STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

Does a new research-funding model ask for a new evaluation framework? The case of Serrapilheira in Brazil[1]

Sergio Salles-Filho[*], Adriana Bin[**] and Ana Maria Carneiro[***]

[*] *sallesfi@ige.unicamp.br*
Department of Science and Technology Policy, University of Campinas, Rua Carlos Gomes 250, Campinas, 13083-855 (Brazil)

[**] *adriana.bin@fca.unicamp.br;*
School of Applied Sciences, University of Campinas, Rua Pedro Zaccaria 1300, Limeira, 13484-350 (Brazil)

[***]*anamaria@nepp.unicamp.br*
Center for Studies of Public Policies, University of Campinas, Av. Albert Einstein 1300, Campinas, 13083-852 (Brazil)

Introduction

The purpose of this manuscript is to present a case report of an on-going experience of impact evaluation applied to a recently created S&T funding agency. The Institute Serrapilheira (ISP) is a private nonprofit organization created to fund new and challenging themes of research.

Funding agencies are more and more creating systematic impact evaluation processes. Internally, managers and policy makers are looking for best practices to enable best designs of policies and programs. Externally, society is demanding a better and clearer understanding of what science and technology can do for different categories of stakeholders.

A common characteristic of impact evaluation in S&T funding agencies has been multidimensionality, here understood as measuring indicators in different – and complementary – dimensions, as for scientific, technological, economic, social, environmental, to mention the most frequent. Another feature that is growing in relevance is the multi-stakeholder perspective, which means adopting the perspectives of different stakeholders when measuring impacts.

More recently, impact evaluation in S&T has explored non-traditional indicators as for the generally known altmetrics, which considers metrics of research from different digital media in order to look for impacts other than citations within the scientific domain (Tweeters, readers, likes etc.).

On top of these trends, new models of funding S&T arise in different countries particularly dedicated to foster young researchers with high-risk / high-return research proposals (Gewin, 2014). In the present research we are particularly interested in measuring traditional and non-

---

[1] This work is being developed by the Laboratory for Studies on the Organization of Research & Innovation (GEOPI) from University of Campinas with the support of Serrapilheira.

traditional ex-post impact and also to identify indicators able to discuss the so-called academic rising stars – young researchers who gather specific characteristics pointing to potential high impacts with time accumulation (Gogoglou and Manolopoulos, 2017; Ding et al., 2018).

Serrapilheira Institute: a new research-funding model in the Brazilian landscape

Serrapilheira is a private nonprofit institution created in March 2017 to promote and increase visibility of science in Brazil. Its mission is to support young and promising scientists that are posing bold scientific questions in their fields of research.

Among the objectives of the Institute there are themes that go beyond traditional funding, including the stimulation of scientific careers, scientific literacy, gender and race diversity in science and the promotion of a culture of science. Serrapilheira funds creative and innovative research projects and science outreach in the fields of hard sciences, including chemistry, computer science, earth sciences, engineering, life sciences, mathematics, and physics.

Serrapilheira brings innovation to the Brazilian scenario of science funding in at least three directions. First, ISP is the first family-philanthropically funded organization totally dedicated to support science in Brazil. This isn't a new model in the world. Some similar initiatives can be found in US, Canada and Europe, such as The Velux Foundation (Denmark), Thiel Foundation (US and Canada), Future of Life Institutes (US) or Chan Zuckerberg Initiative (US), just to mention a few.

Secondly, as a private institution, it can provide researchers with greater flexibility in terms of resource allocation comparing to public funding agencies. The 2018 estimated budget is US$ 4,5 million. Thirdly, in its first call for proposals launched in mid-2017, ISP asked young researchers to present proposals out of the "Normal Science". The peer-reviewed assessment was double blind and oriented not to consider curricula, but moreover original and inspiring proposals. This call was built to target young careers whose scientific curiosity would probably not be welcome in the regular sources of funding.

The first call was quite competitive. It selected 65 out of 1,955 proposals received from 331 different institutions from 26 out of 27 Brazilian states. The candidates had to fulfill the following requirements: must have received a PhD after January 1st, 2007 (with a little adjustment for women with children) and hold a position at a Brazilian research organization. The selection process involved 12 members of the Scientific Committee and more than 300 reviewers, from Brazil and abroad.

The first selection phase was double-blinded, and the reviewers had to score proposals in terms of their originality and scientific challenge. Proposals based on incremental research were considered non-competitive. This first phase narrowed the process to 200 proposals. The final list had predominantly research projects in life sciences (60%), followed by engineering (13%), chemicals (9%), earth sciences (6%), computer science (5%), physical (5%), and mathematics (2%).

This model was inspired by the experience of international public and private institutions. While publically launched in March 2017 ISP started its organization in 2014. In its first three years the executive team analyzed different institutional models and visited organizations that had started innovative models for funding science. ISP also interviewed researchers and

scientific managers to understand how a private institution might best meet the needs of the scientific community.

Just after launching its first call for proposals, ISP initiated a partnership with an external organization to build-up a methodology able to measure the impacts of the overall model, putting emphasis on the first call. As can be inferred, an approach for impact evaluation must be coherent and respond to the institutional model proposed by Serrapilheira, both in the operational, tactical and strategic aspects involved. Traditional and widely accepted indicators of scientific and technological impact are supposed to be used, even for dialogue with communities and scientific institutions and to get means of comparison. Nevertheless, the main challenge is to identify indicators that allow the measurement of variables that directly or indirectly approach the particularities above mentioned.

A new evaluation framework?

The definition of the impact evaluation model of the first call of Serrapilheira was based on both the institutional model proposed by the organization – and its main purposes – as well as on the particular goals of the first call. As a first step, these objectives were discussed with Serrapilheira's staff. The second step consisted of detecting evaluation hypotheses, themes and indicators based on the discussion of the model's objectives.

Three main evaluation hypotheses were proposed:

(i) Serrapilheira's funding model contributes to the creation of new scientific and technological trajectories, and to identify young promising researchers.

(ii) Serrapilheira's funding model modifies scientific and technological production of grantees.

(iii) Serrapilheira's funding model alters the professional trajectory of grantees.

Having these three main hypotheses in mind, the process of identifying evaluation themes and indicators followed the Decomposition Method (Salles-Filho et al. 2010, 2011). This method was developed to support the definition of themes and indicators in policy or program evaluation studies. It proposes an in-depth examination of the objectives of the policy or program (in this case, the model) as well as of the information available about it. This analysis seeks to identify the main terms that appear in the policy or program objectives and, by a deductive as well as inductive effort, to further identify themes and indicators more suitable to cope with its complexity.

Additionally, available evaluations of similar institutions around the world were analyzed, in order to check the consistency of use of certain indicators, as well as new evaluation approaches employed by them.

As a result, six evaluation themes were defined – one oriented to characterizing the grantees both in terms of diversity (age, gender, ethnicity, nationality, socio-economic background) and academic training (under graduation and graduation levels) – and the other five oriented to measure impacts.

The first impact theme is oriented to measure changes in the **professional trajectory** of grantees, with emphasis in career mobility and wages, including fundraising for research projects and peer review roles in journals and research funding agencies. These can be considered typical evaluation indicators commonly used and discussed in the increasing literature about the careers of doctorate holders, as can be seen in the recent publication of Gokhberg, Shmatko and Auriol (2016) that draws evidences from labor market of PhDs in different countries.

In addition to these more common indicators, the theme also addresses the contribution of grantees to academic entrepreneurial initiatives, spin-off creation, knowledge commercialization, technology transfer and consulting activities, categories explored in Cantaragiu (2012).

Finally it addresses the contribution of grantees to human resources training, through supervision in different levels (research carried during under graduation courses, master, PhD and post doc). Beyond measuring the number of advisee, we introduced academic genealogy indicators as a way to study the intellectual ancestry and heritage of grantees (Rossi, Freire & Mena-Chalco, 2017).

The second theme is concerned with **scientific and technology production** and covers conventional scientific impact indicators oriented to measure the quantity and quality of published papers, such as publication and citation counts, journal impact factors and H index. (Thompson et al., 2009; Hicks & Melkers, 2013). In addition it includes intellectual property rights and innovation indicators. Nevertheless, these are considered only complementary indicators, since Serrapilheira's model is fundamentally oriented to basic and applied sciences that not necessarily will generate new products, services and processes.

The third theme – named **night science** – is central to the evaluation model. As proposed by Jacob (1988):

> "Science has in fact two aspects. Day science involves reasoning as articulated as gears, results that have the strength of certainty. Aware of its style, proud of its past, sure of its future, the science of days advances in the light. Night Science, on the contrary, wanders in the dark. It hesitates, stumbles, falls. Questioning everything, it is searching itself endlessly, combining, associating myriads of hypotheses, assumptions still in the form of vague hunches, projects barely taken shape. Nothing guarantees its successes, its ability to survive the tests of logic and experiments, but sometimes thanks to intuition, instinct and the will to discover, as a lightning it illuminates more than a thousand suns…" (Jacob, 1988).

The measurement of how far the science being produced is novel and risky is not a trivial task, precisely because there is a typical time lag for new advancements to be recognized as "Normal Science", identified by Kuhn (1970) as the regular work of scientists theorizing, observing, and experimenting within a settled paradigm or explanatory framework. As a consequence, it cannot be measured by the conventional indicators of the previous theme.

This third evaluation theme is fundamentally based on the new phenomena of rapid diffusion that different social media as well as the preprints provide of scientific outputs, which in some

way overcome the limits of the mentioned time lag for recognizing novel science. The use of altmetrics (Priem et al., 2010) as a strategy to measure the occurrence of these phenomena and identify impacts in a "diverse scholarly ecosystem" is a choice within the night science theme.

Beyond altmetrics, the theme should involve a kind of state-of-art indicator, based on peer-review evaluation of intermediate and final research projects results. Additionally, it will cover some non-typical indicators such as the creation and consolidation of new research areas, new under graduation and graduation courses and new research groups. As proposed by Wilsdon et al. (2015), "Metrics should support, not supplant, expert judgement" and for such a non-conventional issue as night science, the balance may be a right option.

The fourth theme deals with **insertion and prominence** of grantees. Insertion is understood by the collaboration of grantees with national and international partners through publications, research projects and research networks, since collaboration is frequently understood as a proxy of academic productivity and high impact research (Bozeman and Boardman, 2014). By its turn, prominence deals with international mobility, editorial roles on international scientific journals and leadership part in international scientific events. To some extent, this theme can also as part of more conventional research evaluation framework.

Finally, the last theme is oriented to **research culture** and covers indicators to measure scientific communication and dissemination skills and initiatives of grantees to a wider public. This comes in order to increase visibility and accessibility of scientific findings. It also includes citizen science indicators, embracing its two complementary meanings as defined by Cavalier and Kennedy (2016): (i) how much the research is responsive to citizens' concerns and needs; and ii) how citizens themselves are involved in producing reliable scientific knowledge.

The five evaluation themes results in a quite diverse range of conventional and non-conventional indicators that can be considered adequate to lead with the new research funding model proposed by Serrapilheira. Figure 1 shows the correlation between evaluation hypotheses and evaluation themes, designating how much each evaluation theme is relevant to test the defined evaluation hypothesis.

Figura 1: Evaluation Hypotheses and Themes

Hyphotheses	Professional trajectory	Scientific and technology production	Night science	Insertion and prominence	Research culture
Serrapilheira's funding model contributes to the creation of new scientific and technological trajectories, and to identify young promising researchers	✔	✔✔	✔✔✔	✔✔✔	✔✔
Serrapilheira's funding model modifies scientific and technological production of grantees	✔✔	✔✔✔	✔✔✔	✔✔✔	✔
Serrapilheira's funding model modifies the professional trajectory of grantees	✔✔✔	✔	✔✔	✔✔✔	✔✔

✔: low relevance

✔✔: medium relevance
✔✔✔: high relevance

Beyond the definition of hypotheses, themes and indicators, other definitions about the evaluation design were simultaneously done. Concerning data collection, five moments along five years timeline were outlined:

- M1: baseline (beginning of the execution of the 65 research projects)
- M2: end of the first cycle of the first call (one year after the beginning of the execution of the 65 research projects), in order to measure outputs and immediate impacts
- M3: one year after the end of the first cycle, in order to measure impacts (65 projects)
- M4: end of the second cycle of the first call (end of execution of the 10-12 research projects that received additional funding), in order to measure outputs and immediate impacts
- M5: one year after the end of the second cycle, in order to measure impacts (10-12 projects)

Information is being collected through secondary and primary data. Secondary data includes Serrapilheira's database with information about approved and denied proposals, Lattes Curricula[2] and other database of peer-reviewed literature (such as Web of Science, Scopus and Scielo). Primary data is collected through web-based questionnaires. Data collection also involves judgment of experts about outputs and outcomes.

Lastly, the evaluation model is also based on a quasi-experimental design, having grantees as the treatment group and researchers whose proposals were denied as the comparison group. Some additional possibilities to refine the quasi-experimental design are being developed, such as using the best-rejected researchers (as proposed in Bohmer and von Ins, 2009 and Bornmann et al., 2010) and a sample of researchers whose proposals were submitted to ISP and in another similar call from a traditional funding agency. The quasi-experimental design will enable comparative analysis among grantees and the control group in order to identify their propensity towards risky and high-return researches as well as towards successful scientific trajectories.

Conclusion

As a first remark, we understand that the particularities of the new research-funding model proposed by ISP asks for a combination of traditional and new evaluation frameworks, oriented to a broad perspective about outputs, outcomes and impacts in a more diverse and multifaceted research system.

In this current hands-on impact-study a methodology is being built and simultaneously applied to a concrete case. This experience may contribute to the field of impact evaluation to the extent it proposes not only alternative indicators to be mixed with traditional ones, but also because it has a longitudinal perspective of following grantees since the inception (with the advantage of starting with a baseline) till two years after projects end, performing a 5-year long evaluation. Another possible contribution refers to the understanding about the ways funding agencies are evaluating the outcomes and impacts of their initiatives.

---

[2] The Lattes platform is a government-maintained database of résumés and other information on researchers throughout Brazil. It has a web interface, which is used for almost all research grant applications nationwide.

Future studies about the implementation of the proposed evaluation framework are needed within the next 5 years in order to check the acceptance of the alternative indicators and the effective use of evaluation for institutional planning and learning.

References

Bohmer, S., & von Ins, M. (2009). Different - Not just by label: Research-oriented academic careers in Germany. *Research Evaluation*, 18(3), 177–184.

Bornmann, L., Leydesdorff, L., & Besselaar, P.V. (2010). A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. Journal of Informetrics, 4, 211–220.

Bozeman, B. & Boardman, C. (2014). *Research Collaboration and Team Science: A State-of-the-Art Review and Agenda*. Springer.

Cantaragiu, R. (2012). Towards a Conceptual Delimitation of Academic Entrepreneurship. *Management & Marketing*, *7*(4), 683–700.

Cavalier, D., Kennedy, E. B. (2016). *The Rightful Place of Science: Citizen Science*. Tempe, AZ: Consortium for Science, Policy & Outcomes.

Ding, F.; Lui, Y.; Chen, X. & Chen, F. (2018). Rising Star Evaluation in Heterogeneous Social Network. *IEEE Access*, 6, 29436- 29443.

Escobar, H. (2017). Brazil's "doomsday" scenario. *Science* (New York, N.Y.), 355(6323), 334. http://doi.org/10.1126/science.355.6323.334

Gewin, V. (2012). Risky research: The sky's the limit. *Nature*, 487, 395-397.

Gogoglou, A. & Manolopoulos, Y. (2017). Predicting the Evolution of Scientific Output. In: N. T. Nguyen; G. A. Papadopoulos; P. Jedrzejowicz; B. Trawinski; G. Vossen (Eds.). *Computational Collective Intelligence* (pp. 244–254). Switzerland: Springer.

Gokhberg, L., Shmatko, N. & Auriol, L. (2016). *The Science and Technology Labor Force: The Value of Doctorate Holders and Development of Professional Careers*. Switzerland: Springer.

Hicks, D. & Melkers, J. (2013). Bibliometrics as a tool for research evaluation. In A. N. Link & N. S. Vonortas (Eds.), *Handbook on the theory and practice of program evaluation* (pp. 323–349). Cheltenham: Edward Elgar.

Jacob, F. (1988). *The Statue Within: An Autobiography*, translated by Franklin Philip. New York: Basic Books.

Kuhn, T. (1970). The Structure of Scientific Revolutions. Chicago: The University of Chicago Press.

Priem, J., Taraborelli, D., Groth, P.; Neylon, C. (2010), *Altmetrics: A manifesto*. Available at: http://altmetrics.org/manifesto.

Rossi, L., Freire, I.L. & Mena-Chanco, J.P. (2017). Genealogical index: A metric to analyze advisor–advisee relationships. *Journal of Informetrics*, 11(2), 564-582.

Salles-Filho, S.M., Avila, A.F., Alonso, J.E.O.S. & Colugnati, F.A.B. (2010). Multidimensional assessment of technology and innovation programs: The impact evaluation of INCAGRO-Peru. *Research Evaluation*, 19, 361–72.

Salles-Filho, S.M., Bonacelli, M.B., Carneiro, A.M., Castro, P. F. D. & Santos, F.O. (2011). Evaluation of ST&I programs: A methodological approach to the Brazilian small business program and some comparisons with the SBIR program, *Research Evaluation*, 20, 157–69.

Thompson, D.F., Callen, E.C. & Nahata, M.C. (2009). New indices in scholarship assessment. *American Journal of Pharmaceutical Education*, 73(6), 1–5.