

Welcome address of the new co-editors

By Stefania Panebianco and Francesco Zucchini, 15/06/2013

ITALIAN POLITICAL SCIENCE (IPS) is the professional journal of the **Italian political science community**. It was created in 2007 by the generous efforts of **Maurizio Cotta** with **Giliberto Capano** and the IPS editorial staff. In 2013 it passed under our direction and we are honored to fulfill this challenging task. We are pleased to have set up an editorial board composed of experienced researchers who will bring in their knowledge of different sub-areas of the discipline.

So far the **SISP** online journal has been a new and unique tool to inform our scientific community about many 'professional' topics. As new editors, we intend to follow the same path and strengthen the dissemination of the journal nationally and internationally. At the same time, readers will find some new sections.

IPS will issue two numbers per year and will address a broad range of topics. Each issue will **focus** upon a relevant topic. The one selected for the current issue is crucial for the scientific and professional community of political scientists: research assessment. The other sections of the on-line journal provide relevant information on **teaching and research**. The current issue deals specifically with PRIN research projects. A specific '**News**' section is devoted to *calls for*, job placements, new research projects, awards, members' career advancements and events, etc.

Since IPS will go hand in hand with the **Italian Political Science Review/RISP** that is being published entirely in English, IPS will host the completely new section 'Books review' inherited from RISP. We believe that through reviews written in English and published in a completely online journal, books authored by Italian Political Scientists will reach a much wider audience. This notwithstanding, IPS also welcomes reviews of books by international scholars dealing with relevant topics. All interested authors are warmly invited to send their books directly to **Stefania Panebianco**.

By taking advantage of new technological tools, the new IPS web-site will host not only the on-line journal but also multimedia contents such as videos, blogs, etc.

The editors will employ the same quality criteria adopted so far to select the articles to be published. However, alongside traditional scientific contributions the genres '**interview**' and '**debate**' will find due space in the journal in order to foster academic networking and to allow an open and frank exchange of views among scholars. Ideally, IPS intends to address topics that will intrigue any political scientist – Italians or foreign scholars interested in the Italian political science community – and to help strengthen the sense of belonging to the community.

Italian political science has grown quantitatively and qualitatively and deserves a professional journal that is quick, flexible and helpful. We hope to be up to the task. We will give it all. All readers' help is more than welcome!!

Stefania (Panebianco) & Francesco (Zucchini)
IPS co-editors

Evaluating the Evaluation. The pros and cons of 'VQR' in social and political research

By Carla Monteleone, Stefania Panebianco, and Francesco Zucchini, 15/06/2013

ACKNOWLEDGING THE NEXUS between science and economic development, and in the name of democratic control over the management of public resources, governments have progressively gained a role in mechanisms of knowledge production. In several countries, the United Kingdom, Australia, and France being the most well known cases, this has resulted in various evaluation exercises. All have generated wide debate within the scientific community on the most appropriate methods and criteria to be used. Italy, where the evaluation of research came a little later, is no exception. The debate has mostly taken place in the review *Il Mulino* and on the website Roars, with occasional articles appearing on the major national newspapers. Six key issues have dominated the discussion:

1) the definition of quality; 2) the drawbacks of impact indexes; 3) the informed peer review method; 4) the definition of research products; 5) the *ex post* adoption of evaluation criteria; 6) interdisciplinary comparison. In the following chapters we briefly introduce these topics with some comments.

1. The definition of quality

As for the definition of quality, the debate has focused on the inherent tension between quality as an objective fact and quality as a social construction. However, such a theoretical distinction is blurred when we pass from conceptual discussion to the empirical operationalization. In other terms, when we are interested in methods that are used to assess the 'quality'.

The most 'scientific', and seemingly objective, assessment procedures of quality are those based on bibliometric data. Since bibliometric data are numbers, you are unconsciously led to consider such procedures as objective. But these numbers are in some way created by the scientists themselves through the practice of quoting. Therefore these data are in fact inter-subjective evaluations. Moreover, there is a large variety of bibliometric data and different types convey slightly different information such as the different wording in a survey. They measure the impact on a certain scientific community of its members' products. Another distinction could be based upon the traditional tension in social research between qualitative and intensive methods and quantitative and extensive methods. Qualitative methods such as peer review are centered on scientific community evaluation as well. However, they measure the level of '**liking**', or acceptance. An author can be quoted because of his/her mistakes while on the contrary he/she is praised in a review only if his/her contribution is liked. On the one hand, **the impact** is inferred from a very large universe of cases, potentially the whole universe of scholars in a certain research field, but may convey ambivalent information. On the other, the level of '**liking**' inferred from a peer review is much less ambiguous, but it is usually based upon the judgment of a very narrow set of referees that do not represent, necessarily, the prevalent opinion among experts with the same scientific credentials.

In fact, the impact indexes themselves contain some information about the '**liking**' to the extent that these measures relate only to articles published in scientific journals where peer review method is adopted. Citations of these articles are considered simply because these articles are published. They are published because they are liked by a narrow set of experts.

2. The drawbacks of impact indexes

Other drawbacks of these measures have been the focus of discussion. First, the scientific quality is revealed over time. Impact measures can punish prematurely scientific products not yet sufficiently understood and appreciated for their value by the scientific community. In addition, the amount of citations obviously depends on the number of researchers who deal with a particular topic. A publication may have a considerable impact over another because the topic attracts more scholars than another topic and not for its intrinsic qualities. For example, in Political Science students of International Relations form a much wider community of scholars than students of Italian Politics and they are *a priori* likely to be quoted more often. Finally, the impact may ultimately reflect the extension of a network of scholars headed by powerful academics. In other words, the number of citations may reflect the level of subordination of scholars who cite, instead of indicating the degree of innovation, originality and explanatory power of the cited publication.

These difficulties should not be underestimated nor exaggerated. However, they can be mitigated. Some would suggest using the impact index of the journals in which the articles are published instead of the articles' impact, in order to minimize the problems connected with the different popularity of the topics and the local academic power. Nevertheless, the use of journal rankings is as controversial as the method of the informed peer review.

3. The informed peer review method

The third debate has centered around the use of the informed peer review method for Humanities and Social Sciences. This method combines traditional peer review with a classification of journals and it drew criticism from opposing sides of the debate, namely from both those who are against peer review and those who oppose bibliometrics. In this way the worst of two worlds is attained. Critics of the first type have pointed out that it is since Adam Smith that a warning circulates regarding the risk that peer review can be controlled by the most powerful academic groups. Therefore, it is important to know how referees are appointed, and, once the evaluation procedure is over, to have data about the referees, the number of research products that each of them have evaluated and the distribution of their evaluation. It has also been suggested that referees should know in advance that their evaluations will be revealed. It has also been proposed to start a dialogue within the scientific community on whether alternative methods now under discussion, for instance, *ex post* review, peer-to-peer review, etc, can be of any use to improve evaluation exercises in Social Sciences. According to this second type of criticism the experience in other countries seems to suggest that journal classification produces

standardization and opportunistic behavior among researchers, and discourages cross-fertilization and interdisciplinary research. For example, in Australia the original classification into three groups was abolished because it produced distortions that were deemed too serious, and was substituted by a single list that only distinguishes whether a journal can be considered scientific or not.

Giving up classifying scientific journals, however, is a choice based on specific country-based assumptions. For instance, this could make sense if we estimated that the probability of ignoring an important contribution published in a not very diffused journal is greater than the probability that a referee, devoid of information on the journal status, could overstate or understate an article because of personal idiosyncrasies, incompetence or lack of time. The prevalence of one of these dangers depends on size, pluralism, expertise, resources of the panel of reviewers and duration of the evaluation assessment.

In the evaluation process in Political Science (VQR), the informed peer review method has been used only partially. The monographs and edited books in Italian political science are still an important part of scientific production. Almost all are in Italian and there is no reliable information to classify the Italian publishers according to the procedures that are used to select manuscripts to publish. Unfortunately, with perhaps one exception, there are no scientific editorial committees sufficiently broad and plural to guarantee an authentic *ex post* quality control of what is printed. So, contrary to what happens in journals, the nature of the container generally provides poor information as to the referees, and peer review has come to be uninformed.

4. The definition of research scientific products

An even more radical dispute concerned what should be considered as a scientific research product.

The debate has focused on whether it is enough to follow a standardized procedure, for instance peer review or inclusion in the ISI database, for a product to be considered scientific, or whether only by analyzing the content is it possible to tell whether a product is scientific or not. This debate has important implications in terms of costs, because the former position has lower costs than the latter. It also allows for a faster process, and limits the incidence of subjective elements. But, opponents claim, it has the disadvantage of inferring the content of the scientific product, for instance the article, from its container, or the journal. Because of the importance and sensitivity of the issue, the CUN (National University Council) has recently launched a public consultation process in view of defining what should be considered as scientific criteria and research products.

However, the difficulties of the theoretical discussion need not dramatically affect the effectiveness of the practices. Scientific journals are normally read only, or prevalently, by experts in a particular field of knowledge and they both publish articles whose primary purpose is related to the advancement of knowledge and have mechanisms which are as neutral as possible to evaluate the quality of the article with respect to the realization of its primary purpose. Other products do not. It is up to the scientist to associate the medium chosen to distribute his/her scientific work to the proper public.

5. The *ex post* adoption of evaluation criteria

The debate has also concentrated on the *ex post* adoption of criteria. The assessment exercise in Italy started without any prior indication. At the time of the articles' publication, publishing venues were not classified. At the least, they could be formally indifferent to the researchers. Undeniably, the introduction of a new evaluation system always involves some adjustment costs related to contemporary criteria retrospectively applied to previous behavior. However, at least during the VQR in Humanities and Social Sciences, the fundamental role played by peer review should have mitigated these costs. No product has been excluded from the evaluation on the basis of criteria unknown at the time in which the product has been submitted to the VQR. Both supporters and critics presumably agree that these evaluation exercises, rather than certifying the status quo, have a transformative effect, influencing how a scientific community will behave in the future. And the new criteria must be effectively adopted, not just announced, if they have to be credible for the future. This also raises the question of who establishes the criteria and upon which basis.

6. Interdisciplinary comparisons

Finally, critics point also to the fact that non bibliometric units (GEV) could significantly diverge in their classification methods, favoring opportunistic behavior. The risk is that Universities and departments with the highest concentration of scholars evaluated by 'stricter' GEVs may receive less funds, with long-term consequences on their possibility to grow.

In principle, the assessment should always be carried out only within the same disciplines. The publishing practices and assessment criteria vary greatly among different disciplines and make the same bibliometric indicators useless for such comparisons. It makes no sense to imagine that political scientists "contend" with lawyers, philosophers and natural scientists. Therefore, the distribution of funds between disciplines cannot be based on an evaluation exercise and is inevitably a political choice. Political scientists should not be scandalized but rather create a well heard and prestigious advocacy coalition with other social scientists to maintain, and possibly to increase, the proportion of funds diverted for the benefit of our research and studies.

On all of these issues, the debate is still open and is to be welcomed, because evaluation exercises are terribly complex and involve very sensitive issues. After all, if we study these phenomena we should know how to do it.

..A new attempt at evaluating Italian research: the 2012-13 VQR

By Maurizio Cotta (<https://italianpoliticalscience.wordpress.com/author/ipsmauriziocotta/>), 15/06/2013

1. From the old to the new process of evaluation

In 2012 the new Italian Research Evaluation Process (VQR, *Valutazione della Qualità della Ricerca*) was launched. With this process Italy joined – after a long delay – the group of European countries that have decided to regularly assess the quality of their Universities and Research institutions. Under the aegis of ANVUR (the *Agenzia Nazionale per la Valutazione dell'Università e della Ricerca*), that has devised and guided the process, a committee was created for each research area to conduct the evaluation process. The process is now finished and, even if the final results are not public as yet, it is possible to analyse specific aspects of this significant event. Having been part of the group responsible for the scientific area 14 (Political and Social Sciences), I will describe the process and some of the problems encountered. I will also present some preliminary data for that which concerns the sub-area of political science.

The 2012-13 VQR has not been the first Italian attempt at evaluating Universities and research, since it was preceded by the Triennial Evaluation of Research, the VTR 2001-2003, conducted by another body set up by the Ministry of University, the CIVR (*Comitato di Indirizzo per la Valutazione della Ricerca*). This evaluation was completed at the end of 2005 and **the final report** (<http://vtr2006.cineca.it/>) published at the end of 2006. Unfortunately, instead of moving directly from the experiment to a stable process of evaluation, the whole thing stopped. Only some years later, with the creation of a new body – the ANVUR – and the formulation of new rules, could the evaluation process be resumed. This shows the difficulties that the Italian University system, and its highly centralised bureaucracy in particular, has in dealing with problems of evaluation and in setting up a regular process of assessment.

Leaving aside the organisational innovations, it must be noted that the new evaluation process significantly broadened its scope. Instead of a limited selection of research products, which each University or Research institution had to submit in each scientific area for the VTR 2001-2003, this time every permanent member of the scientific staff of Universities and Research institutions, Full Professors, Associate Professors and researchers, had to submit three products, presumably the best, published between 2004 and 2010. Thus, the amount of material to be examined increased dramatically as compared to the previous attempt. The goal was not individual evaluation of academic personnel, but to produce 'institutional scores' for Universities and Departments. In view of this choice individual evaluations will not be made public by ANVUR; only aggregate evaluations will be available.

For each of the 14 scientific areas defined by the Italian University system a group of evaluation experts (GEV) was created. In my case, I was nominated member of GEV 14, the group covering the history of International Relations, the history of political institutions, the history of Asia and America, political philosophy, political science and sociology. The group was composed of 13 members from the different sub-sectors represented. Two of them were from foreign institutions. The President nominated by ANVUR was Professor Ivo Colozzi, a sociologist of the University of Bologna. GEV 14 was then divided into two subgroups, one responsible for the broad sociological area, the other for political science, philosophy and the historical disciplines included in this area. Political science was represented by Jean Pierre Gaudin, Gianfranco Pasquino and myself.

The first step of our work was to articulate the evaluation criteria following the general indications of ANVUR and then to establish how the evaluation should be conducted. The second step was to recruit a sufficient number of referees to do the evaluations in full respect of the normal conflict of interest rules. Then the products had to be assigned to the referees and, when all the evaluations were completed, the institutional scores were to be calculated.

With regard to the evaluation criteria, the choice was to select the following: 1) scientific relevance, 2) innovation and 3) internationalisation. For each of these criteria, which were to receive the same weight, an A-D scale was adopted. The resulting scores were to be: *excellent*, *good*, *acceptable* and *limited*. The debate about the type of evaluation procedure to be adopted – bibliometric or peer review – was quickly solved in our area as only a small minority of products (approximately 280 out of more than 4000) were published in journals for which bibliometric (ISI or Scopus) data were available. All products had then to be submitted to two referees. It was, however, decided that – if possible – a double evaluation (bibliometric and peer) would be implemented with the purpose of comparing the results of the two methods. This exercise was, however, done only for study reasons and it was not bound to affecting the evaluation.

At this stage of the process it was also established which products could be considered acceptable scientific products. This was done both on the basis of the internal characteristics of the product and of the place of publication. For instance, a simple research report, or an article published in a journal of 'general culture' was to be rejected. Such cases were not so frequent, but still existed. For this purpose a list of scientific journals had to be defined and ANVUR established a ranking of their scientific quality.

2. The evaluation process

The recruitment of referees was probably one of the most crucial moments of the whole process. There was to be a sufficient number to handle all the products received, they were to be qualified to do the job, possibly there was to be also a good number of foreigners and finally they had to be willing to complete their task. The process was made more difficult because of the need to respect stringent rules on conflicts of interest. No referee was to evaluate any author coming from the same University, or any collective product where both author and referee were involved. These rules are quite obvious, but in some fields, such as political science, the relatively small number of Italian scholars, and their concentration in some universities, coupled with the need to find scholars expert in the specific sub-fields, made it sometimes very hard to find the suitable referees. Having a good number of foreign referees was potentially a solution to the problem of such conflicts of interest, but it proved to be in practice far from easy. Given the fact that a large majority of products were written in Italian, the referees had to have a sufficient command of this language and on top of this they had to be convinced to do a job for which the remuneration was more or less symbolic and to be willing to overcome some of the practical problems connected with the experimental nature of the ANVUR computer system in charge of handling the whole process of products distribution. In the end, only a relatively small group of generous foreign scholars could be involved. Their contribution was, however, very important.

At a given point of the process the products arrived. it was a bit like the moment when in 'western' movies a herd of recalcitrant cows coming from the prairies have to be corralled by the cowboys into the stables..... The procedure adopted to handle this stage was that each member of the GEV had to accept the responsibility for a share of these products and then proceed with the assignment to the referees. Each product had to be assigned by two different GEV members to a referee. The GEV members too had obviously to respect the conflict of interest rules. This means that they can not assign the products from their own University. The referees were also not to know which of the GEV members had done the assignment. As a consequence queries by the referees about a product were to be handled through a blind mechanism of communication. This part of the process, computer aided and in itself apparently simple, proved in practice much more difficult than expected. It was soon, clear that the pool of referees selected in the beginning was too small for the number of products to be assigned. For the whole GEV 14 the products to be distributed were more than 4000; out of them 573 for political science. Given the need of two evaluations for each product, it meant 8000 evaluations for the whole GEV and 1146 for political science. Just for political science this number would require at least 55 referees and each one to evaluate about 20 products, mostly articles, but very often books. However, it was not only a matter of sheer numbers, it was also a matter of expertise. It is true that a majority of the products were in the mainstream areas of the discipline, but still a substantial number came from rather obscure and peripheral corners of research for which it was often rather difficult to find an expert. In addition to this, referees could obviously refuse an assignment. And, indeed, to a great extent they used this faculty. In the best of cases they refused explicitly, in the worst they did not answer. It became soon clear that the original number of referees selected had to be greatly enlarged, if the task was to be accomplished. And, also, the members of the GEV had in the end to review many products for which it was not possible to find a referee available. It must be mentioned, though, that a number of very dedicated referees accepted to do many more reviews than was originally envisaged.

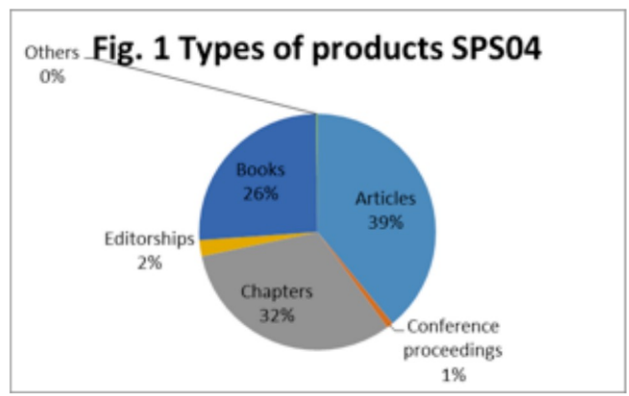
Whenever the evaluations of the two referees coincided the process was concluded, but this situation was far from common and in a good number of cases the evaluations differed not only by one degree, but even by two or three. In these cases a consensus committee had to be established involving three members of the GEV and, eventually, a third referee to define the final score. Only by the end of April 2013 could the whole process be finished.

3. The products in the field of Political Science

Without waiting for a more systematic and detailed assessment, it might be interesting to anticipate a few remarks about the products submitted to evaluation. I will limit my analysis to the subsector of political science that was more directly under my attention. A first positive point that must be stressed is the breadth and variety of the themes covered. In spite of the still relatively small number of political scientists in the Italian academic system, the range of the subfields of the discipline covered by their research efforts is greater than could have been expected. The second point concerns the type of products submitted. The results, which are still provisional but not far from being final data, show that for political science

(SPS04) ‘articles’ have become the largest group, followed by ‘chapters in edited volumes’ and by ‘books’ (Fig. 1). However, the two last categories together still represent the majority of products. Italian political science follows, probably with some delay, the trend prevailing abroad.

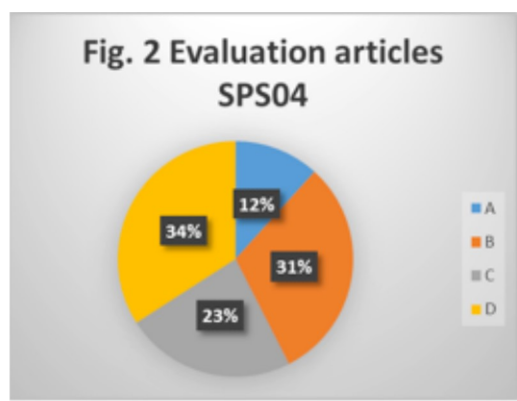
Types of academic products for class SPS/04.



(<https://italianpoliticalscience.files.wordpress.com/2013/06/cotta1.png>)

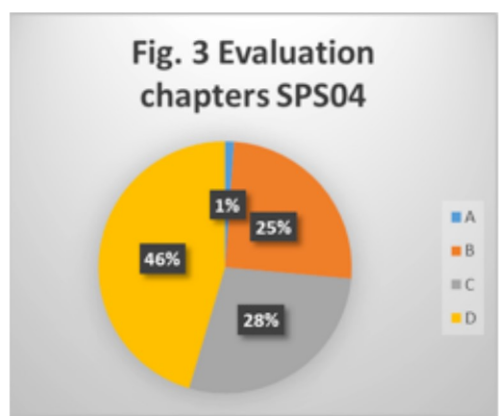
With regard to the evaluation of the products from A to D, where A is excellent and D is limited, it is interesting to notice that articles score on average better than chapters in collective books, but not of books (Fig. 2, 3 and 4).

The evaluation of articles produced in the SPS/04 class.



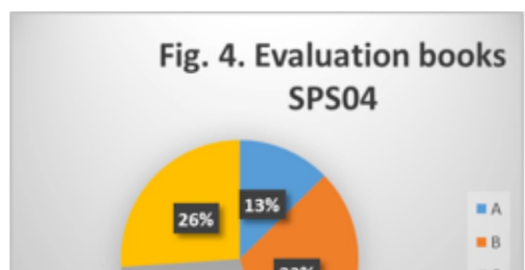
(<https://italianpoliticalscience.files.wordpress.com/2013/06/cotta2.png>)

Evaluation of book chapters in class SPS/04.



(<https://italianpoliticalscience.files.wordpress.com/2013/06/cotta3.png>)

Evaluation of books in class SPS/04.





(<https://italianpoliticalscience.files.wordpress.com/2013/06/cotta4.png>)

In order to assess more carefully the state of the discipline, all these data will require a much more detailed analysis and discussion within the professional organisation of the discipline. It would be, for instance, important to explore more in depth the degree of internationalisation of Italian political science by looking at the language and places of publications.

4. A final conclusion

A final note is required for the whole evaluation process. The process was far from perfect as many have noted. Some conceptual and organisational aspects will require careful review and discussion. It may be, for instance, debated whether it was wise to require such a large amount of products to be evaluated. The advantages and disadvantages of peer review also need to be better assessed. It is clear also that the job done by the referees was not always optimal for objective and subjective reasons such as the difficulty in finding enough expert referees for all products, the criteria of evaluation that were probably not sufficiently clear, the lack of professionalism in some referees, etc.. This being said, I think that the process was far from being redundant as it forced all the members of the academic community to think more carefully about their research achievements and because it will provide the Ministry of University, Universities and Departments with an instrument to assess the quality of academic life. The instrument can be improved, but this requires that evaluation does not remain an isolated case, but becomes a regular exercise.

An 'outside' perspective on the research assessment exercise

By Manuela Moschella, 15/06/2013

With these interviews, we collect an 'outside' perspective on the research assessment exercise. Specifically, we interviewed Professor James Newell, one of the external evaluators in the Research Quality Evaluation 2004-2010 (VQR 2004-2010), and Professor Tony Payne, the 2008 Chair of the panel on Politics and International Studies in the UK Research Assessment Exercise on the strengths and weaknesses of this procedure.

*Interview with
Professor James
Newell,
University of
Salford, UK,
External
Evaluator for
Anvur (April 2,
2013)*

IPS: Could you tell us how were you selected and what was the mandate you received?

JN: I was contacted by email by Anvur staff upon their own initiative. I think I was selected because I had already participated to a number of qualitative assessments of research projects funded by the Italian Ministry of Education/University (i.e. Prin, Firb). I was also contacted because of my research interests. Indeed, I was asked to review the scientific output of Italian scholars whose research agenda covers both Italian politics and history.

IPS: How large was the scientific output you were asked to assess and how much time were you granted for conducting the assessment?

JN: I received around 25-30 research products (i.e. articles, books, book chapters) to be reviewed within around six months (I was originally contacted in the summer 2012). After the first round of evaluation, in January 2013, I was once again contacted to review another set of research papers (around 12 pieces). I have just finished the referee process.

IPS: What are the criteria that Anvur suggested following in the research assessment?

JN: Each evaluator is asked to assess the research output based on three criteria: relevance, originality and internationalization. For each criterion, different scores are provided. For instance, relevance is ranked between 0-3 and refers to an assessment of whether the research product contributes to the advancement of knowledge in the field. Originality refers to the disclosure of new findings that can be taken up by other scholars in the community. Under the criterion of internationalization, evaluators are required to assess the positioning of the research product in the international scenario, in terms of importance, competitiveness, editorial spreading and appreciation from the scientific community. On top of these criteria, reviewers also have the discretion to add a very short commentary on the reviewed product.

IPS: I understand that your evaluation was limited, meaning that it covered scientific outputs that fall within the research area of Italian politics. However, was it possible to discern any pattern/common themes of investigation in the production of Italian scholars?

JN: I found that that one of the major areas of investigation is electoral behavior and political parties. However, as you said, this also reflects my research interests and it may not be indicative of the direction of the Italian political science scholarship at large.

IPS: And in terms of quality, what is your general assessment of Italian scholarship?

JN: I found that quality is extremely high. And I confess I was (as I always am) seriously impressed. The Italian political science community is a relatively 'young' one in that it was born in the 1970s. In spite of this, Italian scholars have been able to develop solid research programs.

IPS: You are also a faculty member of a UK University that is subject to the local research assessment exercise (RAE) where the evaluation process, like the one that is taking place in Italy, is Department-based and driven by ex ante specified criteria. In your view, what are the main problems in the UK RAE?

JN: From my perspective, one of the major problems in the UK is that research is assessed in terms of its utilitarian value. That is to say, one of the key criteria upon which research is ranked is its 'impact'. This is not to suggest that practical implications are unimportant. However, research has also a value on its own.

IPS: What about the criticisms that could be raised at the Italian evaluation? Do you have any?

JN: Reviewers like me are asked to review ex post. That is to say, we review articles/books that have already been published and that have already gone through a referee process. In these cases, and especially in those cases where the output has been published in prestigious journals, the assessment cannot but be positive! Of course this is not just a problem in the Italian evaluation process but also in the UK.

*Interview with Professor Tony Payne, University of Sheffield, UK,
Chair of the panel 'Politics and International Studies', UK RAE 2008 (30 April 2013).*

IPS: The UK RAE has a longer tradition than the Italian one. When did it start and how regularly is it conducted?

TP: It began somewhat tentatively in 1987 but has since settled into an event that re-occurs approximately every five years or so. We had the first full research assessment in 1992, then in 1996, in 2001 and in 2008. The new one is currently taking place and will be completed in 2014. [The 2014 research assessment has been renamed into the Research Excellence Framework (REF)] Everyone expects to be a further REF in, say, 2020, but that has not been announced. In other words, the whole process has grown incrementally, and been adjusted as it went along.

IPS: The results of the research assessment provide a picture of the research quality of Departments as a whole and not of single researchers. Is that correct?

TP: Yes, the assessment is meant to evaluate Departments, or Units to use the RAE jargon, as whole. This means that the results that are made public are those that relate to the research profile of the Department. Of course, the way we get at these results is by accumulation of assessment of single research outputs. In general, each Department submits up four research outputs for each member of the Department. The relevant panel then reads them, assesses them and aggregates the results. However, we do not publish the results for each researcher.

IPS: Is the main purpose of the RAE/REF that of allocating public funding to the Universities with the highest scores in research quality?

TP: Yes, it is the main purpose of the RAE to identify the Universities that should be the recipient of the greatest bulk of government funding for research. The principle is that funding is allocated on excellence. Of course, establishing status, reputation and esteem is an ancillary goal of the research assessment exercise. This matters a lot because it shapes the overall standing of a Department and directly affects its capacity to recruit students both domestically and internationally.

IPS: Could you please tell us something about the procedures through which each panel is created and assess the research output in its field?

TP: The panel members are nominated by the relevant professional association. As for the panel I chaired in 2008 [i.e. Panel Politics and International Studies], members were nominated principally by the Political Studies Association (PSA) and the British International Studies Association (BISA). Associations usually nominate senior members of the profession in whom they confidence. In 2008, the members of the Panel were 16. Once the panel had been created, we wrote and published the subject-specific criteria that were to inform the evaluation drawing on general considerations set at the national level. In 2008 subject panels had considerable room of maneuver and flexibility in adapting the criteria to the specificities of the subject area. The panel members then proceeded with the assessment, which involved panel members reading and assessing all of the research outputs that had been submitted.

IPS: Were all the panel members political scientists? Did it happen that you needed the help of other social scientists (not political scientists) to evaluate the research outputs?

TP: Yes, all panel members were political scientists, assuming for this purpose that this term includes specialists in international relations (IR), some of whom, as we know, prefer to think of IR as a separate field in its own right. The RAE system was, however, very open and accommodating to the assessment of interdisciplinary work and provision existed for the 'cross-referral' of the assessment of the work on the edge of a subject panel's competence to members of another appropriate panel.

IPS: What's your take on your experience as a Chair of an evaluation panel?

TP: I'm very satisfied with the work the panel did. We took the job seriously, read and discussed all research outputs, especially when disagreement arose. In other words, it was a totally 'in-house' evaluation process – a process that totally preoccupied us for some 6-7 months full-time. Furthermore, as anticipated, in 2008 each panel was granted significant flexibility in translating the general criteria into the operational guidelines that informed our research assessment. This means that we interpreted the four star-ranking in a way that was suitable to the research output in the field of political science and international studies [i.e. in 2008, research outputs were graded on a 4* ranking ranging from a 4* score that indicates research quality that is world-leading in terms of originality, significance and rigueur to 1* score that indicates quality that is recognized nationally in terms of originality, significance and rigueur.]

IPS: You stressed twice the flexibility you enjoyed in 2008 in specifying the criteria for the research assessment. Has the situation changed since then?

TP: Yes, one of the major changes that are taking place with the shift from the RAE to the REF is that there will be more centralized control on the activities of the subject panels. They will all have to work to tighter central guidelines and report in emerging outcomes to so-called 'main panels' covering, say, pretty much the whole of the social sciences. The idea is to create greater consistency in the assessment process, but the practicalities of this may just generate a regression to the mean.

IPS: Next to this decreasing flexibility, what are, in your view, the sources of potential problems in the forthcoming REF?

TP: I think that one problem could stem from the formalization of 'impact' as one of the categories that informs the research assessment. The problem here is that this benchmark risks 'unlevelling the playing field.' Whereas it is possible to find common ground for all subjects, from medicine to physics to the social sciences, on what good research is about (including core principles such as rigor, originality and significance), it is much more difficult to compare across subjects based on the criterion of 'impact'. Some subjects are much better placed than others in having an obvious impact. A journal article reporting on a new pharmaceutical product is likely to have a greater impact than a philosophical article. Hence, giving a formal status to 'impact' in the evaluation (and scoring it as the REF is currently doing) could create disparities between subjects and definitely opens up a number of challenges.

IPS: In spite of these potential shortcomings, the research assessment exercise is deeply ingrained in the culture of UK Universities. Why do you think this is the case?

TP: I think that the legitimacy of the research assessment derives from one its major strengths, namely, its peer-review nature. The process is run not by civil servants but by senior academics who have been nominated by the professional associations based on trust. Furthermore, the process is neither mechanical nor mathematical in that it is based on careful assessments carried out among peers. Of course, the fact that it is human beings who carry out the research assessment is no guarantee that no problems will ever arise. However, the fact that researchers own the evaluation process is certainly a positive aspect of the whole procedure and lends legitimacy to it.

IPS: What are the main weaknesses of the research assessment exercise then?

TP: It costs a lot of money and energy. People involved in the evaluation work full time for the assessment for quite a period of time. The other major problem is that, in the UK at least, the assessment increasingly certifies what we already know. The Departments you would have expected to perform better than others tend to demonstrate that better performance. The system rewards the winners and enables them to go on winning! By this I mean that funding for research goes to good Departments which then hire the best young scholars, allocate resources wisely among staff members, and so continue to do well in the RAE and the REF. In contrast, the Departments that before the RAE/REF had perhaps been struggling with the competition will largely continue doing so. They do not gain enough resources to make a real difference to their competitive position. This gives rise to two opposing views on the RAE/REF. On the one hand, it is argued that we should give up completely on the assessment because there is a well-established ranking among UK universities. On the other hand, however, the research assessment is a way to avoid hierarchies of reputation being frozen and thus keeping alive a need to 'perform' in research by the better Departments.

IPS: Do the panel members receive a fee for their service? In the affirmative, it is a 'good' fee?

TP: Yes, panel members were paid a fee, but only for attendance at meetings. The massive amount of 'reading time' involved was unpaid, which means that the RAE was in effect hugely subsidized by UK universities releasing key members of staff to its needs for considerable periods of time. No panel member would have ever taken on the huge amount of extra work involved merely for the fee earned!

IPS: What is your final consideration on the research assessment exercise that could be relevant for a country like Italy that has just started experimenting with its own research assessment?

TP: My overall assessment on research assessment exercises is positive. As usually the case, however, what is good when it is a moderate thing may risk becoming a negative thing when you overdo it. For instance, this is the situation we are confronting in the UK at the moment with a REF that has become increasingly complex and open to game-playing. Under these conditions, the risk is that research will not be driven fundamentally by the joy of producing a new theory or new results but just by a managerial concern on the part of Universities to do well in competitive research assessment.

Italian research funds: escaping from the PRIN labyrinth

By Marco Brunazzo, 15/06/2013

Introduction

In the last few months a number of researchers have been involved in the submission process of the PRIN 2012 projects. PRIN is the acronym of **Progetti di ricerca di interesse nazionale** (Research projects of national interest) and refers to one of the major Italian public sources of funding for academic research. The submission to the new proposals follows the recent publication of the PRIN 2010/2011 projects admitted to and cofounded by the Italian Ministry of Education, University and Research (tab.1).

Table 1. PRIN 2010/2011: Projects admitted to cofounding

National Coordinator	Title	University of the national coordinator
Laura Bazzicalupo	Biopolitical governance: inclusion and happiness. Historical and conceptual genealogy, present challenges and future perspectives of a politics for man and society	University of Salerno
Laura Bovone	Sustainable practices of daily lives in a context of crisis: job, consumption, participation	Università Cattolica del Sacro Cuore
Pierangelo Isernia*	Italian foreign policy and the new challenges of the international system: actors, institutions and policies	University of Siena
Mario Morcellini	Professions of the public space: beyond the crisis	University of Roma La Sapienza
Leonardo Morlino*	Economic crises and quality of democracy in Europe	LUISS University at Roma
Ida Regalia	Old and new modes of job regulation in Italian and European small enterprises	University of Milano
Antonio Schizzerotto	Social belongings, beliefs about instruction and participation to the university: an integrated experiment with a longitudinal survey	University of Trento
Paolo Segatti	Political representation changes in Italy. Voting decision in the 2013-2015 electoral cycle	University of Milano
Carlo Trigilia	University, innovation and regional economies	University of Firenze
Pierluigi Valsecchi	State, plurality, changes in Africa	University of Pavia

Note: Due to the lack of information and transparency of the Miur Internet site, I would prefer not to give the list of the research units, which would be incomplete. At the same time, I have decided to give the complete list of all the projects funded in the research area of Political and Social Science because of the possibility of multidisciplinary research projects with, for instance, political scientists participating in sociologists' research units and vice versa. The names of the coordinators belonging to the disciplinary sectors of Political Science are also emphasized by an asterisk.

After the publication of the PRIN 2012 call for proposals, colleagues have been confronted with the new procedures required for submission. After the path-breaking approach introduced in the last 2011-2012 call, the Miur has promoted different relevant reforms. The main ones are the following:

1. the pre-selection phase—until now carried out under the responsibility of the single universities—will be based on draft projects;
2. the criteria used in the pre-selection phase are now partially defined by the National Committee of Research Guarantors (*Comitato Nazionale dei Garanti della Ricerca* – CNGR) created after the approval of the Gelmini Reform;
3. instead of making reference to the 'traditional' fourteen disciplinary areas, the projects are now divided according to the three ERC domains;
4. there are three lines of interventions, defined according to the time-span between the coordinator's first PhD, specialization or Bachelor and the publication of the 2012 call; specific funds are reserved for younger scholars;
5. there are no limits to the number of operative units participating in projects and no minimum and maximum costs associated with the projects.

I will now, briefly, illustrate the new requirements and procedures established by the 2012 call for proposals, with the aim of rendering them more familiar to the research community. The changes introduced by the 2012 call for proposals have on occasion been the cause of disorientation among researchers and professors. Finding the way out of this labyrinth is not always easy...

Evaluation and selection

The procedure introduced in the 2010-2011 call was confirmed in 2012. The evaluation of the submitted proposals is organized in two different steps: the first is under the responsibility of the single university, the second is organized by the MIUR itself. The Ministry works by means of Selection Committees (*Comitati di Selezione – CdS*) organized according to the three European Research Council domains (*Social Sciences and Humanities – SH; Life Sciences – LS; Physical Sciences and Engineering – PE*), whose members are appointed with Directorial decree after designation by the CNGR (tab. 2 and tab. 3).

Table 2. Composition of the National Committee of Research Guarantors (CNGR)

Name	Institution
Angelos Chaniotis	University of Princeton
Daniela Cocchi	University of Bologna
Anna Maria Colao	University of Napoli
Alberto Sangiovanni Vincentelli	University of California, Berkeley
Francesco Sette	European Synchrotron Radiation Facility, Grenoble
Vincenzo Barone	Scuola Normale Superiore, Pisa
Claudio Franchini	University of Tor Vergata, Roma

Note: CNGR members have been nominated by the Ministerial Decree D.M. 7645 of 26 April 2012.

Table 3. Composition of the CdS Social Sciences and Humanities – SH

Panel	Name	Institution
SH 1 Individuals, Institutions and Markets: economics, finance and management	Massimo Bordignon	Catholic University of Milano
SH2 Institutions, Values, Beliefs and Behavior: sociology, social anthropology, political science, law, communication, social studies of science and technology	Giovanna Colombini	University of Pisa
SH3 Environment, Space and Populations: environmental studies, demography, social geography, urban and regional studies	Francesco Billari	University of Oxford
SH4 The Human Mind and Its Complexity: cognition, psychology, linguistics, philosophy and education	Michela Cennamo	University of Napoli
SH5 Cultures and Cultural Production: literature, visual and performing arts, music, cultural and comparative studies	Lina Bolzoni	Scuola Normale Superiore di Pisa
SH6 The Study of the Human Past: archaeology, history and memory	Tonio Hölscher	University of Heidelberg

Note: CdS members have been nominated by Directorial Decree D.D. n. 648 8 April 2013.

Once they have gathered the evaluations of anonymous referees, the CdSs are invited to express their evaluation according to a 5-point scale (tab. 4).

Table 4. The evaluation scale of the projects

Category	Description	Points
Excellent	Fully convincing, without weakness	5
Very good	Extremely strong with, at most, some minor weakness	4
Good	Strong but with some moderate weakness	3
Fair	Some important weakness	2
Poor	Not very convincing with numerous weakness	1

Note: CdS members have been nominated by the Directorial Decree D.D. n. 648 of 8 April 2013.

Only the projects that have been pre-selected by the single universities are submitted to evaluation by the MIUR. Each University shall pre-select a number of proposals not exceeding 0,75% of the number of professors and researchers of the university or, if superior, not exceeding the double of the mean of the projects financed in the last five PRIN call for proposals (tab. 5).

Table 5. Italian Universities that can pre-select more than 20 projects

University	Max N of pre-selected projects	University	Max N of pre-selected projects
------------	--------------------------------	------------	--------------------------------

University of Rome La Sapienza	122	University of Genoa	32
University of Bologna	75	University of Pavia	32
University of Milan	74	University of Perugia	29
University of Naples Federico II	73	University of Siena	28
University of Padua	73	University of Palermo	26
University of Florence	66	University of Trieste	26
University of Pisa	48	University of Roma Tre	25
University of Turin	42	Polytechnic of Turin	24
University of Rome Tor Vergata	35	Second University of Naples	22
University of Catania	33	Catholic University of Milan	20
Polytechnic of Milan	33	University of Milan Bicocca	20
University of Bari	33	University of Parma	20

Pre-selection is based on the judgment of three anonymous referees that take into account the innovativeness and originality of the proposed project, the methodology and scientific qualifications, in relation to the submitted project, of the scientific coordinator and of the managers of the operative units. Each university is subsequently responsible for the definition of the pre-selected projects to be newly submitted to the MIUR.

Characteristics of the projects

First of all, PRIN projects should be submitted by a 'Principal Investigator' called the PI, who is a Professor or researcher that coordinates several operative units and has the scientific responsibility of the entire project. The operative units are led by a local manager, in charge of organizational responsibility, and are composed of Professors and researchers.

The 2012 call for proposals is organized along three lines of action, according to the time span between the PI's first PhD, specialization or Bachelor and the publication of the call for proposals. The first line (line A) is called 'Starting PRIN – young researchers', the second line (line B) is called 'PRIN Consolidator', and the third line (Line C) is called 'PRIN Advanced' (tab. 6). Mixed projects are allowed: in line A and B researchers can belong to line A and/or B but not C; in line C researchers can belong indifferently to line A or B or C.

Table 6. The temporal limits applied for the definition of the PI's line of belonging

Line A – PRIN starting	PhD/Specialization ≤ 7 or bachelor ≤ 10
Line B – PRIN consolidator	$7 < \text{PhD/Specialization} \leq 12$ or $10 < \text{bachelor} \leq 15$
Line C – PRIN advanced	PhD/Specialization > 12 or bachelor > 15

The MIUR finances the approved projects with 70% of the costs which are considered adequate for the implementation of the research. The approved projects will have at their disposal no less than 80% of the costs considered adequate. The Social Sciences and Humanities (SH) domain has €7,651,978 available for funding, half of the funds provided for the other two domains (€ 15,303,958 each). In the SH domain at least €750,000 is reserved for 'PRIN starting' and €1,000,000 for 'PRIN consolidator'. Results of the national selection are due on 20 October 2013 (tab. 7).

Table 7. Principal deadlines for the 2012 call for proposals

Date	Activity
11/02/2013	Submission of a brief project proposal by the PI to its university
28/02/2013	The CNGR nominates the CdS' members
19/04/2013	End of the peer review for the pre-selection phase
10/05/2013	Communication of the pre-selected proposals
14/06/2013	Submission of the complete detailed proposal to the MIUR
30/09/2013	End of the peer review for the selection phase
20/10/2013	Communication of the selected projects

Conclusion

The PRIN 2012 call for proposals introduced several changes to the selection procedure. It is plausible that the next call for proposals will change yet again. The aim of the MIUR is apparently that of making the national system of research funding more

similar to the one adopted by the EU and more in line with Horizon 2020. For this reason, one may expect that in 2013 scholars will enter into yet another labyrinth.

For more information

General information about the PRIN is available at [this link](#).

The general call for proposals is available [here](#), with amendments introduced by the following [ministerial decree](#).

Some universities have elaborated some operative documents and Q&A. As examples, one can see those available at the [University of Trento](#), [the University of Milan](#), and [the University of Padua](#).