

Evaluation of Quality Research in Italy: a comparison between different models adopted in the past

Paolo Miccoli

Pisa University, email: paolo.miccoli@unipi.it

Published: 2 Sep 2024

Abstract

The fourth research quality evaluation (VQR4) recently launched by the National Agency for the Evaluation of Universities and Research Institutes (ANVUR) is now an integral part of the general evaluation process of Italian universities. After an introduction to the concept of evaluation, the previous VQR, known as VQR3, is critically examined through studies by authors who played a predominant role in it. It appears that, with each exercise, the methodologies have been refined, the scope has significantly expanded, and researchers have become increasingly familiar with the processes.

Keywords: evaluation italian universities; quality of research evaluation.

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

As the start of the next Quality of Research evaluation (VQR) approaches, it is worth very briefly revisiting the topic of evaluation and the experiences gained during the previous one, namely the VQR 2019-2023 (VQR3).

First of all, we still question whether it is truly possible and correct to evaluate Research. Since the beginning, a clear fault line between two very different positions has emerged among scholars, particularly sociologists who have dealt with this issue. We refer especially to the opposition between two models:

- The constructivist model, which describes evaluation as impossible (Pierre Bourdieu)
- The normativist model, which considers the evaluation both possible and desirable, essentially through the mechanism of citations, in its various components referring to Priority Rule, Recognition, and Peer Review (Robert K. Merton).

According to Andrea Bonaccorsi, "the Mertonian competitive model... is not only normatively but also descriptively adequate in all hard sciences. On the contrary, there are good reasons to believe that forms of segregation are at work in the humanities and social sciences (horizontal stratification)."[1]

The philosopher Massimo Mori, in a review of Bonaccorsi's book *La Valutazione possibile*, states the following about evaluation: "Awareness of this historically artificial origin of evaluation should not be alarming: it does not mean rejecting evaluation, but rather understanding its unavoidable necessity. Indeed, recognizing its artificiality can better account for the complexity of the project, the difficulties of implementation, the gaps that distance certain results from the optimal solution, and the inevitable discrepancy between formalization/quantification of evaluation and substantive judgment. But it is also an incentive to exercise tenacity, to invent new solutions and criteria of consensus not yet tried to solve an evaluative problem that is unavoidable and cannot be postponed, and at the same time limit the disadvantages of the solution."[2]

Even though many have believed that one of the limitations of the VQR exercise is represented by the significant disparity in criteria used at different times in the various exercises, leading to a constant realignment by universities, it is worth noting that concepts

related to evaluation inevitably change over time. The National Agency for the Evaluation of Universities and Research Institutes (ANVUR) itself has had to adjust its approach based on some possible aporias that may become evident when analyzing VQR results. To put it in a nutshell, ANVUR is dynamic and capable of quickly intercepting changes in an open and competitive society.

For this reason, comparing the results across different VQR exercises could be helpful, on the one hand, to guide legislators more comprehensively and informatively towards formulating clearer, simpler, and more understandable guidelines and calls by ANVUR for universities. On the other hand, it allows researchers to submit their work as consciously as possible.

An analysis conducted by the Polytechnic University of Turin, particularly by Professor Gianluca Setti, showed that when comparing the first two VQR periods (2004-2010 vs. 2011-2014), the percentage difference between IRAS 1 and dimensional quota revealed that the curve expressed by the second VQR is higher and closer to the average than that of the first. This suggests that universities have significantly converged in terms of research quality (Figure 1). Continuing his analysis in a comparison among the three previous VQRs (Figure 2), Setti demonstrates how the performance distribution of universities in VQR1, VQR2, and VQR3, in terms of the average score of the submitted products normalized by area, shows that, compared to the two previous exercises, the distribution of results in VQR3 is much more concentrated around the value corresponding to 0, i.e., the dimensional quota. Most analysts agree that this convergence may be due to an optimization in the submission of products by universities, which in turn resulted from ANVUR's choice to allow universities more freedom in selecting the products to submit.

Among the potential sources of distortion, ANVUR itself notes, for example, an excessive variability in the distribution of merit classes: particularly in Class A, where values range from 41.6% in Area 2 (Physics) to 13.6% in Areas 8a and 14 (Architecture and Political and Social Sciences) and, somewhat surprisingly, to 16% in Area 6 (Medicine).

A distinctive feature of the third VQR, indeed the most distinctive, was the significant emphasis placed on the so-called Third Mission. Perhaps the biggest surprise in this activity

was the very high frequency with which universities, even those with a distinct STEM or applied focus like the Polytechnic universities, chose public engagement activities as Case Studies, i.e., those with the highest social impact (Figure 3). This field of application had the flattering result of being the most chosen overall with 33% of selections.

I leave the conclusion to the words of Alessandra Celletti, Professor of Mathematics at Tor Vergata University of Rome, Vice President of ANVUR, and Head of the latest VQR (VQR3), who summarizes what she considers the most innovative and positive aspects of this latest evaluation exercise: "The informed peer review represented a significant difference compared to VQR2, in which bibliometrics played a central role, as it limited the automatism of bibliometrics itself. The discretion left to institutions to choose the number of products for each researcher (up to a maximum of 4) allowed for optimized choices but made it difficult to adhere to the percentages foreseen, although indicative, in the different categories. For the first time, significant importance was attributed to the training of PhD students through the R3 or IRAS3 indicator, which measures the research quality of faculty members at an evaluated institution who earned their PhD in the five-year period 2012-2016. The Third Mission represented a significant innovation. Some fields of action included by the call were significantly selected (e.g., public engagement), unlike others (e.g., open access tools)."

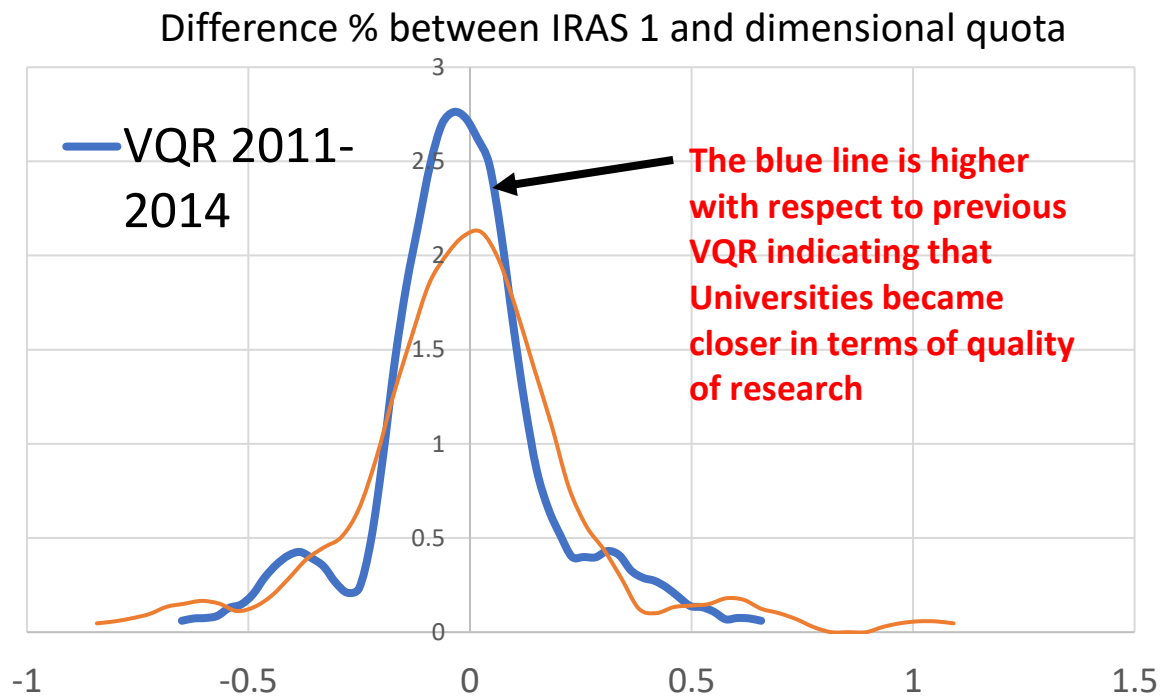
I believe it would be difficult to deny the evidence of what the Vice President stated, and it seems to me that the concept of research evaluation by universities and research institutions has indeed been strengthened by this third VQR. The fourth one, which has just begun, finally ensures that the evaluation process itself, refined during previous exercises, has now become a stable part of the Italian university system.

Declarations

Conflict of Interest

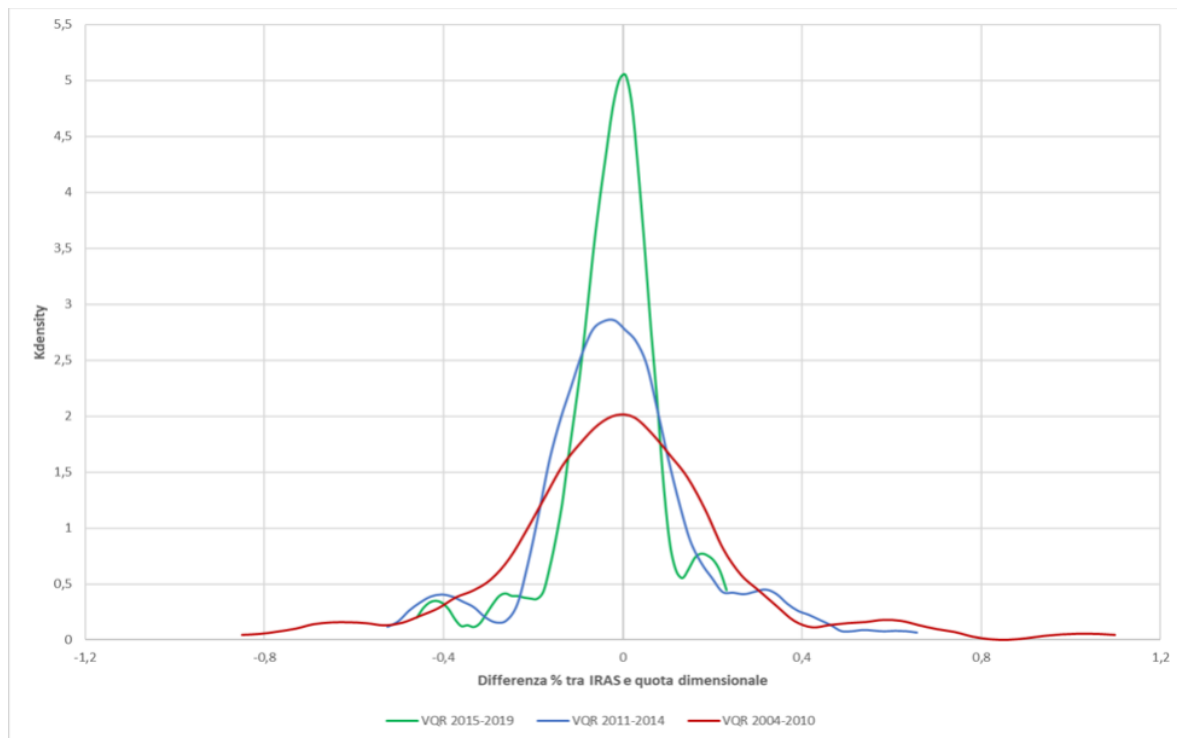
The Author declares that there is no conflict of interest.

Figure 1



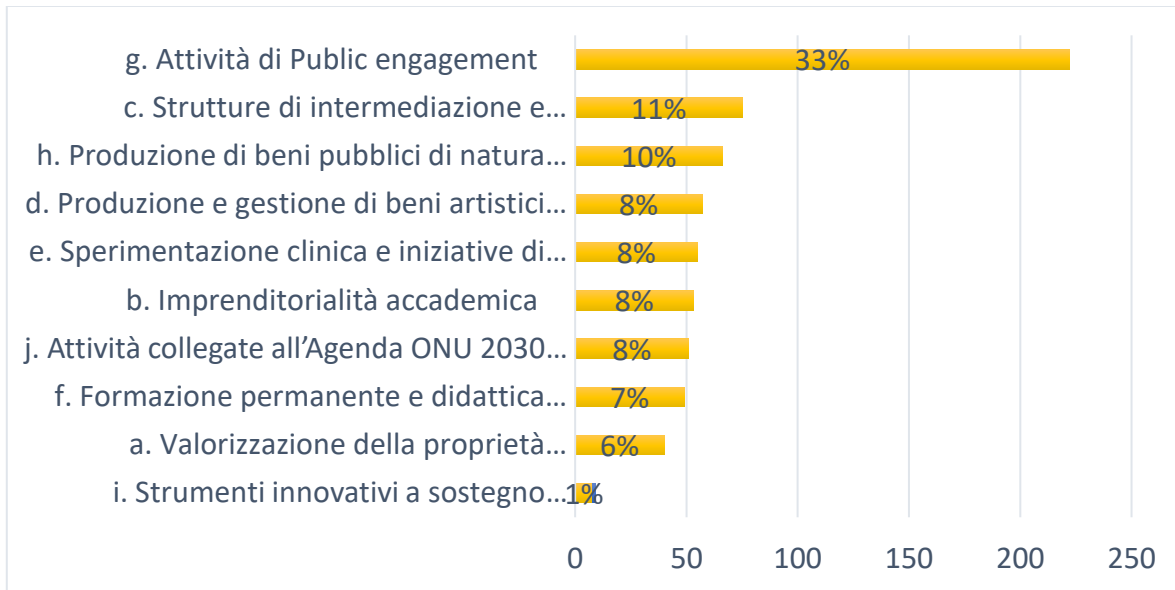
(By appointment to Politecnico di Torino, Professor Gianluca Setti)

Figure 2



- 1) Informed peer review represented an important novelty with respect to previous VQR where bibliometry played a pivotal role, limiting the distortions by pure bibliometry
- 2) Discretionary power given to Institutions to choose the products, until a maximum of 4 items allowed an optimization in the process of choice but made difficult to respect the expected percentages.
- 3) For the first time an important role was attributed to the products coming from the Doctorate Fellows also with respect to the Institutions they were coming from
- 4) The Third Mission represented an important novelty. In particular the importance of public engagement was surprisingly observed, even among Institutions devoted to STEMS

Figure 3



ANVUR data available at the link: <https://www.anvur.it/attivita/vqr/vqr-2015-2019/rapporto-finale-anvur-e-rapporti-di-area/>

Translation items in Figure 3:

- g. public engagement activities: 33%
- c. intermediation and technology transfer structures: 11%
- h. production of public goods of a social, educational nature and policies for inclusion: 10%
- d. production and management of artistic and cultural goods: 8%
- e. clinical experimentation and health protection initiatives: 8%
- b. academic entrepreneurship: 8%
- j. activities related to the UN 2030 Agenda and Sustainable Development Goals (SDG): 8%
- f. lifelong learning and open education: 7%
- a. valorization of intellectual or industrial property: 6%
- i. innovative tools to support Open Science: 1%

References

1. A. Bonaccorsi in: *Valutare la Ricerca?* P. Miccoli, A. Fabris, Ed. ETS, Pisa 2012
2. Massimo Mori *Recensione al Volume: La Valutazione Possibile* A. Bonaccorsi, Il Mulino, Bologna, 2015