Ivo Družić[a]

JEL classification: A2, I2
Original scientific paper
https://doi.org/10.32910/ep.75.5.4

Elements of scientific excellence assessment

Abstract

The dynamics of the international and national research excellence assessment, especially in the last twenty years, is characterized by "tidal waves" in which quantitative and qualitative criteria alternate at the "tops". Analyzing experiences from that relatively long period, the paper argues for a rational and gradual improving of the balance between quantitative and qualitative criteria by passing through the three "levels" of forming a comprehensive evaluation of scientific excellence. In this sense, it is advocating the use of IF as the first necessary "level" that needs to be passed, in order to be able to reach the second "level". This second "level" consists of a qualitative excellence assessment, including the assessment of individual contributions. These are the perquisites of a complete expert review (peer review) assessment at the final third "level" determining research excellence. In this sense, the paper offers an integrated model of a balanced quantitative and qualitative criteria combination, not only for an integral assessment, but also for encouraging scientific excellence.

Keywords: Research Quality Assessment, Quality Indicators, citation base, citation ratings, Impact factor, integrated Model of quality assessment

[a] I. Družić, Ph.D., Professor Emeritus, Faculty of Economics and Business Zagreb, (e-mail: idruzic@efzg.hr). The paper was received on 18.01.2024. It was accepted for publication on 03.02.2024.

1. INTRODUCTION

Due to the growth of global scientific production, a professionally competent and objective evaluation of a researcher's scholarly contribution has become an indispensable element of research findings design and presentation for various purposes. The first is the assessment of the quality of a scientific work/article on the occasion of publication in a scholarly journal. The second is the evaluation of the scientific quality of the scientist himself by an overall evaluation of the quality of all his research presented in scholarly journals, for the purpose of selection for academic titles/research positions. The third is the thematic quality evaluation of the research teams' heads and members in tenders/calls for the award of scientific projects. The fourth purpose is to evaluate the quality of research institutions (universities, faculties, scientific institutes) in the (re)accreditation process, in which the quality of researchers/teachers scientific output plays an important role both individually and at the level of the institution as a whole.

This resulted in the need to develop both quantitative and qualitative criteria to evaluate the quality of scientific publications in which the work was published, then the quality of the work itself, and finally, perhaps most importantly, the quality of the author's input in the single-authored work and the quality of the individual author's contributions in the joint scientific work.

In the last two decades, a whole spectrum of globally accepted bibliometric/scientmetric quantitative indicators has been developed, which indirectly measure scientific quality and productivity. The basis of international quantitative standards of excellence is the citation database of journals/publications. They are used to create indicators of the journal's impact/citation rate in other publications of the same citation database. Then indicators of citations of papers published in the journal included in the citation database in other journals/papers of the same citation database. Finally, there are indicators of citations of authors/names, individual/collective works published in a publication included in the citation database, in other works published in publications included in the same database. But after the "wave" of new quantitative indicators, especially from the mid-1990s and the beginning of the 2000s onwards, a period of "sobering" began in the 2010s, especially as can be seen in the DORA declaration (DORA, 2012) and the Leiden Manifesto (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015). The document of the European Commission "Towards reform of the scientific research evaluation system" (European Commission, 2021) is a kind of a peak of the "sobering" discourse.

This significant turn in the criteria for evaluating scientific excellence is the result of the fact that, through practical application, bibliometricians themselves identified not only advantages, but also limitations in the application of bibliometrics when it comes to the process of evaluating both scientific projects and candidates for academic positions/jobs. This led to the need for a broader professional discussion and the design of a system that would establish a sustainable balance between quantitative and qualitative criteria for evaluating the quality of scientific research.

Investigating the possibilities of a balanced approach, the paper first starts from the analysis of quantitative criteria, which includes the advantages and disadvantages of citation databases and especially the advantages and disadvantages of the impact factor - IF as a primary quality of the publication indicator and through it, the quality of the paper/author published in it.

This, of course, includes the analysis of the criteria and policies of classifying publications into quartiles that have similar shortcomings, since it evaluates/establishes the ranking position of the journal, and is not assessing the article itself.

Then, two groups of quality criteria are introduced into the analysis. The first one refers to the categorization of the paper/article according to an elaborated criterion of authenticity. The second group deals with the quality of individual contributions of authors in collective works. Finally, the integration of qualitative and quantitative criteria forms the appropriate model formulation outline. Its goal is the objectification of the peer review of the quality of both the results of the presented scientific re-

search and the quality of the contribution of an individual researcher.

2. QUANTITATIVE CRITERIA DETERMINED BY THE PUBLICATION PLACEMENT– M

2.1. The citation base quality

The process of scientific contribution quantification normally begins with indexing in international citation databases of recognized/established excellence. The indisputable benefit of the quantitative criterion of indexing scientific information in the form of citation databases, and the explosive growth of both their number and content in each of them, points to the need to carefully distinguish between their envisioned and real role. Three reasons are usually given for their growth in the last 60 or so years.

The first was the need for a faster, easier and more cost-efficient model of access to the enormous growth of scientific information, as a consequence of the significant increase in investment in scientific research after World War II. The previous system of manual subject/author cataloging of research papers "sunk" under the growing influx of published papers.

The second was the researchers' dissatisfaction with the speed or "delay" in the subject/author works cataloging, as well as the terminological inconsistency or the incomprehensibleness of the terms used by researchers from different fields and disciplines.

The third is informatization with the enormous capabilities of computers in generating and processing a large amount of data (Clarivate, 2015).

2.1.1. Advantages

Journals/publications citation databases systematically and in an organized manner monitor research publication, record the scientific periodicals regularity, monitor the quality of the selected editors, and establish the standards of the review process. The criteria for including journals in a certain database include, among other things, the frequency of citations, citations by authors and editors, editorial integrity and expert review of their research profile, and the inclusion of as many international journals as possible in order to balance the representation of authors of different nationalities. By developing, using and improving a number of quantitative indicators, citation databases provide a respectable basis for evaluating the quality of the publications included in them. In this sense, citation databases, especially Wos and Scopus and more recently Google Scholar, along with specialist databases, constitute a necessary/desirable element in compiling evaluation criteria and encouraging scientific excellence.

2.1.2. Disadvantages

The shortcomings of citation databases as criteria for research excellence assessment (especially if they are exclusive criteria) could be classified into three main groups.

The first group of criticisms refers to the exclusivity of the peer review procedure, because the entire process of evaluation, calculation of citations and ranking takes place exclusively between publications within the database in the case of Wos and Scopus, that is, they are closed evaluation systems that exclude "third" publications and the works and authors cited in them. It is therefore not surprising that the biggest controversy, regarding researcher's hiring or obtaining project funding, are scientific excellence criteria depending on the "peer review" quality assessment within the "closed" database such as Wos and Scopus, as the key to the entire process.

Proponents of the Wos and Scopus exclusivity in the closedness/selectivity of these databases (which employ numerous expert teams), see precisely a guarantee of the quality of the peer review process, disputing the quality of Google Scholar as an open database. Proponents of open databases, especially Google Scholar (while acknowledging its imperfections) emphasize its "democracy" and accessibility. In the early stages of Google Scholar's entry into the "citation arena", conceptual differences are noticeable. The similar observation could be applied to specialized web sites like RePEc in eco-

nomics (Baum, K., Zimmerman, Ch., 2023). Wos mainly registers citations of scientific papers in peer-reviewed journals indexed in "its" three (SCI; SSCI and AHCI) ISI databases, which is its main complaint besides not registering citations in scientific books and in languages other than English.

Some of these shortcomings were intended to be dealt with by introducing an additional index: ESCI (Emerging Sources Citation Index) established in 2015 by Thomson Reuters. Initially it encompassed around 3000 journals in the process of editorial evaluation as a (pre)condition for listing in Web of Science citation bases. Since 2017 the index has been produced by Clarivate Analytics (Editor Resources 2023).

In addition, Clarivate also introduced *The Book Citation Index* as an addition to Core Collection with more than 60 000 library books indexed (61% in Social Sciences + Arts & Humanities and 39% in Natural Sciences) with a promise of expanding the index by an additional 10 000 new titles each year (Clarivate 2020).

The Book Citation Index has been immediately recognized as an important means of presentation of those scholarly disciplines for which books are major research products. However, not neglecting the benefits of the Book index to the individual scholars and book publishers, there has been warnings issued. Based on an empirical study for the discipline of political science it claims that the Book Citation Index is not well structured and not well marketed to the essential actors in the publishing industry. It has been mainly so because there has been "no systematic evidence on how indices were constructed or on how comprehensively they capture important scholarly work in books" (Hill, K. Q. & Hurley, P. A., (2022)

Furthermore, some of the journals indexed in Wos citation bases, used to publish papers simultaneously in English and other languages.

Nevertheless, these shortcomings and endeavors to overcome them notwithstanding, and returning to the citation practice, different policies could be observed. Wos transparently lists the publications/journals whose citations it monitors, and the number of works published in them. Scopus does the same. On the other hand, Google Scholar does not list journals/publications from which it registers citations, but also records citations in conference proceedings, doctoral dissertations, workshops, on various websites and the like. This has a different effect on author citations in different disciplines. For example, a comparison of two similar scientific fields; "traditional" library sciences and "modern" information sciences have significantly different citations. An author in library science has twice as many citations in Wos as in Google, while the citation of an author in information science is about the same. At the same time, about 75% of the citations of "librarians" in Google come from peer-reviewed/indexed publications, while 48% of citations for "computer scientists" are peer-reviewed/indexed publications. The most important thing here is the realization that WoS should by no means be the base used in assessing the researcher's output quality, because the "integrated" use of all three bases results in almost 25% higher citations of "librarians" and nearly 40% more citations of "information" scientists (Yang & Meho, 2007). The analysis made for environmental sciences, employing numerous criteria, shows the advantage of WoS or Scopus depending on the level of data aggregation, both indicating that Google can be useful, but only as secondary and/or auxiliary means of scientific quality assessment (Adriaanse & Rensleigh, 2011). A similar conclusion recommending the joint use of alternative subscription (WoS and Scopus) and free (Google Scholar) bases is also reached based on the analysis of 15 journals in the field of economics and business economics (Levine-Clark & Gil, 2008). Google Scholar, by indexing journals outside the English-speaking area, has a unique advantage in showing the real international/global impact of research results (large part of the world is outside the English-speaking area) (Meho & Yang, 2006). Furthermore, beside conferences presentations/papers, it also includes various documents from different (web – online) sources in citation reports. That is also its main weakness, because it includes cites of non-reviewed sources that raise the issue of citation quality. Despite the numerous improvements, revisions and additions made by the creators of all three platforms, the main characteristics or advan-

tages/disadvantages resulting from different conceptual approaches are still valid today. Wos and Scopus are still commercial/subscription databases, which means that access to the data is paid for. Google is free and available to anyone with an internet connection. All three databases collect and index scientific literature in such a way as to indicate the author, title of the work, source (journal, book, conference proceedings, etc.), summary and citations (references cited in the work as well as articles that cite the work/author, i.e. IF and h -index). No database provides full text, but various users, especially libraries, usually provide links to repositories with full texts. The essential difference remains the same. In Wos and Scopus, editors select and determine which journals, books, conference proceedings and other materials will enter the citation base (editors define - citable items). Google Scholar includes everything that can be found through an automated process (software). Therefore, Google Scholar registers far more citations, but at the same time it is far less structured and therefore more prone to errors. Furthermore, it is less selective regarding "gray" literature, presentations, websites, and the like, but it has an advantage regarding non-English literature. Due to the significantly different coverage in citation/indexing, the claim regarding the usefulness and necessity of using all three databases in the assessment of the quality of research input is still valid (Unibiz News, 2020).

Integrating the indicators of all three citation bases into a sustainable and operable process of evaluating scientific excellence is not only a complex, but also a sensitive procedure. It always has a scientific work/article or researcher/scientist as its origin. The publicly presented result of scientific research first of all proves its quality through a peer-review process, in which qualified experts determine the methodological validity and originality of the research through a double-blind review process (the reviewers do not know who the author of the text is, and the author does not know who the reviewers are), i.e. its suitability for public disclosure. All subsequent procedures for evaluating scientific excellence depend on this "beginning". The selection process for research and/or academic position and the related employment is based on the number and quality of published papers. The selection and funding of scientific research projects depends on the number and quality of the published works of the leader and the project team members, in the research topic of the project. The (re)accreditation evaluation of scientific research/higher education institutions depends on the (total) number and quality of its employees works. The sensitivity stems from the fact that each "step" in the quality assessment process is exposed to a greater or lesser possibility of arbitrariness. Who initiates and finances a scientific publication? Does (and how) the publisher/financier influence the editorial policy? How does the editorial board/editor select topics that have publication priority? How does it select reviewers? If the text "sees the light" of (publication) day, a relatively unclear process of measuring both its impact on the specific profession and its (broader) social relevance begins. What is the real quality of the publication that published it, where quality is determined by the number of its mentions in other publications? How many times was the article "clicked"/"read"? How many times was it mentioned? Who and in which publication mentioned it? Who actually read it, analyzed/criticized it? Then there is the question of where (in which publication) did she/he quote it? Of course, the impact factor is not a satisfactory synthetic answer to these and many other questions. That is why, depending on research specifics, number of researchers, social priorities, the public focus towards certain research fields and disciplines, innovative models of "summary" assessment integrating the results of all databases should be found. In doing so, it is possible, depending on the purpose of research excellence assessment (individual promotion/employment, project allocation, institution ranking), in certain cases to give priority to the base that (under equal conditions) gives the most favorable result for the one who/what is the subject of the evaluation.

The second group of criticisms refers to "monetary discrimination", which in principle takes place through two instruments: "subscriber exclusivity" and "reverse compensation".

"Subscriber exclusivity" is particularly cherished by the publicly recognized databases such as Wos and Scopus. These are "closed" private

databases with limited/paid access/membership. By fee paying the user (institution, library, researcher) has the possibility of viewing complete works (not just summaries), their citations, as well as the possibility of indirectly evaluating the relevance of the topic/institution/author. Although, for example, Elsevier claims that 600 of its journals have completely open access (Elsevier, 2022), this does not diminish, and especially does not solve, the problem.

"Reverse compensation" that often characterizes "predatory" publishing is a feature of the editorial policy of journals represented in closed databases as a consequence of the "marketability" of works/authors who publish in journals represented in citation databases. Namely, the high citation of the work and/or the high impact factor - IF of the journal in which the work was published are more and more often the "entry" criteria when applying for various research projects, for research/academic positions, and even when, as possible editor, one is applying for financing new journals. Therefore, "reverse compensation" has become more or less standard practice. Instead of an usual fee for the completed research and quality presentation of the results, the authors often have to; a) pay for the anonymous review process without refund in the case of a negative review, b) bear the costs of publishing the work in electronic and/or printed form, c) pay high fees for the availability of the complete published work in a citation database journal in "open access" that enables citation .

The third group of criticisms refers to "arbitrary selectivity". There are different procedures and criteria for the availability of individual complete texts published in journals represented in the citation database, and thus (indirectly) for potential citations. The whole procedure is established and conducted by professional and/or administrative-technical staff of the base owner. Therefore, in addition to the "open" access fee, there could be different "supplement" criteria employed such as: the attractiveness of the topic and the interest to the public, the relatively high impact of "controversial" topics attracting more readers outside the narrow circle of specialists that often lead to different availability levels, especially of complete texts.

2.2. The publication quality

In recent decades, the Impact Factor – IF, has developed into a fundamental indicator of citation/visibility of scientific publications/journals, and thus indirectly of the quality of all works published in them and their authors. IF is usually calculated as the citation of papers from the previous 2 years (e.g. 2020 and 2021) in a given year (2022)

The Impact Factor standard formula is calculated as;

$$2022IF_x = \frac{A_{x2020,2021}}{B_{x2020,2021}}$$; where;

- 20220IF$_x$ –Impact factor of the journal's x in 2022 year, x = citation rate of journal x in all journals included in the associated citation base,
- $A_{x2020,2021}$ = number of citations of works published in the journal x in the previous 2 years (2020, 2021) in all other journals of the citation base,
- $B_{x2020,2021}$ = total number of papers published in the journal x in the previous 2 years (2020,2021) considered as citable items.

For example, $2022IF_x = 5$ means that (all) works published in the journal x (in 2020, 2021) were cited 5 times on average in the journals of the citation base in 2022. It should be kept in mind that this is a relative indicator which means that different "combinations" of numerator and denominator can give the same - IF. Thus, a "theoretical" journal x, whose total of 50 published papers were cited 250 times, has an IF = 5, as does an "applied" journal y, whose 200 published papers have been cited 1000 times.

Journals in the given database are then ranked according to the IF value. Theoretically, IF can take on infinite values, say from 0.01 (100 published papers that are cited a total of 1 time) to 100 (100 published papers that are cited a total of 10,000 times) in each observed year.

After that, the list of all cited publications according to IF rank in the database is divided into

quartiles (Q1, Q2, Q3, Q4) or "quarters". First quartile (Q1) aligns 25 percent (1/4) of journals/publications that published the most cited papers on average. In the second quartile (Q2) are the next 25 percent of journals ranked by average number of citations per published paper, and so on.

There are also publications in the databases that are not cited anywhere else ("new" emerging that have just started, "old" vanishing that are losing "momentum") and that are periodically "cleaned" or "delisted" from the individual database based on these and other criteria. In the case of "delisting", it seems advisable to stick to "historical" data at the time of publication ranking, which is more correct for the authors, who could not have known in any way that a journal that was, for example, in the first quartile at the time of publication of their article in 2017 year, is going to be delisted from the database in 2022.

2.2.1. Advantages

IF is a relatively simple and verifiable indicator that provides the first "objective" insight into the quality of the publication publishing the paper, in which the candidate whose research excellence has been evaluated, is a (co)author.

The citation rate of a publication/journal relative to thousands of publications selected in a certain database, and especially relative to the subject category in which the journal is located (Subject Category List), in principal is a good quality indicator. It is based on (anonymous) review process conducted by qualified experts specialized in a particular research topic (peer review). Upon this professionally qualified procedure relevant papers are published. Formatted in this way the peer review has a significant effect on the public interest that individual (and all published in that publication) papers actuate which is circumstantially evident from the citation rate (IF) of the journal in which it was published, and indirectly by its impact on related scientific research.

IF is widely used in university ranking, in criteria for individual headway at numerous world universities, criteria and tenders for competitive projects and centers of excellence in various EU funds and programs, especially in natural and (bio)technical sciences and biomedicine. A special example is Sweden and China, which take an IF greater than 15 as a special "bonus" criterion in the allocation of financial resources for research projects. Finally, distinguished members of the scientific community known by their very critical opinion on the exclusiveness of the IF recognize it as a useful, but not decisive tool in a process of integrated research quality assessment (Hicks et al., 2015).

2.2.2. Disadvantages

The fundamental shortcoming stems from the fact that IF is primarily intended for ranking journals, i.e. determining/stimulating the scholarly journals quality growth, their promotion/"-profitability", not institutions or individuals. At the same time, this makes it insufficiently suitable for quality ranking individual papers and even more unsuitable for assessing the research quality of individual author(s) during the selection process for academic positions or during research projects application and/or allocation process.

Since the publication citation/quality is measured, it can also be achieved by high citation of a smaller number of works. An extreme example of high citations is the journal Nature, which for example had 89 percent of its IF in 2005, was generated by citations of only 25 percent of the articles published in it (Van Noorden, 2010).

Translated to the encouraging excellence criteria and by them indirectly defining most of the academic positions selection process in terms of citations, one should bear in mind that looking for a high IF (Q1, Q2) simply means that the paper was published in that year of the journal that is ranked in the upper quartiles. Therefore, a research quality assessment process defined in this way classifies it more as a "lottery effect" and less as a qualified scientific opinion. As a rule, this does not provide any insight into the attached work quality itself, and even less insight into the quality of the candidate's own research contribution in that work.

Focusing on citation in the editorial policy of the journal/publication encourages "creative" citation methods in "increasing" the numerator and "reducing" the denominator:

- "stimulating" the numerator by publishing a larger number of review papers instead of original research. Review papers can generate more interest because of the broad insight into the topic and eventual controversy, therefore providing researchers with the necessary bibliography for their own papers. For example, in 12 indexed American pathology journals in the period 1991-2006, the number of original scientific papers published annually increased by 2-3 times, while the number of review papers increased by 5-6 times (Ketcham & Crawford, 2007).
- "stimulating" the numerator by publishing a larger number of works by older and/or better-known authors whose relevance "automatically" causes greater interest/citations, and/or by publishing "trendypapers", i.e. highlighted topics currently in the professional/public attention focus (Abbott, Cyranoski, Jones, Maher, Schiermeier, & Van Noorden, 2010),
- "stimulating" the numerator by publishing a larger number "citable items" (shorter works, executive summaries, editorial comments, notes and reviews) at the beginning of the year, thus opening a wider time horizon for citing individual works which "blows up" the IF as in the case of the British Journal of Sports medicine (Himmelfarb Library News, 2022),
- "stimulating" the numerator by "coerced" citation, that is "justified by the editorial request to cite certain authors, as observed in the case of 20% of published texts in social sciences (Van Noorden, 2010),
- "stimulating" the numerator by creating "citation cartels", that is, connecting authors in specific areas in a network of mutual "conscious and intentional" citation, regardless of the "cited colleague" paper real relevance for one's own research. There is a well-known case of the "medical citation cartel" in Wos's JCR, in which three medical journals participated; Cell Transplantation, Medical Science Monitor and The Scientific Science Monitor. Cell Transplantation journal's IF jumped in the period 2006-2010. from 3,482 to 6,204. A subsequent investigation by the then owner of JCR (Reuters Thomson) showed that in 2010, a review article was published in Medical Science Monitor that cited 490 articles, of which 445 were published in Cell Transplantation. Three of the four authors of the article were members of the Cell Transplantation editorial board, two were executive editors, and one was a founding editor. In the same year, two of those four authors published a review article in The Scientific World Journal. It cited 126 articles, 96 of which were published in Cell Transplantation, also in 2008 and 2009. What triggered the investigation, which ended in the expulsion of the journal from the JCR, was the fact that until 2010 neither of these two journals had cited Cell Transplantation, precisely the year IF was calculated for, on the basis of citations in the previous two years - 2008. and in 2009 (Oransky & Marcus, 2017),
- "stimulating" the numerator with "overlapping" members of the journal's editorial boards who mutually cite "their" journals in their texts in "friendly" publications. At the same time, the possibility that such texts bypass the standard peer review procedure and be published as "editorial material" is particularly worrying. Although the JRC owner at the time did not have an algorithm for identifying citation cartels, as it had for detecting self-citations, Thomson Reuters, based on its own investigation, excluded and canceled the IF for six journals in the field of business economics, in addition to these three medical journals (Davis, 2012),
- "destimulating" the denominator by rejecting articles with expected lower citations due to the editor's professional/subjective assessment, evolving in discretely "extended" review time for researchers engaged in less interesting topics and/or located in less interesting regions. Since the publishing is often official requirement keeping/improving academic position, they are necessarily "stimulated" to focus on lower-ranked publications (Dahal, Shrestha, Singh, Joshi, Sitaula, Sharma, & Chalise, 2022),
- "destimulating" the denominator by modifying it so that the "abstracts" at the beginning and the "bibliography" at the end of a

certain number of papers are omitted so that (those "additions") are not part of the "citable items" (Van Noorden, 2010).

Naturally, this kind of orientation also results in the "adaptation" of the author. Namely, IF is based on the "peer (expert) review" procedure, or more precisely, it is based on the works that have been reviewed. Highly expert reviewers for certain topics make it possible to publish quality papers in the journal, which will attract scholarly attention, i.e. citations, which places the journal in the "upper quartiles", as a proof of its high quality. In the desire for as many papers as possible to be of high quality, journals try to define "benchmark", i.e. the standards that the texts must meet in order to be potentially highly cited, and therefore eligible for publication. Such text's "normalization/standardization" is indicated to potential authors in the form of "technical norms" representing the text required technical format and table of content, and to reviewers in the form of an elaborate review procedure.

The perceived "accounting" review procedure and "standardization" shortcomings, especially in the field of sociology, were analytically interpreted (paraphrasing the famous German sociologist A. Munch) by M. Petrić (Petrić, 2013). Among the numerous elements of this critical discourse, it is worth highlighting two, relatively closely related to some aspects of research excellence criteria, which is particularly important for younger scholars.

The first is the rapid young researchers' adaptation to the reviewers' expectations by concentrating on "main stream" topics, i.e. "following the beaten track". Most of the time is "spent" on the construction of a large database suitable to be "sliced" in compact statistical patterns eligible for the future numerous variants of (slightly) different quantitative (econometric) analysis and interpretations, resulting in equally numerous papers properly formatted for publication in highly cited journals. Relevant scholarly questions thus become a secondary motive for research. The features of the collected data come to the fore dictating the research questions or scientific topic, which is then impeccably processed using sophisticated methodological procedures. Potential "publishability" or "citation potential" by itself, directs younger scholars to such a "normalized" approach of avoiding the risks of "out-of-date" topics, outside the "main-stream" and outside "real science". Strongly suggesting to young scientists to avoid "exiting" the boundaries of the defined thematic and disciplinary framework, creates so-called "normalized" knowledge, discouraging theoretical and research curiosity. Adherence to "prescribed" forms and "generally accepted methods" imposing "normalized standards" can lead authors to publish professional instead of scholarly papers (Petrić, 2013).

The second is the "forced" adaptation to the unethical "gaming of the system, that is, the maximization of the publishable/citation effect of works, which has several forms. The "salami slice" technique is well-known, when theoretically and empirically rounded research is "cut into slices" or "artificially" divided into several articles. Then there is the creation of "citation cartels", i.e. a network of authors who quote each other according to the "returned favors" principle. The system does not exclude even strategic (of established authors and/or potential reviewers) citation, occasionally ending up in direct fraud and falsification of results (Petrić, 2013).

Finally, numerous "specific" shortcomings were also observed, such as; lack of metrics for "non-English" literature, insufficient public exposure and insufficient transparency of methodology, criteria and indicators used by database owners, regional/continental imbalance of citation databases, and the like.

2.2.3. Efforts to overcome shortcomings

It should be noted that bibliometrics invests great efforts in improving the quality and informative reach of indicators by redesigning existing and creating new indicators. Thus, for example, the perceived indicator invalidity due to the neglect of the certain research fields and topics specifics is attempted to be corrected by creating the aforementioned journals "Subject Category List". Citations (IF and Q 1-4) are thus calculated separately for each category - scientific field.

The owner of Wos does this through the Scientific Journal Impact Factor - JIF (Journal Citation

Report) as part of the "Journal Citation Report". There are different journals in the natural sciences (JCR Science edition), the citations of which are tracked by the Science Citation Index - SCI. Journals in the field of social sciences are classified in a separate Social Science Edition– JCR, whose citations are recorded by the Social Science Citation Index – SSCI. It is similar in the field of humanities whose citations are recorded by the Arts&Humanities Citation Index – AHCI.

Conceptually, the same approach is being developed by Elsevier through the Subject Journal Report (SJR).

The indisputable and constant improvements of both databases and their indicators enabled the calculation of the Aggregate Impact Factor for Subject Category - AIFSC (Aggregate Impact Factor for Subject Category).

The "slowness", i.e. the time gap between the paper's publication date and the "indexation" date of the journal, is attempted to being reduced by the introduction of the " Imedeacy Index ". It measures the number of citations that paper(s) published in the journal have in the current year in relation to the total number of published papers in the current year.

"Obsolescence", or the objection that earlier published works accumulate citations over a longer period of time and therefore are cited more, is being corrected by the introduction of the "Cited Half-life Index". This is the median age of articles cited in JCR or SJR. For example, if the Cited Half-life of a journal in 2020 = 5, this means that half of the citations are from the period of the last 5 years (2015-2020) and the other half are from earlier periods.

The "comparability" of citations between and within individual fields of science is being improved by giving greater weight to works published and cited in journals with a higher IF, so JCR Wos introduces Eigenfactor, and SJR or Elsevier introduces Scimago.

In overcoming the shortcomings of quantitative criteria, a special effort will require (non) comparability within/between certain scientific areas, fields and disciplines. For example, the quality of a Q1 journal in theoretical physics does not have to be the same as a Q1 journal in clinical medicine, that is, the same impact factor in different subject categories does not have to be of the same "weight". This leads to the need of quality assessment (and competitiveness) within individual scientific areas or subject categories, i.e. the avoidance of comparisons between different areas, fields or disciplines, including the comparison between different subject categories.

The current widespread (despite numerous criticisms and "benefits of doubt") triple function of quantitative criteria for research excellence evaluation in;

- assessing the quality of the individual scientist's research output,
- assessing the quality of research projects funding applications,
- evaluating the quality of research and higher education institutions,

asks for careful additional addressing the still open questions that might have risen even after efforts to overcome the perceived shortcomings.

3. QUALITATIVE CRITERIA DETERMINED BY QUALITY OF WORK - Vr

Analyzed "pros&cons" of quantitative criteria is a good starting point emphasizing the need for qualitative criteria introducement. Bearing that in mind, the next step should be evaluation process enrichment by supplementing with quality specific criteria. In this sense, standard scholarly papers quality classification based on originality and comprehensiveness could be of significant help.

Scholar and research communities' bylaws distinguish three "regular" and one "conditional" types of scholarly and/or research papers:

"regular";

- original scientific/research paper,
- preliminary communication,
- review article,

"conditional";

- conference paper/communication

3.1. Original scholarly paper/article

An original Scientific/Research Paper is an original scientific work in which new results of fundamental and/or applied research are presented. This category of works include new, still unknown scientific facts, scientific laws, scientific regularities, scholarly knowledge and/or theories that represent a contribution to global science. The methodology and findings should be presented in a manner that any qualified researcher, based on the information provided, can: repeat the experiment and achieve the described research results with equal accuracy, or within the margins of experimental error, as defined by the author (Knjižničarstvo, 2012a). It should be noted that an original scientific paper defined in this way is more suitable for natural, medical and technical sciences, which of course does not mean that it cannot be created in social sciences and humanities as well (Zelenika, 1988). When it comes to social (and humanistic) sciences, it should be noted that the very useful overall use of mathematical-statistical methods and models in the digital environment is a particularly powerful analytical tool in examining the hypotheses and findings sustainability. At the same time, it is (only) one of the possible (equal) forms of scientific insight and findings formulation that should be considered in the assessment of ontological/logical constructs. Ontology is understood here in the modern analytical sense as a theory of general categories such as an object, trait or more precisely a fundamental philosophical discipline that studies what is and at the center of which are essential questions such as; What does it even mean to be or to exist? What is existence? What are fundamental properties of existence? What are the relationships between objects that exist? (Veljak, 2013). In a simplified way, we understand ontology as the science of the fundamental causes of everything that exists.

3.2. Preliminary Communication

A preliminary communication or scholarly note is a working paper that contains new knowledge and/or new research findings of such importance that it requires immediate publication. This kind of paper contains one or more scientific information, but without sufficient details that would allow the qualified reader to verify the presented findings, research information and results, in the manner described for the "original research paper" (Knižnjičarstvo, 2012). This type of scientific work is equally appropriate/applicable in all scientific fields (Zelenika, 1988).

3.3. Review article

A review article is a research work containing an extensive and complete overview of so far compiled published information about specific research problem investigations already conducted by other authors. However, it's specifics imply that the known facts have been collected, analyzed and especially synthesized in a novel way and presented from a fresh critical point of view. It differs from the original scientific article in that it does not have to contain original (new) research results. The author is obliged to list all bibliographic items used in the topic investigation (Sveučilište u Zadru, 2014). Due to the comprehensive synthesizing power stemming from the systematic and critical analysis of one's own and others' research, this type of work could be interesting not only to the scholars directly engaged in the given research subject of research, but to those from related fields as well (Zelenika, 1988). It is also important to note that a review paper/article can be categorized as both scholarly and professional, depending on the degree of innovation that the author brings to the interpretation of known findings, i.e. it is possible that it generates new hypotheses that lead to new research (Žugaj, Dumičić, & Dušak, 1999).

It is certainly worth pointing out the specific connection between review articles and citations in connection to IF. Namely, papers that have an elaborate overview of the topic they deal with, citing numerous authors, their contribution and shortcomings who have dealt with a similar topic in the previous period, have a special interpretive power. Every researcher who is interested in that topic must, in order for his findings to be relevant process history, that is, to critically analyze the majority of (past) works related to it. Thus, review papers, as a rule, rich in literature sources, are regularly/necessarily

cited by new researchers, achieving in that manner citations that can be significantly higher than the original scholar/research paper.

3.4. Conference Paper/Communication

A Conference Paper/Communication presented at an academic conference can be classified as a scholarly paper in any of the previous 3 categories depending on the peer review and/or professional panel/review outcome at the conference, and must be printed as a full paper in conference proceedings. In other cases, it is a professional conference work (Zelenika, 1988). Contemporary practice and experiences gained in the organization and participation of Croatian researchers at international conferences have resulted in heterogeneous treatment of full papers there published, and sometimes to their practical "exclusion" from the research excellence quality assessment. On the other hand, the scholarly and professional benefit that such gatherings bring in strengthening the international visibility of Croatian scientists, networking in relevant research topics/projects, and establishment of communication channels and exchange of experiences is unquestionable. The professional attention that the presented works receive, as a rule, has a positive effect on improving the quality of the final version of the work. Therefore, it would be rational for the peer reviewed fully published conference paper to be classified as a Preliminary Communication

A professional paper published in a scholarly journal or a professional monographic paper with partial scholar characteristics. In special cases that define scientific fields, especially professional work published in a highly cited scholarly journal, can be scored with 0.5 points in the corresponding category of quantitative criteria - M, which are described in more detail below. A monographic work with partial scholarly characteristics (for example dictionaries) can be treated in a similar way compared to academic books.

3.5. The role of the peer review (hiring) committee

In special cases the peer review (hiring) committee determines the quality/category of a candidate's scholarly work through direct analysis of published items.

It does so compulsorily in the case when the publication/magazine does not categorize the published papers, or does so in a different way, from the above presented categorization.

But it also does so in the case when the publication determines the category of scholarly work in one of the following ways:

a) can accept the categorization (original scholarly paper, preliminary communication, review article, and article in a group of mixed scholarly and professional papers) as specified by the publication in which the work was published and
b) can, (based on professionally qualified explanation) as an expert group, establish its own categorization different from the one specified by the publication in which the work was published.

The analysis and proposals in this paper are based on the " two-stage" peer review system assessing the quality of the overall individual researcher scientific profile.

At the first stage the expert committee appointed by the institution (university, faculty, institute) itself (with a mandatory at least one outside member), compiles the written report assessing the research excellence of the candidate's published scholarly output, recommending her/his hiring/promoting or non-hiring/demoting.

At the second stage, or round two, an outside peer-review board of the highest expert level (in the case of Croatia, the Field Board composed of 9-11 scholars with the highest ranking in a particular research field) analyzes the first stage report and makes a final assessment of a candidate's research output with authority to confirm or discard the first stage report. The Field Board also has the authority to reclassify a paper's category made at the first stage. That is, among other things, an additional instrument of independent checking and promoting both research austerity and research excellence of individual published work as well as an overall scholarly profile of the candidate.

4. QUALITATIVE CRITERIA DETERMINED BY THE QUALITY OF THE AUTHOR'S INDIVIDUAL CONTRIBUTION - Va

There are three main purposes of research excellence assessment. It could/would be incorporated in the procedures of academic hiring/promoting, or getting research projects, or it could be part of institution's' accreditation process. The so far established criteria for research excellence assessment, was in essence acquired indirectly thorough - "dual intermediation".

In the first level of "intermediation", as we saw earlier, the procedure was essentially based on the quality of the publication (its Impact factor - IF), claiming that a highly cited publication, with its thus determined quality, implies that the work published in it will "automatically " be of high quality. The "indirect" determination of quality here refers to the fact that the paper of the respective author itself did not have to have a single citation, but it was enough that the issue (year) of the journal to be highly ranked based on the total citations of all published works in it.

The second level of "intermediation" stems from the fact that once the quality of the work was "indirectly" established, the contribution of an individual author (in collective works) was thereby "automatically implied". Therefore, the research excellence of the author (in a collective work) is "indirectly" determined by the quality of the published paper, which in turn is itself "indirectly" determined by the quality (number of scholarly citations - IF) of the publication in which the paper was published.

4.1. Research peculiarities

The character of research in particular sciences, as well as the methods and models of research results preparation and presentation (mainly in English) resulted in the dominance of collective/joint authorship. More and more frequent and desirable international research collaboration additionally affects the need for authors to collaborate.

Therefore, it is necessary to respect the fact that different scientific disciplines and mutual collaboration of researchers from nominally different fields requires joint research efforts, which result in joint works/publications. The fact that the different theoretical and experimental focus affects the different form of presentation of scientific papers in certain sciences is also clear. In this sense, two determinants should be emphasized.

First, it is clear that several authors can give equal and full author's input to a published work, considering the complexity of individual research topics. But there is also a finite number of authors whose scholarly/research (not technical) input can be identified in a single paper/article. The fundamental question is what is the maximum number of authors who can be given 100 percent contribution in one paper, where the current practice records 3-5 (co)authors depending on the discipline/research field. There is also a different question of how many authors (K) and with what participation percentage is it possible to have in order to be able to apply the contribution formula 100: K and get an individual share. Scientific fields should argue expert positions, regarding the discipline specific maximum number of recognizable individual inputs and their relative shares.

The second are single-author scholar/research works. The demonstrated ability to independently conceive and realize scientific research of the identified phenomenon is reflected differently in different fields and disciplines. In more theoretically oriented research, possibly, it is easier to assess the overall individual input of the author. This is particularly (but not exclusively) true for the social sciences and humanities, where the ability to conduct independent scientific research can be an important factor in the scholarly profile of an individual researcher. On the other hand, especially (but not exclusively) in natural, biomedical and technical sciences, focused on experiments and identifying patterns based on extensive observations and complex analytical procedures and layered synthesis, the emphasis is on collective works. Realistically defining and measuring the necessary independence and desirable cooperation in collective research is the task of peer-review

experts in individual scientific areas, fields, branches and disciplines.

Altogether, it creates an additional challenge in ascertaining the excellence of a candidate competing for an academic and/or research position, or competing as a head researcher of research grants and projects. Facilitating an objective and valid assessment of the quality and intensity of a candidate's scholar/research output in a dominant co-author environment is one of the fundamental tasks of establishing excellence assessment criteria. It is advisable to start from the key elements of possible author participation, i.e. research input in joint work.

4.2. Identification of individual author's contribution category

In understanding and qualifying individual contribution in collective work, it is possible to distinguish several categories/levels of composition and presentation of scholarly research:

i. Conceptual-methodological level/category
- Development of the idea and elaboration of the research concept
- Defining the form and selection/formulation of the research method

ii. Analytical – content level/category
- Critical analysis/review of the research subject available studies
- Collection and processing (qualitative, quantitative, model) of available documentation and empirical data
- Analysis/discussion and interpretation of research results

iii. Synthesizing level/category of consistent and rounded up formulation of the topic
- Designing and writing a scholar/research paper
- Critical revision, verification and confirmation/approval of the work basic findings and the form of public presentation

Co-authorship in a scientific work would in principle imply the possibility of identifying the contribution of an individual author in at least one element in each of the three listed categories/levels. However, the specifics of modern research in certain fields and disciplines, and the exponential growth of knowledge and research findings, point to the need and possibility of a closer definition of the minimal input that qualifies an individual researcher as a co-author of a scholarly work, within the above levels.

One of the possibilities of qualifying an individual research input in joint works is as follows:

- maximum scholarly/research contribution – the author achieves this if it is possible to separate and identify his contribution in all three levels (conceptual-methodological, analytical-content and synthesizing),
- partial scholarly/research contribution - this is achieved by the author if it is possible to separate and identify his contribution in at least two (out of three) levels of research/scholar input,
- minimal scholarly/research contribution - the author achieves this if it is possible to separate and identify his contribution in at least one (of three) levels of scholar/research input.

Additionally, the mere collection and processing of data without participating in their analysis/interpretation is usually not enough to recognize a scholarly/research contribution. The possible exceptions here depend on the specifics of a particular scientific field, type and number of experiments needed, the possibly required additional knowledge in the assessment of relevance, and sample collection hindrance, i.e. the selection and data processing complexity important for the quality of research work.

5. INTEGRATED STIMULATIVE QUALITY ASSESSMENT CRITERIA DETERMINATION - Ui

It is clear so far that (with all limitations), the integrated criteria on which scholarly/research excellence is based on must be a sustainable combination of quantitative and qualitative criteria in a clear functional relationship. It should

be applied in thorough their concretization in the selecting/hiring procedures for academic/research positions, in quality assessment of research projects propos work Ni established als, and in institutions quality assessment during (re)accreditation process.

5.1. Criteria functional interconnection

The quantitative criteria interrelationship defined by the placement (M) of the published work -*i* that is determined for each category of work (N) depending on the impact factor (IF), and the qualitative criteria of work excellence (VR) and the quality of the individual author's contribution in the work (VA), can be displayed in the following manner:

$$U_i^n = m_i^n vr_i^l va_i^k \quad (1)$$

In expression (1), U_i^n represents the total number of points that the author achieves on the basis of a particular work that belongs to one of the three main categories of the quantitative contribution of the work in the evaluation of its scientific excellence. The three main categories of work are; $N = (I^j, II^j, III)$, whereby the first and second categories of work are divided into two subcategories, i.e. j=(1,2).

Expression (1) is a "multiplicative" form of representation of the functional relationship, which calculates the number of points of a particular paper as a product of three coefficients. This format can be appropriate for scientific niches in which there are a large number of indexed journals, so a relatively large point difference is intended to stimulate publication in journals in higher (especially Q1) quartiles. For example, in the related categories Economics and Econometrics (Scimago Journal & Country Rank, 2023a), and Economics, Econometrics and Finance (Scimago Journal & Country Rank, 2023b), WoS indexes a total of 732 journals, of which 228 were in Q1. On the other hand, in the category Acoustics and Ultrasound, Wos indexes 31 journals in all four quartiles, of which 10 are in Q1 (Scimago Journal & Country Rank, 2023c). In such small scientific niches, it is much more difficult to publish a paper in a Q1 journal, so "additive" scoring, in which the number of points is obtained by adding three coefficients, would probably be more appropriate.

In the final selection of the functional form, depending on the specifics of individual areas, branches or scientific disciplines, three things are always important. First, the quantitative contribution of M (however defined) determines the category of work. Second, whatever functional connection is established (multiplicative, additive), it must include all three contributions, whereby the qualitative individual input is clearly recognized, not only in single-author but also in collective/team work. Third, the values (points) of both coefficients of the qualitative contribution va and vr have the same values and the same mutual relationships for all categories and subcategories of work N_i established on the basis of the quantitative contribution M.

The coefficient m_i^n represents the coefficient of **the quantitative share** *(M)* in research quality assessment of the - n^{th} category paper. Based on the quantitative contribution, the category of work for each area (larger field) is defined. It is most often determined by the impact factor IF, then by citations, ranking of journals in relevant databases, h-index or some other (aggregate) measure or their combination.

The coefficient vr_i^l represents the **coefficient of the qualitative share** *(VR)* in research quality assessment of the - n^{th} category paper.

The coefficient va_i^k represents the coefficient of the **author's individual qualitative share** *(VA)* in research quality assessment of the - n^{th} category paper.

The coefficient of quantitative contribution in the evaluation of the scholar paper excellence can take the following values depending on the category of the work:

$m^{I^1} = 3$ for **I category** papers defined by a high Impact factor IF (Q1,) and it can also be defined by high citation score, journal indexing in WoS/Scopus and/or relevant specialist databases, high h-index or some other (aggregate) measure or their combination that roughly corresponds to the a1 category of work of the current national scientific excellence criteria.

m^{I^2} **= 2,5** for **I category** papers defined by a high Impact factor IF (Q2,).

m^{II^1} **= 2** for **II category** papers that defined by the Impact factor IF (Q3), or journals indexing in databases other than WoS/Scopus and may also include complete papers published in the proceedings of an international scientific conference and chapters in scholar books, roughly corresponding to the a2 category of the current national scientific excellence criteria.

m^{II^2} **= 1,5** for **II category** papers that can be defined by the Impact factor IF (Q4), and the rest equal to

m^{III} **= 1** for **III category** papers, that may include works published in national and international scholar journals that are not indexed in previously included databases, complete papers published in the proceedings of domestic scientific conferences/workshops, patents, projects or the like and roughly corresponds to a3 category of the current national scientific excellence criteria.

The coefficient of the qualitative share (vr^l) assessing the paper scientific excellence takes the following values depending on the qualitative category of work - *i*, where L = (I, II, III), that is:

vr^I **= 2** for an original scholar/research paper that is considered to be the I category of qualitative contribution to the paper quality,

vr^{II} **= 1,5** for the preliminary communication that is considered to be the II category of qualitative contribution to the paper/work quality, and

vr^{III} **= 1** for a review paper that is considered a category III of qualitative contribution to the paper quality.

The coefficient of the qualitative individual author's contribution to the paper excellence (va^k) depending on the number of paper authors -*k*, where-*k* can take on any reasonable positive value, while at the same time, papers up to 3 authors can take the following values:

va^I **= 2** for single author paper, i.e. the maximum qualitative individual author's contribution to the quality of the collective work is assumed,

va^{II} **= 1,5** for papers with two co-authors, i.e. a partial qualitative individual author's input to the quality of the work is assumed, and

va^{III} **= 1** for three co-authors paper, i.e. a minimal qualitative individual author's input to the quality of the work is assumed.

For four or more paper co-authors (for k > 3) applies $vr_i^k = 1/k_i$.[1]

In accordance with all of the above, **the overall number of points** achieved by an individual author on the basis of all of his scientific papers ($I = I^I + I^{II} + I^{III}$) broken down into three main categories of the quantitative contribution of the paper (N) in the assessment of scientific excellence amounts to:

$$U = \sum_{i=1}^{I^I} U_i^I + \sum_{i=1}^{I^{II}} U_i^{II} + \sum_{i=1}^{I^{III}} U_i^{III} \qquad (2)$$

In expression (2), - *U* denotes the **overall number of points** of an individual author based on all of her/his scientific papers (quantitative and qualitative) categories, while ($I = I^I + I^{II} + I^{III}$) represents the **total number of points** achieved on the basis of paper **quantitative** contribution category (I, II or III) in the scientific excellence assessment process.

[1] In this paper, the starting point is the usual evaluation of individual qualitative contribution in joint/team work, where the work of up to three authors is recognized for each author's full (one hundred percent) contribution. For papers with four or more authors, the individual contribution is determined by dividing the total maximum contribution - 1 by the number of authors. However, depending on the characteristics of research in different scientific areas/branches/disciplines, it is also possible to define (increase) the number of authors with full contribution differently. However, it is necessary that for each author it is possible to single out or attribute to her/him at least one of three types of qualitative contribution (conceptual-methodological, analytical-content, synthesizing) which is specific and differs from the inputs of other authors.

5.2. Description of the criteria functional relations

In this description of functional connections, the following assumptions are made;

1. **Primacy of quantitative criteria**
   - the category of work - U^I, U^{II} **and** U^{III} is determined by the quantitative criterion - M
   - scores in the first category - U^I can only be achieved based on the quantitative criterion M, i.e. the coefficient of the first category quantitative contribution - $\mathbf{m^1}$,
   - at least 1/3 of the total number of points required for the research excellence assessment (selection into a research -teaching position, employment, research project allocation) must be achieved in the I category (U^I)

Quantitative criteria, as can be seen, are set as a necessary but not sufficient condition. Trying to consider the observed shortcomings and reasoned criticisms, they are taken as a starting point, but not as a final assessment. We see them as a certain base - a threshold that must be crossed in order to be able to analyze the quality. Qualitative criteria start to "play/dominate" only when/after the corresponding quantitative criteria are met.

2. **Importance of qualitative criteria**
   - the meaning of scientific research is to discover and expand scholarly knowledge, and identifying the quality and originality of that work contribution to science is the meaning of the quality assessment process,
   - the categorization of the qualitative contribution to the recognized/proven original paper, the previous communication that "waits" for its recognition after the necessary refinement and additional confirmation, and the review paper that is the original "processing" of "other people's" findings, is the **first key (Vr)** to a quality and professionally qualified peer- review process related to the work itself,
   - defining the academic status or research qualification for the project's allocation and/or the employment of scholars, rests on the evaluation of individual research quality, which often needs to be discerned in predominantly team/group research. In order to achieve this, it is necessary to single out or attribute at least one of three types of qualitative input (conceptual-methodological, analytical-content, synthesizing) for each author which is specific and differs from the inputs of other authors. It is the **second key (Va)** of a quality and professionally qualified peer-review process that refers to the "allocation/attribution of credit" for the published work to its "creators",
   - citation databases, publication indexing/ranking and impact factor, do not address either the first or the second "key" for deciphering the quality of scientific research in published work, which emphasizes the importance of qualitative criteria.

The criteria defined in this way can be presented in a table (Table 1).

Table 1. Points attributed to the quantitative and qualitative criteria categories

Category of the paper Ni	Coefficients m^n	Coefficients vr^I	Coefficients va^I
I^1 I^2	3 2,5	2; 1,5; 1 2; 1,5; 1	2; 1,5; 1 2; 1,5; 1
II^1 II^2	2 1,5	2; 1,5; 1 2; 1,5; 1	2; 1,5; 1 2; 1,5; 1
III	1	2; 1,5; 1	2; 1,5; 1

The Category - I number of points (U^I) is obtained by adding up the product of category I quantitative coefficient m^{i^1} (3 points) or m^{i^2} (2.5 points)-, multiplied by the points of the corresponding category of the paper quality coefficient (vr^I) and by the points of the corresponding category of the individual author's contribution quality coefficient (va^k) of the category- I of the papers - i^I.

The Category - II number of points (U^{II}) is obtained by adding up the product of category - II

quantitative coefficient - m^{II1} (2 points) or - m^{II2} (1.5 points), and the points of the corresponding category of the paper quality coefficient (vr^l) and by the points of the corresponding category of the individual author's contribution quality coefficient (va^k) of the category- II of the papers - i^{II}.

The Category - III number of points (U^{III}) is obtained by adding up the product of category - III quantitative coefficient - m^{III} point), and the points of the corresponding category of the paper quality coefficient (vr^l) and by the points of the corresponding category of the individual author's contribution quality coefficient (va^k) of the category- III of the papers - i^{III}.

The points obtained in - I category I (U^I) can substitute the necessary points in – II category (U^{II}) and - III category III (U^{III}). The reverse is not allowed.

The points obtained in the - II category (U^{II}) can be used to replace the necessary points in the - III category (U^{III}). The reverse is not allowed.

It should be taken into account that the share of the quantitative criterion increases when the qualitative criteria are achieved in lower categories that is, the lower the VA and VR, the higher the proportion of quantitative criterion in the total number of points. This can be explained by two examples.

Let's first take a paper in the I category with 3 points according to the quantitative criterion - M, 2 points in the qualitative criterion - Vr and 2 points according to the qualitative criterion - Va, that is, with the maximum number of points in all three categories. Then their product is 12 (3x2x2=12), i.e. the relative ratio of coefficients for this paper is estimated as 40%:30%:30% = 100%. This assessment results from the mutual relative ratio of the number of points (3+2+2=7 where 3 is 40% of 7, and 2 is 30% of 7). Here we see that the quantitative criterion - M makes up 40% of the scientific excellence assessment.

Now let's take a paper in the I category with also 3 points according to the quantitative criterion - M, but 1 point in the qualitative criterion - Vr and 1 point according to the qualitative criterion - Va. Then their product - 3 (3x1x1), i.e. the relative ratio of coefficients for this paper is estimated as 60%:20%:20%. The assessment results from the mutual relative ratio of the number of points (3+1+1=5 where 3 is 0.6 or 60% of 5, and 1 is 0.2 or 20% of 5). Here we see that the quantitative criterion - M constitutes 60% of the evaluation of scientific excellence.

Since hypothetical examples are employed here, that is, arbitrary (based on experience) estimates of the relationship between quantitative and qualitative components, other approaches are also possible, for example, narrowing the point range between the coefficients or the like. Accordingly, a careful analysis of the definition of "normalized" and/or "expected/desirable" values of the coefficients and their relationships by research fields/branches/disciplines is in order.

5.3. A hypothetical example

The application of one of the possible variants, i.e. a combination of quantitative and qualitative criteria, is shown on a (hypothetical) example of social sciences (field of economics) (table 2).

As can be seen, the "multiplicative" form of the relationship between the criteria was used here, with a significant point/range difference between the categories. This was done to take into account the specifics of the Croatian definition of the scientific field of economics, which in WoS fully or partially "covered" by 17 subject categories with a total of 2,288 indexed journals, of which 866 are Q1 (see Supplement 1). As we have already emphasized, the specifics of different research fields will probably require different "multiplicative/additive" combinations of relations on the one hand and point/score ranges on the other.

In the field of social sciences (regardless of the number of points achieved), applicants for higher scientific degrees must produce scholarly books with 100% author's contribution (up to 3 authors) as follows:

- senior research associate – 1 scholarly book
- scientific advisor – 1 scholarly book after acquiring previous academic level

Table 2. Hypothetical example – social sciences (economics)

	Research Associate	Senior Research Associate (after receiving previous academic position)	Research advisor - 1 (after receiving previous academic position)	Research Advisor -2 – Tenure (after receiving previous position)
U^{I}	20	30	30	20
$U^{II} + U^{III}$	15	20	20	15
U	35	50	50	35

- scientific advisor -2 tenured position– 1 scholarly book after acquiring previous academic level.

A book that has been reviewed by at least 2 (peer) review experts in the discipline/field and edited by a publisher specialized in publishing works from the discipline/field to which the book's topic mainly belongs to can replace up to 10 points in the U^{II} category. The expert committee determines whether the book meets the specified criteria.

A particularly high-quality book internationally published in one of the world's languages and by a field/discipline specialized and internationally recognized publisher can replace up to 12 points in the II category scholar papers - N. The expert committee determines whether the book by its quality and the publisher meet the stated criteria.

Editorial books can replace up to 10 points in III category scholar papers - N. A particularly high-quality editorial book published by field/discipline specialized and internationally recognized publisher can replace up to 12 points in II category N.

The issue of the scholarly books impact in the social sciences can be compared to the meaning of publishing papers in Q1 or Q2 journals for the natural sciences, i.e. establishing a balance between two forms of scientific communication, essential for the research legitimacy of an intellectual in the social sciences. In this hypothetical example, the character of economic/social research requires the ability to elaborate the findings of specific research subjects (articles), as well as the ability to synthesize separate findings into relevant scientific "scrolls "of a specific research subject (books).

Finally, the criteria of research/scholar excellence should, especially in a relatively small country, address the issue of the use of the national language in scientific communication. Here, it is important to achieve a balanced approach between universal global scientific communication that takes place in major world languages (English at the moment) and the need to include the national language in the conceptual and terminological development of contemporary science. This could be achieved by requiring at least one paper to be published in national (Croatian) language in the research excellence assessment criteria (especially in the social sciences and humanities).

It should be emphasized that (by applying this minimum) there is neither a positive nor negative discrimination in the assessment of research excellence of either Croatian or foreign scientists interested in grants and/or research/academic positions in the Republic of Croatia. This does not even require that an international researcher knows the Croatian language, but that she/he publishes in it, just as a Croatian researcher does not need to know, for example, English (but must provide a translation) when he publishes in a highly cited/ranked international journal. There are intentions that seek a language issue possible solution in "pushing" the Croatian language papers to the professional rather than scholar journals. Such intentions do not recognize the significantly different features of the scholar terminology formation and the specifics of scientific knowledge transformation into the norms of individual professions. It is about the contribution that the criteria

should make to the development of Croatian as a living and dynamic world language of science instead of its "sinking" into a local dialect.

5. CONCLUSION

The dynamics of the international and national research excellence assessment, especially in the last twenty years, is characterized by "tidal waves" in which quantitative and qualitative criteria alternate at the "tops". Analyzing experiences from that relatively long period, the paper argues the need for a balance between "tide" and "ebb", that is, the practical use of combined quantitative and qualitative research assessment indicators.

Among the mass of numerous indicators measuring scientific performance, two quantitative indicators are undoubtedly in the center of attention, therefore requiring additional critical treatment. The first, Impact factor - IF, which "indirectly" valorizes research/scholarly works. The second is the h-index, which "indirectly" assess an individual's scientific output.

The insistence on IF as a key indicator of scientific excellence, and its intensive use in the late 1990s and early 2000s, especially in some natural sciences, revealed more and more limitations, and thus caused more and more obvious questioning of IF as an exclusive indicator of scientific quality.

Since the 2010s, this has resulted in increasingly strong professional criticism on a conceptual and practical level.

At the conceptual level, the relative weight of criticism is particularly noticeable when it comes from expert bibliometricians, creators and promoters of documents such as DORA and the Leiden Manifesto, who insist on the uselessness/harmfulness of using IF in judging the scientific quality of works, and especially the individual quality of scientists.

But the practical level, the concrete research policy guidelines of individual countries, probably has an even greater influence. Thus, the German Science Foundation (DFG) in its "Science Policy Guidelines" from 2010 (DFG, 2010), which were additionally elaborated in 2022 (DFG, 2022), almost completely excludes the application or use of IF and h-index in the quality assessment of research project proposals.

Such policy on the national level is further supported by a firm determination of the European Commission, which in recent documents after a thorough analysis, treats IF exclusively as an auxiliary/supplementary tool in the comprehensive assessment of scientific excellence (European Commission, 2021). In the USA and the UK, university administrations, for example Oxford, Stanford or MIT, claim that IF has very little or no role in hiring/recruiting researchers/teachers. The main criterion is a letter of recommendation from an external expert (outside the institution) in the candidate's scientific field, followed by an interview with the Hiring Committee (Abbott et al., 2010).

At the same time, in Croatia, the exclusive position of IF in assessing the researcher individual quality (acquiring professorship, tenured positions etc.), in the negotiation process when signing institution program financing contracts, and in allocating research government grants, has largely been maintained.

The paper argues for a rational and gradual improving of the balance between quantitative and qualitative criteria by passing through the three "levels" of forming a comprehensive evaluation of scientific excellence. In this sense, it is advocating the use of IF as a necessary initial quantitative datum that contributes to the integrity of the research excellence assessment process. It is the first necessary (in some cases it can also be an elimination) "level" that needs to be passed, in order to be able to reach the second "level". This second "level" consists of a qualitative excellence assessment, and the assessment of individual (co)author contributions. These are the perquisites of a complete expert review (peer review) assessment at the final third "level" determining research excellence. In this sense, the paper offers an orientational synthesized model of a balanced quantitative and qualitative criteria combination, not only for an integral assessment, but also for encouraging scientific excellence.

References

Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Van Noorden, R. (2010). Metrics: Do metrics matter? *Nature, 465*, str. 860-862. doi:10.1038/465860a

Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & Van Noorden, R. (2010). Metrics: Do metrics matter? *Nature, 465*, str. 860-862. doi:10.1038/465860a

Adriaanse, L., & Rensleigh, C. (2011). Comparing Web of Science, Scopus and Google Scholar from an Environmental Sciences perspective. *South African Journal of Libraries and Information Science, 77*, pp. 169-178. doi:10.7553/77-2-58

Baum. K., Zimmermann, Ch. (2023.) RePEc *General Principles* RePEc: Research Papers in Economics

Clarivate. (2015). *History of citation indexing*. Retrieved from https://clarivate.com/webofsciencegroup/essays/history-of-citation-indexing/

Dahal, S., Shrestha, G., Singh, S., Joshi, D., Sitaula, R., Sharma, M., & Chalise, P. (2022). Perceived Barriers to Publication in Scholarly Journals by Faculty from Maharajgunj Medical Campus. *Journal of Nepal Health Research Council 20(55)*, pp. 366-371. doi:10.33314/jnhrc.v20i02.4238

Davis, P. (2012). *The Emergence of a Citation Cartel*. Retrieved from The Schlary Kitchen: https://scholarlykitchen.sspnet.org/2012/04/10/emergence-of-a-citation-cartel/

DFG. (2010). *Quality not Quantity.* Njemačka naklada za znanost. Retrieved from https://www.dfg.de/en/service/press/press_releases/2010/pressemitteilung_nr_07/

DFG. (2022). *Academic Publishing as a Foundation and Area of Leverage for Research Assessment.* Njmeačka naklada za znanost. Retrieved from https://www.dfg.de/en/service/press/press_releases/2022/press_release_no_15/

DORA. (2012). *San Francisco Declaration on Research Assessment.* (American Society for CellBiology, Ed.) Retrieved from https://sfdora.org/read/

Elsevier. (2022). *Why choose Scopus*. Retrieved from https://www.elsevier.com/solutions/scopus/why-choose-scopus

European Commission. (2021). *Towards a reform of the research assessment system – Scoping report.* Ured za publikacije Europske unije. Retrieved from https://data.europa.eu/doi/10.2777/707440

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. (Nature, Ur.) *COMMENT*, str. 429-431. doi:10.1038/520429a

Himmelfarb Library News. (2022). *Journal Impact Factor Manipulation and Citation Cartels*. Retrieved from Himmelfarb Library News: https://blogs.gwu.edu/himmelfarb/2022/05/25/journal-impact-factor-manipulation-and-citation-cartels/

Ketcham, C., & Crawford, J. (2007). The impact of review articles. *Laboratory Investigation, 87*, pp. 1174–1185. doi:10.1038/labinvest.3700688

Knižnjičarstvo. (2012b). *Smjernice za autore*. Retrieved from Knižnjičarstvo: https://nova.knjiznicarstvo.com.hr/smjernice-za-autore/

Knjižničarstvo. (2012a). *Smjernice za autore*. Retrieved from Knjižničarstvo: https://nova.knjiznicarstvo.com.hr/smjernice-za-autore/

Levine-Clark, M., & Gil, E. (2008). A Comparative Citation Analysis of Web of Science, Scopus, and Google Scholar. *Journal of Business & Finance Librarianship, 14:1*, pp. 32-46. doi:10.1080/08963560802176348

Meho, L., & Yang, K. (2006). A New Era in Citation and Bibliometric Analyses: Web of Science, Scopus, and Google Scholar. *Journal of the American Society for Information Science and Technology*, pp. 25-26. Retrieved from https://arxiv.org/abs/cs/0612132

Oransky, I., & Marcus, A. (2017). *Gaming the system, scientific 'cartels' band together to cite each others' work*. Retrieved from STAT: https://www.statnews.com/2017/01/13/citation-cartels-science/#:~:text=But%20people%20looking%20closely%20at%20the%20scientific%20literature,to%20be%20more%20important%20and%20worthy%20of%20attention.

Petrić, M. (2013). Akademski kapitalizam i uloga sociologije. *Revija Za Sociologiju 43(3)*, pp. 273-288. doi:10.5613/rzs.43.3.4

Scimago Journal & Country Rank. (2023a). *Economics and Econometrics*. Retrieved from Scimago Journal & Country Rank:

https://www.scimagojr.com/journalrank.php?wos=true&category=2002

Scimago Journal & Country Rank. (2023b). *Economics, Econometrics and Finance (miscellaneous)*. Retrieved from Scimago Journal & Country Rank: https://www.scimagojr.com/journalrank.php?category=2001&wos=true&page=2&total_size=224

Scimago Journal & Country Rank. (2023c). *Acoustics and Ultrasonics*. Retrieved from Scimago Journal & Country Rank: https://www.scimagojr.com/journalrank.php?category=3102

Sveučilište u Zadru. (2014). Retrieved from Znanstvena, znanstveno-stručna i stručna djela: http://www.unizd.hr/Portals/4/nastavni_mat/1_godina/metodologija/ZNANSTVENA_DJELA.pdf

Unibiz News. (2020). *Scopus, Web of Science and Google Scholar*. Retrieved from https://www.unibz.it/en/news/135792-scopus-web-of-science-and-google-scholar

Van Noorden, R. (2010). Metrics: A profusion of measures. *Nature, vol. 465*, pp. 864-866. doi:10.1038/465864a

Veljak, L. (2013). Ontologija i metafizika. *Logos, 1*, pp. 9-19. Retrieved from https://www.bib.irb.hr/678247

Yang, K., & Meho, L. (2007). Citation Analysis: A Comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for Information Science and Technology 43(1)*, str. 1-15. doi:10.1002/meet.14504301185

Zelenika, R. (1988). *Metodologija i tehnologija izrade znanstvenog i stručnog djela.* Beograd: Fakultet za pomorstvo i saobraćaj Sveučilišta "Vladimir Bakarić" ; Fakultet prometnih znanosti Sveučilišta ; Saobraćajni fakultet Univerziteta.

Žugaj, M., Dumičić, K., & Dušak, V. (1999). *Temelji znanstvenoistraživačkog rada.* Varaždin: Fakultet organizacije i informatike - Sveučilišta u Zagrebu.

Supplement 1.

WoS subject categories in Croatian research/scholar field - Economics		
Subject categories	**Number of WoS journals**	**Q1 journals**
Accounting	122	42
Business and International Management	256	98
Business, Management and Accounting	206	86
Demography	74	25
Development	175	63
Economic Geology	26	9
Finance	221	75
Health Policy	187 (Econ 38)	64 (Econ 18)
History	907 (Econ 41)	295 (Econ. Hist. 32)
Leadership and Management	24	9
Management Information Systems	65	27
Management, Monitoring, Policy and Law	236	83
Management, Technology and Innovation	157	60
Management, Science and Operations Research	119	44
Marketing	126	51
Strategy and Management	304	110
Tourism, Leisure and Hospitality Managment	98	34
Economics -total	2288	866

Source:WoS subject category list, subject category list - Search (bing.com) 2022

Note; This is not a complete list of "economic" journals, because economic topics are addressed for example, by journals in the subject categories of medicine, informatics, transport, and the like.

Elementi vrednovanja znanstvene izvrsnosti

Sažetak

Dinamika međunarodnog i nacionalnog sustava prosudbe znanstvene izvrsnosti, poglavito u posljednjih dvadesetak godina, ima obilježje „plimnih valova" u kojima se na „vrhovima" smjenjuju kvantitativni i kvalitativni kriteriji. Propitujući iskustva iz tog relativno dugog razdoblja, u radu se argumentira potreba građenja ravnoteže između kvantitativnih i kvalitativnih kriterija uspinjanjem na tri „stepenice" oblikovanja cjelovite ocjene znanstvene izvrsnosti. U tom je smislu zalaganje za korištenje IF kao potrebne početne kvantitativne informacije koja doprinosi objektivizaciji procesa prosudbe znanstvene izvrsnosti. To je prva neophodna „stepenica" koju treba proći, da bi se moglo popeti na drugu „stepenicu". Drugu „stepenicu" čine, i ocjena kvalitativne vrsnoće kako prezentiranog znanstvenog rada, i ocjena individualnih doprinosa (su)autora. To su pretpostavke cjelovite istorazinske stručne recenzije na konačnoj „trećoj" stepenici utvrđivanja znanstvene izvrsnosti. U tom je smislu u radu ponuđen i orijentacijski sintetski model ravnotežne kombinacije kvantitativnih i kvalitativnih kriterija ne samo cjelovitog vrednovanja, nego i poticanja znanstvene izvrsnosti.

Ključne riječi: vrednovanje kvalitete istraživanja, kvalitativni indikatori, citatno rangiranje, čimbenik odjeka (IF), integrirani model vrednovanja kvalitete u znanosti