CHAPTER 12

Exploring research evaluation from a sustainable development perspective

Diego Chavarro

Introduction

Evaluation studies produce 'value judgements about the quality, worth, or value of intervention programmes' (Mouton 2014: 64). These value judgements usually rely on standards, theories and ideals as points of reference against which the assessment is performed. They are lenses through which an intervention is analysed, and therefore determine what is valued and how. For instance, an evaluator using a conventional economics perspective will likely value policy interventions for their contribution to solving market failures (Dollery & Worthington 1996), whereas a constructivist will likely value more the process of inclusion and debate in policy co-construction (Guba & Lincoln 1989). The perspective chosen has implications for the type of analysis performed, the unit of analysis chosen, the assessment criteria used and the methodology for the evaluation. Therefore, discussing the theories underpinning evaluations helps understand their rationales, usefulness and limitations. This is relevant for policy-making because, when some evaluation frameworks become dominant or popular, they are used indiscriminately for purposes that do not match their aims.

Research evaluations are based on theories about the value of scientific knowledge production (Molas-Gallart and Ràfols 2018). Countries such as Australia, UK, Brazil and Colombia have developed national Research Evaluation Systems - RES (Chavarro 2017), which can be understood as 'organised sets of procedures for assessing the merits of research undertaken in publicly funded organisations that are implemented on a regular basis, usually by state or state-delegated agencies' (Whitley & Gläser 2007: 6). RES can have different theoretical underpinnings, but one that is common to many of them is the sociology of science perspective, which values research for its 'scientific impact' and 'scientific quality' (Chavarro et al. 2018). Frequently, RES use quantitative indicators such as citation and bibliographic output counts to measure such concepts. These will be referred to as production and citation indicators. The popularity of these indicators has given scientometrics, or the quantitative study of science (Wouters 1999), a preponderant role in RES.

However, the evaluation of scientific knowledge production based only on 'intrinsic', 'scientific' criteria is being challenged by alternative perspectives that value knowledge production for 'extra-scientific' criteria, such as its impact on society, institutions and environment (Orozco et al. 2007). Sustainable development, which can be understood as a balance between economic, environmental and social development (Gallopín 2001), is one of those extra-scientific criteria by which research can be evaluated. This perspective challenges RES based on the sociology of science because sustainability demands the social accountability of knowledge construction.

In this chapter I focus on the examination of some of the theoretical underpinnings of conventional RES, focusing on RES that use production and citation indicators to produce rankings, discussing their limitations in capturing the features of knowledge construction in a context of sustainable development. I also discuss, from a policy-making perspective, some of the reasons why transforming national RES faces resistance, suggesting alternatives that can be explored by research councils and other policy organisations wanting to develop evaluations in the context of sustainability.

Some foundations of scientometric indicators for research evaluation

Scientometrics is a discipline devoted to the quantitative study of science, technology and innovation (ST&I). One of its main areas of research is the development of indicators for research evaluation. Scientometric research evaluation focuses mainly on the assessment of two properties[1] of scientific knowledge: production and quality (Molas-Gallart and Ràfols 2018). Usually, these properties are expressed as some measure of quantity of scientific outputs (articles, books, patents) and citations to scientific literature respectively.

For Merton, the function of science is the production of certified knowledge (Merton 1973)[2]. This knowledge is communicated in journals that are recognised by scientific communities as the gatekeepers of quality research, overseen by a system of peer reviewers who evaluate the soundness of scientific contributions based on disciplinary standards. For Merton, the production of certified knowledge is considered a social value in itself, so considerations about the social utility of that knowledge are extra-scientific and not necessary in order to justify funding decisions.

Derek de Solla Price and Eugene Garfield were fundamental in diffusing Merton's sociological description of scientific norms and Mertonian ideas of scientific production. Price set the foundations of scientometrics and Garfield operationalised these ideas and analyses through the development of the Science and Social Sciences Citation Indices and his invention of the citation indicator (Wouters 1999; Godin 2006a). Although initially the intention of Garfield was not to produce indicators for rankings, rankings of researchers, journals, organisations and countries became the main use of production and citation indicators (Chavarro 2017). This use fits the Mertonian conception of science because he depicted the scientific system as a hierarchical structure. In this hierarchical structure, some scientists receive more recognition than others due to their experience and the significance of their contributions to science, which brings them reputation and scientific authority. In the end, RES based on production and citation indicators reproduce this emphasis on reputation, because many of them aim to distribute funds and public recognition to individuals and organisations in a competitive manner.

The above understanding reproduced by production and citation indicators embedded in RES[3] fits into a certain framing of research policy, referred to as frame 1 ST&I policy (Schot and Steinmueller 2016) or linear model (Godin 2006b). This frame is portrayed as a sequential process in which research activities are an input for technology development, and technology development is the engine for economic growth. In this conception, the role of the state is to fund researchers and innovators, leaving the decision on what to research to them because there is an implicit assumption that all science and innovation brings positive outcomes in terms of economic growth (Schot and Steinmueller 2016). In this sense, reputational indicators such as the ones discussed earlier fit a policy framework that conceives science as a means for economic growth. The outcomes of many RES, especially rankings of scientists and organisations, are then the recognition by the state to those who contribute the most to economic growth.

To summarise, production and citation indicators are frequently used in a way that considers the production of certified knowledge an aim in itself. This conception of science rewards individuals and organisations who excel at publishing their research, creating a system in which scientists compete for public recognition and reputation. Questions about the utility of science are not covered by this conception because it is assumed that all knowledge is beneficial to society and scientific development produces positive (economic) outcomes. In the next section, I will present the limitations of this conception to address new societal demands essential to sustainable development.

Sustainable development challenges traditional scientometric indicators

The concept of sustainable development does not have a single definition. However, the historians of the concept place its origins in the environmental movement and in environmental economics (Meadows et al. 1972). Although there are different definitions, it can be concluded that sustainable development seeks social, environmental and economic balance (Chavarro et al. 2017). This balance, however, is a 'dynamic equilibrium' (Gallopín 2001) because the three systems are in constant renovation.

In current public policy, sustainable development has become a framework of wide acceptance, mainly due to the thrust of multilateral organisations such as the United Nations (UN) and the Organisation for Economic Co-operation and Development (OECD). These organisations have managed to convene several countries over time and around joint programmes of cooperation. One of these programmes is the 2030 Sustainable Development Agenda, which seeks to achieve 17 social, economic and environmental goals known as the Sustainable Development Goals (SDGs) (see UN 2015).

Within this sustainable development agenda, science is not considered only a source of new certified knowledge or the engine of economic growth, but also a contributor to solving social and environmental issues. The Scientific Advisory Board (SAB) of the UN Secretary General has acknowledged the 'crucial role of science for sustainable development' (SAB 2014), and published a report discussing some of the ways in which science is related to the achievement of SDGs. According to this Board, science has the following roles:[4]

- Provides the basis to identify and face global challenges;
- Offers a mechanism to cross the 'national, cultural and mental barriers', which is necessary to work collaboratively in the challenges of sustainable development;
- Through scientific literacy, provides education and helps build the capacity to use science to solve everyday problems;
- Can strengthen democratic practices if it is treated as a public good;
- Thanks to its ability to integrate knowledge from different disciplines, it helps to face challenges that are interdependent (for example poverty, economic growth, clean water and clean energy);
- Provides evidence to formulate ST&I public policy and to inter-relate other public policies;

- Helps to monitor progress in the different sustainable development objectives; and
- Education with a strong scientific component prepares societies to respond creatively to the challenges presented to them.

From the list above it can be derived that the role of science in sustainable development differs from its role according to the sociology of science. The main difference is that in sustainable development the production of new knowledge is not valued on its own, but on its relevance to the environmental, social and economic challenges of the world.

With regard to policy framing, sustainable development is not adequately represented by the linear model or even by the national systems of innovation model. For this reason, scholars are developing specific policy frameworks to conceptualise the role of science in sustainable development. One of these frameworks is the Transformative Innovation Policy (TIP), which proposes that to transition towards sustainable development there is a need to exert profound changes in the systems of consumption and provision of goods, as well as in culture and economy (socio-technical systems) (Schot and Steinmueller 2016). TIP understands ST&I as cross-cutting to all SDGs, and as a means to achieve socio-technical changes towards sustainability. Based on the above points, I present a list of four principles, understanding principles as properties that research evaluation should consider when appraising science in a context of sustainable development.[5] I also summarise some of the key challenges posed by sustainable development to traditional research evaluation:

- *Transformation*: Sustainable development requires transformation of socio-technical systems, not only the production of new knowledge. Therefore, sustainable development demands involvement of scientists in solving environmental, economic and social challenges. Increasingly, research councils are requiring scientists to show the social impact of their research; see for instance, the UK's Research Excellence Framework or the NSF's broader impact criteria for research proposals, which show a movement towards this direction. However, social impact is only a part of sustainable development, which requires 'transformative impact'. This implies the development of evaluation frameworks to account also for the environmental and economic impacts of research, as well as their inter-relationships.

- *Collaboration*: Sustainable development is about collaboration, not competition as promoted by RES focused on ranking systems. In the context of sustainable development, collaboration includes international and intra-national collaboration, interdisciplinarity and inter-sectorial cooperation.
- *Directionality*: Sustainable development requires directionality, in the sense that not all scientific production or innovation will bring positive sustainable effects. For instance, from a sustainability perspective, research on chemical weapons with the aim to produce them on a large scale is not desirable, even if it is highly cited.
- *Participation*: Sustainable development requires the participation of government, citizens and entrepreneurs, in addition to scientists (which is not favoured by the sociology of science or linear model approaches).

The list is based on my interpretation of the role of science from a sustainable development point of view, in comparison to the sociology of science or linear model conceptions. It is presented, then, as a list for discussion. However, even in this preliminary elaboration, it can be seen that traditional ST&I indicators do not account for the properties of sustainability. Despite their unsuitability, similar indicators are being used to measure progress on sustainable development. For instance, the UN elaborated a set of indicators to measure progress on goals related to sustainable innovation; the ones proposed to measure progress on research and innovation capacity are research and development (R&D) expenditure and number of researchers per inhabitants.[6] These indicators are based on a linear model perspective, which assumes that more is better, but does not address issues such as on what subjects the researchers are working, what research is being funded, who are being funded, or how interdisciplinary and transformative the research is. Therefore, the use of these indicators as proxies for sustainability is a case of 'streetlight effect' or 'drunkards search', meaning that their use can only be justified because they are the indicators at hand (Molas-Gallart and Ràfols 2018). In this way, cases such as the UN's choice of traditional ST&I indicators as proxies for sustainable innovation fail to capture the sustainability concept they intend to measure. A question arises as to what can be done, then, to include sustainability concerns in research evaluation. In the next section, I suggest some ideas to consider in the design of indicators to address this issue.

Ideas for research evaluation systems (RES) in the context of sustainable development

RES are an important component of ST&I policy, because they help to steer research in desired ways through recognition and funding (Whitley & Gläser 2007). Although it is not clear how effective RES are in actually shaping research agendas and setting priorities towards desired goals (Rijcke et al. 2016), they make visible what is valued by research councils and funding agencies. For this reason, if sustainable development is to be supported by a country, its RES needs to incorporate clear criteria in the direction of sustainability. Such criteria, however, are absent from many RES that continue to apply an evaluation model based on production and citation indicators. Here I provide some ideas on why this continues to happen and suggest points to be considered when designing research evaluations in the context of sustainable development, attending to the aforementioned principles. For this I use the case of Colombia, a country with an RES which is heavily based on production and citation indicators. I start by describing the Colombian RES of research groups, attempt to understand why the foundations of this RES have resisted structural modifications, and then suggest some of the points that could be included in it if the country is to promote science for sustainable development.

In Colombia, as well as in other countries such as Mexico, Brazil, Chile and Spain, RES based on production and citation indicators affect directly or indirectly the distribution of funds and recognition to researchers and organisations (Chavarro 2017). Colciencias manages two large research assessments, one for research groups (GrupLAC) and the other for journals (Publindex). Recently, Colciencias has also started to evaluate individual researchers. Groups are evaluated according to a quantitative index composed mainly of a weighting scheme applied to their bibliographic output. This RES underwent several changes from 2000 to 2015 (Nupia 2018). The first change was to introduce the calculation of a score to measure different types of bibliographic outputs according to quality criteria, mainly citations to research papers. Another change was the introduction of other outputs of importance to disciplines such as arts and architecture, for instance, evidence of concerts, performances, paintings, novels and blueprints. Other modifications include the criteria to endorse research groups, or weights given to different types of outputs, and recently the measurement model has included scores for non-bibliographic results such as spin-offs.

However, most of the changes introduced to the Colombian RES have not questioned profoundly the principles of production and quality under which it is built, and the indicators used. I queried experienced colleagues who have worked on the development of Colciencias' research evaluation model as to why they think its principles seem to be accepted, complementing their answers with my own experience and with literature search. My findings are summarised in the following list:

- *Stability*: since 2000 the criteria for assessing research have remained relatively stable, as well as the way to measure these properties (a quantitative model that produces a score);
- *Routine*: once established, the procedure of measuring research groups was codified in software. The software is run each year and this has become part of the organisations' routine with an allocated yearly budget;
- *Predictability*: experience has taught Colciencias how to deal with software errors and complaints from researchers and institutions; the software even offers simulations to predict the ranking of the research group before the actual evaluation is performed;
- *Co-construction*: the criteria have been debated between Colciencias and representatives of the academic community (Nupia 2018);

- *Flexibility*: even though the model is based on a formula, the discussions with the academic communities have allowed the incorporation of new products into the measurement;
- *Link with distribution of funds*: in public universities, scores obtained from scientific production indicators represent a salary increase for teachers. Also, the classification of a research group may determine its eligibility for funding;
- *Isomorphism*: the research evaluation model, which directly affects public universities, is being reproduced in private universities, some of which give economic incentives to researchers for their scientific production indicators; and
- *The model works*: citations and production counts are readily available and have become a standard. In comparison, other indicators are less developed in terms of their reliability and their interpretation is even less clear (e.g. altmetrics).

The current RES may have been able to allow for minor changes to the calculation of rankings, but the work that Colciencias is doing on the design of an ST&I policy for sustainable development (Colciencias 2018) has shown that to incentivise the contribution of ST&I to solving grand challenges, the current research evaluation model needs to be renewed. In order to include the principles of sustainable development in evaluation for sustainability, Colciencias' RES could:

- Incorporate ways of appraising scientific collaboration and the participation of diverse social groups in research activities (citizens, entrepreneurs, NGOs, etc.) – participation and collaboration principles;
- Increase efforts to use content analysis in evaluations of research outputs instead of citation counting only, because sustainable development research is directed towards specific goals. Assessing research related to the SDGs, for instance, requires identification of the specific subjects being researched. Semantic analysis can give relevant evidence to policy-makers to steer research funding and promotion – directionality principle;

- Prioritise the communication function of science over the use of research evaluation indicators for career development and individual or organisational reputation. This may imply unlinking research evaluation from direct funding of researchers and institutions, and instead allocating funds to subjects or topics of national or local interest, as well as problem-oriented research – transformation and directionality principles;
- Incorporate ways of appraising environmental and social impact in addition to scientific and economic impact – transformation principle;
- Reward novelty and relevance of contributions, not only accumulation (e.g. of citations) – transformation principle;
- Incentivise science as a public good, as opposed to science as a private endeavor. This may imply giving more weight to research that can be openly distributed, but also to research that uses or builds open infrastructures, etc. – collaboration and participation principles.
- Incentivise interdisciplinarity, because it is needed to address local issues – transformation principle; and
- Include non-traditional research outputs, such as technical manuals or other products that are difficult to codify in standard bibliographic outputs and can have transformative impacts – transformation principle.

However, based on conversations with colleagues in Colciencias and on my own experience in research evaluation, I find that there are different barriers to implementing radical changes. For instance, changing an established RES requires huge investments both in terms of funds and time, thorough discussion with academics and other stakeholders if the principle of participation is to be put into practice, and restructuring areas of Colciencias that are devoted to managing the current RES. Connected with this, a new RES requires creating internal capacity. This capacity refers not only to technical skills, but also to the suitability of the legal framework to accommodate a new research evaluation model, which could have an impact on the salaries of researchers and the distribution of funds. Also, some may see modifications to the current RES as a threat to 'scientific quality', given that, in the case of sustainability, the 'relevance' criterion is just as important. Some may even argue that the basic sciences would be disadvantaged by a 'utilitarian' understanding of knowledge production. Therefore, from a policy-maker's perspective, it is not an easy decision to radically change the current RES. However, some alternatives could be explored.

One alternative is to adjust the current quantitative model to include and give weight to some of the principles of sustainable development. This is the approach that has been followed in the past 15 years to introduce changes. Another alternative is to have two separate measurement models, one to award 'quality' and the other to award 'relevance'. This would require two research evaluation systems, which would be very costly and operationally demanding. A third one could be to have a multidimensional model, with one of the dimensions being 'quality' and the other being 'relevance'. The question here is how to weight the dimensions: which is the more important, quality or relevance? In summary, any option implies trade-offs between different valuations of science.

A recent attempt to bridge the 'quality' and 'relevance' gap, addressing the issues above, is the Research Quality Plus (RQ+) framework designed for the evaluation of research for development (Ofir et al. 2016). The idea behind RQ+ is that research quality is a multidimensional concept, which goes beyond scientific merit. The framework is highly customisable, and offers ways to include key influences that constrain research, the different dimensions of quality beyond citations (integrity, legitimacy, importance and positioning for use) and rubrics for assessing each component. By employing this framework, researchers at IDRC have found that research for development produced in the Global South outperforms research for development produced in the Global North, contradicting most of the studies of scientific production based on production and citation indicators. This shows that evaluations of research are dependent on how 'quality' and 'relevance' are defined.

Including sustainability principles such as transformation, directionality, collaboration and participation in the quality dimension of RQ+ and developing rubrics for assessing them is something that is allowed by the framework. How the inclusion of these principles would change quality assessments of research in the Global South in different disciplines is a question worth exploring. Although promising for evaluations of research in a context of sustainable development, novel frameworks for research evaluation have yet to be tested in countries such as Colombia that rely heavily on production and citation bibliographic indicators. There is, then, a great opportunity to conduct pilots to learn how concepts such as sustainable development can be included in research evaluation and the acceptance of, or resistance to, novel ways of research evaluation by research communities and other social groups.

Conclusion

The purpose of this chapter is to contribute to a better understanding and use of scientometric indicators and to help develop principles to guide the design of new indicators and research evaluation in the context of sustainable development policy. By examining the theories that underpin indicators' development and use, as suggested by Mollas-Gallart and Ràfols (2018), it was possible to see why some research evaluations fail to convey the properties that they try to measure. This was done by exploring some of the assumptions underpinning RES, based on scientometric indicators and comparing these assumptions to those of sustainable development. My argument is that the conventional scientometric indicators used by RES cannot evaluate research in the context of sustainable development, mainly because they are based on a theory of science that regards the production of certified knowledge as a social value in itself, whereas sustainable development values it in relation to its relevance for social, environmental and economic issues. For this reason, if policy-makers want to develop research evaluations to support sustainability, there is a need to understand this radical difference and design alternative indicators and evaluation frameworks that reflect sustainability more accurately. I have also suggested some changes that could help to produce more sensible research evaluations that meet their stated objectives. Basically, RES that want to better represent the concept of sustainability could include criteria to address transformation, collaboration, directionality and participation criteria, which are absent from conventional scientometric evaluation.

Despite the need for alternatives, transforming a RES is challenging for a research policy organisation. By using the case of Colombia as an example, this chapter provides a concrete account of how discussions on research evaluation materialise in the decisions that a research policy organisation must make and why changes, which appear to be relatively 'simple' from an academic point of view, are complex in practice: the resources devoted to the development of an RES, the time required to establishing it, the routines developed around its implementation, the regulations and funding linked to it, the human capacity needed in order to operate it, and potential criticisms are constraints that policy-makers face when taking the decision to embark on a new RES.

Despite the above constraints, evaluation frameworks such as RQ+ offer a way to test alternative understandings of research quality and incorporate new criteria, such as sustainability, in a way that bridges the gap between 'quality' and 'relevance'. Conducting evaluation pilots in countries such as Colombia will contribute to establishing the usefulness and limitations of these novel frameworks, and their complementarity to conventional research evaluation. Although it can be seen as a costly exercise, the benefits of experimenting and learning will outperform its cost, which is applying a conventional evaluation instrument that does not fit the new societal demands from science.

Acknowledgements

This chapter is based on my experience as an advisor to the Policy Evaluation and Design Unit of Colciencias, which is the organisation in charge of ST&I policy in Colombia. I have drawn also on conversations with colleagues and discussions in different workshops on the development of an ST&I policy for sustainable development. I am especially indebted to my colleagues who participated in the SGCI/IDRC workshop on 'Perspectives on Research Excellence'. My special thanks go to Johann Mouton, Erika Kraemer-Mbula and Matthew L. Wallace for their insightful comments.

Notes

1 I use the word ‘property’ in the sense pointed out by Molas-Gallart and Ràfols (2018), who state that indicators are ways to approximate the measurement of properties that are not directly observable. Quality, for instance, is a vague term to indicate a property that makes a scientific outcome more valuable than another. As quality cannot be observed directly, an indicator such as the number of citations can give an idea of the level of quality of that product, of course assuming that researchers cite others based on quality considerations. In this case, the property is quality and the indicator the number of citations.

2 Further references to the work of Robert Merton, especially his views on the notion of ‘excellence’, are found in Chapter 4 ‘Re-valuing research excellence: From excellentism to responsible assessment’.

3 Production and citation indicators can be used in a variety of ways that differ from the one pointed out here. I specifically refer to the use of these indicators for ranking purposes, a practice that has become popular in different research evaluation systems (Chavarro, Tang and Ràfols 2017).

4 This list is reproduced from Chavarro et al. 2017, and its sources are the reports SAB (2014) and SAB (2016).

5 These principles are not exhaustive and are given here only as an example to show some of the properties that are not addressed by dominant scientometrics evaluation.

6 https://sustainabledevelopment.un.org/sdg9

References

Chavarro D (2017) Universalism and particularism: Explaining the emergence and growth of regional journal indexing systems. Doctoral thesis, University of Sussex. http://sro.sussex.ac.uk/66409/

Chavarro D, Tang P and Ràfols I (2017) Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling. *Research Policy* 46(9): 1666–1680

Chavarro D, Vélez MI, Tovar G, Montenegro I, Hernández A and Olaya A (2017) *Los Objetivos de Desarrollo Sostenible en Colombia y el aporte de la ciencia, la tecnología y la innovación*. Documento de trabajo 01, Colciencias – UDEP http://www.colciencias.gov.co/sites/default/files/objetivos_de_desarrollo_sostenible_y_aporte_a_la_cti_v_3.5.pdf

Chavarro D, Ràfols I and Tang P (2018) To what extent is inclusion in the Web of Science an indicator of journal ‘quality’? *Research Evaluation* 27(2): 106–118

Colciencias (2018) *Libro verde 2030. Política de ciencia e innovación para el desarrollo sostenible*. Bogotá: Colciencias. http://libroverde2030.gov.co/wp-content/uploads/2018/05/LibroVerde2030.pdf

Dollery BE and Worthington AC (1996) The evaluation of public policy: Normative economic theories of government failure. *Journal of Interdisciplinary Economics* 7(1): 27–39

Gallopín G (2001) *Science and Technology, Sustainability and Sustainable Development*. CEPAL. Retrieved from http://www.istas.ccoo.es/escorial04/material/dc12.pdf

Godin B (2006a) On the origins of bibliometrics. *Scientometrics* 68(1): 109–133. http://doi.org/10.1007/s11192-006-0086-0

Godin B (2006b) The linear model of innovation: The historical construction of an analytical framework. *Science, Technology, & Human Values* 31(6): 639–667. ftp://ftp.ige.unicamp.br/pub/CT010/aula%202/GODIN-The%20Linear%20Model%20of%20Innovation-STHV2006.pdf

Guba EG and Lincoln YS (1989) *Fourth Generation Evaluation*. Thousand Oaks, CA: Sage

Meadows DH, Meadows DL, Randers J and Behrens III WW (1972) *The Limits to Growth: A Report to the Club of Rome*. New York: Universe Books

Merton R (1973) The normative structure of science. In: N Storer (ed.) *The Sociology of Science*. Chigago and London: The University of Chicago Press. pp. 267–278

Molas-Gallart J and Ràfols I (2018) Why bibliometric indicators break down: Unstable parameters, incorrect models and irrelevant properties. *SSRN Electronic Journal.* January. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3174954

Mouton J (2014) Programme evaluation designs and methods. In: F Cloete, B Rabie and C de Conning (eds) *Evaluation Management in South Africa and Africa*. Stellenbosch: Sound Media. pp. 163–202

Nupia C (2018) La medición de la producción científica de los grupos de investigación en Colombia: del diálogo de expertos a la incorporación de prácticas más representativas. In: G Dutrénit and JM Natera (eds) *Procesos de diálogo para la formulación de políticas de CTI en América Latina y España*. Buenos Aires: Clacso https://www.researchgate.net/publication/325541711_La_medicion_de_la_produccion_cientifica_de_los_grupos_de_investigacion_en_Colombia_del_dialogo_de_expertos_a_la_incorporacion_de_practicas_mas_representativas

Ofir Z, Schwandt T, Duggan C and McLean R (2016) *Research Quality Plus (RQ+): A Holistic Approach to Evaluating Research*. Canada: IDRC. https://www.idrc.ca/sites/default/files/sp/Documents%20EN/Research-Quality-Plus-A-Holistic-Approach-to-Evaluating-Research.pdf

Orozco LA, Chavarro-Bohórquez DA, Olaya DL and Villaveces JL (2007) Methodology for measuring the socio-economic impacts of biotechnology: A case study of potatoes in Colombia. *Research Evaluation* 16(2): 107–122

Rijcke SD, Wouters PF, Rushforth AD, Franssen TP and Hammarfelt B (2016) Evaluation practices and effects of indicator use—a literature review. *Research Evaluation* 25(2): 161–169

SAB (2014) *The Crucial Role of Science for Sustainable Development and the Post-2015 Development Agenda. Preliminary Reflection and Comments*. UNESCO. http://en.unesco.org/un-sab/sites/unsab/files/Preliminary%20reflection%20by%20the%20UN%20SG%20SAB%20on%20the%20Crucial%20Role%20of%20Science%20for%20the%20Post2015%20Development%20Agenda%20-%20July%202014.pdf

SAB (2016) Science for sustainable development. Policy brief. UNESCO. http://unesdoc.unesco.org/images/0024/002461/246105E.pdf

Schot J and Steinmueller E (2016) *Framing Innovation Policy for Transformative Change: Innovation policy 3.0*. http://www.johanschot.com/wordpress/wp-content/uploads/2016/09/Framing-Innovation-Policy-for-Transformative-Change-Innovation-Policy-3.0-2016.pdf

UN (United Nations) (2015) *Transforming our World: The 2030 Agenda for Sustainable Development*. United Nations Organisation

Whitley R and Gläser J (eds) (2007) *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*. Dordrecht, the Netherlands: Springer

Wouters P (1999) The citation culture. Doctoral thesis, Universiteit van Amsterdam. https://pure.uva.nl/ws/files/3164315/8231_13.pdf