
Criteria for the Evaluation of Research Outputs

Group of Evaluation Experts

for Area 11b -Psychological Science (GEV11b)
1. INTRODUCTION ................................................................. ERRORE. IL SEGNALIBRO NON È DEFINITO.

2. DELIMITATION OF GEV AREA ........................................... ERRORE. IL SEGNALIBRO NON È DEFINITO.

3. ORGANIZATION OF THE GEV ............................................. 5
   3.1 ASSIGNMENT OF RESEARCH OUTPUTS WITHIN GEV ........................................... 5
   3.2 OPERATING RULES OF THE GEV ........................................... ERRORE. IL SEGNALIBRO NON È DEFINITO.

4. THE EVALUATION OF RESEARCH OUTPUTS ................................. 6

5. PEER REVIEW EVALUATION ....................................................... 6
   5.1 THE SELECTION OF EXTERNAL PEER REVIEWERS ........................................... ERRORE. IL SEGNALIBRO NON È DEFINITO.
   5.2 PEER EVALUATION ................................................................. ERRORE. IL SEGNALIBRO NON È DEFINITO.

6. BIBLIOMETRIC ANALYSIS ....................................................... 6
   6.1 DATABASES ................................................................. ERRORE. IL SEGNALIBRO NON È DEFINITO.
   6.2 CITATION’S TIME WINDOW ................................................................. 10
   6.3 SELF-CITATIONS ................................................................. 10
   6.4 BIBLIOMETRIC INDICATORS ................................................................. 10
   6.5 THE ALGORITHM FOR PRODUCTS CLASSIFICATION ............................. ERRORE. IL SEGNALIBRO NON È DEFINITO.
   6.6 CALIBRATION PROCEDURE ................................................................. 11

7. OTHER PRODUCTS ................................................................. ERRORE. IL SEGNALIBRO NON È DEFINITO.

8. CONFLICTS OF INTEREST .......................................................... 16

9. FINAL SYNTESIS ........................................................................... 19
1. Introduction

This document describes the organization of the Group of Experts for Evaluation in the Psychological Science Area (from now on, the GEV) and the criteria the group will use in evaluating research outputs. The document is divided in eight parts. Section 2 lists the Scientific Sectors (SS), the Academic Recruitment Field (ARF) and the ERC sectors relevant for the GEV. Section 3 summarises 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 selected by the GEV, the algorithm and the calibration procedure. Section 7 describes the evaluation criteria for specific research outputs. Section 8 describes how the GEV plans to solve potential conflicts of interest between GEV members and authors of research outputs. Finally, section 9 provides a summary of the document.

2. Delimitation of the GEV Area

The Group of Experts for Evaluation in the Area 11b- Psychological Science (from now on, the GEV11b) will carry out the evaluation of the products submitted by researchers belonging to the Sectors (SS), the Academic Recruitment Field (ARF) and the ERC Sector (ERC) listed in Tables 1-3.

<table>
<thead>
<tr>
<th>Area 11b – Psychological Science</th>
<th>Relevant Scientific Sectors (SS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-PSI/01</td>
<td>General Psychology</td>
</tr>
<tr>
<td>M-PSI/02</td>
<td>Psychobiology and Physiological Psychology</td>
</tr>
<tr>
<td>M-PSI/03</td>
<td>Psychometry</td>
</tr>
<tr>
<td>M-PSI/04</td>
<td>Developmental and Educational Psychology</td>
</tr>
<tr>
<td>M-PSI/05</td>
<td>Social Psychology</td>
</tr>
<tr>
<td>M-PSI/06</td>
<td>Psychology of Work and Organizations</td>
</tr>
<tr>
<td>M-PSI/07</td>
<td>Dynamic Psychology</td>
</tr>
<tr>
<td>M-PSI/08</td>
<td>Clinical Psychology</td>
</tr>
</tbody>
</table>

Table 1. Relevant Scientific Sectors for Area 11b- Psychological Science
### Area 11b – Psychological Science

#### Relevant Academic Recruitment Fields (ARF)

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/E1</td>
<td>General Psychology, Psychobiology and Psychometry</td>
</tr>
<tr>
<td>11/E2</td>
<td>Developmental and Educational Psychology</td>
</tr>
<tr>
<td>11/E3</td>
<td>Social Psychology, Psychology of Work and Organizations</td>
</tr>
<tr>
<td>11/E4</td>
<td>Clinical and Dynamic Psychology</td>
</tr>
</tbody>
</table>

### Table 2. Relevant Academic Recruitment Field (ARF) for Area 11b - Psychological Science

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH1_7</td>
<td>Behavioural economics; experimental economics; neuro-economics</td>
</tr>
<tr>
<td>SH1_9</td>
<td>Industrial organisation; strategy; entrepreneurship</td>
</tr>
<tr>
<td>SH1_10</td>
<td>Management; marketing; organisational behaviour; operations management</td>
</tr>
<tr>
<td>SH3_2</td>
<td>Inequalities, discrimination, prejudice, aggression and violence, antisocial behaviour</td>
</tr>
<tr>
<td>SH3_3</td>
<td>Social integration, exclusion, prosocial behaviour</td>
</tr>
<tr>
<td>SH3_4</td>
<td>Attitudes and beliefs</td>
</tr>
<tr>
<td>SH3_5</td>
<td>Social influence; power and group behaviour; classroom management</td>
</tr>
<tr>
<td>SH3_6</td>
<td>Diversity and identities, gender, interethnic relations</td>
</tr>
<tr>
<td>SH3_9</td>
<td>Health, ageing and society</td>
</tr>
<tr>
<td>SH4_1</td>
<td>Cognitive basis of human development and education, developmental disorders; comparative cognition</td>
</tr>
<tr>
<td>SH4_2</td>
<td>Personality and social cognition; emotion</td>
</tr>
<tr>
<td>SH4_3</td>
<td>Clinical and health psychology</td>
</tr>
<tr>
<td>SH4_4</td>
<td>Neuropsychology</td>
</tr>
<tr>
<td>SH4_5</td>
<td>Attention, perception, action, consciousness</td>
</tr>
<tr>
<td>SH4_6</td>
<td>Learning, memory; cognition in ageing</td>
</tr>
<tr>
<td>SH4_7</td>
<td>Reasoning, decision-making; intelligence</td>
</tr>
<tr>
<td>SH4_8</td>
<td>Language learning and processing (first and second languages)</td>
</tr>
<tr>
<td>SH6_14</td>
<td>History of Science, Medicine and Technologies</td>
</tr>
<tr>
<td>PE1_13</td>
<td>Probability</td>
</tr>
<tr>
<td>PE1_14</td>
<td>Statistics</td>
</tr>
<tr>
<td>PE1_20</td>
<td>Application of mathematics in sciences</td>
</tr>
</tbody>
</table>
### Table 3. Relevant ERC Sectors for Area 11b - Psychological Science

<table>
<thead>
<tr>
<th>Sector Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE6.7</td>
<td>Artificial intelligence, intelligent systems, multi agent systems</td>
</tr>
<tr>
<td>PE6.12</td>
<td>Scientific computing, simulation and modelling tools</td>
</tr>
<tr>
<td>LS5.4</td>
<td>Sensory systems (e.g. visual system, auditory system)</td>
</tr>
<tr>
<td>LS5.5</td>
<td>Mechanisms of pain</td>
</tr>
<tr>
<td>LS5.6</td>
<td>Developmental neurobiology</td>
</tr>
<tr>
<td>LS5.7</td>
<td>Cognition (e.g. learning, memory, emotions, speech)</td>
</tr>
<tr>
<td>LS5.8</td>
<td>Behavioural neuroscience (e.g. sleep, consciousness, handedness)</td>
</tr>
<tr>
<td>LS5.9</td>
<td>Systems neuroscience</td>
</tr>
<tr>
<td>LS5.10</td>
<td>Neuroimaging and computational neuroscience</td>
</tr>
<tr>
<td>LS5.11</td>
<td>Neurological disorders (e.g. Alzheimer's disease, Huntington's disease, Parkinson's disease)</td>
</tr>
<tr>
<td>LS5.12</td>
<td>Psychiatric disorders (e.g. schizophrenia, autism, Tourette's syndrome, obsessive compulsive disorder, depression, bipolar disorder, attention deficit hyperactivity disorder)</td>
</tr>
<tr>
<td>LS8.7</td>
<td>Animal behaviour</td>
</tr>
</tbody>
</table>

### 3. Organisation of the GEV

The GEV's coordinator is Professor Roberto Cubelli from the University of Trento, the GEV's assistant assigned by ANVUR to GEV11b is Doctor Serena Mastroberardino.

#### 3.1 Assignment of research outputs within the GEV

The assignment of research outputs to the GEV is based on author's SS. Depending on the SS specified in the research output descriptive form associated with the research output, the GEV may decide to assign the research output to another GEV, if the topic of the research output is judged to be more appropriate to the latter. In this case the evaluation of the research output is based on the criteria of the GEV of destination.

The SS assigned to the research output can be different from the one the author belongs to, because it is referred to the GEV and to the SS that the author/institution believes to be more appropriate for the research output evaluation. The GEV representing the SS of the author may decide to ignore the author/institution's indication if it is considered inadequate to best evaluate the research output.
Within the GEV11b, the attribution of the research outputs to the members appointed to perform the evaluation will be done taking into account the SS of the research output indicated by the author/institution.

If a research output is assigned to more than one GEV (for instance because co-authors indicated different SS belonging to different GEVs), the research output will be evaluated according to the VQR Guidelines for the Groups of Evaluation Experts. If necessary, the Coordinators of the involved GEVs will create specific Inter-Area Consensus Groups.

### 3.2 Operating rules of the GEV

The operating rules of the GEV are the following:

- a GEV meeting is called with an at least 15 day 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 appointed by ANVUR to the GEV attends the GEV’s 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 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 following methods are used:

- Peer review evaluation by (normally two) external reviewers selected independently by two different GEV members.
• Direct evaluation by the GEV, which can conduct an *internal peer review* according to the same procedure described for *external peer review* (that is, two GEV members will be involved).

• Bibliometric analysis, to be conducted using indicators and algorithms described below in this document. Research outputs subjected to bibliometric analysis are not assigned *automatically* (that is using automatically the final class of merit provided by the algorithm) to the classes of merit established by the Ministerial Decree (MD) and by the VQR call. The allocation is based instead on the expert judgment of the GEV, which will employ any element of evaluation beside bibliometric indicators, such as the expertise of its members and the information described in forms associated with the research outputs.

• Direct evaluation by the GEV, that will perform an *internal peer review* using the same modality of the ones performed by external reviewers.

**5. Peer review evaluation**

Each research output to be evaluated in *peer review* will be sent to two external reviewers, independently chosen by the two GEV members to whom the output has been assigned. Alternatively the research output will be evaluated within the GEV using 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, both Italian and foreign experts, given its relevant purpose in the public interest, follows the principle of honest institutional cooperation and it is based on the criteria of correctness, objectivity and impartiality.

A great deal of attention will be devoted to maintain the anonymity of reviewers, both at the stage of preparation of the reviewers’ list, and at the operational stage of the evaluation. The results of the evaluation for individual research outputs and the identity of the reviewers that evaluated them will not be disclose to the public. The list with the reviewers’ names will be published by ANVUR within 30 days of the publication of the VQR Final Report.
Reviewers will be selected among expert scholars and the most authoritative and scientifically qualified specialists in the disciplines relevant to the to be examined research outputs. They are also expected to be scientifically productive during the time period covered by the VQR.

Starting from the list provided by ANVUR, the GEV will prepare an updated list of external reviewers comprising those who, according to GEV's judgement, adequately satisfy the scientific requirements and evaluation expertise. The list will be extended with new reviewers selected by the GEV. Specifically the Coordinator will ask GEV's members to recommend a significant number of experts who satisfy the required parameters and are available for the evaluation. The GEV’s Coordinator will collect the suggestions together with information about the reviewers’ qualifications using a shared form, and will modify the initial list with integrations and/or cancellations of experts.

It will be possible to extend the reviewers' list throughout the evaluation procedure, on the basis of the needs that might emerge after the research outputs are transmitted by authors/ institutions.

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

Rather than internal reviewers chosen among GEV's members, the GEV will preferably employ, whenever possible, external reviewers.

5.2 Peer evaluation

The evaluation by external or internal reviewers will be based on a reviewer evaluation form and guidelines created by the GEV, taking into account, if needed, the indications provided by the research groups on the evaluation process organized by the ANVUR in the period of time before the VQR. The evaluation form is structured so that it allows the reviewer to assign a score to each one of the three evaluation criteria established by the MD and the VQR call, that is, originality, methodological rigor, and attested or potential impact. The form will also include a space with a limited word count where the reviewer will have to write a brief resume motivating the judgements.

The GEV will transform the indications contained in the evaluation form into one of the five classes of merit established by the VQR call.
In presence of non-converging evaluations or controversy between reviewers, the GEV may request a third expert evaluation or create, within the GEV itself, a Consensus Group (or arbitration commission) with the assignment to provide the GEV with a final score for the research output under examination, using the consensus report methodology.

In case of direct evaluation, the procedure to be followed by the GEV’s member will be the same as the external reviewers’ one and in case of disagreement between two GEV’s member, the GEV’s coordinator will assign the review to a third member.

6. Bibliometric analysis

The following research outputs, indexed in the citational databases ISI, WoS and Scopus are subject to bibliometric evaluation:

- Scientific papers, also in the form of Letters\(^1\)
- Scientific critical review papers

In case of research outputs "of emerging research fields or with strong specialisation or with interdisciplinary characteristics, or without adequate citational data" the author/institution can ask for an informed peer review evaluation.

Within the articles classified with the bibliometric algorithm, 10% will be randomly selected and undergo a peer review process in order to measure the correlation between the two evaluation methods. In this case, the peer evaluation will not influence in any way the assignment to the class of merit of single research outputs.

6.1 Databases

The GEV will employ the Thomson Reuters’ Web of Science (WoS) Elsevier's Scopus (Scopus) databases, according to indications provided by the author/institution in the research output descriptive form.

---

\(^1\) Publications on journals that include scientific papers in the form of letter. Letters to the editor that comment on papers published on the journals are excluded.
6.2 Citation’s time window

In calculating the bibliometric indicators, the GEV will use citation updated to 29th February 2016.

6.3 Self-citations

The opportunity to include or exclude self-citations in the bibliometric evaluation is currently under debate in the scientific community. The GEV11b has decided, following suggestions from the working group on bibliometric evaluation created during the first plenary meeting of GEV’s Coordinators, not to exclude self-citations but to carefully examine articles with a number of self-citation above 50% of the total, eventually turning it to informed peer review.

6.4 Bibliometric indicators

For all articles published on journals indexed in WoS and Scopus, the evaluation will employ, an algorithm that takes into account, in relation to publication date, both number of citations and the impact indicator (Journal Metric, JM) of the hosting journal. Consistent with the orientation of the scientific community in the bibliometric sector, and taking into account the differences with which the manifold indicators measure the impact of a journal, the GEV11b, following indications of the working group on bibliometric evaluation, has established to use more than one indicator of JM. For each database two indicators are going to be used, one to measure popularity of the hosting journal (where the definition of the received citations are considered independently of the origin of each citation) and one used to measure journal prestige (where the definition of the citations are weighted on the basis of the prestige of the journal hosting the publication). Specifically the following indicators will be used:

- For WoS (https://www.webofknowledge.com): 5-year Impact Factor (5YIF), as popularity index and Article Influence (AI), as prestige index\(^2\).
- for Scopus (http://www.journalmetrics.com): Impact per Publication (IPP), as popularity index, and SCImago Journal Rank (SJR), as prestige index\(^3\).

\(^2\) The 5YIF was selected instead of the more popular Impact Factor (IF) because of larger stability when varying the year of publication and because the time window (5 years) is the same as the AI.

\(^3\) In this case the time window where citations are considered is 3 years for both indicators. The definition of IPP is the same as that of 5YIF; the definition of SJR is similar but not identical to the definition of AI.

\(^4\) In WoS recent journal may not have a 5YIF or AI. In this case, if the author or the institution indicated WoS as preferred database for the presented research output, the IF will be used ad default indicator.
In the research output descriptive form the author/institution will be asked to indicate the preferred database (WoS or Scopus) and one of the two impact indicator associated to the database that will be used for the evaluation.

6.5 The algorithm for research outputs classifications

The algorithm used for assigning the articles to one of the five classes of merit defined in the VQR call is based on the combined use of the bibliometric indicators relative to the impact of the journal where the article has been published (JM) and the citational index that measures the impact of the single article (CIT). Depending on the year of publication, the first or the second indicator may have a larger relative impact. Each article is evaluated within a specific category of reference (more details in the following sections) and within the year of publication. The evaluation procedure within the category of reference is previously calibrated so to ensure that the probability ex-ante at worldwide level for each article of a given category and a specific year to fall in one of the evaluation classes is the one defined by the VQR call:

- Excellent [top 10% of the distribution of the international scientific production of the reference area];
- Good [10% - 30% of the distribution of the international scientific production of the reference area];
- Fair [30% - 50% of the distribution of the international scientific production of the reference area];
- Acceptable [50% - 80% of the distribution of the international scientific production of the reference area];
- Limited [80% - 100% of the distribution of the international scientific production of the reference area].

The indication of the percentiles in relation to the class of merit does not refer to the percentage of attended results of the products presented to the VQR. The evaluation of the single article is not comparative, each article will be allocated to the class of merit independently of the allocation of the other research outputs.

The first step towards the evaluation of an article is the identification of the scientific area of reference represented by the reference category known as Subject Category (SC) in WoS and All Science Journal Classification (ASJC) in Scopus (from now on SC). The SC, that includes journals belonging to the same research or disciplinary area, will have to be indicated in the
research output descriptive form by the author/institution that presented the article. Normally, a journal belongs only to one SC. However, since a journal may belong to more than one SC of different disciplinary areas, the unambiguous identification of the SC will depend on the author/institution indication. The choice is in any case is subject of approval by the GEV.

A multidisciplinary category is present both in WoS (Multidisciplinary Sciences) and Scopus (Multidisciplinary) and includes journals, such as Nature, Science and so on, characterised by a variety of scientific topics. Articles published on journals under such category will be re-assigned to another SC on the base of (i) the citations included in the article and (ii) the citations received by the article. Specifically, the publication will be assigned to the SC that includes the larger number of cited or citing journals. In this way, the publication will be compared with publications of the same thematic/disciplinary area. In the allocation to the new SC, the article will keep the JM of the journal and the number of received citations, without changing the distribution of the final SC.

The same procedure will be applied to psychological journals solely present in the Psychology, multidisciplinary of WoS and Psychology (Miscellaneous) and Psychology (All) of Scopus category. The articles published in journals belonging to such SC will be assigned to a new psychological SC according to the content of the majority of cited and citing journals.

As already mentioned, the allocation of an article to one of the five classes of merit indicated in the VQR call, is made after a calibration of the thresholds for the identified SC and for the specific year of publication. Such procedure allows to have in each class of merit, whatever the analysed category and year, the percentage of research outputs defined by the MD and by the VQR call.

6.6 Calibration procedure

The calibration of the bibliometric algorithm is a function of the specific SC in the specific year in analysis. The algorithm distinguish also typology journal article and letter from the typology review, calculating separate empirical cumulative distributions caused by different numbers of citations usually associated with such publications.

First, the empirical cumulative distribution of the bibliometric indicator JM is calculated for journals belonging to the selected SC, for the year of publication of the article under evaluation and for each journal a percentile is assigned. Afterwards the empirical cumulative distribution function of the number of citations (CIT) of all the article published in journals belonging to the selected SC and a percentile is assigned to each article. At the end of the procedure two
percentile will be associated to each article (journal percentile and citation percentile). The two percentiles obtained this way identify a point in the region of the Cartesian plane \( Q = [0,1] \times [0,1] \), delimited by the JM percentile of the journal (X axis) and by the percentile of the citations (Y axis). \( Q \) is then divided in five zones or regions that follow the percentage of articles belonging to each region as defined in the VQR call.

Such partition is realized using simple straight lines identified by the following linear equation:

\[
CIT = A \cdot JM + B_n
\]

The angular coefficient of the lines that delimits the zone (A) is imposed as equal to all the lines in order to increase homogeneity of the adopted criterion. The intercept \( B_n \) are calculated by ANVUR, depending on the specific distribution of the SC, to ensure that the percentage indicated in the VQR call are always respected. An example of \( Q \) division in the five zones is presented in Figure 1. Despite the articles’ distribution vary between categories and years, the algorithm allows to obtain an evaluation calibrated with respect to the chosen set.

The gradient A of the threshold lines is established by the GEV. The gradient has a relevant role since, according to A value, the final classification will be more based on the citation percentile (for gradient with absolute value smaller than 1) or on the percentile metric of the journal (for gradient with absolute value larger than 1). As an example, an horizontal line corresponds to an evaluation uniquely based on citation percentile. Taking into account both what is reported by the state of the art of the literature in the bibliometric field, and by the different statement on the correct use of bibliometry for evaluation purposes\(^5\), the use of high gradients should be avoided, given the absolute impossibility to use only the JM of a journal as a proxy of the impact of a single article published in that journal. In other words, values of A smaller that 1 will have to be used if possible in order to favour the information provided by CIT that is an impact measure at the level of the single research output under assessment (article level metric). Such choice however is not unconditional but depends on the different citation routine of the different disciplines/communities, as well as the sample size and composition of the SC, that makes the information provided by the citational data, with the decrease of the publication year, more or less reliable.

\(^5\) See as an example the IEEE's Statement on Appropriate use of Bibliometric Indicators- https://www.ieee.org/publications_standards/publications/rights/bibliometrics_statement.html
Based on the simulations made by the working group on bibliometric evaluation, the GEV11b has decided to establish different gradients for different years of publication, so that the more the research outputs are recent the more the evaluation will be based on the journal metrics (angular coefficient of the threshold line with values progressively higher from 2011 to 2014).

The values that will be used are the following:

- 2011: -0.4
- 2012: -0.6
- 2013: -1.0
- 2014: -1.5

In order to avoid unwanted situations, gradients may vary up to a maximum of 30% for the years 2011, 2012 and 2013. For 2014, being the citational information less stable, the gradient will be included in the range $[-2.0, -1.0]$.

Articles published in 2014, for which the citational information may be unreliable, will undergo informed peer review (IR), that will take into account both of the bibliometric indications and of the expert judgement of the GEV members to which the paper has been assigned.

---

6 For example, the possibility to classify in the excellent class research outputs that did not receive any citation.
Thresholds: \[ CIT = A \cdot JM + B_n \]

**Figure 1.** Representation, in percentiles, of all the articles published in a specific SC in a determined year. Each publication is positioned on the plane according to the percentile of the impact indication of the journal JM (line) and of the percentile of the citations CIT (column). The plane is divided in five zones according to the percentages reported in the VQR call. The angular coefficient A of the lines that circumscribe the zones is imposed as equal for all the lines. The intercepts B_n are calculated by ANVUR, according to the distribution of the particular SC, in order to guarantee that the VQR call’s percentages are respected.

As an example, in Figure 2 the calibration of a SC through four parallel lines is reported. It has been chosen an angular coefficient of -0.6 in order to privilege the weight of the citations in the final evaluation. As shown in the figure, the points that represent the articles in the SC, are distributed irregularly in Q. By opportunoely choosing the values of the intercept, one can guarantee that the percentage of the VQR call are respected with accuracy larger than the tenth of one percentage point. In other words, when the bibliometric algorithm is applied to the "worldwide" production the percentage defined in the MD and in the VQR call are respected. As a consequence the specific article sent to the VQR will be evaluated according to a percentile in the “international scientific production of the reference area”. 
Figure 2. An example of the application of the bibliometric algorithm to a sample SC. The partition of the subspace Q through the parallel lines allows to respect the percentage defined in the VQR call when the algorithm is applied to the worldwide population of a specific SC.

Once the calibration procedure is completed, the attribution of a research output presented for the VQR works as follows. The percentile of JM for the journal where the article has been published and the percentile of the received citation are calculated and the resulting point is allocated in the above described space. According to the area where the point falls, the evaluation of the research output is obtained following the bibliometric algorithm.

There are some borderline cases where articles are published in high prestige journals but are scarcely cited (area low right in Figure 2) or published in low impact journals but with high citation impact (area high left Figure 2). In such uncertainty cases the evaluation will be conducted using informed peer review procedure either within the GEV or using external reviewers if GEV's members are not expert in the field. In order to identify such articles, the GEV11b, in accordance with the other bibliometric GEVs, has decided to draw two straight lines with positive gradient to create two triangles that will identify research outputs with contrasting values relatively to citations and impact indicators (see for example Figure 3). The triangle in the upper left portion is determined by the left superior side of Q and by the segment that connects the point (0, 0.5) with the intersection of the boundary line in the "Excellent" classification zone and Q's upper side. The right isosceles triangle in the right inferior part allows to identify the 5% of the research outputs for 2011 and 2012 and the 7% for 2013.
Given the inadequate numeric consistency of the citational information, all articles published in 2014, for which the classification of bibliometric algorithm will not result as a final evaluation of "Excellent", will undergo an informed peer review.

![Figure 3](image.png)

**Figure 3.** Example of definition of uncertainty zones handled using informed peer review (IR).

Specific attention will be paid to research outputs that are allocated in the regions immediately next to the boundary lines between classes of merit. In this case, taking into account the bibliometric information and eventually turning to informed peer review, the GEV may change the allocation assigned by the algorithm.

### 7. Other products

Among those indicated in the VQR call, other products category are admitted to evaluation:

1. Scientific monograph and analogous products: (a) Research monograph; (b) Congruent collection of one's research essays (if published before 2011 are not to be considered); (c) Critical editions; (d) Publication of unpublished work accompanied by introduction and commentary; (e) Critical manuals, not expressly of didactic content; (f) Grammars and scientific dictionary.
2. Essay in volume: (a) Essay in volume (book chapters or essays); (b) Scientific papers in peer reviewed conference proceedings; (c) Essay like Preface/Postface; (d) Volume editing with introductory essay; (e) Critical dictionary or encyclopaedic entry.

3. Other kind of scientific products (only if accompanied with official elements suitable for identification of publication date): Databases, software and tests.

Only products where the person under assessment is the author are evaluable, while if the products are only edited by the person are not taken into consideration. Didactic products intended for students of a University course, self edited products and contribution on journals in the form of Author reply, Commentary, Editorial or review are labelled as "not evaluable".

In sum, products submitted to peer review are monographs and essays in volume as long as ISBN index is provided, articles published on journals non indexed in WoS e Scopus, database, software and test.

In the evaluation of the attested and potential impact of monographs and book chapters, products with international dissemination (published by publishing houses involving authors, readers, editors and reviewers from different countries), volumes indexed on WoS and Scopus and volume that have been translated in other languages will be privileged. Monographs with international dissemination will be allocated to a maximum of "excellent"; nationally distributed ones will reach a maximum of "discrete". Concerning essays in volume, those published in volumes with international dissemination will be allocate to a maximum of "discrete", while those with national dissemination to a maximum of "fair". Databases, software and psychological tests, will reach a maximum of "discrete". Not indexed scientific papers will reach a maximum of "acceptable".

8. Conflicts of interest

GEV’s members will abstain to evaluate or to assign to other GEV’s members or to other external experts the following products:

- products in which they are co-authors;
- products in which spouses, family or other relatives within the fourth degree of kin are authors or co-authors;
• products presented by institutions where GEV members have or have had a working relationship or where they had appointments or official collaborations, including affiliation to research institutes in the period of time starting from 1/1/2011;

• products presented by research institutions monitored by MIUR and other public or private institution that voluntarily participate in the VQR where GEV's members have or have had a working relationship or where they had appointments or official collaborations, including affiliation to research institutes in the period of time starting from 1/1/2011.

There is a conflict of interest for the following products:

• in case the institution has a permanent internal partition at territorial or disciplinary level (for example local sections of research institution, institutes, departments) limited to products presented by the same aggregate;

• in case the institution does not have a permanent partition at territorial or disciplinary level (for example local sections of research institution, institutes, departments) limited to products presented by the same aggregate;

• in case the permanent internal partition is based on numerous hierarchical levels (for example, several institute under one department) the conflict of interest arise at a lower level (for example, GEV's members affiliated to different institutes of a same department, are in conflict of interest with respect to the products presented by authors belonging to the same institute).

In case of conflict of interest the GEV's coordinator will appoint another member of the GEV with no conflict of interest to proceed with the evaluation.

In case the conflict of interest involves the GEV's Coordinator, the assignment of the products will be done by the VQR coordinator or by someone directly appointed by him.

9. Final synthesis

In the research output descriptive form, the author/institution will have to indicate the Scientific Sector (SD) of the research output, that may or not correspond to the SD of the author. Since the
SD of the research output identify the GEV that will evaluate it, if the selected GEV is different from the one the author belongs to, the decision to confirm the choice of the author/institution or to evaluate the research output belongs to the GEV of reference.

For each research output, any award, translation, review or other information relevant to underline the impact of the research output, may be indicated.

For the bibliometric products the author/institution will have to indicate also the SC of reference or, in case of more the one SC, the one to which the journal that hosts the article under evaluation belongs to.

The author/institution will have to indicate also the SC of reference or, in case of more the one SC, the one to which the journal that hosts the article under evaluation belongs to.

The distribution of bibliometric indicators (and its calibration) will be calculated for each year of publication, for each Subject Category (ISI Wos) and ASJC (Scopus) and for type of publication (article/letter o review).

The types of products that will undergo peer review evaluation (exclusively or together with the bibliometric algorithm) are the following:

1. Not indexed products
2. Bibliometric products that are in the uncertainty areas or in the boundary line of the classes of merit
3. A statistic sample of 10% for the study of the correlation of the evaluation methods.

The bibliometric products that may be evaluated also using a peer review process are:

1. published in 2014 with scarce citational data;
2. with a number of self citation above 50% of the total;
3. suggested by the author/institution as products "of research in an emerging area or with a strong specialisation or with interdisciplinary characteristics or for which citational data are scarce".