

RESEARCH

Ethical Theories in Research Evaluation: An Exploratory Approach

Maria Teresa Biagetti¹, Aldis Gedutis² and Lai Ma³

¹ Facoltà di Lettere e Filosofia, Sapienza Università di Roma, IT

² Centre for the Studies of Social Change, Klaipeda University, LT

³ School of Information and Communication Studies, University College Dublin, IE

Corresponding author: Maria Teresa Biagetti (mariateresa.biagetti@uniroma1.it)

Research evaluation encompasses the practices of assessing research quality and impact at various stages of research. The processes and criteria of research evaluation vary depending on the nature and objectives of the assessment. Different research evaluation systems influence the research strategies of universities and institutes. There are, however, some known issues of research evaluation with regards to the peer review and, most prominently, the use of citation-based metrics, which lead to recent calls for responsible use of metrics. In this paper, we argue that there is a need for ethical theories for considering research evaluation and that research evaluation ethics, as an overlapping area between research ethics and evaluation ethics, deserve its own treatment. The core of the article consists of a discussion of the most influential ethical theories in the context of the research evaluation, including the deontological ethics, the consequentialist ethics and the virtue ethics. The aim is to highlight the need to assume an ethical view that combines the deontological and the consequentialist concepts, adopting 'common good' as the most likely pillar for the research evaluation procedures. We propose that the mixed approach would be useful for developing a framework for research evaluation ethics and for analysing ethical approaches and ethical dilemmas in research evaluation.

Policy Highlights

The misuses and abuse of evaluative metrics have been discussed and debated in many high-profile publications including the San Francisco Declaration on Research Assessment (DORA), The Metrics Tide, the Leiden Manifesto, and the Hong Kong Principles. There are also many studies stating the limitations of and bias in peer review.

The debates and discussions, however, have not been explored in light of ethical theories. The article considers also good practices in evaluation, including the American Evaluation Association Guiding Principles for Evaluators (AEA 2018), the Australasian Evaluation Society Guidelines for the ethical conduct of evaluations (AES 2000, 2010, 2013) the UK Evaluation Society Guidelines for Good Practices in Evaluation (UK 2019) and the United Nations Ethical Guidelines for Evaluation (UNEG 2008).

The paper argues that ethical theories are useful in understanding ethical assumptions and ethical dilemmas in research evaluation and are pertinent in future design and development of research evaluation processes and criteria.

Ethical theories that can construct ethical principles for research evaluation, including deontological and consequentialist ethics, taking into account the Mertonian normative theory, have been examined.

In order to address the issues of research evaluation, we propose a mixed approach that combines the deontological and the consequentialist concepts that is able to infringe the boundaries of the rivaling theories and provide basis needed for research evaluation ethics.

Keywords: Research evaluation; Ethical theories; Ethics of Research Evaluation; Peer review; Evaluative metrics; Bibliometrics

1 Introduction

Research evaluation encompasses the practices of assessing research quality and impact of scholarly works *ex ante* and *ex post*. *Ex ante* research evaluation usually refers to the evaluation of research proposals for grant funding, where the quality, feasibility and potential contributions of funding proposals are assessed. *Ex post* research evaluation, on the other hand, is used to assess scientific-scholarly and sometimes economic and societal impacts, after a research project has been conducted. At the individual level, the assessments are often used in the decision-making process of hiring

and promotion of scholars and of their career advancement, as well as the evaluation of grant proposals and awards. At the university level, research evaluation is sometimes used for allocating block grants to universities. While research evaluation is considered necessary to assess the research performance of individuals and universities, there have not been ethical guidelines in the drafting of evaluation processes or criteria notwithstanding the constitutive effects of evaluation (Dahler-Larsen, 2012). A survey of national evaluation systems (Ochsner, Kulczykcki and Gedutis 2018) reveals that different systems have incompatible priorities for the research evaluation. As a result, their evaluation criteria vary, e.g. metric and non-metric require diverse approaches, including the ethical ones.

Whitley (2007) argues that universities and centers of research are often in competition for favorable assessment and that strong research evaluation systems can limit intellectual autonomy and the ability to implement research strategies that challenge current orthodoxies. Moreover, research evaluation may impact on the development of disciplines and limit novelty and inventiveness (Whitley, Gläser and Laudel 2018). In many scientific fields highly cohesive scientific elites may influence the organization of strong Research Evaluation Systems (RES) in agreement with their conceptions of quality. Where elites hold the consensus on central topics of their disciplines, strong RES may reinforce their authority, as they might decide the quality standard for the discipline. Research evaluation plays a fundamental role in both the development of disciplines and the career advancements of researchers. It is expected to impact on the development of scientific fields, as it may limit novelty and inventiveness of emerging researchers, which must conform to the dominant elites to achieve academic consensus.

There are also known problems and issues concerning research evaluation—both in the peer review process and the criteria used, including citation-based metrics. The assessment in Social Sciences and Humanities (SSH), for instance, may be based on emotions and on the interactions between individuals; the social identity of reviewers and their membership to a scientific-scholarly community, rather than the neutral judgment, can play a fundamental role (Lamont 2009). For one, bias in peer review has been discussed with respect to gender, race, language, career stage and interdisciplinarity (see, for example, Helmer, 2017; Lee, et al., 2013). Peer reviewers also tend to be conservative and risk-averse in their evaluation of innovation methods and approaches (Luukkonen, 2012), not to mention the inconsistent reliability and validity of peer review notwithstanding the availability of innovative procedures and platforms (Bornmann, 2011; Horbach & Halfman, 2019).

Furthermore, studies have shown that the use of citation-based metrics have led to the misuse and gaming of evaluative metrics (see, for example, Biagioli & Lippman 2020) as well as changes in research practices and knowledge production (de Rijcke et al., 2016). It is understood that the use of metrics induces competition, rather than collaboration, between researchers. The drive to publish in high JIF journal also prompts researchers and scholars to publish in international journals, leading to decreased number of publications in local/national languages that are important especially for the SSH. Some have argued that the use of metrics, which eventually has led to ‘misuses’ and ‘abuses’ of metrics, is due to the audit culture, in which accountability is at its core. Ma and Ladisch (2019) have suggested that evaluation complacency and evaluation inertia are a cause, as well as an effect, of the use of metrics in research evaluation. Recently, there are increasing pressures for institutions to reconsider and reconfigure the use of citation-based metrics in response to DORA (ASCB, 2013), The Metrics Tide (Wilsdon et al., 2015), The Leiden Manifesto (Hicks, et al., 2015) and the Hong Kong Principles (Moher, et al., 2020).

Taking into account the complexities of research evaluation, we must consider whether to look “from above” and seek universal ethics, or to calibrate our optics for an empirical case study. At the initial stage of our endeavour it is more viable to seek for a theoretical horizon than limit ourselves to empirical case studies. Merton (1973) has proposed scientific ethos often known by the name of CUDOS: communalism (originally communism), universalism, disinterestedness, and organized skepticism. While his conceptions are closely related to the goal of science and scientific method, there seems to be a lack of ethical justification. Therefore, CUDOS are open to criticism for being too general, not reflective enough and rather inefficient if compared with the particular practices of scientific research in their diversity. For this reason, a more detailed study of the ethical field is needed. In the rest of the paper, we will argue that ethics of research evaluation lies in the overlapping area of research ethics and evaluation ethics, followed by a discussion of three ethical theories: deontological ethics, consequentialist ethics, and virtue ethics. Finally, we propose that best practices of research evaluation can be based on a mixed approach.

2 Research Ethics and Evaluation Ethics: Guidelines and Principles

In this section, we review major documents concerning research ethics and integrity, on the one hand, and evaluation ethics, on the other, to situate ethics of research evaluation in the overlapping area of these two domains (**Figure 1**).

2.1 Research ethics and research integrity

The European Code of Conduct for Research Integrity (ALLEA 2017) is a comprehensive document illustrating the principles of research ethics, including reliability, honesty, respect, and accountability. It also describes good research practices in different scenarios. Of particular interest to this article is the section on *reviewing, evaluating and editing*, where it states:

- Researchers take seriously their commitment to the research community by participating in refereeing, reviewing and evaluation.

- Reviewers or editors with a conflict of interest withdraw from involvement in decisions on publication, funding, appointment, promotion or reward.
- Reviewers maintain confidentiality unless there is prior approval for disclosure.
- Reviewers and editors respect the rights of authors and applicants, and seek permission to make use of the ideas, data or interpretations presented.

Although the good practices prescribe what a reviewer should do, there is little guidance as to how to develop research evaluation processes and criteria, or how to deal with oft-debated issues of bias and conservatism in peer review and the negative impacts of the use of citation-based metrics. In other words, there is a lack of principles guiding the processes and criteria of research evaluation in and of itself.

2.2 Evaluation ethics

Evaluation ethics has been discussed and debated in the context of international development. The American Evaluation Association (AEA), Australian Evaluation Society (AES, AES2), Canadian Evaluation Society (CES), UK Department of International Development (DFID), and United Nations (UN), for example, have published guidelines and best practices of evaluation (**Table 1**). In Table 1 we list the topics about the ethics of evaluation drawn from the current ethical guidelines and good practices for evaluation by the aforementioned institutions.

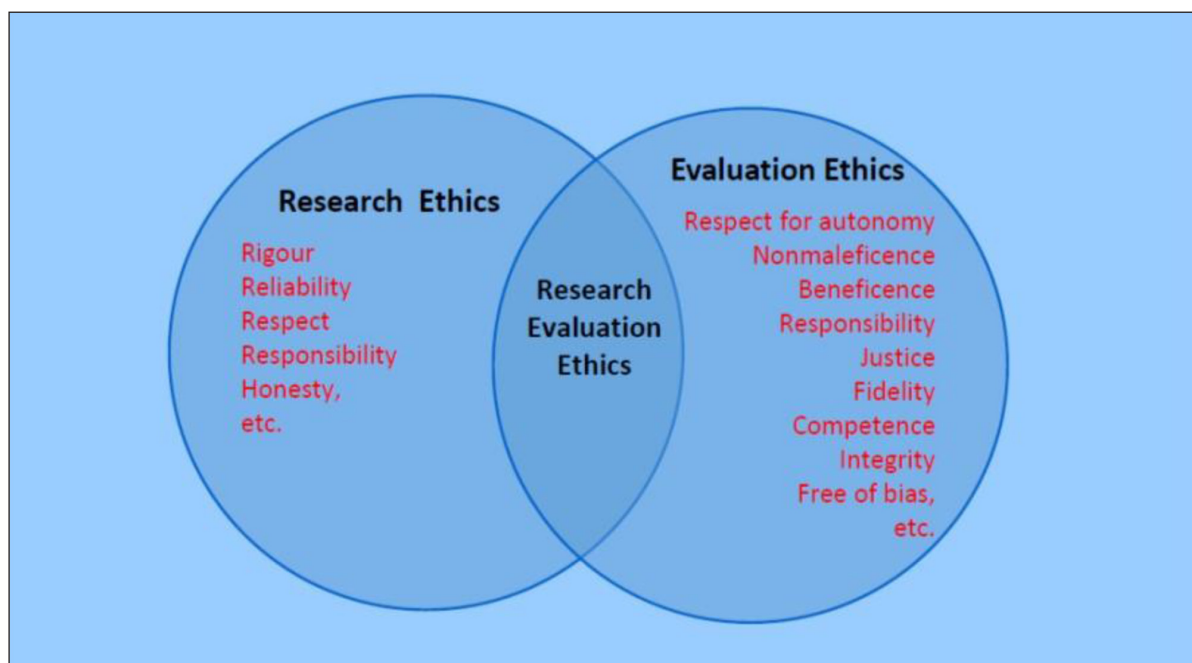


Figure 1: Research Evaluation Ethics between Research Ethics and Evaluation Ethics. (Adaptation from Biagetti, Gedutis 2019).

Table 1: Topics addressed in guidelines of ethics of evaluation.

Topics	Organisations
Responsibility	AES2, UNEG
Systematic inquiry	AEA, AES
Free of bias	AEA, UNEG
Avoid conflict of interest	AEA, AES2, CES, DFID, UNEG
Competence and honesty	AEA, AES, AES2, CES, UNEG
Accountability	AES2, CES, UNEG
Respect for dignity and diversity	AEA, AES, CES, DFID, UNEG
Avoidance of harm	AEA, AES2, DFID, UNEG
Common good	AEA, AES2
Disclose evaluation results	AEA, AES

However, these guidelines and best practices are not usually justified or supported by ethical theories. For instance, Helen Simons, a plenary speaker at The Framing Ethics in Impact Evaluation workshop (Barnett & Munslow, 2014) argues that the current ethical guidelines are mostly principles of intentions and they are often about methodology of evaluation and about the quality of the evaluation product, while ethical guidelines should instead focus on whether research evaluation is good and right, pointing to the need for an ethical theory to guide behaviour and choices of evaluators. Further, the importance of socio-political contexts is highlighted by another speaker at the workshop, Laura Camfield: "Therefore, instead of having an absolute minimum standard, it seemed to be useful to put in place a process whereby standards may be arbitrated in relation to the specific socio-political context" (Barnett, Munslow (Eds.) 2014: 11).

Similarly, publications such as The Leiden Manifesto has drawn attention to the important topics for the ethical entailments for research evaluation. However, it is worth noting that they do not present deep discussions about the issues and challenges. Groves Williams (2016) argues that ethical advices tend to be different for evaluation and for research, as these differ in purposes and follow different processes; despite this, some consider the evaluation a kind of research activity, as it generates knowledge. Hence, it is important to consider the ethical theories for research evaluation, which considers specifically evaluation ethics in the context of research.

3 Ethical Theories

To address ethical issues of research evaluation, it is necessary to look at the ethical theories and instruments they provide. Studying traditional ethical approaches, Doris and Stich (2007) attempt at broadening the methodological scope of philosophical ethics. They argue that traditional ethical theories can benefit from an empirical approach toward the solution of ethical issues. It is thus possible that new ethical approaches, as well as introduction of implicit and tacit ethical knowledge, can provide us with explanatory mechanisms needed in our pursuit for considering ethics of research evaluation. Looking back at the theoretical tradition in ethics, which is applicable to research in general and research evaluation in particular, the general ethical choices are limited to certain number of theoretical approaches. The 'hard-core' of normative ethics, is comprised, first of all, of deontological ethics and consequentialist ethics, and, to a lesser extent, virtue ethics.

3.1 Deontological ethics

Deontological ethics places special emphasis on the relationship between duty and the morality of human actions. In deontological ethics an action is considered morally good because of some characteristic of the action itself, not because the product of the action is good. Deontological ethics holds that at least some acts are morally obligatory regardless of their consequences for human welfare. It might be even labeled as 'Duty for duty's sake'. The most typical examples: Thou shalt..., thou shalt not... (Old Testament); Love thy neighbor (New Testament); Good is to be done and evil is to be avoided (Thomas Aquinas); Act as if the maxim of your action were to become through your will a universal law of nature (Immanuel Kant).

Individuals are subject to absolute and universal rules and duties, which are defined independently of them. Individuals find themselves in the realm of duties given by extra-individual entities such as God, Humanity, Rationality, *Weltgeist* etc. These duties are claimed to be universal, therefore, they cannot be altered by an individual. Moral behavior is the one, which sticks to the rules without exceptions. Ethical instructions are reduced to clearly enumerated and limited number of norms. These norms and their explanations are straightforward directions. They leave no space for moral ambiguity and further discussions. This type of ethical theories faces at least two challenges: (i) as a rule, they are too general, and/or (ii) too rigid. Therefore, they are hardly applicable in concrete situations, e.g. one must know what is good in advance and this kind of knowledge tend to neglect individual differences in complex situations or in ethically desired cases? What if universal moral law cannot be applied in everyday action of a peer reviewer? Or what if being a researcher became a universal norm? In the context of research evaluation, deontological ethics can guide the list of norms, regulating behavior of research evaluators. The norms are to be formulated in a way of BEs and DOs or DON'Ts, for example: do not harm, respect, be objective.

3.2 Consequentialist ethics

Consequentialist ethics concerns universal values, for example, life, freedom, property, and so on. Moral behavior is defined by the values 'saved'. In other words, the only important criterion of moral action is the increase of the amount of common good in society. There is not a defined list of norms, as predefined norms are not relevant in consequentialist ethics. Consequentialist ethics is a theory of morality that derives duty or moral obligation from what is good or desirable as an end to be achieved. Morally good action is the one having the best possible consequences if compared with other actions. The leading principle in consequentialist is so-called social principle: The greatest good for the greatest number. Other consequentialist principles include: principle of consequences, principle of utility, principle of hedonism and principle of universality.

In the context of research evaluation, we can ask questions such as: What greater social good is created during the evaluation process? What values does the evaluation process refer to? How to maximize happiness in scientific community? Consequentialist ethics have at least three unresolved issues: (i) unpredictability of consequences, e.g. not every

situation is as simple and transparent, as providing clear-cut assurance of its outcomes, (ii) hedonist approach, e.g. it would be rather wrong to suppress or censor the results of a research study, which would definitely make some politicians unhappy, and (iii) difficulties to measure and to compare the consequences of ethical action, e.g. how to provide a strong case in measuring the consequences of two or more conflicting values, such as truth and happiness, freedom and security, scholarly integrity and solidarity etc.

3.3 Virtue ethics

Virtue ethics serves as an alternative to both deontological and utilitarian ethics. Virtue ethics concentrate on what kind of person one should be and become, and what virtues should she possess. That is to say, ethics is not about rules or actions but about personal character and traits. From this point of ethical view, it is crucial to define the virtues of both researcher and evaluator in the context of research evaluation. Having in mind that research evaluation is not only about the evaluator but also about the evaluated, virtue ethics faces possible dilemma: What if the virtues of the former and the latter are incompatible? And what if personal character traits might be incompatible with general research ethos? With the focus on what kind of person an evaluator should become, virtue ethics does not provide a strong case for universality of ethical norms and principles.

3.4 A mixed approach of ethics in research evaluation

In order to avoid the shortcomings of the ethical theories as discussed above, we propose a mixed approach to tackle the issues of research evaluation. We assume it is possible to combine deontology and consequentialism. This kind of middle-way approach has the capacity to transgress the boundaries of the rivalling theories and provide a basis needed for research evaluation ethics.

When one talks about principles, simultaneously she might state the values. In this case the utilitarian (or similar) values might be expressed as norms in a commandment way: e.g. universalism as a value might be transformed into a commandment or imperative. Certain values and virtues, such as honesty, responsibility, respect etc., might be transformed into duties adding verbs: be honest, stay responsible, respect others etc. In this sense both evaluation ethics and research ethics are mainly based on norms, rules and principles, manifested in statements, initiatives and manifestos. For example, the San Francisco Declaration on Research Assessment aka DORA (ASCB, 2013), The Leiden Manifesto for Research Metrics (Hicks, et al., 2015), The Hong Kong Principles for Assessing Researchers (Moher, et al., 2020) are formulated as imperative deontological claims, for example, 'do not use metrics as a surrogate measure', 'be open and transparent'. In **Table 2** are presented some basic characteristics of principal ethical theories along with some examples of their use in research evaluation.

One of the major issues in deontological ethics is the source and authority of the principles: Who issues them? And why those subjected to them are supposed to comply with the principles? Eliminating entities (God, karma, Kant's

Table 2: Ethical theories and their characteristics.

Ethical theories	Basic characteristics	Typical examples from research evaluation
Deontological	<ul style="list-style-type: none"> Given rules, norms and principles Morality is based on individual duties and obligations The notions of right and wrong are preconceived, they are not subject to change Formalism: priority of rules over consequences (e.g. common good) 	The duty of the evaluator is to comply with the rules and norms, which are not under her control. Precise application of the rules is a priority of the evaluation procedures. Research is evaluated <i>per se</i> and not on its social consequences. As in Leiden Manifesto: protect excellence in locally relevant research; allow those evaluated to verify data and analysis; scrutinize indicators regularly etc.
Consequentialist	<ul style="list-style-type: none"> Given values Morality is based on the consequences of action The notions of right and wrong are not clearly defined in advance Teleology: priority of consequences (e.g. common good) over rules Context-dependence 	The evaluator seeks to maximise common good assessing potential impact of the research under scrutiny. Research is not evaluated <i>per se</i> , its importance is revealed rather via its consequences. As in objectives of Cardiff Statement (2019): "the first is to restate and champion the fundamental role that the SSH play in society and the second is to call for an expanded role for the social sciences and humanities in tackling problems through interdisciplinary research."
Mixed approach (of this paper)	<ul style="list-style-type: none"> Norms and principles are related to common good Some notions of right and wrong are given but they might be changed by ethical subjects Common good means taking stakeholders into consideration Context-dependence 	The evaluator needs to take into consideration all the stakeholders that research under evaluation deals with. The norms might be more flexible and not that rigid as in deontology. Still the norms are present (which is denied by traditional consequentialism).

universal rationality), we have to deal with different types of rule-giver or try to justify the rules in different ways. In this case we can consider the notion of common good in utilitarian ethics, for it is not necessary to interrogate the question as to who exactly creates the principles. Rather, it is more crucial to elaborate such a set of principles that increases common good. For example, Collins and Evans (2017) argue that science is a moral enterprise, guided by values that matter to all, that is, the idea of common good. Thus, if a researcher participates in the creation of common good, then research evaluation is supposed to take part in it as well. For the mission of the evaluators is to assure and to control the quality of research, by which valuable knowledge along with societal impact are fostered. In this case evaluation procedures could both enable societally relevant and block irrelevant research.

The next question – how to define common good? How not to neglect theoretical knowledge, which cannot provide impact immediately on the spot? A social contract is needed, meaning discussion *inter pares* and not the dictate used by the powerful to the powerless. If one delegates certain rights, she must receive something important in return, e.g. security, freedom of thought, expression, research etc. As a temporary solution the notion of the veil of ignorance as a precondition to the original position might be borrowed from John Rawls (1971, 2001). If any stakeholder making an ethical list of principles does not know which part (including the most unprivileged one) she is going to perform during the evaluation process, then the principles might be more just or justified than otherwise.

All the above means that discussing the principles for research evaluation ethics it is necessary to organise them around the notion of common good or collective good (as Kitcher 2001 alternatively labels it) which would not solely fall under the criterion of utilitarian notion of societal happiness. By identifying the stakeholders and moral responsibility to them the research evaluation ethics gains a legitimate ground to construct further research evaluation principles and norms.

4 Conclusion: Ethical Considerations in Research Evaluation

One of the crucial issues of the research evaluation ethics is its borders and scope of the field. In other words, it is the question of what is ethical. Or, how to clearly distinguish between, let's say, the epistemological and ethical? According to Mustajoki and Mustajoki (2017), recognition of ethical questions comprises of a three-fold way: (a) identification of stakeholders (e.g., individuals, groups, communities, animals, ecosystems, future generations etc.), (b) understanding rights and responsibilities of the stakeholders and for the stakeholders, and (c) definition of options, i.e. looking for the win-win situation, or at least proximity to it, for the stakeholders involved. Although Mustajoki and Mustajoki (2017) do not refer to the idea of common good, the three criteria of the ethical presuppose common good as a horizon of ethical deliberation.

If the ethical aim of any research is to increase the amount of common good or impact, then what is the ethical aim of the research evaluation? Or to paraphrase the latter question, whom the research evaluator is responsible to? An abstract notion such as common good serves as a horizon, which provides a both thematic and problematic framework for the ethical consideration in the research evaluation. Also Social Sciences and Humanities contribute to the common good, notwithstanding they are usually difficult to trace, track, and measure. Therefore, in order to make it operational 'common good' is to be divided into smaller realms inhabited by different collective stakeholders. In practice it means the following: every evaluator should consider the target groups (stakeholders) that research is dealing with. The authors of *ENRESSH Policy Brief on Research Evaluation* (Ochsner et al., 2020) claim that one can find four major categories—research production, research consumption and use, research policy and administration, evaluation services—and three intermediary categories of stakeholders, resulting in a taxonomy of twenty different types of stakeholders: from researchers to business, from cultural institutions to research councils, from taxpayers to learned societies, from funders to data providers, and so on. And this diversity of potential stakeholders must be taken into consideration when ethical issues are discussed.

Evaluation procedures should start with acknowledging responsibility to the disciplinary and academic communities. Only truthful research of high quality might be of any societal value. Thus, at the initial stage of an evaluation scientific integrity is needed to be checked and evaluated. Research with considerable flaws is incapable to benefit the collective good in any meaningful sense. If and only if the research under consideration is both epistemologically and methodologically plausible and valid, it has potential to benefit the broader amount of common good in extra-academic communities or the larger society as a means of practical problem solving. Therefore, setting the ethical principles for the research evaluation we must concentrate on these, which increase the amount of common good for different groups of stakeholders.

Schwandt (2015) urges general guidelines for evaluation grounded on the requirement of 'critical thinking', which involves the absence of political, personal, cultural and disciplinary biases and refuses the group-centered perspective and prejudices. More recently, Schwandt (2018) urges the necessity of a 'professionalism in evaluation', that is an ethical culture of evaluation, which concerns interpersonal relationships, the right conduct of evaluative process, the social responsibility, the necessity to serve the common good, to respect the dignity and cultural values of individuals and groups. Furner (2014) has suggested a conceptual framework for bibliometric ethics with the following essential tasks:

- Identification of relevant subgroups
- Identification of the kinds of *actions* taken by the members of each subgroup in the course of bibliometric evaluators

- Identification of the *values* held by the members of each subgroup
- Identification of the *principles* for which the members of each subgroup advocate
- Identification of *holes* in the ethical systems analysed
- Identification of *violations*

The different systems of assessment can affect scientific production in academic and research institutions (Whitley 2011) and the crucial point concerns the potential limit, due to the fear of the assessment, of the universities' independence in pursuing research that follows unorthodox methodologies, or in developing innovative fields of research in disagreement with dominant approaches. The assessment of research activities, either *ex ante* or *ex post*, entail very important ethical issues. Pursuing the 'common good' in the research evaluation means that in case of multidisciplinary or interdisciplinary research every stakeholder with which the research is dealing with must be considered (ESF 2011). In this article, we consider three ethical theories—deontological, consequentialist and virtue ethics—and propose a mixed approach for developing a framework in the design and development of research evaluation. Moreover, the ethical theories can be deployed in analysing empirical findings for understanding the ethical approaches, as well as ethical dilemmas, in research evaluation.

Acknowledgements

This study is supported by COST Action 15137 European Network for Research Evaluation in the Social Sciences and Humanities (ENRESSH).

The authors would like to thank the Department of Letters and modern Cultures of Sapienza Rome University, which held two Short-Term Scientific Meetings on the topic in 2019.

Competing Interests

The authors have no competing interests to declare.

References

- AEA.** (2018). American Evaluation Association Guiding Principles for Evaluators. <https://www.eval.org/p/cm/ld/fid=51>
- AES.** (2013). Australasian Evaluation Society Code of ethics, 2000 and 2013. https://www.aes.asn.au/images/stories/files/About/Documents%20%20ongoing/code_of_ethics.pdf
- AES.** (2013). Australasian Evaluation Society Guidelines for the ethical conduct of evaluations, 2010 and Revision 2013. <https://www.aes.asn.au/images/stories/files/About/Documents%20-%20ongoing/AES%20Guidelines%2010.pdf>
- ALLEA.** (2017). The European Code of Conduct for Research Integrity. <https://allea.org/portfolio-item/the-european-code-of-conduct-for-research-integrity/>
- ASCB.** (2013). San Francisco Declaration for Research Assessment. <<http://www.ascb.org/files/SFDeclarationFINAL.pdf>> accessed 15 June 2017.
- Barnett, C., & Munslow, T.** (Eds.). (2014). Framing Ethics in Impact Evaluation: Where Are We? Which Route Should We Take? IDS Evidence Report 98, Brighton: IDS. <https://www.cdimpact.org/publications/workshop-report-framing-ethics-impact-evaluation-where-are-we-which-route-should-we>
- Biagetti, M. T., & Gedutis, A.** 2019. Towards Ethical Principles of Research Evaluation in SSH, *The Third Research Evaluation in SSH Conference*, Valencia, 19-20 September 2019. Slides.
- Biagioli, M., & Lippman, A.** (Eds.) (2020). *Gaming the Metrics: Misconduct and Manipulation in Academic Research*. Cambridge, Massachusetts: The MIT Press. DOI: <https://doi.org/10.7551/mitpress/11087.001.0001>
- Bornmann, L.** (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45, 197–245. DOI: <https://doi.org/10.1002/aris.2011.1440450112>
- Cardiff University.** (2019, July). Research Integrity and Governance Code of Practice. https://www.cardiff.ac.uk/_data/assets/pdf_file/0004/937021/Research-Integrity-and-Governance-Code-of-Practice-v3-PDF.pdf. Accessed December 4, 2019.
- CES – Canadian Evaluation Society.** Guidelines for Ethical Conduct. <https://evaluationcanada.ca/ethics>
- Collins, H., & Evans, R.** (2017). *Why Democracies Need Science*. Cambridge: Polity Press
- Dahler-Larsen, P.** (2012). *The Evaluation Society*. Stanford, California: Stanford Business Books. DOI: <https://doi.org/10.2307/j.ctvqsdq12>
- de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B.** (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, 25(2), 161–169. DOI: <https://doi.org/10.1093/reseval/rw038>
- DFID.** (2011). Department for International Development. *Ethics principles for research and evaluation*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/67483/dfid-ethics-prcpls-rsrch-eval.pdf

- Doris, J. M., & Stich, S. P.** (2007). As a Matter of Fact: Empirical Perspectives on Ethics. In F. Jackson and M. Smith (Eds.), *The Oxford Handbook of Contemporary Philosophy*. Oxford: Oxford University Press, pp. 114–152. DOI: <https://doi.org/10.1093/oxfordhb/9780199234769.003.0005>
- ESF** (2011). European Science Foundation, Member Organization Forum European Peer Review Guide Integrating Policies and Practices into Coherent Procedures. https://repository.fteval.at/148/1/2011_European%20Peer%20Review%20Guide.pdf
- Furner, J.** (2014). The Ethics of Evaluative Bibliometrics. In B. Cronin & C. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*. Cambridge, Massachusetts: The MIT Press.
- Groves Williams, L.** (2016). Ethics in international development evaluation and research: what is the problem, why does it matter and what can we do about it? *Journal of Development Effectiveness*, 8(4), 535–552. DOI: <https://doi.org/10.1080/19439342.2016.1244700>
- Helmer, M., Schottdorf, M., Neef, A., & Battaglia, D.** (2017). Gender bias in scholarly peer review. *eLife*, 6. DOI: <https://doi.org/10.7554/eLife.21718>
- Hicks, D., Wouters, P., Waltman, L., et al.** (2015). The Leiden Manifesto for research metrics. *Nature*, 520. pp. 429–431. DOI: <https://doi.org/10.1038/520429a>
- Kitcher, P.** (2001). *Science, Truth, and Democracy*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/0195145836.001.0001>
- Horbach, S. P. J. M., & Halffman, W.** (2019). Journal Peer Review and Editorial Evaluation: Cautious Innovator or Sleepy Giant? *Minerva*, 58(2), 139–161. DOI: <https://doi.org/10.1007/s11024-019-09388-z>
- Lamont, M.** (2009). *How professors think. Inside the curious world of academic judgement*. Cambridge, London: Harvard University Press. DOI: <https://doi.org/10.4159/9780674054158>
- Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B.** (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2–17. DOI: <https://doi.org/10.1002/asi.22784>
- Luukkonen, T.** (2012). Conservatism and risk-taking in peer review: Emerging ERC practices. *Research Evaluation*, 21(1), 48–60. DOI: <https://doi.org/10.1093/reseval/rvs001>
- Ma, L., & Ladisch, M.** (2019). Evaluation complacency or evaluation inertia? A study of evaluative metrics and research practices in Irish universities. *Research Evaluation*, 28(3), 209–217. DOI: <https://doi.org/10.1093/reseval/rvz008>
- Merton, R. K.** (1973). The normative structure of science. In R. K. Merton (Ed.), *The Sociology of Science: theoretical and empirical investigations*, edited and with an introduction by Norman W. Storer (pp. 267–278). Chicago: The University Chicago Press.
- Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., et al.** (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLOS Biology*, 18(7), e3000737. DOI: <https://doi.org/10.1371/journal.pbio.3000737>
- Mustajoki, H., & Mustajoki, A.** (2017). *A New Approach to Research Ethics: Using Guided Dialogue to Strengthen Research Communities*. New York: Routledge. DOI: <https://doi.org/10.4324/9781315545318>
- Ochsner, M., Kancewicz-Hoffman, N., Ma, L., Holm, J., Gedutis, A., Šima, K., et al.** (2020). ENRESSH Policy Brief Research Evaluation. *figshare*. Online resource. DOI: <https://doi.org/10.6084/m9.figshare.12049314.v1>
- Ochsner, M., Kulczycki, E., & Gedutis, A.** (2018). The Diversity of European Research Evaluation Systems, in Science, Technology and Innovation indicators in transition. *23rd International Conference on Science and Technology Indicators (STI 2018)*, 12–14 September 2018, Leiden (pp. 1235–1241). <https://openaccess.leidenuniv.nl/handle/1887/65217>
- Rawls, J.** (1971). *The Theory of Justice*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.
- Rawls, J.** (2001). *Justice as Fairness: A Restatement*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.
- Schwandt, T. A.** (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford: Stanford University Press. DOI: <https://doi.org/10.1515/9780804795722>
- Schwandt, T. A.** (2018). Acting together in determining value: A professional ethical responsibility of evaluators. *Evaluation*, 24(3), pp. 306–317. DOI: <https://doi.org/10.1177/1356389018781362>
- UK Evaluation Society Guidelines for good practice in evaluation.** <https://www.evaluation.org.uk/app/uploads/2019/04/UK-Evaluation-Society-Guidelines-for-Good-Practice-in-Evaluation.pdf>
- UNEG – United Nations Ethical Guidelines for Evaluation.** (2008). <http://www.unevaluation.org/document/detail/102>
- Whitley, R.** (2007). Changing Governance of the Public Sciences. The Consequences of Establishing Research Evaluation Systems for Knowledge Production in Different Countries and Scientific Fields. In R. Whitley & J. Gläser (Eds.), *The Changing Governance of the Sciences. The Advent of Research Evaluation Systems* (pp. 3–27). Dordrecht: Springer. DOI: https://doi.org/10.1007/978-1-4020-6746-4_1
- Whitley, R.** (2011). Changing Governance and Authority Relations in the Public Sciences. *Minerva*, 49(4), 359–385. DOI: <https://doi.org/10.1007/s11024-011-9182-2>
- Whitley, R., Gläser, J., & Laudel, G.** (2018). The Impact of Changing Funding and Authority Relationships on Scientific Innovations. *Minerva*, 56, 109–134. DOI: <https://doi.org/10.1007/s11024-018-9343-7>
- Wilsdon, J., Allen, L., Belfiore, E., et al.** (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. DOI: <https://doi.org/10.4135/9781473978782>

How to cite this article: Biagetti, M. T., Gedutis, A., & Ma, L. (2020). Ethical Theories in Research Evaluation: An Exploratory Approach. *Scholarly Assessment Reports*, 2(1): 11. DOI: <https://doi.org/10.29024/sar.19>

Submitted: 25 May 2020

Accepted: 08 September 2020

Published: 25 September 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Scholarly Assessment Reports is a peer-reviewed open access journal published by Levy Library Press.

OPEN ACCESS The Open Access logo, consisting of the words 'OPEN ACCESS' followed by a circular icon containing a stylized padlock with an open keyhole.