to inaccessible sources in them (from the point of view of editors). But such assessments are related to the publisher's arbitrariness, they are *quality assessment* performed by the editors, and they have nothing to do with the *value of the cited literature*. *Krulev* just confuses value with quality–i.e. the “degree to which a set of inherent characteristics <...> of an object <...> fulfills requirements” (ISO, entry 3.6.2).

Also, in this section *Krulev* discusses the technical problems of scientometric citation studies related to the retraction of publications and the fact of possible changes in the content of the officially published documents as compared with that of having been published as preprints. However, after the publication of the same material in the form of an article the cited preprint would *not* lose its value, confirmed by the citations that it has already gained. Also, if someone used a paper that was subsequently retracted, the fact that it has been retracted does not change the fact that it was *used*. In this regard, the claim of *Khrulev* “used but not valuable” is also devoid of all reason.

Also, *Krulev* claims (in the section of his paper entitled “**Valuable but not used**” – see pp. 87-88) that a valuable document may remain unused. This is, of course, another logical nonsense because value (in contrast with quality) is being manifested (as was already shown) *only through use*. So, when *Krulev* claims that conscious refusal to use of valuable material is often practiced (2000, p. 87), he, in fact, can mean only the *refusal to cite*. However, there are no evidence-based examples of such facts in his paper. The fact that different specialists of the same profile may just have *different* approaches to the same scientific document (and, correspondingly, use it or not) just does not occur to *Krulev*. In addition, a relevant potentially valuable document could remain simply unknown to the scientists. It is common knowledge that *all* documents containing information relevant to a particular study *never reach* the authors that perform such a study. However, this argument has never been used to refute scientometrics.

The argumentation presented in the section of his paper entitled “**Cited but not used**” (2000, p. 81-84) can hardly be attributed to methodology at all: the *general principle* of citing is opposed here by particular... not even always the peculiarities of its manifestation, but by peculiarities of human activity in *falsifying* it. This argument is similar to the statement that the monetary system is doubtful due to the presence of counterfeiters! But the successful forgery of banknotes, providing the falsifier with real benefits (before it is disclosed), indicates just the *perfection* of the current monetary system; along with this counterfeit banknotes are *not* part of the monetary system.

If the percentage of works that were cited but not read actually could reach 70% or more, as it follows from (Simkin & Roychowdhury, 2005), this would mean that the entire scientometrics is built on an absolutely unreliable foundation. Rejecting the possibility of such a share of “nominal citations” I rely on my own experience and common sense. I mean that in 1979, in my first publication, I once naively cited a paper knowing only its abstract, but when I had a chance to read its full text (which completely refuted the abstract written by an abstractive journal employee) I found out that I had just misled my readers… and I have never done such things since (Lazarev, 2017, p. 12). Such a behavior (not to cite the paper if you are not aware of

https://doi.org/10.15802/unilib/2020_220647

its whole content) seems natural and typical, while the one described in (Simkin & Roychowdhury, 2005) seems to be unbelievable (Lazarev, 2019b, p. 131) because it is clear that the conscious practice of citing unread (and even underread) research papers pose a very significant risk "to be trapped". So, would the authors so often neglect this risk and consciously resort to such practices? It is difficult for me to imagine their significant prevalence among scientists.

Krulev states that sometimes bibliographic descriptions included in the lists of references are not part of the cited literature, but part of the materials of scientometric research, so they do not indicate the documents that had been actually read and used in writing the citing article. This is an imitation of citing by malice, but by negligence. Of course, everyone understands the wrongness of this practice – but *does it still exist*? For example, three years ago, some journals did not include in my reference lists even citations to the sites that I studied to make clarifying inquiries about the materials of my research. They *were read*, so the citations were not even "nominal"! But still, they were considered to be part of the "materials" and not of the "sources used". And, as it seemed to me, it was quite fair decision. So, the inclusion a list of bibliography that is part of materials of metric research to the reference lists seems even more unlikely. Usually journals present such materials in a form of tables or Annexes.

Indeed, there is problem of "nominal citing", but it is a technical problem, not the problem of methodology. There are indeed practices that are aimed against the manifestation of the very principle on the basis of which citing is carried out. But they do not abolish it. As a "rhetorical analogy" I shall give another example: once in the United States there were a lot of corrupt police officers who act de facto *against* the police system; but could they be a reason to talk about ambiguity, or–even more–about the inadequacy of the very idea of the police?

Some of my articles–not the methodological ones–are devoted to this problem. Thus, in my works (Lazarev, 2019d, p. 17-18; Lazarev, 2019e) citations to unread documents that were made involuntarily are mentioned. It happened due to the fact that some publishers and journal editorial offices refused to accept articles with less than a certain number of references. With my co-authors we have also encountered a glaring case, viz. the requirement of an American journal to replace Russian-language references with English-language ones (Lazarev, 2019d, p. 17-18; Lazarev, 2019e). It is clear that following such requirements would distort possible future scientometric research based on citations in such a journal and would simply lead to the fact that citations would no longer serve as a reliable means of searching for publications by the reader of the corresponding papers (Lazarev, 2019e, p. 93). Being reluctant to produce fake citations, we responded to the requirement to replace Russian-language citations with English-language ones simply by stopping correspondence with this journal and submitting our paper to another one. As for "norming" the number of citations by some journals (both according to the "no less than…" or "no more than…" rule), I simply refrain from submitting papers to such journals – except in cases when the "norm" does not prevent me from making as many citations as I myself consider necessary. In other words, the technical problem created by publishers can be overcome by the authors themselves through technical solutions! As for scientometric research, it is clear that the journals which do not accept papers containing, for example, less than ten or more than twenty citations should not be adopted as part of scientometric research materials, because inclusion of such journals will obviously lead to the fact that some of the documents actually used by the authors will *not* be reflected in the study sample, while some "nominal" citations will be included in the sample. But why should scientometricans use such (dubious) sources?

There *is* a problem of interference of publishers and editors in the structure of citations–interference not motivated by the content of a citing paper and its internal associations with the

https://doi.org/10.15802/unilib/2020_220647

cited material–and this is a serious problem. But this problem is *technical, organizational*, and it is a problem of incompetent and irresponsible interference of people in the self-regulating system of science, and not of the system of science itself. The reasons for this interference, by the way, confirm the effectiveness of the "cited because used" principle: after all, an editor who requires English-language citations or citations to his own journal is concerned that its country (region) and its journal look better in terms of scientometric evaluation: he fakes the values which, if being natural, would indicate the demand for publications of a particular region and journal. At the same time, when submitting the initial manuscript, the author still cites what he used (Lazarev, 2017, 2019a, 2019b, 2019c), and false citations are added to the list (or even replace part of the original list) *not by his will or choice*. However, the authors may not accept such ultimatums–although, since this practice exists, it is not always easy. These vicious practices have nothing to do with the nature or essence of bibliographic citation. Such problems require technical, rather than conceptual solutions…

It was shown that (*The "phantom reference"...*) careless writing caused by the use of a template that was not subsequently removed from the list of cited references, followed by careless quality control, can generate hundreds of citations to a non-existent paper without violating the principle "Cited, because it is used, and used because it is valuable". The citations to a non-existent paper occurred to be just a paragraph of a template with "how-to-cite" specimens that had not been deleted by the authors when preparing a paper! However, the phantomness of a highly cited non-existent paper was easily detected in a scientometric study (*The "phantom reference"...*). But in any case, this example shows the need for accurate work of reviewers and editors–and nothing more. The situation described in (*The "phantom reference"...*) has nothing to do with the regularities of scientific communication, one of which is that of "cited because used".

Krulev also insists that sometimes some authors deliberately do not cite the used documents (The "**Used but not cited**" section of his paper–see pp. 84-85). "Sometimes they are forced to exclude sources, following the editorial requirements to minimize the bibliography", he adds (Krulev, 2020, p. 84).

We have already commented the situation with "the editorial requirements to minimize the bibliography". But *Krulev* also thinks that the reasons for not citing documents that were actually used can be: plagiarism, "errors and carelessness in preparation a paper", unwillingness to cite the opponents. Herewith he does *not* provide *any* reference or data confirming his idea about abundance of not citing the used documents. Maybe they are single cases, "statistical error"?

Indeed, the *lack* of evidence that some materials were allegedly used, but not quoted, is given by *Krulev* almost as proof of the *prevalence* of such a phenomenon. I have already noted above that all documents containing information relevant to particular research *never reach* the authors performing this relevant research. That means that the absence of citations to the "most obvious" publications does not mean that they were used, but were not cited: the authors just might not know about them! Or the authors really did *not* use these documents when writing the a citing paper.

"Errors and carelessness in preparation a paper", which *Krulev* writes about, do occur, and sometimes they do cause the fact that not everything that was really used in the research would be cited in the paper describing it. This, by the way, is very well known (Kara-Murza, 1981). Moreover, not everything used in the performance of individual specific works is cited in the articles describing its results–*even out of connection* with the errors and malicious intent of the citers (Kara-Murza, 1981). But citations are based on the opinions of a large number of

https://doi.org/10.15802/unilib/2020_220647

scientists, i.e. when attracting large arrays of citations, individual "distortions" are likely to make up an insignificant proportion. "If one looks at references in an individual paper, many peculiarities may be found, such as missing references to specifically important papers, or to the work of authors who have generally made essential contributions to the field, or an exaggerated attention to a specific author. Indeed, if just this one paper with its peculiar references would be analysed, a seriously mistaken picture of the field concerned will be obtained. But as soon as further papers are added, similar but also other irregularities will be discovered in their reference lists. Does this mean that one would never be able to get any sensible idea of the most important work in that field? This is statistically only the case if all researchers refer to earlier work completely arbitrarily. But nobody can seriously maintain that the references in, for instance, this paper are totally unreasonably and completely arbitrary" (van Raan, 1998, p. 134). When the arrays of citations are large, these peculiarities are all supposed to be smoothed out.

Furthermore, if not all the sources used in the creation of the citing scientific paper are cited in it (Kara-Murza, 1981), the "direct" statement remains true: everything that was cited by the scientist was used by him in the creation of the citing scientific paper (Kara-Murza, 1981). One can regret the "inevitable" incompleteness of its reflection, but–incompleteness compared to what? There is no method that is more accurate than citation analysis to reflect the use of documents when performing a specific scientific work.

Finally, let us consider the statement by *Krulev* that various citation practices in review articles, original journal articles, and conference proceedings allegedly cause to "citation losses" (2000, p. 84).

What actually follows from the fact that "the document can be used and cited in the review article, but not cited in the proceedings of the scientific conference" (p. 84)? Only the need for accurate planning of scientometric research and nothing else–which is absolutely obvious without Krulev's paper. In general, scientometric research of citations in conference proceedings has its own specifics, and for generalizing judgments about scientific activities, it is used much less often than the study of citations in journals. It is good to use conference proceedings as material for scientometric research when there are no journals specialized in a corresponding subject (Lazarev, Roath, Yunusova, & Safonenko, 1999; Lazarev, Safonenko, & Yunusova, 2001). The specificity of conference proceedings and review journals as sources of bibliographic references for scientometric research is also not a methodological issue; it is a technical one. A proper understanding of its specifics is indeed important for adequate research planning, but this understanding is not in the least related to the general principle of citing.

Thus, we have to admit that, despite the bold attempt to question the existence of causal relationships between use and citation expressed in the paper by Krulev (2020) constructive approaches in this paper are actually absent.

REREFENCES

Dictionary of the Terms of the Information Science. (1975). Moscow: International Center for Scientific and Technical Information. (in Russian, with the applied lists of terms in 13 more languages)

ISO 9000:2015(en) Quality management systems – Fundamentals and vocabulary. In: ISO Online Browsing Platform (OBP). Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en (in English)

Kara-Murza, S. G. (1981). Citation in science and approaches to evaluating scientific contributions. *Vestnik AN SSSR [Herald of USSR AS],* 5, 68-75. (in Russian)

https://doi.org/10.15802/unilib/2020_220647

Krulev, A. A. (2020). Citation as a form of science communication. *Scientific and Technical Libraries*, *3*, 79-92. doi: https://doi.org/10.33186/1027-3689-2020-3-79-92 (in Russian)

Lazarev, V. S. (2017). Scientific documents and their ordered totalities: citation, use, value. *Mezhdunarodnyy forum po informatsii, 42*(1), 3-16. Retrieved from http://lamb.viniti.ru/sid2/sid2free?sid2=J15472040 (in Russian)

Lazarev, V. S. (2019a). Insufficient definitions or a vaguely grasped notion? On definitions of "impact". *Scholarly Research and Information*, *2*(1), 63-78. doi: https://doi.org/10.24108/2658-3143-2019-2-1-63-78 (in Russian)

Lazarev, V. S. (2019b). The property that is factually being evaluated when they say they evaluate Impact. *Scholarly Research and Information, 2*(2), 129-138. doi: https://doi.org/10.24108/2658-3143-2019-2-2-129-138 (in Russian)

Lazarev, V. S. (2019c). On the possibilities of evaluating properties of scientific documents on the basis of their citations count (Or again: What property is reflected by citations count par excellence, after all?). Part 1: Value. *University Library at a New Stage of Social Communications Development. Conference Proceedings*, *4*, 28-36. doi: https://doi.org/10.15802/unilib/2019_187405 (in English)

Lazarev, V. S. (2019d). The power of bibliometric illusions over the lazy, the profanation of fruitful ideas and the curse of parabibliometric evaluation of science. *Science Editor and Publisher, 4*(1-2), 12-20. doi: https://doi.org/10.24069/2542-0267-2019-1-2-12-20 (in Russian)

Lazarev, V. S. (2019e). Manipulation of bibliometric data by editors of scientific journals. *European Science Editing, 45*(4), 92-93. doi: https://doi.org/10.20316/ESE.2019.45.19011 (in English)

Lazarev, V. S., Roath, S., Yunusova, D. A., & Safonenko, O. K. (1999). Biomedical application of magnetic fluids as a specific interdisciplinary problem. In G. C. Cornelis, & J. P. Van Bendegem (Eds.) *Metadebates on Science: The Blue Book of "Einstein Meets Magritte"*. Dordrecht; Boston; London; Brussels: Kluwer Academic Publishers, VUB. Retrieved from https://rep.bntu.by/handle/data/25411 (in English)

Lazarev, V. S., Safonenko, O. K., & Yunusova, D. A. (2001). Bibliometric research of international scientific conferences proceedings in the context of current objectives of library and information service. *Journal of Interlibrary Loan, Document Delivery and Information Supply, 11*(4), 103-118. doi: https://doi.org/10.1300/J110v11n04_09 (in English)

Minkina, V. A. (1983). Investigations of documentary flows for assessment the characteristics of value of technical literature. In *Documentary Flows on Natural Sciences and Technology and Bibliographical Problems: Collected Papers*. Leningrad, 1983. (in Russian)

Simkin, M. V., & Roychowdhury, V. P. (2005). Stochastic modeling of citation slips. *Scientometrics, 62*(3), 367-384. doi: https://doi.org/10.1007/s11192-005-0028-2 (in English)

The "phantom reference": How a made-up article got almost 400 citations. In *Retraction Watch*. Retrieved from https://retractionwatch.com/2017/11/14/phantom-reference-made-article-got-almost-400-citations (in English)

van Raan, A. F. J. (1998). In matters of quantitative studies of science the fault of theorists is offering too little and asking too much. *Scientometrics, 43*(1), 129-139. doi: https://doi.org/10.1007/BF02458401 (in English)

https://doi.org/10.15802/unilib/2020_220647

ISSN 2707-0476 (Online)
University Library at a New Stage of Social Communications Development. Conference Proceedings, 2020, No V
UniLibNSD-2020

ЛАЗАРЄВ В. С.
Наукова бібліотека, Білоруський національний технічний університет (Мінськ, Республіка Білорусь), e-mail: vslazarev@bntu.by, ORCID 0000-0003-0387-4515

ПРО НОВІ ЕКЗОТИЧНІ СПРОБИ ПОСТАВИТИ ПІД СУМНІВ ВЗАЄМОЗВ'ЯЗОК МІЖ РІВНЕМ ЦИТОВАНОСТІ ТА ВИКОРИСТАННЯМ / ЦІННІСТЮ ЦИТОВАНИХ НАУКОВИХ ДОКУМЕНТІВ

Метою даної статті був аналіз ряду понять, що відносяться до наукометричного методу «цитат-індекс». Деякі фахівці, як і раніше, вважають «негативне цитування» та «непропорційно велике самоцитування» «недоліками» наукометричного методу «цитат-індекс», що, на погляд автора, свідчить про віру в «нормативну теорію цитування», яка передбачає свідомий вибір посилань, принциповий відбір найбільш якісних робіт для списків цитованих джерел, «бажання повернути інтелектуальний борг» цитованим авторам. Але виникають і нові спроби поставити під сумнів адекватність методу «цитат-індекс». Так, в одній з публікацій 2020 року причинно-наслідковий зв'язок цитованості наукових документів з їх використанням, а використання – з їх цінністю ставиться під сумнів, в основному на тій підставі, що дії деяких людей, які не відповідають етичним нормам, можуть бути довільно спрямовані проти її прояви. Тут мається на увазі, зокрема, такі нібито дуже поширені явища: свідома відмова від використання та цитування необхідного матеріалу; відмова від цитування використаного матеріалу (включаючи і плагіат, і відмову від цитування з технічних причин); практика оцінки маловідомих і недоступних джерел як недостатньо цінних (цінність при цьому переплутана з якістю); виробництво фальшивих посилань на непрочитані роботи. Ця стаття містить спростування точки зору на розглянуті явища як на свідчення відсутності даного причинно-наслідкового зв'язку і вказівка на те, що, навпаки, деякі з них є підтвердженням адекватності методу «цитат-індекс».

Ключові слова: цитованість; використання; цінність; якість; науковий документ; наукометрія; маніпулювання посиланнями

https://doi.org/10.15802/unilib/2020_220647