National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

Evaluation of Research Quality 2011-2014 (VQR 2011-2014)

Criteria for the Evaluation of Research Outputs

Group of Experts for Evaluation

Area 01 – Mathematics and Computer Science

(GEV01)

November 2015

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

1. Introduction

This document describes the organization of the Group of Experts for Evaluation (from now on, the GEV) for the Area of Mathematics and Computer Science, and the criteria the Group will use in evaluating research outputs. The document is divided in eight parts. Section 2 lists the Scientific Sectors, the "Academic Recruitment Field" (*Settori concorsuali*) and the ERC Sectors relevant for the GEV. Section 3 summarizes the internal operating rules of the GEV. Section 4 describes the evaluation criteria for the research outputs. Section 5 describes the peer review process and the guidelines for the selection of external reviewers. Section 6 describes the bibliometric criteria: the databases, the bibliometric indicators chosen by the GEV, the algorithm and calibration procedure. Section 7 indicates which are the typologies of products admissible to the evaluation for the GEV 01. Finally, Section 8 describes how the GEV plans to solve potential conflicts of interest between GEV members and authors of research outputs.

2. Delimitation of the GEV Area

The Group of Experts for Evaluation of the Area 01 will take care of the evaluation of the products submitted by researchers belonging to the *Scientific Sectors* (SSD), Academic Recruitment Field (SC) and ERC Sectors (ERC) listed in Tables 1-3.

	Area 01 – Mathematics and Computer Science Reference Scientific Sectors (SSD)
INF/01	Informatica (Computer Science)
MAT/01	Logica Matematica (Mathematical Logic)
MAT/02	Algebra (Algebra)
MAT/03	Geometria (Geometry)
MAT/04	Matematiche complementari (History of Mathematics and Mathematical Education)
MAT/05	Analisi Matematica (Mathematical Analysis)
MAT/06	Probabilità e Statistica Matematica (Probability and Mathematical Statistics)
MAT/07	Fisica Matematica (Mathematical Physics)
MAT/08	Analisi Numerica (Numerical Analysis)
MAT/09	Ricerca Operativa (Operational Research)

Table 1. Scientific Sectors (SSD) related to the Area 01

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

	Area 01 – Mathematics and Computer Science
	Reference Academic Recruitment Field (SC)
01/A1	Logica Matematica e Matematiche Complementari (Mathematical Logic, History of Mathematics and Mathematical Education)
01/A2	Geometria e Algebra (Geometry and Algebra)
01/A3	Analisi Matematica, Probabilità e Statistica Matematica (Mathematical Analysis, Probability and Mathematical Statistics)
01/A4	Fisica Matematica (Mathematical Physics)
01/A5	Analisi Numerica (Numerical Analysis)
01/A6	Ricerca Operativa (Operational research)
01/B1	Informatica (Computer Science)

Table 2. Academic Recruitment Field (SC) related to the Area 01

	Area 01 – Mathematics and Computer Science
	Reference ERC sectors
PE 1	**Mathematics:** All areas of mathematics, pure and applied, plus mathematical foundations of computer science, mathematical physics and statistics
PE 1_1	Logic and foundations
PE 1_2	Algebra
PE 1_3	Number theory
PE 1_4	Algebraic and complex geometry
PE 1_5	Geometry
PE 1_6	Topology
PE 1_7	Lie groups, Lie algebras
PE 1_8	Analysis
PE 1_9	Operator algebras and functional analysis
PE 1_10	ODE and dynamical systems
PE 1_11	Theoretical aspects of partial differential equations
PE 1_12	Mathematical physics
PE 1_13	Probability
PE 1_14	Statistics
PE 1_15	Discrete mathematics and combinatorics
PE 1_16	Mathematical aspects of computer science
PE 1_17	Numerical analysis
PE 1_18	Scientific computing and data processing
PE 1_19	Control theory and optimization
PE 1_20	Application of mathematics in sciences

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

PE 1_21	Application of mathematics in industry and society life
PE 6	**Computer Science and Informatics:** Informatics and information systems, computer science, scientific computing, intelligent systems
PE 6_1	Computer architecture, pervasive computing, ubiquitous computing
PE 6_2	Computer systems, parallel/distributed systems, sensor networks, embedded systems, cyber-physical systems
PE 6_3	Software engineering, operating systems, computer languages
PE 6_4	Theoretical computer science, formal methods, and quantum computing
PE 6_5	Cryptology, security, privacy, quantum crypto
PE 6_6	Algorithms, distributed, parallel and network algorithms, algorithmic game theory
PE 6_7	Artificial intelligence, intelligent systems, multi agent systems
PE 6_8	Computer graphics, computer vision, multi media, computer games
PE 6_9	Human computer interaction and interface, visualization and natural language processing
PE 6_10	Web and information systems, database systems, information retrieval and digital libraries
PE 6_11	Machine learning, statistical data processing and applications using signal processing (e.g., speech, image, video)
PE 6_12	Scientific computing, simulation and modelling tools
PE 6_13	Bioinformatics, biocomputing, and DNA and molecular computation
SH 3	**The Social World, Diversity, Population:** Sociology, social psychology, demography, education, communication
SH 3_10	Social aspects of learning, curriculum studies, educational policies
SH 3_13	Science and technology studies
SH 4	**The Human Mind and Its Complexity:** Cognitive science, psychology, linguistics, philosophy of mind
SH 4_13	Philosophy of science, epistemology, logic
SH 6	**The Study of the Human Past:** Archaeology and history
SH 6_14	History of science, medicine and technologies

Table 3. The ERC sectors (ERC) relevant for the Area 01

3. Organization of the GEV

The GEV is organized as follows.

Coordinator: Marco Abate

Vice-coordinators: Nicola Fusco and Pierangela Samarati

Assistant: Silvia Columbu

3.1 Sub-GEV composition

	Coordinator	Components
Computer Science (INF/01)	Chiara Petrioli	Pierre Alliez, Vincenzo Loia, Fabio Paternò, Pierangela Samarati, Gianluigi Zavattaro
Fundamental Mathematics (MAT/01, 02, 03, 04)	Aldo Conca	Maria G. Bartolini, Matteo Viale, Stefano Vidussi
Analysis and Probability (MAT/05, 06)	Piermarco Cannarsa	David Arcoya, Veronica Felli, Nicola Fusco, Laura Lea Sacerdote, Giuseppe Savaré
Apllied Mathematics (MAT/07, 08, 09)	Luigi Preziosi	Massimo Fornasier, Nicola Mastronardi, Giovanni Rinaldi, Giuseppe Toscani

Table 4. Sub-GEV, relative scientific sectors (SSD), coordinators and components

3.2 Assignment of research outputs within the GEV

To each research output are assigned two SSD: the author's SSD, and the research output's SSD. The latter, assigned by the author, may differ from the former because it indicates the SSD that, according to the author, represents better the content of the product.

The assignment of outputs to a GEV depends on the author's SSD. If the SSD assigned to the research output is different from the author's SSD, the GEV, on the basis of the research output's SSD, may decide to pass the product to another GEV more closely related to the content of the research output; in this case the evaluation will be performed according to the criteria decided by the GEV of destination.

The GEV01 will subdivide research outputs by type of publication and by research area, and will assign them to the appropriate Sub-GEV according to the research outputs' SSD, unless the content of the research output suggests the assignment to a different, more competent, Sub-GEV. The Sub-GEV Coordinator will then assign them to two competent Sub-GEV members.

If a research output is assigned to more than one GEV (for instance because co-authors indicate different SSD belonging to different GEVs), it will be evaluated according to the VQR Guidelines for the Groups of Evaluation Experts (Section 3.2): if the two (or more) GEV Coordinators involved in the evaluation of a specific research output think that it is interdisciplinary and it provides a comparable scientific contribution in the different areas, they

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

agree upon an interGEV evaluation procedure forming specific Inter-Area Consensus Groups in order to give it a unique overall evaluation. If instead the two (or more) Coordinators involved think that the scientific contribution of the research output is substantially different in the respective areas, each of them will evaluate the research output by following the procedures of his/her own GEV. In this way the evaluation could be different for each of the areas involved (for instance, this might happen if the research output gives a significant contribution to the biology field by applying well known mathematical methodologies). For this reason, it is suggested to the Institutions to associate to the submitted research output the SSD most relevant for the evaluation, independently, whenever necessary, of the author's SSD.

3.3 Operating rules of the GEV

The operating rules of the GEV are recalled as follows:

- a GEV meeting is called with an at least 15 days notice. The meeting is called by the Coordinator, who also sets the agenda;

- decisions within the GEV are made by simple majority rule among members who attend a meeting. In order to vote, physical presence is not required if presence is assured via web or phone connection;

- the Assistant assigned by ANVUR to the GEV attends to the GEV meetings, with secretariat functions and without voting rights. At the end of each meeting, minutes and a synthetic report outlining the main decisions will be drafted, circulated among GEV members, approved by the Coordinator and by the members, and then sent to ANVUR to be filed.

4. The evaluation of research outputs

The evaluation of research outputs by the GEV follows the *informed peer review* methodology, which consists in employing different, and if possible mutually independent, evaluation methods, to be harmonized within the GEV, which ultimately remains responsible for the final evaluation.

The available evaluation tools are:

- *External peer review* evaluation by (usually two) external reviewers selected independently by two different GEV members.

- *Internal peer review* by the GEV, which can conduct a (completely or in part) internal peer review according to the same procedure described for external peer.

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

- *Bibliometric analysis*, to be conducted according to the procedure described below in this document (Section 6). Research outputs subject to bibliometric analysis are not assigned *automatically* (that is automatically using the class suggested by the application of the bibliometric algorithm) to the classes of merit established by the Ministerial Decree (DM) and by the VQR Call. The allocation is instead based on the expert judgment of the GEV, which will consider all possible evaluation elements beside bibliometric indicators, such as the expertise of its members and the information described in the form associated to the research outputs.

Among the typologies of research outputs listed in Section 2.3 of the ANVUR VQR 2011-2014 Call considered eligible for the evaluation by the GEV 01 (listed in Section 7 of this document), the outputs of typology 1 (Scientific monograph and related products), 3 (Contributions in books), 4 (Other scientific research outputs) and 5 (Patents) will be evaluated by *peer review,* internal or external. Among research outputs of typology 2 (Contributions in journals) at least 10% of them will be evaluated in *peer review,* as described in Section 5; the remaining research outputs will be evaluated using bibliometric analysis within the Sub-GEVs as described in Section 6.

The selection of research outputs of typology 2 to be assessed by *peer review* will be done by the competent Sub-GEV on the basis of the following criteria:

- Availability of bibliometric data for that research output;

- Indications coming from the application of the bibliometric analysis (see Section 6.6);

- Eventual indications provided by the submitting Institutions in the CINECA interface (outputs for which the request of evaluation in *peer review* is well justified will be more likely evaluated by *peer review*);

- Publication date of the research output (recent outputs are more likely to be evaluated by *peer review*);

- Research output typology (outreach articles will be evaluated by *peer review;* reviews and surveys are more likely evaluated by *peer review*);

- As a rule, uniform distribution of the percentage of outputs sent to *peer review* among the different institutions;

- As a rule, uniform distribution of the percentage of *peer review* among the different SSDs (with the exception of the outputs associated to the SSD MAT/04; since this sector is

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

more akin to humanities for research and publications characteristics, its outputs will be evaluated exclusively by *peer review*).

In addition, research outputs published by journals known for having adopted tricks to artificially increase their bibliometric indicators will be assessed by *peer review.*

Particular attention will be devoted to research outputs classified as interdisciplinary by the submitting institutions, with the aim of choosing the most appropriate evaluation procedure for each of them.

Finally, conference proceedings (in particular those associated to the SSD INF/01) will be assessed by *informed peer review,* taking into consideration the quality of the research output, the relevance of the conference, and, if there are any, citation data.

5. Peer review evaluation

Each research output to be evaluated by *peer revie*w will be sent to two external reviewers, independently chosen by the two GEV members to whom the research output was assigned.

Alternatively, a research output might be evaluated within the GEV according to the same procedure, provided that the necessary expertise is available and that no conflict of interest is present.

5.1   The selection of external peer reviewers

The selection of external reviewers, among Italian and foreign scholars, given its relevant public interest goals, follows the principle of loyal institutional cooperation and is based upon the criteria of correctness, objectivity and impartiality.

Great attention will be devoted to maintaining the anonymity of the reviewers, both at the stage of preparation of the list of reviewers and at the operational stage of the evaluation. The results of the evaluation of individual research outputs and the identity of the associated reviewers will not be made public. The list of the reviewers' names will be published by ANVUR within 30 days of the publication of the VQR Final Report.

Reviewers will be selected among the most authoritative and scientifically qualified scholars and specialists in the disciplines relevant to the products to be examined. They are expected to have been active in research during the period covered by the VQR.

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

Starting from the MIUR reviewer archive REPRISE, the GEV will prepare an updated list of external reviewers such to adequately satisfy the standards set by the GEV in terms of scientific quality and experience with evaluation. The list will be extended with new reviewers selected by the GEV. In particular, through the Sub-GEV Coordinators, the Coordinator will invite GEV members to suggest a significant number of experts who satisfy the required standards and are available for the evaluation. The GEV Coordinator will collect suggestions together with information about the reviewers' qualifications, as summarized in a specific form, and s/he will update the initial list with integrations and/or cancellations.

It will be possible to extend the reviewer list throughout the evaluation procedure, on the basis of the needs that may emerge after the research outputs are transmitted by institutions.

In order to reduce potential conflicts of interest, the GEV will use, whenever possible, reviewers that are active in foreign universities and institutions.

5.2   Peer evaluation

The evaluation, by external or internal reviewers, is based on an evaluation form and a set of instructions for the reviewers to be prepared by the GEV, following, if necessary, the guidelines provided by the evaluation research group set up by ANVUR during the months before the opening of the VQR. The evaluation form will allow the reviewer to assign a score to the three evaluation criteria established by the DM and the VQR Call (that is, originality, methodological rigor, and attested or potential impact); the sum of the three scores will give the evaluation in one of the five classes of merit indicated in the VQR Call. The form will also include an empty field where a brief motivation of the assigned scores should be entered.

In case of non-converging evaluations by the reviewers, the Sub-GEV may request the opinion of a third expert or creates a Consensus Group with the task of proposing to the GEV the final score for the research output under examination, using the consensus report methodology. In case of conflict among the members of the Consensus Group, it will be integrated with the Sub-GEV Coordinator or with the Coordinator of the GEV.

In any case the GEV ultimately remains responsible for the final evaluation.

6. Bibliometric analysis

The research outputs that will be evaluated in bibliometric analysis are those indexed in the citation databases Web of Science by Thomson Reuters (WoS) and Scopus by Elsevier (Scopus) and in particular:

- Scientific articles presenting original results published in scientific journals, that in principle could also be *Conference Papers;*

- Scientific articles of critical review of the literature (*Reviews* or *Surveys*).

A random sample of at least 10% of the journal articles evaluated with bibliometric analysis will be also sent to peer review, in order to identify the degree of agreement between the two methods of evaluation. The sample will be formed by randomly choosing papers among all Sub-GEVs, and it will not include papers for which the bibliometric algorithm indicated the needs of an *informed peer review*. For the papers in the sample, the peer review will not be used for the final evaluation, that will exclusively depend on the bibliometric analysis.

6.1 Databases

The GEV will use the databases WoS and Scopus following the indications given by the author/institution in the research output description form. Only for measuring the impact of journals, the GEV will also use the database MathSciNet of the American Mathematical Society, as illustrated in Sections 6.5, 6.6 and 6.7. For calculating the number of citations only the databases WoS and Scopus will be used. In fact, for a correct evaluation is necessary to have collected the number of citations up to a fixed and known date, that must be the same for all research outputs; that requirement can be achieved with the databases WoS and Scopus, but not with MathSciNet because the American Mathematical Society, when asked, communicated that such an use of its database would contravene the agreements it has with publishers, that provide the reprints of the papers under the condition that they are used for the institutional aims of the database only, aims which are informative and not pertaining evaluation procedures.

6.2 Citation time span

To compute bibliometric indicators, the GEV will use citations up to February 29, 2016.

6.3 Self-citations

The treatment of self-citations in bibliometric evaluations is still an object of debate within the scientific community. In the VQR 2011-2014, the GEV01 has decided, as suggested by the

ANVUR workgroup on bibliometric evaluation formed during the first plenary meeting of GEV coordinators, to not exclude self-citations in the application of the bibliometric procedure. However, papers showing a number of self-citations higher than the 50% of the total number of citations will be analyzed more carefully. The final decision about the class of merit of such research outputs will also take into account the information provided by the author in the research output description form, recurring, whenever deemed necessary, to (external or internal) informed peer review.

6.4 Bibliometric indicators

The evaluation will use, for all articles published in journals indexed in the databases WoS and Scopus, an algorithm which takes into account, in a different way depending on the date of publication of the paper, both the number of citations and a measure of the journal impact (Journal Metric, JM) of the hosting journal.

Coherently with the orientation of the international scientific community in the bibliometric field, and keeping into account the differences among the ways different indicators measure journal impacts, the GEV 01, by following the indications coming from the ANVUR workgroup on bibliometric evaluation, has decided to use more than one measure of impact. There exist two kinds of impact indicators. The first kind has the goal of measuring the *popularity* of the hosting journal (therefore the citations received are considered independently from their origin); the second kind intends to measure the *prestige* of the hosting journal (therefore the citations received are weighted on the basis of the scientific authoritativeness of the source journal). Examples of popularity indicators are the *Impact Factor* (IF) in WoS, the *Impact per Publication* (IPP) and the *Source Normalized Impact per Paper* (SNIP) in Scopus; examples of prestige indicators are the *Article Influence* (AI) in WoS and the *SCImago Journal Rank* (SJR) in Scopus. The *Mathematics Citation Quotient* (MCQ) in MathSciNet is a popularity indicator that collects citations only coming from a selected set of journals (*reference journals*), and not from the whole set of indexed journals. More information about these indexes can be found in the following websites: http://www.webofknowledge.com for WoS, http://www.journalmetrics.com for Scopus, and http://www.ams.org/mathscinet/help/citation_database_help_full.html for MathSciNet.

The GEV 01 has decided to use the following impact indicators:

- AI for WoS;

- SNIP and SJR for Scopus;

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

- MCQ for MathSciNet.

IF and IPP have been excluded since it has been verified that the measuring coming from pure impact indicators, not normalized on the area (SNIP is) or computed without a careful selection tailored on Mathematics of source journals (thit is what happens for the MCQ), can not give an adequate measure of the impact of journals in Mathematics or Computer Science.

The procedure used to select the impact indicator to be used in the evaluation of each research output, which depends on the SSD, on the characteristics of the research output, as well as on the database chosen by the author and on possible extra indications given in the research output description form, will be described in Section 6.7.

6.5   The algorithm for classifying the research outputs

The algorithm used for the classification of the articles in the 5 classes of merit defined in the VQR Call is based on a combined use of an indicator (JM) of the impact of the journal in which the article was published and of the number of citations (CIT) that measure the impact of the single article. Depending on the publication year the first or the second indicator can have a higher relative weight. Each article is evaluated within a specific reference category (more details below), and relatively to the publication year. The evaluation procedure in the reference category is previously calibrated in order to ensure that the probability *ex ante* at the world level of each article, of a particular category and of a given year, to fall into one of the evaluation classes is the one defined by the VQR Call (for the GEV 01 the sub-area is identified by the SSD):

- Excellent [top 10% of the distribution of the international scientific research outputs in the relevant sub-area];

- Good [10% – 30% of the distribution of the international scientific research outputs in the relevant sub-area];

- Fair [30% – 50% of the distribution of the international scientific research outputs in the relevant sub-area];

- Acceptable [50% – 80% of the distribution of the international scientific research outputs in the relevant sub-area];

- Limited [80% – 100% of the distribution of the international scientific research outputs in the relevant sub-area].

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

It is important to highlight that the indication of the percentiles in the definition of the classes of merit does not refer to the percentage results expected from the assessment of the research outputs presented for the VQR, but to the world production in the area. The final results of the evaluation could have (and it is expected to have) a different percentage distribution. Furthermore, the assessment of single items is not comparative: each research output will be assigned to a class of merit independently of the positioning of the other research outputs.

The first step in the evaluation of a given article is the identification of the *reference category*, that depends on the journal hosting the research output. For the Area 01, the GEV 01 has identified a set of reference categories, one for each SSD, that integrates the *Subject Categories* (SC) in WoS and the *All Science Journal Classification* (ASJC) in Scopus, because SC and ASJC do not correctly represent the internal subdivisions of the Mathematics and Computer Science research fields.

As for subject categories in WoS and the ASJC in Scopus, to each reference category is associated a list of journals, indexed in WoS and/or Scopus and relevant for the corresponding research area. The lists (attached to this document) were compiled as described in Section 6.7.

The reference category of a given scientific paper is usually determined by the SSD assigned to the paper by the institution, unless the GEV, according to the content of the paper, decides that it is more properly assigned to a different category (whose corresponding list of journals contains the journal hosting the paper). Analogously, it will be the GEV, always according to the content of the article, to choose the reference category to be used if the journal list associated to the SSD of the research output does not include the journal hosting it.

The bibliometric procedure of assessment is build as to use all the available instruments in order to evaluate in most correct way articles addressing core topics in the area. For a correct and consistent evaluation of articles on the boundary of the research field or interdisciplinary, it is fundamental to take into account and respect the specific characteristics of each single item; for this reason these articles will be evaluated in *informed peer review*, meaning that the bibliometric analysis might be associated to a *peer review* (internal or external) in order to detect their specific Mathematical or Computer Science contribution. Whenever a paper submitted to the GEV is published in a journal not included in any of the journal lists of the reference categories of the GEV 01, the competent Sub-GEV will choose the WoS *subject category* and the Scopus ASJC containing the journal and

14

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

more appropriate to the content of the paper. The quality of that research output will be then assessed with bibliometric analysis performed using the chosen subject category and/or ASJC, rather than the reference categories of the GEV01.

Finally, papers published in journals that are not indexed in any of the WoS or Scopus databases, will be only evaluated by *peer review*.

The bibliometric evaluation of each article will be based, applying the procedure extensively described in Section 6.6, on the number of citations received by the article and on an impact indicator of the hosting journal. The impact index to be used depends on the characteristics of the research output, on the reference category and on the database in which the research output is indexed, as illustrated in Section 6.7. The number of citations is given by the database (WoS and/or Scopus) chosen by the author among those in which the article is indexed when filling the research output description form (see Section 6.8 below). For research outputs associated to SSD INF/01 or MAT/01, the GEV recommends the use of Scopus, which supplies a better coverage of the areas of Computer Science and Mathematical Logic.

6.6   Evaluation procedure and its calibration

As already mentioned, the scientific articles will be assigned to one of the five classes of merit established in the VQR Call, after a calibration of the thresholds in each reference category and in each specific publication year. This procedure allows to achieve, in the world production, regardless of the sector and the year, the percentage distribution of articles in classes of merit defined in the DM and in the VQR Call.

The calibration of the bibliometric algorithm depends on the particular reference category and on the specific year analyzed. The algorithm makes a distinction between the typology of *journal article* and that of *review*, by calculating separate cumulative empirical distributions for the citations for the two typologies. In the following we shall describe the procedure only for the case of the *journal article*; an analogous procedure will be applied to *review* papers.

For each reference category, and for each impact measure JM associated to that category, the empirical cumulative distribution of JM for journals within the given reference category, and within the publication year of the article to be evaluated, is calculated; from that distribution, a percentile is assigned to each of the journals. Next, the algorithm computes the empirical cumulative distribution function of the number CIT of citations of all articles published in that year by the journals contained in reference category considered; a percentile is assigned to each of the papers. At the end of the procedure there will be two percentiles associated to each article (journal percentile and citation percentile). The two percentiles obtained with this procedure

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

identify a point in the subset $Q = [0,1]\mathrm{x}[0,1]$ of the Cartesian plane, where the *x*-axis is given by the percentile of the JM of the journal, while the *y*-axis by the citation percentile CIT. The subset Q is then divided in five regions designed so as to achieve the percentages of articles per region established in the VQR Call. These percentages are given by the ratio between the number of research outputs belonging to each region and the total number of research outputs published in journals associated to that particular reference category in that specific year.

This partition is made by plotting *threshold segments* defined by the following linear relation:

$$CIT = A \cdot JM + B_n$$

The slope (*A*, that will be negative) of the threshold segments must be the same for the entire set of segments in order to improve the homogeneity of the criterion adopted. The intercepts B_n are calculated by ANVUR, on the basis of the distribution of the particular reference category, to make sure that the percentages in the VQR Call are achieved (Figure 1 illustrates an example of the partition of Q into the five regions). Thus the algorithm allows to calibrate the evaluation in the selected set, even though the distribution of the articles changes accordingly to the category and the year.

The slope *A* of the threshold segments is decided by the GEV. The final classification is strongly dependent on the choice of the *A* value. In fact, slopes smaller than 1 in absolute value give more weight to the citation percentile, whereas slopes greater than 1 in absolute value give more weight to the percentile of the impact index of the journal. For instance, horizontal lines in Figure 1 would correspond to an evaluation entirely based upon the citation percentile, whereas vertical lines would correspond to an evaluation entirely based upon the journal impact measure.
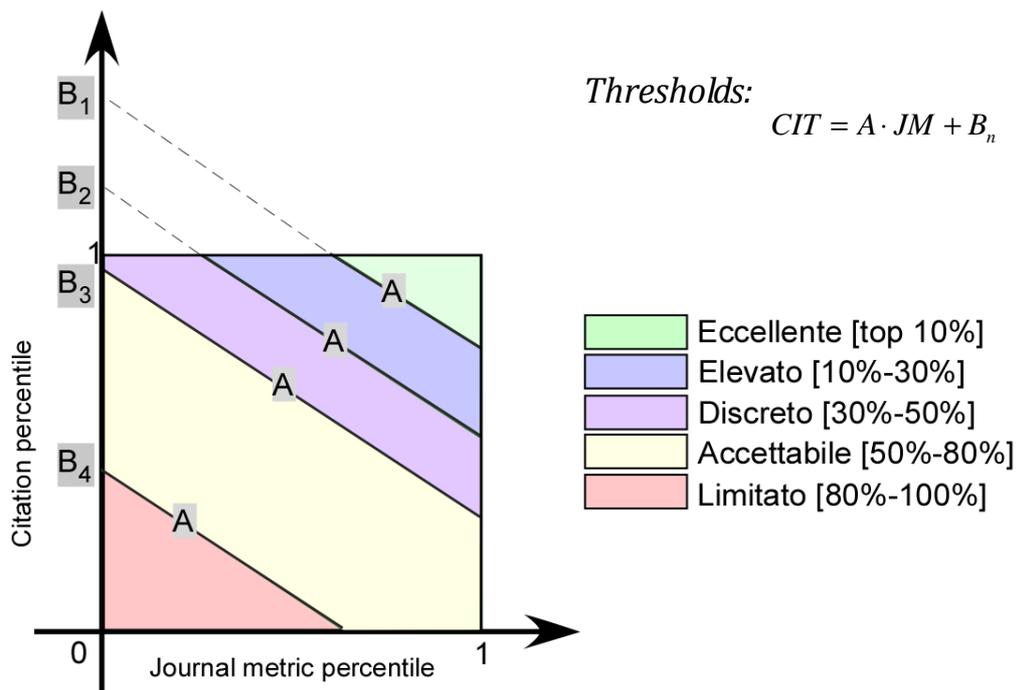
Figura 1. Percentile representation of all articles published in a given category for a particular year. Each research output is located in the plane depending on the percentile of the journal impact index JM (*x*-axis) and on the percentile of the number of citations CIT (*y*-axis). The plane is divided into five regions according to the percentages established in the VQR Call. The slope of the threshold segments delimiting the 5 areas must be the same for the entire set of segments. The intercepts B_n are calculated by ANVUR, on the basis of the distribution of the particular reference category, to ensure that the percentages in the VQR Call will be achieved.

The slopes will depend on the publication year, but, in general, they will not depend on the reference category. Considering that in Mathematics and Computer Science citations are collected quite slowly and the peer review procedures adopted by the principal journals are very strict, ensuring thus a strong correlation between the quality of the paper and the quality of the hosting journal, and after having extensively examined the results of many simulations carried out by the ANVUR work group for the bibliometric evaluation, the GEV 01 has decided to choose slopes of the threshold segments higher than one in absolute value. This choice is made in order to give more weight to the impact index of the journal hosting the articles than to the number of citations received by the paper.

More precisely, the slopes identified for the years 2011-2013 are:

Year	Value
2011	$-1.1 \pm 30\%$
2012	$-1.4 \pm 30\%$

2013	$-1.7 \pm 30\%$

The variation range has been defined as to leave open the opportunity to modify the slopes in the final calibration of the algorithm for avoiding degenerate situations. In particular, the slopes will be chosen as to avoid that articles with 0 citations could automatically achieve an "Excellent" classification.

Simulations have proved that papers published in 2014 are too recent to have collected a meaningful number of citations. For that year the slope will be chosen in the range [-2.5, -1.5], always being careful to avoid that articles with 0 citations could automatically achieve an "Excellent" classification. Furthermore, all papers published in 2014 that are not classified as "Excellent" by the algorithm will be evaluated by *informed peer review*.

Figure 2 illustrates an example of the calibration of a reference category by means of four threshold parallel lines. The slope coefficient was set equal to -0.6 in order to assign a higher weight to the citations in the final assessment. From the figure it is possible to observe that the points, identifying the articles in the reference category, are unevenly distributed. A suitable choice of the intercept values ensures that the percentages indicated by the Call are obtained, with a precision greater than one-tenth of one percent. In other words, when the bibliometric algorithm is applied to "world" production of journal articles the percentages established in the DM and in the Call are obtained. It follows that each specific paper submitted to the VQR will always be evaluated with respect to the percentile of the "scientific international production within the area it belongs to".
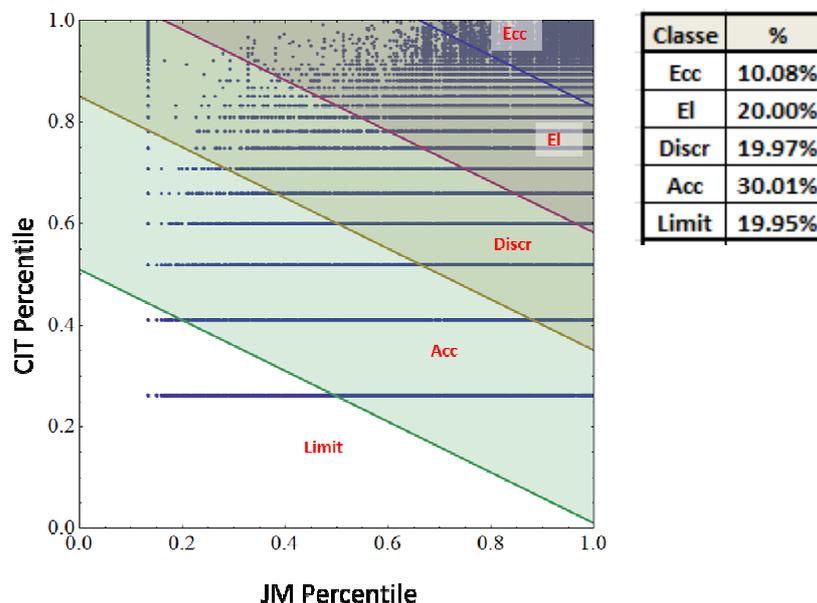
Classe	%
Ecc	10.08%
El	20.00%
Discr	19.97%
Acc	30.01%
Limit	19.95%

Figura 2. Example of application of the bibliometric algorithm to a sample reference category. The subdivision of the subset *Q* by means of parallel lines (in this example the slope chosen is -0.6) allows to conform to the percentages in the Call when the algorithm is applied to the world population in the specific reference category.

There are borderline cases of articles having received a low number of citations but published in journals characterized by a high impact (lower right region in Figure 2) or of articles having received a high number of citations but published in journals with low impact (top left region on Figure 2). In these situations of uncertainty (*IR* papers) the evaluation will follow an *informed peer review* procedure which contemplates also the use of peer review, internal or external depending on the presence of suitable expertise inside the GEV. The areas including IR articles will be detected by two triangles (compare Figure 3); one of them will be positioned in the upper left-hand corner (articles with many citations published in paper of low impact), the other in the lower-right corner (articles with few citations published in paper of high impact).

Coherently with what decided together with the other bibliometric GEVs, for the years 2011 and 2012 the top left triangle will be delimitated by the point of coordinates (0,0.6) and by the intersection of the "Excellent" threshold segment with the top border of *Q;* the lower right triangle will be an isosceles right triangle including 5% of the papers. For the year 2013, the top-left triangle will be instead delimited by the point with coordinates (0,0.5) and by the intersection between the "Excellent" threshold segment and the top border of *Q;* the lower right triangle will be an isosceles triangle including 7% of the papers.
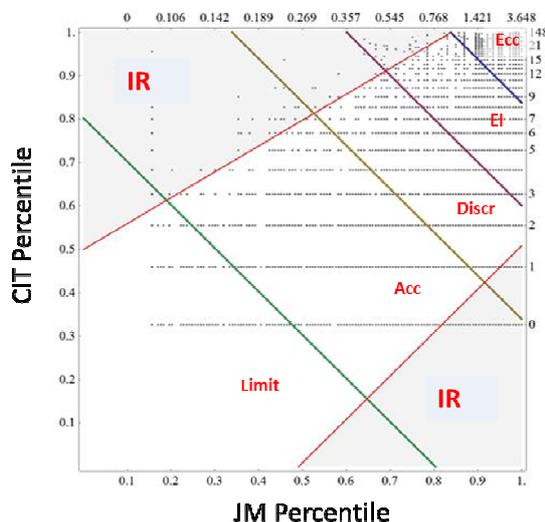
Figura 3. Example of definition of the uncertainty regions (IR) to be assessed by informed peer review.

When the calibration procedure has been completed, the articles submitted to the VQR exercise are assigned to one of the classes of merit as follows. The GEV, as explained in Section 6.5, identifies the reference category (or the subject category or the ASJC class) of the paper. Furthermore, the author has indicated during the submission process (see Section 6.8) the database from which the number of citations has to be extracted; then the journal measure of impact will be chosen on the basis of the reference category and of the characteristics of the research output (see Section 6.7). The last step is then the computation of the percentiles of the measure of impact of the journal that has published that paper and of the number of citations received in the database selected by the author. The point individuated by these two values is then positioned in the subset Q described above; the evaluation of the research output—or the indication that the research output must be evaluated through *informed peer review*—is given by the region to which the given point is assigned.

6.7 Journals in the reference categories and impact indicators

As already stated before in this document, the evaluation procedure was conceived with the intent of using all the available tools to provide the better possible evaluation of papers dealing with the core topics of the area 01. To do so, the GEV 01 has deemed necessary the identification of a set of reference categories, one for each Scientific Sector (SSD) related to the GEV, to integrate the Subject Categories used in WoS and the All Science Journal Classification (ASJC) used in Scopus, that do not faithfully represent the internal subdivisions of the Mathematics and Computer Science research fields.

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

Analogously to what happens for WoS Subject Categories and Scopus ASJC classes, to each reference category has been associated a list of journals that are considered relevant to that area. It should however pointed out that:

- There can be articles of a given SSD that are published in journals not included in the corresponding reference category; this does not give any indication about the content or the quality of that article or of the journal. In fact, the intent of the GEV 01, when compiling the lists, was not to cover the whole body of knowledge in the Mathematics/Computer Science research fields, but simply to identify journals publishing a significant amount of papers dealing with the core topics of each sector at an international level. In particular, the lists will not include many interdisciplinary journals, since, as already mentioned, interdisciplinary papers have the right to be evaluated respecting their own specific peculiarities.

- Papers published in indexed journals that were not included in the lists will be evaluated by applying the bibliometric algorithm to the subject category or ASJC class which is the most consistent with the content of the article among all those containing the journal in which the article has been published.

- These reference lists are to be considered valid only for this specific edition of the VQR exercise. They cannot be considered representative for the future situation of the research in the area.

- These lists are built only to be used together with the algorithm previously described; the GEV 01 does not endorse any other use of these lists. In particular, it is important to recall that the VQR results must *not* be used to any extent for evaluating the scientific production of single researchers.

These lists have been compiled by the GEV on the basis of its own expertise, starting from the journals indexed in MathSciNet, those classified in WoS Subject Categories relevant for the area 01, and those classified by Scopus in Mathematics or Computer Science. In particular the WoS *Subject Categorie*s and the ASJC classes considered are the following:

- For Mathematics:
  o WoS Subject categories: *Astronomy & Astrophysics* (only for Celestial Mechanics)*; History & Philosophy of Science; Logic; Mathematical & Computational Biology; Mathematics; Mathematics, applied; Mathematics, interdisciplinary applications; Mechanics; Operations Research & Management Science; Physics, Mathematical; Statistics & Probability.*

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

o ASJC classes: *1207 History and Philosophy of Science; 1702 Artificial intelligence* (only for Mathematical Logic); *1703 Computational Theory and Mathematics; 1803 Management Science and Operations Research; 1804 Statistics, Probability and Uncertainty; 2600 Mathematics; 2601 Mathematics (miscellaneous); 2602 Algebra and Number Theory; 2603 Analysis; 2604 Applied Mathematics; 2605 Computational Mathematics; 2606 Control and Optimization; 2607 Discrete Mathematics and Combinatorics; 2608 Geometry and Topology; 2609 Logic; 2610 Mathematical Physics; 2611 Modeling and Simulation; 2612 Numerical Analysis; 2613 Statistics and Probability; 2614 Theoretical Computer Science; 3103 Astronomy and Astrophysics* (only for Celestial Mechanics).

- For Computer Science:
    o WoS S*ubject categories*: *Computer Science, Artificial Intelligence; Computer Science, Cybernetics; Computer Science, Hardware & Architecture; Computer Science, Information Systems; Computer Science, interdisciplinary applications; Computer Science, Software Engineering; Computer Science, theory & methods.*
    o ASJC classes: *1700 Computer Science; 1701 Computer Science (miscellaneous); 1702 Artificial intelligence; 1703 Computational Theory and Mathematics; 1704 Computer Graphics and Computer-Aided Design; 1705 Computer Networks and Communications; 1706 Computer Science Applications; 1707 Computer Vision and Patter Recognition; 1708 Hardware and Architecture; 1709 Human-Computer Interaction; 1710 Information Systems; 1711 Signal Processing; 1712 Software.*

For all the SSD MAT (with the exception of the SSD MAT/04, see below) it has been adopted the following procedure:

- First of all it has been identified the journals indexed in the subject categories and in the ASJC categories aforementioned and that are also indexed in MathSciNet;
- For each SSD, the journals kept in the list are those contained in subject categories or ASJC classification relevant for the SSD, excluding only journals mostly publishing papers outside the core topics of the SSD (or journals misclassified in WoS or Scopus);
- For each SSD, the journals listed in generic subject categories or generic ASJC classes have been included in the list if they publish a significant amount of papers in the core topics of the SSD. Mathematical journals not specialized in a specific sector are thus included in more than one list. As a consequence, the lists are not disjoint, as it happens for the subject categories and the ASJC classes.

Because of the peculiarities of the areas of History of mathematics and of Mathematical education included in SSD MAT/04, from several points of view closer to humanities than

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

sciences, the list associated to the SSD MAT/04 contain also non-indexed journals. The selection of journals to be included in the list has been made by the Sub-GEV members and has also been based on indications of the main scientific societies of the two fields. As in the previous VQR, this list will not be used for the bibliometric algorithm, but to keep updated, the list of international reference journals in the field, highlighting the most prestigious ones, but without making other distinctions among the remaining ones (contrarily to what happened in the previous VQR).

For the Computer Science sector (SSD INF/01) the procedure adopted was the following:

- First of all, the journals present in the *subject categories* and in the ASJC classes mentioned above have been individuated;
- *Magazines* and journals that publish only *surveys/tutorials* have been excluded
- The Sub-GEV 01.1 members, using their expertise, have identified the interdisciplinary journals that are not focused on the core topics of the sector;
- The final list does not include those journals that have unanimously been identified as interdisciplinary in the Sub-GEV, that are contained in at least another subject category or ASJC class not included among those considered, and that contained none or just a few papers devoted to the core topics of the field. Less than the 10% of journals included in the starting list have been excluded from the final list.

The lists associated to each reference category can be found in the attachments.

Finally, the GEV 01 has decided to use the following measures of impact of journals:

- For the reference category *Computer Science* (INF/01): AI of WoS, and SNIP and SJR of Scopus. The choice betwwen SNIP or SJR will be based on the indications given by the author in the research output description form (if the author does not indicate any preference, as a norm SNIP will be used, since its normalization, given the heterogeneity of the discipline, usually makes it more correlated with the research in the Computer Science field).
- For the reference categories *Mathematical Logic* (MAT/01), *Algebra* (MAT/02), *Geometry* (MAT/03), *Mathematical Analysis* (MAT/05): MCQ of MathSciNet.
- For the reference categories *Probability and Mathematical Statistics* (MAT/06), *Mathematical Physics* (MAT/07), *Numerical Analysis* (MAT/08), *Operational Research* (MAT/09): AI of WoS, SJR of Scopus, MCQ of MathSciNet. The choice between MCQ and AI or SJR will be made by the competent Sub-GEV on the basis of the characteristics of the article and by considering the indications given by the author in the research output description form. In particular, papers that are more focused on applications will be evaluated using AI or SJR (depending on the database selected by the author), because published by journals that can collect a significant number of citations from non-

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

mathematical journals that are not taken into account by MathSciNet. The remaining papers will be evaluated using MCQ, in analogy to what happens for the other mathematical sectors.

Articles published by journals that are in one of the lists but are not indexed in MathSciNet will be evaluated by using AI or SJR; articles published in journals for which the AI is not available will be evaluated by using MCQ or SJR (or SNIP in the case of INF/01).

Finally, the GEV 01 is aware of the limitations of an evaluation of research outputs based on bibliometric indicators; therefore in the definition of the bibliometric evaluation criteria, described in this document, indications and suggestions given by the principal, national and international, scientific societies for Mathematics and Computer Sciences have been taken into account. In particular, the GEV 01 is well aware that it is impossible to provide a total order of the journals in a given area that could be considered fully representative of their "scientific quality", because this last concept is inherently multidimensional and to some extent subjective. The use of the measures of impact is then only considered in relation to the implementation of the bibliometric algorithm here described.

6.8   Informations to be inserted in the research output descriptive form

The authors must write in the form associated to a research output submitted to the GEV 01 the following information (the information from c. to i. have to be given only for the journal articles, the remaining ones for all research outputs):

a. Area, Academic Recruitment Field, ERC sector and SSD associated to the research output (this last indication is mandatory, and will decide, together with the SSD of the author, the GEV and the Sub-GEV that will assess the research output);

b. In case of a mathematical SSD, a primary code (mandatory), and eventually a secondary code (facultative), giving the MSC 2010 classification of the research output;

c. If the research output is indexed in both databases (WoS and Scopus), which one should be considered for calculating the number of citations (mandatory; to help in this choice the system will show the number of citations received by the article in each of the databases); for research outputs in the SSD INF/01 the GEV 01 suggests using Scopus;

d. The WoS subject category or the Scopus ASJC class (depending on the choice of the database) more closely related to the content of the article among those that contains the

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

journals in which that paper has been published; this indication is usually facultative, and it is mandatory only for papers published in journals that are not included in the lists of the reference categories of the GEV 01;

e. If the SSD associated to the article is MAT/06, MAT/07, MAT/08 or MAT/09, the choice (facultative) of a measure of impact between MCQ and AI (for WoS) or MCQ and SJR (for Scopus); as described in Section 6.7, this indication must be based on the intrinsic characteristics of the research output: for papers that mainly address applicative problems the measure to be indicated is AI or SJR; for the remaining papers the measure to be indicated is MCQ;

f. If the SSD associated to the article is INF/01, and the database selected is Scopus, an impact measure chosen between SNIP and SJR;

g. If the article is the result of a research activity in emerging areas at the international level (facultative, it suggests a *peer review* evaluation);

h. If the article is the result of a highly specialized research activity (facultative, it suggests a *peer review* evaluation);

i. If the article is the result of an interdisciplinary research activity (facultative, it suggests a *peer review* or *informed peer review* evaluation); in this case it has to be indicated which other GEV can be involved in the evaluation;

j. If the research output is an original outreach research output;

k. Any other information to be considered useful for the valorization and evaluation of the research output.

7. Typologies of admissible research outputs

Among the typologies of research outputs indicated in the VQR 2011-2014 Call, the GEV01 identifies as eligible the following ones:
1. Scientific monographs and related research outputs:
   a. Research monographs
   b. Critical editions
   c. Critical manuals (including outreach books, but excluding basic textbooks)
2. Journal contributions, limited to:

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

    a. Scientific articles (included original outreach articles)

    b. Review essays or surveys

3. Book contributions, limited to:

    a. Chapters or essays

    b. Conference proceedings with peer review

    c. Preface/Afterword essays

    d. Editing of volumes with an introductory essay

    e. Critical entry in dictionary or encyclopedia

4. Other scientific research outputs (only if accompanied by documents allowing the identification of the publication date):

    a. Expositions

    b. Data bases and software

5. Patents granted within the evaluation period (from 1/1/2011 to 12/31/2014)

The others typologies of research outputs listed in the VQR Call are not considered eligible for evaluation in the Area 01. Furthermore, the GEV does not consider admissible for the evaluation:

1. New editions and translations of works published before 2011
2. Introductions and/or afterword to new editions of publications appeared before 2011
3. Abstracts

Furthermore, as established in the VQR 2011-2014 Call, the following research outputs are not considered as eligible for evaluation:

1. Educational Manuals
2. Simple reviews, without an analysis of the related literature
3. Short encyclopedic or dictionary entries without original content
4. Short notes to judgement (*Note a sentenza*) without original content
5. Short parts of catalogues without original scientific content

As indicated in Section 2.3 of the ANVUR VQR 2011-2014 Call, research outputs of typology 4 (Other scientific research outputs) can be considered for evaluation only if they are accompanied by official data certifying the date of research outpution; in addition, it is suggested to indicate in the research output descriptive form all the information considered useful for a complete evaluation. Among all the research outputs of this typology, databases and software will be evaluated in peer review. The other admissible research outputs of typology 4 (Expositions) will be directly examined by the competent Sub-GEV.

Finally, research outputs of typology 5 (Patents granted) can receive an "Excellent" or "Good" evaluation only if international or already assigned or subject to the granting by an enterprise.

National Agency for the Evaluation of
Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del
sistema Universitario e della Ricerca

Evaluation of Research Quality

VQr

Valutazione Qualità della Ricerca

8. Conflicts of interest

GEV members will not evaluate or assign to external reviewers or other GEV members:

- research outputs they have authored or co-authored;

- research outputs which have been authored or co-authored by spouses and relatives up to the fourth degree of kinship;

- research outputs submitted by universities of which they have been employees or from which they have received official assignments or have had formal collaborations (including the affiliation to research centers) since 1/1/2011;

- research outputs submitted by research centers controlled by MIUR, or by other public or private entities voluntarily subjected to the VQR, of which they have been employees or from which they have received official assignments or have had formal collaborations (including the affiliation to research centers) since 1/1/2011;

For the above research outputs there exists a conflict of interest when:

- the institution in question has a permanent internal, territorial or disciplinary, subdivision (e.g., local units of a research center, institutes, departments), only with respect to the research outputs presented by the same internal unit;

- the institution in question does not have a permanent internal, territorial or disciplinary, subdivision (e.g., local units of a research center, institutes, departments), with respect to all the research outputs presented by the institution;

- if the internal organization is based on several hierarchical levels (e.g., several institutes within a single department) a conflict of interest emerges at the lowest level only (e.g., GEV members who are affiliated with different institutes belonging to the same department have a conflict of interest only with respect to the research outputs presented by authors belonging to the same institute) .

In case of conflicts of interest, the GEV Coordinator, or the Sub-GEV Coordinator when appropriate, will assign the research output to be evaluated to another GEV member for whom no conflict of interest is present.

In case of conflicts of interest involving the GEV Coordinator, the corresponding research outputs will be assigned by the VQR Coordinator or by a person designated by the VQR Coordinator.