

Determining and Prioritizing the Evaluation Criteria of Humanities Scientific Outputs: A Case Study of Language and Literature Fields

Elahe Ebrahimi Dorcheh¹, Ali Mansouri^{1,*}, Mitra Pashootanizadeh¹, Ali-Asghar Mirbagherifard², Ahmad Shabani¹

¹Department of Knowledge and Information Science, University of Isfahan, Isfahan, IRAN.

²Department of Persian Literature, University of Isfahan, Isfahan, IRAN.

ABSTRACT

With regard to the specific nature and variety of the humanities fields and disciplines and the need to evaluate the humanities research outputs according to their nature and intrinsic characteristics, two questions have been posed and answered in this study as follows: "What are the criteria and indicators for evaluating the research outputs of humanities?" and "What is the prioritizing of the evaluation criteria according to the research approaches and goals in humanities?" Considering the differences in the fields of humanities, a case study of language and literature was conducted. This research was done with a mixed method (qualitative and quantitative stages). The first stage was carried out using a library research method to extract the criteria and indicators for the evaluation of the research outputs in the fields of language and literature. In the second stage, in order to finalize and prioritize the criteria, a questionnaire was designed and distributed among a number of experts in the fields of language and literature in two rounds of fuzzy Delphi. In the first stage, 42 indicators were identified and divided into 8 categories of criteria: 1) platform for creation, presentation and publication, 2) writing structure, 3) content, 4) impact in online environment, 5) scientific impact, 6) social impact, 7) economic impact, and 8) cultural impact. The prioritizing of the criteria was also based on their average obtained in the second round of fuzzy Delphi, which shows the impact of research approaches and goals on the priority of using the criteria.

Keywords: Evaluation of research outputs, Evaluation criteria, Evaluation indicator, Humanities, Language and Literature.

Correspondence:

Ali Mansouri

Department of Knowledge and Information Science, University of Isfahan, Isfahan, IRAN.

Email: a.mansouri@edu.ui.ac.ir

ORCID: 0000-0003-4047-2697

Received: 14-09-2023;

Revised: 19-11-2023;

Accepted: 16-01-2024.

INTRODUCTION

With the introduction of the information world, the examples of power have changed, and competition over physical resources has given its place to the competition for information, knowledge and innovation. In this context, different sciences, from engineering and basic sciences to Humanities and Social Sciences (HSS), have adopted different ways of affecting and interacting with the information society.^[1] As such, humanities can be considered the software and the soul of knowledge, and if all human knowledge is assumed as a system, humanities will be the core and the center of gravitation of the system, and how it is formed and oriented will have a direct effect on the formation and direction of micro and macro social systems.^[2] As Lanzillo^[3] says, due to their versatile functions as broad as the subject of man and society, humanities are likely to facilitate the application of other sciences, provide the

basis for scientific development, and pursue the social-cultural maturity of man as their main goal and their obvious and hidden concern. In practice, humanities are intertwined with different dimensions of individual and social life of people and can be very efficient and effective than other sciences.^[4]

Despite the great importance and function of the humanities, based on the evaluations, it is commonly accepted that humanities research is not effective in society and has a small share of science in different countries of the world.^[5] Several reasons have been stated for the underdevelopment of humanities in society,^[6-7] which are all valid and can be discussed, but we can also think about the problem from the perspective of how to evaluate research outputs – an issue that has been less attended despite its great importance.^[8] Evaluation of the research and the researcher has many functions, such as creating an opportunity for the researcher to defend his/her performance and compensate for his/her weaknesses, gathering information needed for decision-making, policy-making and planning in research,^[9] improving the performance and quality of research by determining strengths and weaknesses, optimizing the allocation of research



DOI: 10.5530/jscires.13.1.13

Copyright Information :

Copyright Author (s) 2024 Distributed under
Creative Commons CC-BY 4.0

Publishing Partner : EManuscript Tech. [www.emanuscript.in]

financial resources,^[10] providing guidance for designing and modifying evaluation criteria,^[11] determining research priorities and designing relevant policies,^[12] directing researches in line with the real demands of society, Evaluating different dimensions of research effectiveness, identifying effective research projects for commercialization, determining the quality of programs developed for research activities, and deciding on how to improve the quality of research activities.^[13] These can be accomplished if we evaluate research outputs in proportion to each scientific field and use criteria and indicators that are compatible with the nature of these fields.

However, humanities research outputs that are part of soft sciences are often evaluated with criteria and indicators designed for hard sciences, including basic, natural and engineering sciences,^[14] which do not have the priority and competence to be used in humanities and do not indicate the real position of the humanities research outputs. Accordingly, by comparing two different phenomena with the same criterion, which is not acceptable, the prioritizing of humanities is low compared to other sciences. Although humanities and other sciences are similar in being science, their inherent differences and the specific language of each should be considered in evaluations and the use of a similar model for all fields of science should be avoided.^[15] These differences can be found in the four dimensions of research subject and objectives, research method, citation behavior, and science coverage in citation databases. Maybe if the evaluation is done in accordance with the nature and characteristics of humanities, the progress and current position of this field will be better displayed and the level of its research will increase.^[16,6] If the evaluation is done with respect to the research outputs of the humanities and their fields, many researches that have been evaluated before without considering the specific characteristics of each field can be evaluated again and appropriately. By using appropriate criteria and indicators in the evaluation of humanities research outputs, the accuracy and validity of the evaluations will increase, their real position will be demonstrated, and stakeholders can reliably use the evaluation results in the prioritizing, planning, policy-making and decision-making in the fields of humanities.

Humanities consist of various fields and disciplines, each of which has a significant contribution to the development and progress of society. Since it is not possible to examine all the fields of humanities at once, this study has focused on the evaluation of language and literature due to its importance, position, and diverse outputs considering its range of audiences, which can be regarded as an introduction to the flow of specialized evaluation of humanities. The fields of language and literature deal with different aspects of life and most importantly with the identity and intellectual and cultural maturity of a nation and are considered as a support for the richness of a country's civilization. These fields are entertaining and pleasing and are also a reflection of society and present an image of what people think, say and do in society.

These disciplines are a valuable treasure for understanding the values, customs and historical background as well as the future developments of the society, which explain and teach the do's and do not's in the form of a mixture of knowledge and art. Literature is the mirror of every age; it shows the turning point of every period and helps the sociology and anthropology of that time. Based on the reflection theory, literature can contain information about social behavior and values and document the social world for the reader in a transparent way.^[17] It is possible to achieve the goals of language and literature fields through the publication of research results in various research outputs such as journal article, conference article, book, research project, dissertations/thesis, and literary creations/creative literature. Ebrahimi Darcheh *et al.*^[18] discussed the harms and strategies for evaluating the research outputs of humanities, particularly language and literature, and presented three approaches and goals for research in these fields: 1) production of science and promotion of knowledge foundations, 2) applicability and responsiveness to society's problems, and 3) literary creation/creative literature.

In evaluations, the criteria and indicators that fit the research outputs under study are usually chosen and applied. However, it is here suggested that the evaluation and application of criteria and indicators should be done according to the researcher's approach and goal of research and publication of research output, since it is by research goal that the most suitable research output for results publication is selected. On the other hand, it seems that when the research goal changes, the importance of the evaluation criteria changes accordingly. By using the evaluation criteria, it is possible to determine to what extent the desired goal has been achieved. Since the objectives of each discipline are defined according to the discipline itself, using an objective-based evaluation, it can be ensured that more attention will be paid to the characteristics of each discipline. Evaluation based on approaches and goals can be considered as the main innovation of this research. Based upon the stated approaches and goals for language and literature fields, the present study is an attempt to answer the following questions:

What are the criteria and indicators for evaluating the research outputs of humanities, especially in the field of language and literature?

What is the prioritizing of the evaluation criteria according to the research approaches and goals in the fields of language and literature?

LITERATURE REVIEW

By the growing production of scholarly works in different fields of science, further importance has also been attached to the evaluation of research and research on research. Based on the recorded literature and the sources in the databases, scientific research evaluation dates back to the 1970s. A review of the literature shows that it has not been a long time since the evaluation of humanities research became part of the issues

and concerns of experts in the field of research evaluation. In what follows, some recent studies on the humanities research evaluation are reviewed.

Despite the importance and necessity^[19] and the application and effectiveness of the humanities,^[20] there are challenges regarding the evaluation of the research outputs of these sciences, the most obvious of which is the different nature of the humanities from other sciences and the need to conduct evaluation in proportion to the nature of the humanities.^[21,18]

Due to the humanities differences, compared to other sciences, in terms of research subject,^[22] diversity of audience,^[23] methodology and data collection approach,^[24] dependence on the native language,^[25] platform and channel of publication of findings,^[26] national geography of publications,^[27] number of authors,^[23] citation behavior,^[28,29] coverage in database citations,^[30] and method of affecting society,^[31-33] special attention should be paid to the evaluation of the research outputs in the fields of humanities. As the need is strongly felt for the evaluation of the humanities research output, researchers have sought to compare the humanities with other sciences and show the specific features of the fields of humanities. The number of such studies is relatively large and they have generally come to the conclusion that the evaluation of the humanities through the technical-engineering style criteria and indicators is not suitable and do not show what it should. These studies highlighted the importance and necessity of a specific evaluation of the humanities fields and disciplines, which is in agreement with the research problem in the present study.

A review of the literature shows that studies on the evaluation of humanities research have adopted a descriptive approach to investigate issues such as: publishing behavior of researchers, research and citation databases analysis,^[34] citation analysis,^[35,36] core sources,^[37] thematic trend of research,^[38] drawing of a map of science,^[39] review of journals, articles, books, and dissertations,^[40] metadata analysis of information resources,^[41] performance of researchers and faculty members,^[42] research visibility,^[43] and creativity measurement.^[44] In the aforesaid works, there are deficiencies in the evaluation of humanities research, as humanities are generally not differentiated from other sciences and a similar approach has been used to evaluate and compare all the fields disregarding their specific characteristics and nature. Some studies have pointed to the special evaluation of the humanities research, but have not elaborated any further on the way of this evaluation. In these studies, the proposed criteria and indicators are often quantitative, and qualitative criteria and indicators are rarely used. Nevertheless, the problem of the incompatibility of the indicators with humanities remains. Ochsner *et al.*^[45] have also looked for agreed-upon concepts of quality in humanities and believe that a research assessment by means of quality criteria presents opportunities to make visible and evaluate humanities research, while a quantitative assessment

by means of indicators is very limited and is not accepted by scholars. However, indicators that are linked to the humanities scholars' notions of quality can be used to support peers in the evaluation process (i.e. informed peer review).

Concerning the fields of humanities, suggestions have been made for evaluating the research performance of faculty members,^[14] educational departments,^[46] and researchers.^[47] The scope of some research has been narrowed down to the issues such as the study and design of the bibliometric, scientometric and altmetric evaluation criteria and indicators^[48-56] and research impact measurement.^[57-65] The purpose of Thelwall and Delgado^[66] research was to make an explicit case for the use of data with contextual information as evidence in humanities research evaluations rather than systematic metrics; Data are already used as impact evidence in the arts and humanities, but this practice should become more widespread. Humanities researchers should be encouraged to think creatively about the kinds of data that they may be able to generate in support of the value of their research and should not rely upon standardized metrics.

Despite the large number of the aforesaid studies, scant attention has already been paid to the design and investigation of evaluation criteria and indicators related to a specific aspect or field of humanities. Such studies are limited and usually not up to date. With respect to the language and literature fields, Hug, Ochsner and Daniel^[67] proposed criteria to evaluate research quality in three fields: German Literature Studies (GLS), English Literature Studies (ELS), and Art History. Ochsner, Hug and Daniel^[68] ranked the criteria obtained from the previous research. D'Souza^[55] investigated the characteristics and method of evaluation of creativity in story writing. In the present study, different criteria and the Priority and position of each of these criteria in the fields of language and literature are identified. The prioritization of the criteria is based on the research approaches and goals in these fields, which is a distinct feature in the present study.

METHODOLOGY

The current research has been conducted in a mixed method (qualitative and quantitative stages), as described below:

Review of documents

In order to obtain criteria and indicators for evaluating the research outputs of language and literature fields, documents and studies on the subject were analyzed using the library research method. Purposeful sampling of documents at this stage was done, and studies that were compatible with the research subject and question were selected. To this end, sources were examined from the specific to the general: first, studies on the language and literature, then studies on the humanities, and finally in the general dimension, those resources related to other scientific fields were explored. To review the documents, an advanced

search of articles (journal and conference articles) was conducted in databases using the search strategy of Table 1. To increase the comprehensiveness and precision of the search, synonyms and related words, Boolean operators, truncation, and phrase searching were included.

By entering a search query in each database and retrieving sources, duplicate records (in terms of title and subject) or items with irrelevant titles were removed. Among the sources that had similar results, the most appropriate and up-to-date ones were selected. Then, the records that had related titles (377 records) were reviewed in terms of abstracts, and in this step, unrelated items were removed. In the next step, sources with relevant abstracts (82 records) were reviewed in full text. Finally, among 35 related and suitable records, the evaluation criteria and indicators of research outputs were extracted for use in the next stage. The databases searched and the number of documents retrieved and used are given in Table 2. In the library search, note-taking was used as the data collection tool. MaxQDA, version 2020, software was used for documents analysis.

Experts' panel creation

The Delphi method is a search method with the characteristics of iterating different rounds and controlled feedback (data analysis in each round) based upon the anonymous statistical group response of experts.^[69] To get closer to the real-world and overcome the problem of ambiguity and uncertainty in the judgment of decision makers, the classic Delphi is replaced by the fuzzy Delphi method.

In order to finalize the criteria and indicators obtained in the previous stage based on the opinion of experts (adding, subtracting or changing categories) and prioritize them, a researcher-made Delphi questionnaire with a seven-point Likert scale (including verbal expressions) was designed. The questionnaire had 4 open and closed questions and included various items to get the experts' opinion about the importance of the criteria and their suggestions in this regard. The initial questionnaire was revised and finalized using the opinions of 3 experts (in the fields of language and literature, and library and information science), and its face and content validity was confirmed. In order to determine the reliability, the questionnaire was answered by 22 researchers of language and literature fields (English, French, German, Arabic, and Persian) from different universities in Iran.

By analyzing the answers in SPSS software, it was found that the Cronbach's alpha of the questionnaire is 0.95, which indicates the homogeneity and internal consistency of the questions and their items and the reliability of the questionnaire.

The statistical population of the research is the faculty members of the language and literature fields of Iranian universities. According to Hogarth,^[70] six to twelve members are ideal for the Delphi method. Clayton^[71] also posited that if the respondents are a combination of experts with different specialties, between five and ten members are sufficient. Usually, in different sources, at least 10 people are considered as a suitable number. The research sample (17 people) was selected purposefully and by snowball method from among the expert faculty members in the fields of language and literature (English, French, German, Arabic, and Persian) from 8 universities in Iran. Criteria such as experience or involvement in the subject area of research (authoring and translating books and publishing articles in the field of research evaluation, or membership in associations and working groups related to the fields of language and literature in research institutions at different levels) and having diverse outputs in the fields of language and literature were effective in selecting the sample.

Excel software was used to analyze the data collected from the questionnaires. After collecting the questionnaires of the first round, the fuzzification of the verbal expressions was done based on the Triangular Fuzzy Numbers (TFN). Each verbal expression was assigned a triangular fuzzy number consisting of three values of the lower limit or the minimum (l), the middle limit or the most probable value (m), and the upper limit or the maximum (u), the details of which are given in Table 3. Fuzzification allows answers to be defined qualitatively, which is the advantage of the fuzzy Delphi method over the classic Delphi method.

Table 1: Search strategy.

TITLE= ((evaluat* OR measur* OR assess*) AND (research OR article OR book OR monograph OR 'reference source' OR thesis OR dissertation OR 'research output' OR 'scientific output' OR 'research project' OR 'research activit*' OR 'research work' OR 'research performance' OR 'research impact' OR research effect OR 'research application' OR 'creative literature' OR 'literary work' OR artwork OR fiction OR poetry OR 'dramatic literature' OR 'social activit*' OR 'economic activit*'))

Table 2: Databases and the number of documents reviewed.

Databases	Number of records retrieved	Remove duplicates	Number of records with related title	Number of records with related abstract	Number of records with related text
Scopus	35612		277	63	25
Web of Science	23358		100	19	10
Total	58970		377	82	35

Table 3: Triangular fuzzy numbers for the seven-point scale.

Definite number	Verbal expression (importance level)	Fuzzy number $F = (l, m, u)$
1	Very Low (VL)	(0, 0, 0.1)
2	Low (L)	(0, 0.1, 0.3)
3	Medium Low (ML)	(0.1, 0.3, 0.5)
4	Medium (M)	(0.3, 0.5, 0.75)
5	Medium High (MH)	(0.5, 0.75, 0.9)
6	High (H)	(0.75, 0.9, 1)
7	Very High (VH)	(0.9, 1, 1)

Table 4: Coefficient of variation and consensus.

Coefficient of variation	Decision rule
$0 < V \leq 0.5$	Good degree of consensus-no extra rounds required.
$0.5 < V \leq 0.8$	Less than satisfactory degree of consensus-possible need for an additional round.
$V > 0.8$	Poor degree of consensus-definite need for an additional round.

In the next step, using the assigned fuzzy numbers, the fuzzy average of each boundary was calculated (Formula 1). Then, the de-fuzzification of the values was performed, which means calculating the average of the fuzzy averages of the limits for each criterion (Formula 2). The de-fuzzified number is between zero and one. In the comparison that was made between the de-fuzzified number of the criteria and the threshold limit of 0.7 (suitable for the seven-point scale)^[72] all the criteria, except four, scored above the threshold. In view of this, the second round of Delphi was also conducted, although no new criterion and indicator had been proposed in the first round, and the Coefficient of Variation (CV) of all criteria was acceptable (between 0 and 0.5) based on Table 4, which indicated the consensus/agreement of the opinions presented.^[73] The coefficient of variation is obtained by dividing the standard deviation by the mean.

$$F_{AVE} = \frac{\sum l}{n}, \frac{\sum m}{n}, \frac{\sum u}{n} \quad (1)$$

$$\text{If } \tilde{F} = (L, M, U), \text{ then } F = \frac{L+M+U}{3} \quad (2)$$

The second-round questionnaire was also given to the experts (11 people of the first round) with the same criteria and indicators, with the difference that in this round, a column containing the de-fuzzified numbers of each criterion in the first round was added to the questionnaire tables. Respondents were asked to state their opinions on the importance of each criterion and, if necessary, change or modify their first-round answers according to the notified numbers.

In the second round, fuzzification, de-fuzzification, and comparison of the obtained average with the threshold limit were performed. In this round, only five criteria (four items like the first round and one new item) scored below the threshold of 0.7. The result of comparing the averages of the two rounds showed that the opinions did not differ much and there was consistency in the answers of the two rounds. These two cases show the stability of the answers, which is a key factor in deciding to stop iterating the Delphi rounds. Another necessary factor in deciding whether to continue or stop rounds is the degree of consensus and solidarity of opinions. Various mechanisms such as subjective criteria, descriptive statistics and inferential statistics are used to check the consensus, which Von der Gracht^[69] collected in a review study.

In the present study, three methods were used to determine whether the experts reached a consensus or not: 1) the difference between the averages of the two rounds for all criteria based on formula 3 was smaller than 0.1, which indicates the consensus of the experts' opinions; 2) the Kendall's W Coefficient of the second round (0.351) compared to the first round (0.280), both with a significance level of lower than 0.01, did not grow much, which indicates that opinions have not changed significantly (in this coefficient, the value 0.1 indicates a very weak agreement, while the value 0.7 indicates a very strong agreement); and 3) the coefficient of variation of all criteria was less than 0.5 in the second round, which is acceptable according to Table 4. Finally, based on the stability and consensus reached, it was decided to stop distributing the questionnaire in the second round.

$$F_1 = (l_1, m_1, u_1)$$

$$F_2 = (l_2, m_2, u_2) \quad (3)$$

$$\text{Gap} = |(l_1 + m_1 + u_1) - (l_2 + m_2 + u_2)|/3$$

Since the purpose of the fuzzy Delphi method at this stage of the research was to determine the importance and prioritizing of the evaluation criteria of the research outputs of the language and literature fields, no criterion was removed. Rather, using the averages of the second round, the prioritizing of the criteria was done. The closer the average of the criterion is to 1, the higher its importance and prioritize according to the experts.

Findings

What are the criteria and indicators for evaluating the research outputs of humanities, especially in the field of language and literature?

The items that were mentioned in the documents and resources for the evaluation of research outputs were divided into 8 categories of criteria (components) and 42 indicators (subcomponents) according to their consistency with the fields of language and literature, as shown in Table 5. In the questionnaire distributed based on the fuzzy Delphi method, the respondents were asked

through an open-ended question to state their suggestion for adding, removing or moving criteria and indicators. The issues raised by the respondents in the first round were either included in the criteria and indicators obtained from the review of documents or were unrelated to the research topic. No new items were proposed in the second round. Therefore, what was obtained from the review of the documents was used as the criteria and indicators for evaluating the research outputs of the fields of language and literature.

What is the prioritizing of the evaluation criteria according to the research approaches and goals in the fields of language and literature?

Tables 6-8 show the averages (de-fuzzified numbers) of the first and second rounds of fuzzy Delphi, their difference and coefficients of variation. Then, based on these data, in Table 9, the priority of each of the criteria for evaluating the research outputs of the language and literature fields are presented based on the approach and goal of the research. The purpose of the implementation of fuzzy Delphi was not to screen the criteria, but to determine the importance and priority of each one in the evaluation according to the goal of the research. The goals of research in the fields of language and literature are: 1) production of science and promotion of knowledge foundations, 2) applicability and responsiveness to society's problems, and 3) literary creation/creative literature.

According to Table 6, when the goal of research is production of science and promotion of knowledge foundations, criteria such as scientific impact and the content of research outputs have a higher priority for experts, because such research is supposed to help expand the theoretical scope of science. Accordingly, as the criteria go beyond the limit of the scientific environment and become influential in the society, their weight decreases.

Based on the experts' opinion, as shown in Table 7, when the goal of the research is applicability and responsiveness to society's problems, social impact is more important than other evaluation criteria. Perhaps because in this goal, the researcher has more connection with the society, the field effect of his/her activities in the form of research outputs is considered and evaluated more important than the scientific impact.

According to Table 8, when "literary creation/creative literature" is the main goal of a language and literature specialist's research, cultural impact and its evaluation through content validity criteria are more important, because such works deal with the cultural identity of the society.

According to Table 9, the priority obtained for the criteria and their differences in approaches and goals indicate the importance and dominance of targeting in research. The priority of all criteria except two fluctuates with respect to different objectives. The content of the research outputs has been assigned the second priority in all cases, which indicates the high importance of the

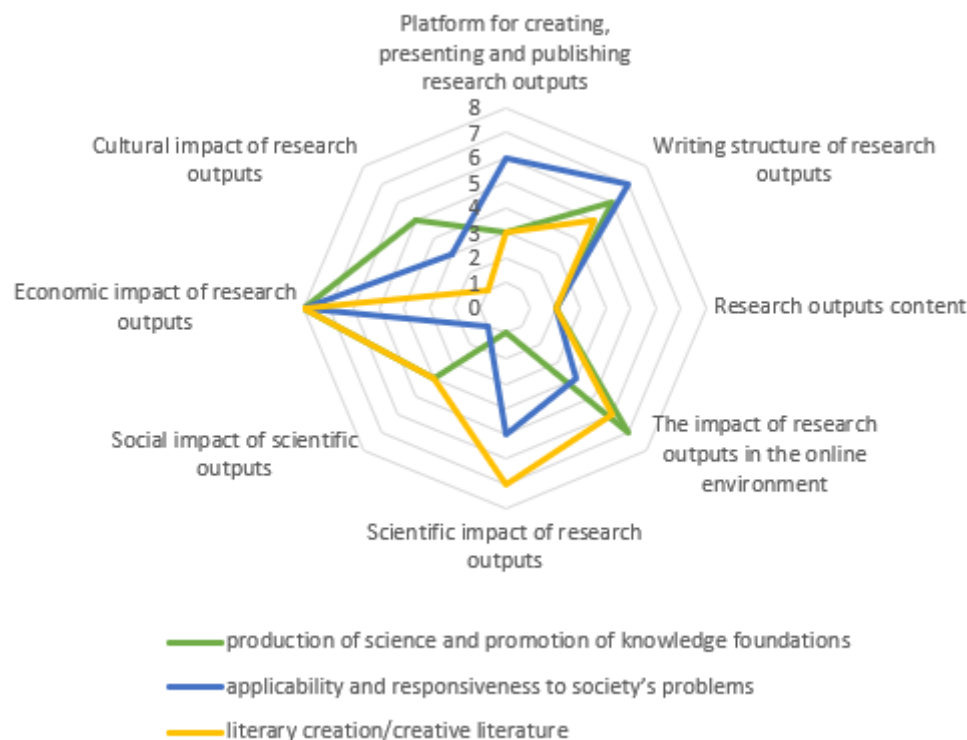


Figure 1: The prioritizing of the evaluation criteria of the research outputs of the fields of language and literature.

Table 5: Criteria and indicators for evaluating the research outputs of language and literature fields.

Components (evaluation criteria)	Sub-components (evaluation indicators)
Platform for creating, presenting and publishing research outputs.	Valid publisher. Valid journal. indicators based on the journal. Validity of the organizer. Validity of the supervising organization. Validity of the mother organization.
Writing structure of research outputs.	Proper and coherent writing. Readability to the audience (ease of reading and understanding).
Research outputs content.	Content coherence and uniformity. Specialization and thematic purposefulness. Reliability of method and accuracy of data and results. Having an Interdisciplinary nature. Positive review by peer reviewers. Newness and originality of the output. Theoretical and methodological innovation. Being practical and community-oriented. Being creative. Having scientific and theoretical support.
The impact of research outputs in the online environment.	Online citations on web pages (Citations). The amount of file storage or adding to favorites (Captures). Discussing the output in social networks (Mentions). The amount of output sharing in social media (Social media).
Scientific impact of research outputs.	Indexing in citation databases. Attracting scientific cooperation at national and international levels (capacity building). Citation-based indicators. Streamlining in a specialized subject. Use in curriculum. Being cited in policy documents, policy-makings and policy guidelines. Authority and scientific reputation of the output. Winning awards from festivals. Attracting research credits. Being accepted and referred to by educational, research and executive centers.
Social impact of scientific outputs.	Publication and presence in the media. Popularity (people's acceptance and reference to the output). Promotion and generalization of research findings. Being accepted and referred to by educational, research and executive centers.
Economic impact of research outputs.	Usage and applicability. Income generation from application. Entrepreneurship.
Cultural impact of research outputs.	Using literary capacities in different branches of art (for example, turning literary works into dramatic works). Output acceptance by mass media. Use of output in cultural centers.

Table 6: Criteria and production of science and promotion of knowledge foundations.

Criteria	First round average	Second round average	Difference of two rounds	First round coefficient of variation	Second round coefficient of variation
Platform for creating, presenting and publishing research outputs.	0.76	0.73	0.05	0.20	0.16
Writing structure of research outputs.	0.71	0.91	0.02	0.19	0.19
Research outputs content.	0.92	0.80	0.01	0.08	0.09
The impact of research outputs in the online environment.	0.74	0.73	0.01	0.18	0.19
Scientific impact of research outputs.	0.91	0.92	0.02	0.11	0.10
Social impact of scientific outputs.	0.78	0.79	0.01	0.17	0.17
Economic impact of research outputs.	0.47	0.48	0.01	0.35	0.18
Cultural impact of research outputs.	0.74	0.79	0.05	0.18	0.15

Table 7: Criteria and applicability and responsiveness to society's problems.

Criteria	First round average	Second round average	Difference of two rounds	First round coefficient of variation	Second round coefficient of variation
Platform for creating, presenting and publishing research outputs.	0.77	0.75	0.02	0.17	0.19
Writing structure of research outputs.	0.70	0.69	0.02	0.19	0.21
Research outputs content.	0.88	0.87	0.01	0.12	0.13
The impact of research outputs in the online environment.	0.83	0.80	0.04	0.13	0.16
Scientific impact of research outputs.	0.80	0.80	0.00	0.16	0.16
Social impact of scientific outputs.	0.94	0.92	0.02	0.09	0.10
Economic impact of research outputs.	0.66	0.62	0.04	0.22	0.27
Cultural impact of research outputs.	0.88	0.84	0.04	0.11	0.14

Table 8: Criteria and literary creation/creative literature.

Criteria	First round average	Second round average	Difference of two rounds	First round coefficient of variation	Second round coefficient of variation
Platform for creating, presenting and publishing research outputs.	0.84	0.85	0.02	0.16	0.15
Writing structure of research outputs.	0.78	0.78	0.01	0.12	0.12
Research outputs content.	0.83	0.87	0.03	0.15	0.13
The impact of research outputs in the online environment.	0.74	0.73	0.01	0.25	0.19
Scientific impact of research outputs.	0.67	0.69	0.03	0.16	0.23
Social impact of scientific outputs.	0.80	0.79	0.02	0.16	0.17
Economic impact of research outputs.	0.53	0.50	0.03	0.31	0.39
Cultural impact of research outputs.	0.89	0.88	0.02	0.12	0.14

Table 9: The prioritizing of the evaluation criteria of the research outputs of the fields of language and literature.

	Production of science and promotion of knowledge foundations	Applicability and responsiveness to society's problems	Literary creation/ creative literature
Evaluation criteria	Priority (based on the average of each criterion in the second round of fuzzy Delphi)		
Platform for creating, presenting and publishing research outputs.	3	6	3
Writing structure of research outputs	6	7	5
Research outputs content.	2	2	2
The impact of research outputs in the online environment.	7	4	6
Scientific impact of research outputs.	1	5	7
Social impact of scientific outputs.	4	1	4
Economic impact of research outputs.	8	8	8
Cultural impact of research outputs.	5	3	1

content of the outputs. Besides, the economic impact of research outputs in all goals has the lowest priority, which can be due to the less inclination of the field of humanities and especially the fields of language and literature towards the economic fields.

The findings show that the goal of the research both determines the audience and is effective on the rating of the evaluation criteria. Academic and non-academic audiences have different needs, and therefore the research outputs of each should be evaluated differently. In Figure 1, the prioritizing of the criteria is shown schematically. In each goal, as we approach from the center of the diagram to its margins, the priority and importance of the evaluation criterion decrease.

DISCUSSION AND CONCLUSION

In this study, first the criteria and indicators for evaluating the research outputs of language and literature fields were determined and prioritized using references and experts' opinions. The literature review shows that the use or non-use of each of the criteria and indicators and the priority of each one is different in various fields.^[67] This difference in the current research can be due to the nature of language and literature fields, on the one hand, and their classification based on the goal of the research, on the other hand. Therefore, it is highly emphasized to plan for research evaluation based on the nature of the fields and the research goal.

From the point of view of research evaluation, the main stages of research from the beginning to fruition consist of input, process, output, outcome and impact. As the findings of the current research show, according to the goal of the research, the evaluation method of each of these stages can be changed. When the goal changes, the expectations of the stakeholders also change, which in turn requires an appropriate evaluation method. For example, when working in a scientific environment with a

specific audience, citation-based indicators are more important, but in a public space with a general audience, the use of research output in cultural centers and acceptance by mass media is more important.

When the approach, goal and audience of scientific research and study change, its evaluation criteria will also change. To publish and share the findings of any research, there are various outputs that should not be evaluated the same. These three aspects (goal, output, criteria) must be consistent to achieve effectiveness. Currently, researchers are looking for the impact of their research in areas that are easier to document, but if the evaluation is done from different aspects and prioritizing is done, it will encourage and guide researchers to research and scientific studies that interest them and have diversity. Because; because it is ensured that their efforts will not be in vain and it is effective in evaluation for promotion.

The approach and goal of *production of science and promotion of knowledge foundations* includes theoretical studies, basic and fundamental research for better understanding of subjects, idea generation, innovation, freedom of thought, thinking and presenting original thought, creativity, discourse formation and realm creation, and cultivating or formulating theory. These are the factors of scientific progress and expanding the boundaries of knowledge that can lead to the increase of science and knowledge and reduction of consumerism. If such a goal is considered for research, the evaluation of research outputs from the perspective of scientific impact is a priority. The obtained priorities not only show the importance of each criterion, but also indicate their sequence. The research output in any format should have the indicators of a suitable content in order to be scientifically effective. If the output has high-quality content, it can be provided in a suitable platform for social and cultural use. The writing structure suitable for the audience makes the research output

acceptable and allows it to be used in the online environment and create its economic impact.

The approach and goal of *applicability and responsiveness to the society's problems* includes applied research, focusing on solving a specific problem and its practical aspect in economic, scientific, cultural, social and political fields, applying science and knowledge in line with social responsibility to solve the problems and needs of society, the use of language and literature in the real world and the environmental application of scientific awareness, and activism in serving the society. If the subject of the research output is to solve a social concern, one should expect a social impact from it and consider evaluation in this regard as a priority. The publication of popular and culturally effective content in the online environment causes generalization and promotion of research findings. The publication of applied research with the desired writing structure in a reliable and accessible platform will provide the basis for its entry into the economic arena and generate income.

Literary creation/creative literature means to be inspired by artistic and literary feelings and emotions in interpreting the human and social world to meet aesthetic needs and promote values. Unlike the other two approaches and goals, which have a more general aspect and are applicable in all fields of humanities, the approach of literary creation/creative literature is specific to the fields of language and literature and can be a support for the first and second approaches. According to the experts, literary creations/creative literature, which can be in the form of poetry, prose (fiction and non-fiction), and dramatic literature, should first be evaluated in terms of cultural impact. In the next priority, the content and platform of their creation, presentation and publication is considered important and can have a social impact. Proper writing and sharing and publishing research in the online environment can bring scientific and economic impact.

The obtained criteria and indicators are in harmony and similarity with the related items in the background section of the research and can generally be divided into two groups: factors related to the academic environment and general factors. Among them are indicators of academic value, innovation, and social value for humanities papers.^[48] Besides, criteria such as innovation and originality, impact on the research community, productivity, relation to and impact on society, connection between research and teaching, fostering cultural memory, and connection to other research can be mentioned in the fields of German literature, English literature and art history.^[68] At a more specific level, factors such as meaning and relevance, reader's immersive experience, development and control, distinctiveness, voice and originality in the evaluation of the story^[55] can be mentioned. As can be seen from the findings, the quality of research is a multidimensional concept; the criteria are influenced by each other and complement each other. Each criterion does not consider all aspects of the research, so one criterion should not be used in isolation, but

several criteria should be used in a complementary manner.^[74] However, what is necessary and worthy of sufficient attention is the consistent and quality content of the research output, which if not observed, will disrupt the process of research use and effectiveness.

Although Vanholsbeeck and Lendák-Kabók^[64] proposed that effectiveness is related to the concept of accountability, this concept is currently not much considered and dominant in the university evaluation culture; however, in the present study, it has taken a significant part. Hinrichs-Krapels and Grant^[60] posed some questions as follows: whether the research has been effective or not? Does research input lead to output, outcome and impact? Has it had sufficient efficiency? To what extent the research input lead to the research output? Has the research achieved its desired goals? These questions show the importance of paying attention to research effectiveness, which is mentioned in various sources. For example, Gibson and Hazelkorn^[61] focused on the prioritization of arts and humanities research based on their potential for production, economic growth, and job creation to overcome the economic crisis. In this regard, researches that are interdisciplinary and related to social issues are preferred. Paying attention to the knowledge economy helps redefine the goal of higher education and emphasizes the need to design policies for the future of research. Oancea, Florez Petour and Atkinson^[51] emphasized the need for qualitative criteria and indicators to configure the cultural value and impact of research. Sörlin^[62] pointed to the humanities of transformation, suggesting that research policies should be changed to expand the values of humanities and increase research effectiveness. This can be enabled by integrative humanities using interdisciplinary research.

The findings confirm that the impact of humanities has gone beyond the scientific dimension and has been directed to various economic, social, cultural and political aspects. From the obtained criteria and indicators, it is clear that the evaluation of research outputs in humanities has gone beyond the dependence on the academic environment and has tended to influence in wider dimensions. This lays the groundwork for the application of humanities in society and increases the need for evaluation in order to obtain the rate of progress and transformation of research. In other words, it is not only the publication of the research that is important, but its use and effectiveness should also be considered in order to bring about the well-being of the target society. Therefore, according to the nature of each field and the goal of research, it is necessary to choose the appropriate criteria for evaluating research outputs.

Among the beneficiaries of this study are policymakers in the field of research, who can adopt a new approach in decision-making and establishing policies for the evaluation of research outputs of humanities and related fields according to the results. Giving importance to impact as an evaluation component has created

a major evolution in research evaluation systems. Creating an impact discourse and understanding its importance requires creating an impact infrastructure in the research process, on the one hand, and stabilizing the impact position through applying and updating the impact and eliminating weak points, on the other hand.^[32] Understanding the concept of research value from the perspective of its stakeholders is also effective in evaluation and clarifies the nature of expected effects.^[75] The criteria for research effectiveness make the researcher responsible in different fields. Paying attention to the impact is the stimulus and reinforcement of research, and following the quantitative and qualitative increase of research, its impact and efficiency will increase. If the problem of impact and impact evaluation is to remain stable, a wider range of internal motivations should be considered and the ability of academics to coordinate with them should also be improved.^[76] Criteria and indicators act like a filter; not every research can adapt to them, and therefore low-quality studies are excluded from the cycle. Over time, this mechanism is hoped to make humanities research more targeted and more efficient. In order to make this planning more practical, it is suggested to examine the research outputs according to each field, to assign suitable criteria and indicators for each research output (these two issues are under study for the fields of language and literature), to determine the maximum and minimum standard limits for each indicator, and to localize the indicators according to the conditions of each research center.

The movement of universities from the first generation to the fourth generation requires providing its infrastructure and ecosystem in each country. Changing the approach of evaluation criteria and indicators and their compatibility with the paradigms of education, research, entrepreneurship, and influencing society are among the things affected by this generational change in universities. This change of paradigms should be considered in all fields, including humanities, which play a decisive role in society. The criteria and indicators examined in this research tend to the characteristics of the fourth generation of universities and evaluation in it, which can be a suitable suggested model for those involved in the evaluation of humanities and especially language and literature fields. According to Sivertsen,^[26] publication patterns are different between humanities fields, while these patterns are similar for each field in different countries. This is a positive point for using the results of this study in the fields of language and literature in other regions.

In short, according to the opinion of experts in language and literature fields, what should be considered in the policy of evaluating the scientific outputs of these fields is the difference of scientific outputs and their evaluation criteria according to the nature of the field and its research approaches and goals. The diversity of approaches and goals, which is the starting point of the method proposed in this study for evaluation, will cause the expansion of research to the environment outside the university,

the application of research and scientific studies in society, and be useful to the public. This research aimed to introduce the evaluation criteria of scientific outputs in different dimensions in order to take an effective step in changing the view of language and literature and improving the status of this discipline and humanities. The realization of such a view on evaluation in humanities requires an effort to spread this style of evaluation and provide the necessary infrastructure for it at micro and macro levels. It is hoped that the conducted study will be an introduction to this matter.

FUNDING

This work was supported by the Iran National Science Foundation (INSF) [grant number 99030052].

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

- Ginting B, Chiari W, Duta TF, Hudaa S, Purnama A, Harapan H, *et al.* COVID-19 pandemic sheds a new research spotlight on antiviral potential of essential oils – A bibliometric study. *Heliyon*. 2023;9(7):e17703. doi: 10.1016/j.heliyon.2023.e17703, PMID 37456016.
- Kazemini MH, Honarvar MA. The subject and nature of humanities and its place in the system of sciences with emphasis on Farabi's views and ideas, Qabasat. 2017; 19/74: 77-88.
- Lanzillo L. The social impact assessment in social sciences and humanities: methodological issues from the Italian experience. In: Bonaccorsi A, editor. *The evaluation of Research in social sciences and humanities: lessons from the Italian experience*. Springer; 2018. p. 345-59.
- Blasi B, Romagnosi S, Bonaccorsi A. Do SSH researchers have a third mission (and should they have)? In: Bonaccorsi A, editor. *The evaluation of Research in social sciences and humanities: lessons from the Italian experience*. Springer; 2018. p. 361-92.
- Olmos-Peñuela J, Benneworth P, Castro-Martinez E. Sci Public Policy. 2014 'Are "STEM from Mars and SSH from Venus"?': Challenging disciplinary stereotypes of research's social value';41/3: 384-400.
- Galleron I, Ochsner M, Spaapen J, Williams G. 'Valorizing SSH research: Towards a new approach to evaluate SSH research' value for society. *fteval Journal for Research and Technology Policy Evaluation*. 2017;44:35-41. doi: 10.2163/fteval.2017.274.
- Moosa IA. Publish or perish: perceived Benefits *versus* Unintended Consequences. *Publish or Perish: perceived Benefits versus Unintended Consequences*. Edward Elgar Publishing; 2018.
- Hudaa S. Optimalisasi Bahasa: penggunaan Bahasa yang Baik, Logis, dan Santun di Media Massa. *Dialektika J Bahasa Sastra Pendidikan Bahasa Sastra Indones*. 2018;5(1):62-74. doi: 10.15408/dialektika.v5i1.5953.
- Brutscher P-B, Wooding S, Grant J. *Health Research Evaluation Frameworks: an international comparison*. United Kingdom: Rand Corporation; 2008.
- Hazelkorn E. Pros and cons of research assessment: lessons from rankings. *Dublin Institute of Technology*; 2009.
- Milzow K, Reinhardt A, Söderberg S, Zinöcker K. Understanding the use and usability of research evaluation studies. *Res Eval*. 2019; 28/1: 94-107.
- OECD. Enhancing public research performance through evaluation, impact assessment and priority setting. Retrieved from; 2010.
- Coryn CLS. *Evaluation of researchers and their research: toward making the implicit explicit*. Western Michigan University; 2007.
- Bonaccorsi A, Daraio C, Fantoni S, Folli V, Leonetti M, Ruocco G. Do social sciences and humanities behave like life and hard sciences? *Scientometrics*. 2017;112(1):607-53. doi: 10.1007/s11192-017-2384-0.
- Rovira-Esteva S, Orero P. Evaluating quality and excellence in translation studies research: publish or perish, the Spanish way. *Babel*. 2012;58(3):264-88. doi: 10.1075/babel.58.3.02rov.
- Wilsdon J, Allen L, Belfiore E, Campbell P, Curry S, Hill S, *et al.* *The metric tide: report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Higher Education Funding Council for England; 2015.
- Corse S. M. Literature and society. In: Borgatta Edgar F, Montgomery Rhonda JV, editors. *Encyclopedia of sociology*; 2000.

18. Ebrahimi Dorcheh E, Mansouri A, Pashootanizadeh M, Mirbagheri Fard A, Shabani A. Harms and strategies for evaluating humanities research outputs: A case study of language and literature. *Scientometr Res J*. 2023;9(2):73-96. doi: 10.22070/RSCI.2022.16133.1583.
19. Heinen JT. The importance of a social science research agenda in the management of protected natural areas, with selected examples. *Bot Rev*. 2010;76(2):140-64. doi: 10.1007/s12229-010-9043-y.
20. Pedersen DB, Grønvad JF, Hvidtfeldt R. Methods for mapping the impact of social sciences and humanities—A literature review. *Res Eval*. 2020;29(1):4-21. doi: 10.1093/reseval/rvz033.
21. Borovik MA, Shemberko LV. The challenges of information retrieval in social sciences and humanities and ways to overcome information barriers. *Sci Tech Inf Process*. 2016;43(2):99-105. doi: 10.3103/S0147688216020039.
22. Faggiolani C, Solimine G. Mapping the role of the book in evaluation at the individual and department level in Italian SSH. A multisource analysis. In: Bonaccorsi A, editor. *The evaluation of Research in social sciences and humanities: lessons from the Italian experience*. Springer; 2018. p. 33-53.
23. Verleysen FT, Weeren A. Clustering by publication patterns of senior authors in the social sciences and humanities. *J Inf*. 2016;10(1):254-72. doi: 10.1016/j.joi.2016.01.004.
24. Bonaccorsi A. Towards an epistemic approach to evaluation in SSH. In: Bonaccorsi A, editor. *The evaluation of Research in social sciences and humanities: lessons from the Italian experience*. Springer; 2018. p. 1-29.
25. Kulczycki E, Guns R, Pölönen J, Engels TCE, Rozkosz EA, Zuccala AA, et al. Multilingual publishing in the social sciences and humanities: A seven-country European study. *J Assoc Inf Sci Technol*. 2020;71(11):1371-85. doi: 10.1002/asi.24336, PMID 33288998.
26. Sivertsen G. Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*. 2016;107:357-68. doi: 10.1007/s11192-016-1845-1, PMID 27122643.
27. Ochsner M, Hug SE. Indicators for research performance in the humanities? The scholars' view on research quality and indicators. *International Conference on Science and Technology Indicators*. València, Spain; 2016. p. 1-8.
28. Leydesdorff L. Caveats for the use of citation indicators in research and journal evaluations. *J Am Soc Inf Sci Technol*. 2008;59(2):278-87. doi: 10.1002/asi.20743.
29. Bunia R. Quotation statistics and culture in literature and in other humanist disciplines. In: Ochsner M, Hug SE, Daniel H-D, editors. *Research assessment in the humanities: towards criteria and procedures*. Cham: Springer; 2016. p. 133-48.
30. Hammarfelt B, Haddow G. Conflicting measures and values: how humanities scholars in Australia and Sweden use and react to bibliometric indicators. *Asso for Info Science and Tech*. 2018;69(7):924-35. doi: 10.1002/asi.24043.
31. Van den Akker W. Yes we should; research assessment in the humanities. In: Ochsner M, Hug SE, Daniel H-D, editors. *Research assessment in the humanities: towards criteria and procedures*. Cham: Springer; 2016. p. 23-9.
32. Wróblewska MN. Research impact evaluation and academic discourse. *Humanit Soc Sci Commun*. 2021;8(1):1-12. doi: 10.1057/s41599-021-00727-8.
33. Lauronen JP. Tension in interpretations of the social impact of the social sciences: walking a tightrope between divergent conceptualizations of research utilization. *SAGE Open*. 21582440221089970. 2022;12(2). doi: 10.1177/21582440221089967.
34. Sile L, Pölönen J, Sivertsen G, Guns R, Engels TC, Arefiev P, et al. Comprehensiveness of national bibliographic databases for social sciences and humanities: findings from a European survey. *Res Eval*. 2018;27(4):310-22. doi: 10.1093/reseval/rvy016.
35. Lin CS. An analysis of citation functions in the humanities and social sciences research from the perspective of problematic citation analysis assumptions. *Scientometrics*. 2018;116(2):797-813. doi: 10.1007/s11192-018-2770-2.
36. Wei M, Noroozi Chakoli A. Evaluating the relationship between the academic and social impact of open access books based on citation behaviors and social media attention. *Scientometrics*. 2020;125(3):2401-20. doi: 10.1007/s11192-020-03678-0.
37. Nixon JM. Core journals in library and information science: developing a methodology for ranking LIS journals. *Coll Res Libr*. 2014;75(1):66-90. doi: 10.5860/crl12-387.
38. Schwekendiek DJ. Trends in Korean studies: A content analysis of Korea-related articles published in the Arts and Humanities Citation Index, 1990-2015. *Int Area Stud Rev*. 2020;23(4):325-34. doi: 10.1177/2233865920934954.
39. Leydesdorff L, Rafols I. A global map of science based on the ISI subject categories. *J Am Soc Inf Sci Technol*. 2009;60(2):348-62. doi: 10.1002/asi.20967.
40. Cicero T, Malgarini M. On the use of journal classification in social sciences and humanities: evidence from an Italian database. *Scientometrics*. 2020;125(2):1689-708. doi: 10.1007/s11192-020-03581-8.
41. Ku M-C. A comparative analysis of English abstracts and summaries of Chinese research articles in three library and information science journals indexed by the Taiwan social science citation index. *J Libr Inf Stud*. 2019; 17/1: 37-81.
42. Abramo G, Aksnes DW, D'Angelo CA. Comparison of research performance of Italian and Norwegian professors and universities. *J Inf*. 2020;14(2):1-23. doi: 10.1016/j.joi.2020.101023.
43. Gumpenberger C, Sorz J, Wieland M, Gorraiz J. Humanities and social sciences in the bibliometric spotlight – research output analysis at the University of Vienna and considerations for increasing visibility. *Res Eval*. 2016;25(3):271-78. doi: 10.1093/res eval/rvw013.
44. Zemits BI. Representing knowledge: assessment of creativity in humanities. *Arts Humanit Higher Educ*. 2017;16(2):173-87. doi: 10.1177/1474022215601862.
45. Ochsner M, Hug SE, Daniel HD 2016. Humanities scholars' conceptions of research quality. *Research assessment in the humanities: towards criteria and procedures*, pp. 43-69.
46. Krampen G. The evaluation of university departments and their scientists: some general considerations with reference to exemplary bibliometric publication and citation analyses for a department of psychology. *Scientometrics*. 2008;76(1):3-21. doi: 10.1007/s11192-007-1900-z.
47. Rowlands J, Wright S. Hunting for points: the effects of research assessment on research practice. *Stud Higher Educ*. 2019; 46/9: 1801-15.
48. Ren Q, Gong X. Evaluation index system for academic papers of humanities and social sciences. *Scientometrics*. 2012;93(3):1047-60. doi: 10.1007/s11192-012-0790-x.
49. Pajić D. Globalization of the social sciences in Eastern Europe: genuine breakthrough or a slippery slope of the research evaluation practice? *Scientometrics*. 2015;102(3):2131-50. doi: 10.1007/s11192-014-1510-5.
50. Kulczycki E, Rozkosz EA. Does an expert-based evaluation allow us to go beyond the Impact Factor? Experiences from building a ranking of national journals in Poland. *Scientometrics*. 2017;111(1):417-42. doi: 10.1007/s11192-017-2261-x, PMID 28386153.
51. Oancea A, Florez Petour T, Atkinson J. Qualitative network analysis tools for the configurational articulation of cultural value and impact from research. *Res Eval*. 2017;26(4):302-15. doi: 10.1093/reseval/rvx014.
52. Sile L, Vanderstraeten R. Measuring changes in publication patterns in a context of performance-based research funding systems: the case of educational research in the University of Gothenburg (2005-2014). *Scientometrics*. 2019;118(1):71-91. doi: 10.1007/s11192-018-2963-8.
53. Haddow G, Hammarfelt B. Quality, impact, and quantification: indicators and metrics use by social scientists. *J Assoc Inf Sci Technol*. 2019;70(1):16-26. doi: 10.1002/asi.24097.
54. Doğan G, Taşkın Z. Humanities: the outlier of research assessments. *Information*. 2020;11(11):1-11. doi: 10.3390/info11110540.
55. D'Souza R. What characterises creativity in narrative writing, and how do we assess it? Research findings from a systematic literature search. *Thinking Skills Creativity*. 2021;42:1-28. doi: 10.1016/j.tsc.2021.100949.
56. Yang S, Zheng M, Yu Y, Wolfram D. Are Altmetric.com scores effective for research impact evaluation in the social sciences and humanities? *J Inf*. 2021;15(1):1-21. doi: 10.1016/j.joi.2020.101120.
57. Spaapen J, van Drooge L. Introducing "productive interactions" in social impact assessment. *Res Eval*. 2011;20(3):211-8. doi: 10.3152/095820211X12941371876742.
58. Bornmann L. What is societal impact of research and how can it be assessed? A literature survey. *J Am Soc Inf Sci Technol*. 2013;64(2):217-33. doi: 10.1002/asi.22803.
59. Benneworth P. Tracing how arts and humanities research translates, circulates and consolidates in society. How have scholars been reacting to diverse impact and public value agendas? *Arts Humanit Higher Educ*. 2015;14(1):45-60. doi: 10.1177/1474022214533888.
60. Hinrichs-Krapels S, Grant J. Exploring the effectiveness, efficiency and equity (3e's) of research and research impact assessment. *Palgrave Commun*. 2016;2(1):1-9. doi: 10.1057/palcomms.2016.90.
61. Gibson AG, Hazelkorn E. Arts and Humanities Research, redefining public benefit, and research prioritization in Ireland. *Res Eval*. 2017;26(3):199-210. doi: 10.1093/reseval/rvx012.
62. Sörlin S. Humanities of transformation: from crisis and critique towards the emerging integrative humanities. *Res Eval*. 2018; 27/4: 287-97. doi: 10.1093/reseval/rvx030.
63. Janinovic J, Pekovic S, Vuckovic D, Djokovic R, Peic M. Innovative strategies for creating and assessing research quality and societal impact in social sciences and humanities. *Interdiscip Descript Complex Syst*. 2020;18(4):449-58. doi: 10.7906/ind ecs.18.4.5.
64. Vanholsbeeck M, Lendák-Kabók K. Research impact as a "boundary object" in the social sciences and the humanities. *Word Text*. 2020;10:29-52.
65. Sani'ee N, Nemati-Anaraki L, Sedghi S, Noroozi Chakoli A, Goharinezhad S. The effective trends and driving forces in the future of research performance evaluation: A qualitative study. *Med J Islamic Repub Iran*. 2022; 36/1: 393-406.
66. Thelwall M, Delgado MM. Arts and humanities research evaluation: no metrics please, just data. *J Doc*. 2015;71(4):817-33. doi: 10.1108/JD-02-2015-0028.
67. Hug SE, Ochsner M, Daniel H-D. Criteria for assessing research quality in the humanities-A Delphi study among scholars of English literature, German literature and art history. *Res Eval*. 2013;22(5):369-83. doi: 10.1093/reseval/rvt008.
68. Ochsner M, Hug SE, Daniel H-D. Setting the stage for the assessment of research quality in the humanities. Consolidating the results of four empirical studies. *Z Erziehungswiss*. 2014;17(S6):111-32. doi: 10.1007/s11618-014-0576-4.
69. Von der Gracht HA. Consensus measurement in Delphi studies. Review and implications for future quality assurance. *Technol Forecasting Soc Change*. 2012; 79/8: 1525-36.
70. Hogarth RM. A note on aggregating opinions. *Organ Behav Hum Perform*. 1978;21(1):40-6. doi: 10.1016/0030-5073(78)90037-5.

71. Clayton MJ. Delphi: A technique to harness expert opinion for critical decision-making tasks in education. *Educ Psychol.* 1997;17(4):373-86. doi: 10.1080/0144341970170401.
72. Habibi A, jahantigh F, F, Sarafrazi A. Fuzzy Delphi technique for forecasting and screening items. *Asian J Res Bus Econ Manag.* 2015; 5/2: 130-43.
73. English JM, Kernan GL. The prediction of air travel and aircraft technology to the year 2000 using the Delphi method. *Transp Res.* 1976;10(1):1-8. doi: 10.1016/0041-1647(76)90094-0.
74. Aksnes DW, Langfeldt L, Wouters P. Citations, citation indicators, and research quality: an overview of basic concepts and theories. *SAGE Open:* 2158244019829575. 2019;9(1). doi: 10.1177/2158244019829575.
75. Molas-Gallart J. Research evaluation and the assessment of public value. *Arts Humanit Higher Educ.* 2015;14(1):111-26. doi: 10.1177/1474022214534381.
76. Chubb J, Reed MS. The politics of research impact: academic perceptions of the implications for research funding, motivation and quality. *Br Polit.* 2018;13(3):295-311. doi: 10.1057/s41293-018-0077-9.

Cite this article: Ebrahimi Dorcheh E, Mansouri A, Pashootanizadeh M, Mirbagherifard AA, Shabani A. Determining and Prioritizing the Evaluation Criteria of Humanities Scientific Outputs: A Case Study of Language and Literature Fields. *J Scientometric Res.* 2024;13(1):148-60.