*Article*

A&H

Arts & Humanities in Higher Education
2015, Vol. 14(1) 111–126

DOI: 10.1177/1474022214534381
ahh.sagepub.com
SAGE

# Research evaluation and the assessment of public value

**Jordi Molas-Gallart**
INGENIO (CSIC-UPV), Spain

**Abstract**

Funding organisations are increasingly asking academics to show evidence of the economic and social value generated by their research. These requests have often been associated with the emergence of a so-called 'new social contract for research' and are related to the implementation of new research evaluation systems. Although the research evaluation rhetoric is similar across countries and organisations, in practice evaluation can fulfil very different purposes. Additionally, the assessment of the public value of research poses different challenges depending on the academic field under analysis. This paper distinguishes three main research evaluation goals: to inform the distribution of public resources among competing objectives or performers, to help improve the implementation of policies and programmes, and to control the use of public funds. It then argues that assessing the value of research in the arts and humanities calls for a research methodology capable of providing a fine-grained understanding of the variety of, often diffuse, ways in which arts and humanities research can generate value. The methods that we need to do this are better suited to fulfil the improvement goal of evaluation, and require a 'formative' approach to evaluation supporting the social engagement of academic researchers.

**Keywords**

Arts and humanities social value, evaluation functions, formative evaluation, impact assessment, research evaluation

**Corresponding author:**
Jordi Molas-Gallart, INGENIO (CSIC-UPV), Universitat Politècnica de València, Camí de Vera s/n, Valencia 46022, Spain.
Email: jormoga@ingenio.upv.es

## Introduction: The changing context for research in the arts and humanities

### *A new 'social contract' for research*

The growing interest in the assessment of the value of research outcomes is linked with a policy practice that has been spreading over the last three decades: funding organisations are increasingly asking for evidence of the results and effects that their 'research investments' eventually generate. Although the way in which they do so, and the type of evidence they ask for, varies across countries and organisations, there is a common interest in finding evidence of the economic or social value that research helps generate. As other articles in this special issue show, the arts and humanities (A&H) are not immune to this trend and present distinctive challenges to the assessment of research value.

The growing interest in obtaining evidence of economic or social impact has often been associated with the emergence of a 'new social contract' between science and society. The new 'contract' would emphasize the need to obtain economic and social benefits from research investments and would not take it for granted that such benefits would naturally follow from academic enquiry and scientific discoveries (Guston, 1992; Martin, 2004; Ravetz, 1988; Rip, 2003). This new environment helps explain the generalised diffusion of research evaluation practices and their institutionalisation (Whitley and Glässer, 2007).The expectation that science would generate social and economic returns is not new. Debates on the social and economic value of research and the extent to which they may guide the work of scientists can be traced to the early 20 th century (Kline, 1995). As Pielke argues, 'in accepting the importance of science, politicians never deviated from their expectation that science would be instrumental' (Pielke, 2012). Yet, what changed towards the end of the 20 th century was the way in which these expectations were translated into policy practice. The position that extraneous interference in the work of science in the name of practical application would be counterproductive to this same objective is also very old and was dominant in the decades following the Second World War. It is perhaps best reflected in Mme Curie's famous quote:

> We must not forget that when radium was discovered no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it. It must be done for itself, for the beauty of science and then there is always the chance that a scientific discovery may become like the radium, a benefit for humanity. (Curie, 1921)

Many scientists will still endorse these words spoken more than 90 years ago, but funding organisations may not be satisfied by assurances that 'there is always a chance' that benefits will ensue from scientific research. The attention paid to 'research impact' and its assessment is based on the belief that we should verify

that such benefits take place, and that we should learn how they come about in order to increase their future likelihood.

From some quarters, the lack of evident economic impact was used to criticise public investment in research and to question whether research generated 'value for money'. For instance, in the late 1970 s a critique of UK public sector waste targeted UK Government research establishments, arguing that expenditure in such organisations should be set 'alongside an appraisal of what this expenditure has bought for the community', and that such appraisal should involve 'that section of industry most concerned with the end product of research' (Chapman, 1979: 180).

This view implicitly considered the economic returns to industry as the only valuable result of research and was based on a simple image of knowledge generation and application: research leads to results that are then transferred to industry and turned into economic benefits. From this perspective, many, although not all, research activities in the A&H would not generate 'value for money'. Indeed, for many years the analysis of how research could generate valuable impacts focused on the economic returns on 'technology transfer' achievable through commercial transactions like patent licensing and spin-offs (Radosevich and Kassicieh, 1994).

Partly as a reaction to the emphasis placed on economic returns, some analysts started to focus instead on the contributions of science to the generation of public or social value. Bozeman and Sarewitz (2011) argue that, while the economic and scientific impact of research has received substantial attention leading to the development of sophisticated impact assessment methodologies, science policy research has not focused with equal energy on the problem of the social value of research. This is important because the economic and scientific effects do not capture all the dimensions that are important to policymakers and the public at large. Science policy often explicitly pursues values other than the merely economic or scientific. Such public values can be defined as those 'providing normative consensus about (1) the rights, benefits, and prerogatives to which citizens should (and should not) be entitled'; (2) the obligations of citizens to society, the state and one another; and (3) the principles on which governments and policies should be based' (Bozeman, 2007). The ways of measuring such values are 'as diverse as the aspirations of users, including not only pricing and profits, but status, curiosity and mastery of the physical worlds' (Bozeman et al., 1999: 9).

The differentiation between economic returns and some wider concept of social value has reached the official policy literature. The Ricci Report on the assessment of the 'social and environmental impact of European research' distinguishes social and environmental effects from economic impacts and makes the well-known observation that measurement of the former is much more challenging (European Commission, 2005). As we move into the second decade of the 21 st century it would seem that the notion of research impact has definitely become multifaceted.

Yet, despite repeated warnings from scholars and evaluation experts that research impact assessment should not limit itself to the calculation of economic benefits, this practice has become entrenched. Benneworth and Jongbloed (2010)

argue, for instance, that studies of how universities generate value have 'largely become narrowly understood in terms of universities' economic contributions' and that 'this emergent restrictive definition [. . .] is a worrying development, overlooking the potential of universities' knowledge in the Humanities, Arts and Social Sciences'. How 'value for money' is defined in practice and therefore how the 'social contract' is implemented becomes crucial for the future of the A&H. Too narrow a focus on economic returns may lead to a perception of valuable return that is biased against the type of contributions that we may expect from the A&H.

## *Implications for A&H research*

Scholars in the A&H see this new environment with concern, and to a certain extent such concern is justified. A study of the A&H in the UK conducted by the Centre for Business Research at the University of Cambridge found that a large number of A&H academics saw their research as 'basic' and that they were more likely than their counterparts in other disciplines to feel that it was of no relevance to external organisations (Hughes et al., 2011).[1] Such self-perception as 'basic researchers' with little applied or economic interest is accompanied by concerns about the way in which A&H scholars are *perceived* by policy-makers. For instance, Martha Nussbaum, while writing in defence of the A&H and against what she sees as 'education for growth', states of the A&H:

> Seen by policy-makers as useless frills, at a time when nations must cut away all useless things in order to stay competitive in the global market, they are rapidly losing their place in curricula . . . nations prefer to pursue short-term profit by the cultivation of the useful and highly applied skills suited to profit making. (Nussbaum, 2012: 2)

Although Nussbaum does not use the term 'social contract', her position can be restated in the language of the scholars that see current developments in the management of academic activities as the outcome of the new 'social contract' between science and society: the concern is that the new 'social contract' will be narrowly expressed as the need to pursue economic returns and that, as a consequence, it introduces a single logic to justify research investments focused on economic growth. Nussbaum appears to assume implicitly that the A&H do not develop 'highly applied skills suited to profit making' and that all work that does not do so is increasingly marginalised by the new policy trends. Yet this is an oversimplification built upon the assumption that the A&H constitute an homogeneous category. Empirical research has shown that there are very substantial differences within the A&H in the way that researchers engage with the potential users and beneficiaries of their research. These differences are important for the way in which evaluation strategies can be defined and implemented.

## Diversity in A&H research and its implications for evaluation

The Centre for Business Research study mentioned above finds important differences within the A&H when it comes, for instance, to the extent and the way in which they relate to research users and beneficiaries outside academia. While academics in the 'Creative Arts and Media' see their research as 'user-inspired' and applied, are highly connected and engage in relatively high levels of commercialisation, scholars in 'languages' tend to see their research as being of little relevance to external organisations. It is likely that further differences would have appeared within these groups had the study descended to a more detailed level of disaggregation.

It could be argued that these internal differences make it difficult to treat the area as a single whole. But internal diversity is not unique to the A&H. There are areas in the 'hard' sciences that are highly theoretical and whose link to application can only be long-term and circuitous, while others (say, pharmaceutical research) are explicitly linked with application though a series of formalised and documented steps (Benneworth, 2015) that make it easy to trace the path between research and application. Such diversity makes it difficult to treat as a unit any grouping of disciplines or intellectual endeavours we may come across. We could then debate, for instance, whether we should refer to the social sciences and humanities as a single area or whether we should always and under every circumstance treat them as separate entities.

Yet the research context will often force us to deal with groups we know are not homogeneous; from a practical perspective, the 'subject' of any evaluation will have to be defined to suit its specific policy and management objectives and these will, in turn, be framed by the existing, relevant organisational boundaries. In the UK much research has been carried out on the impact of the A&H (Arts & Humanities Research Council, 2012; Hughes et al., 2011; Levitt et al., 2010) not because there is any profound epistemological reason to bring these broad bodies of knowledge together and treat them separately from the social sciences, but because such evaluation research has been funded by the Arts and Humanities Research Council, the other social sciences being funded by a different Council. Our own studies of the social engagement of researchers from the Spanish National Research Council (CSIC) treat the humanities and the social sciences together because CSIC organises these disciplines under the same research area (Olmos et al., 2013). These subjects are internally diverse and the studies show that seldom do such differences fit with established large disciplinary aggregates.

Therefore, regardless of the way in which we may outline our research subject, there is likely to be a great deal of internal diversity. The A&H are not a solid, homogenous block and therefore the threat (perceived or otherwise) posed by the 'new social contract' will be felt and interpreted differently within different sections of the A&H research communities. How can we then deal with such internal diversity when analysing the ways in which academic disciplines generate value?

Aggregate quantitative methods aim to uncover common patterns, characteristics and general trends that a diverse population shares, and which we would not have been able to see if we had focused only on the variety of situations and actors and the importance of contextual factors. Quantitative studies can therefore help us find commonalities instead of identifying differences. Several studies have, for instance, found that in many important aspects the social sciences and the humanities are not so different from the natural or hard sciences. Hughes and his colleagues at the Centre for Business Research, University of Cambridge, found that the forms of connection and the external partners with whom such connections were established were similar between the A&H and other disciplines in the 'hard' sciences (Hughes et al., 2011). There is no doubt that such broad areas of knowledge are both internally diverse and different across them, and yet aggregate analysis is able to find common characteristics bringing together diverse populations. On the other hand, however, there are whole evaluation 'schools' (Pawson and Tilley, 1997) and impact assessment techniques (Spaapen and Dijstelbloem, 2005) that seek to identify in detail how policy mechanisms interact with different contexts in order to generate different types of outcomes.

We therefore encounter a tension between diversity and similarity: depending on our level of analysis, subject definition, and research methods, our assessments of the value generated by scientific work may offer common cross-cutting views or emphasize contextual factors explaining differences down to the group or individual level. By using quantitative techniques that allow us to focus on aggregates we can find common trends and characteristics, while some qualitative studies will highlight differences and the 'contextual' conditions that lead to such differences.

This tension between, on the one hand, seeking common causes and generalisable trends, and, on the other, emphasizing the role of contextual conditions and of differences across groups of subjects and even individuals, is crucial to the study of whether and in which ways academic work generates social value. The choice among approaches should be informed by the situation we are trying to analyse and the specific questions we are attempting to answer. If such study is conducted with the purpose of contributing to an evaluation of the performance of an individual, group or organisation, the perspective we will adopt will depend on the objectives of such an evaluation.

## The variety of evaluation goals

Researchers, like many other individuals, get involved very often in evaluations, typically as their subjects. A smaller group of individuals will be actively involved in proposing, discussing, designing, and carrying out evaluations, or in using their results. Evaluation has become such an established part of organisational life that we have even ceased to question its purpose; evaluations are starting to be taken for granted as part of the normal life of an organisation (Dahler-Larsen, 2012).

Academia has embraced evaluation with fervour. Until recently, evaluation in an academic context was limited to a small number of tasks: individual track

records were evaluated prior to hiring or promotion decisions, and projects were assessed before funding; the purpose here was clearly to support decisions on the future distribution of resources. In some universities students assessed the performance of their lecturers (explicitly to allow lecturers to improve on their practices), and everywhere lecturers would assess the work of their pupils. Lately, however, the practice of evaluation has extended to the performance of many concurrent, highly formalised, periodic evaluations of individual academics or groups of academics (departments, institutes, universities, research centres, national aggregates of universities and research organisations).

The rhetoric justifying such practices is similar across countries and revolves around the need to support 'excellent' research and to provide 'accountability' to society for the investment it makes in such research. Accountability is needed because it is no longer taken for granted that research investments will necessarily result, sooner or later, in beneficial outcomes. In other words, the emphasis on evaluation is an expression of the new social contract for research we have discussed above. Yet in practice, there is a lot of variability in the roles played by the wide range of evaluations undertaken. What is supposed to be done and what is eventually done with the results of such evaluations is seldom questioned. In some cases the purpose is obvious. Some major exercises such as the UK Research Assessment Exercise (RAE), now rebranded Research Excellence Framework (REF), are tools for the assignment of research funds (Barker, 2007; Martin, 2011). Yet in many other cases the purpose of the evaluation is less clear. This is important because, as I will move on to argue, our approach to evaluation depends on the goal of the evaluation.

There is an extensive literature on the diverse uses and goals of evaluation. I have elsewhere argued that for the purposes of analysing different research evaluation practices we can distinguish three main evaluation purposes (Molas-Gallart, 2012): distribution of resources, improvement of practices and control of the activities undertaken. In the case of the RAE and REF an evaluation exercise can directly and explicitly be linked to a decision about the distribution of resources; individual evaluations prior to hiring and promotion decisions can also be considered within this category. In research systems where project-funding or adjustable long-term contributions are an important element of the funding structure, evaluation will tend to seek distributive goals.

Another purpose of evaluation is to help in the implementation of research initiatives or programmes by assessing how an activity is working and proposing improvements. In our area of discussion an evaluation could, for instance, be used to help devise and improve ways to increase the social value of research. Such an improvement goal has attracted much attention from evaluation researchers: it is at the core of the approach commonly known as 'formative evaluation' (Scriven, 1967). In formative evaluation the improvement goal is combined with a collaborative relationship between evaluator and evaluand. Instead of aiming at closing down the debate with definitive answers to policy and evaluative questions (Stirling, 2008), evaluation becomes a medium for self-organised learning and

can even be construed as a way to structure and 'moderate' the debate in contested policy areas (Kuhlmann, 1998), enabling negotiation and making actors more aware of the implications of different policy approaches (Lepori and Reale, 2012).

Finally, when there is little scope in the system for change and adaptation, evaluation typically pursues a 'control' purpose: it becomes an audit process checking that resources have been invested according to plan and following established rules and regulations. Research systems based on tenure and stable core funding structures will lean towards a 'control' role for evaluation practice.

Different evaluation goals will often result in the application of different evaluation methods and techniques. For instance, when impact assessment becomes part of an evaluation strategy oriented to control or distributive goals, it is common to narrow the analysis to a set of indicators that are expected to be easily audited and, when used for distributive purposes, comparable across subjects. In both cases the evaluator will try to simplify the object of analysis and find results that can tell us in a synthetic way something very specific about what is being evaluated.

Impact assessments conducted for evaluation pursuing an improvement goal (that is, to improve practice) need to focus, instead, on details. We can encounter, for instance, case stories attempting to analyse the context and conditions within which investments yield results and the ways in which this happens. In these situations, the main objective of the evaluator is to obtain detailed knowledge of the processes through which social value is generated. It must be stressed that the results that such an approach will generate are necessarily different from the results produced by approaches deployed to fulfil a controlling or distributive goal; while improvement-oriented evaluations will focus on the details of how a researcher, group or programme has worked, evaluations with a controlling or distributive purpose will typically seek results that are simple and synthetic.

## Towards an evaluation model for the A&H

Different research subjects are differently suited to different research approaches. There are contexts in which it can be easier and more productive to carry out aggregate quantitative analysis, whereas in others it may be better to rely on qualitative detailed case studies. If the subject of research is internally very diverse we may find that quantitative studies seeking for patterns and regularities will struggle to offer conclusive and relevant results. I will conclude this paper by arguing that, as the roles of evaluation are themselves linked with different research methods, the A&H yield themselves more easily to a specific type of evaluations: those pursuing an improvement goal.

### *The diffuse nature of valorisation processes*

There is an already sizeable stream of work that has analysed through case studies and other qualitative methodologies[2] the impact of social sciences and humanities research.[3] These studies have found abundant examples of academics engaging

with potential users and beneficiaries of their research and remarkable diversity in the ways in which the results of academic research have found practical application. Some are pervasive and have been identified in many studies and evaluations, like, for instance, the work of A&H researchers as curators of museums and exhibitions and their contribution to the production of historical theatre shows and films. Other contributions are less apparent, like the involvement of linguists in police investigations (Olmos et al., 2013), the role of historical research in political debates (Benneworth, in press), and the work of A&H scholars in road safety education projects (Cassity and Ang, 2006). Sometimes, the nature of these contributions has been questioned inside and outside the academic disciplines involved; for instance, the political utilisation of archaeology has a long history that has generated substantial debate and triggered detailed studies (McGuire, 2008).

Together with the different types of use and application comes diversity in the processes of application. Survey-based quantitative studies have confirmed the multidimensional nature of the valorisation processes. The study conducted by the Centre for Business Research surveyed 125,900 academics and 25,015 firms (obtaining a response rate of 17% and 11%, respectively) to find 23 different types of interactions between academics and external organisations classified under four broad activities (people-based, community-based, problem solving and commercialisation) (Abreu et al., 2009).

The pathways through which research results become socially valuable are not only diverse but also diffuse. Instead of being used in a direct and instrumental fashion, research knowledge can shape private and public activities and policies by providing a background of ideas that 'creep' (Weiss, 1980) into social and political attitudes and opinions. These diffuse paths contrast with the more straightforward links between research and application that can be found, for instance, in the formally staged process by which biomedical research acquires value (Benneworth, in press). This observation has methodological implications. Processes that can follow different paths at the same time and are diffuse, engaging different participants in different ways and over different periods of time, are difficult to encapsulate within a limited number of quantitative indicators. Enumerating and counting the connections, transactions and instances of use behind the valorisation processes is unlikely to provide a valid approximation to the social and economic value generated through such processes.

Nevertheless, there have been several attempts at developing scales of knowledge use, typically focusing on the 'utilisation process' and interpreting it as composed of different stages. The view adopted is that of a process by which a clearly bounded piece of knowledge (a theory, a method…) moves from researchers to users in cumulative steps.[4] Yet, such utilisation process is only one of the forms in which the knowledge generated by the A&H creates value; in other words, direct 'utilisation' of a specific piece of knowledge is only one among many 'valorisation processes'. Co-production and iterative processes of knowledge generation and use, together with the multiplicity of pathways and the diffuse nature of other

valorisation processes make synthetic measurement of these processes very difficult and partial, when not simply impossible.

### *The interactive nature of valorisation processes*

Despite the variety of channels through which valorisation takes place, there is an issue where substantial consensus is emerging: the importance of social interactions as the main foundation through which research results find applications that generate social value. This argument has been made in different ways and from different perspectives. Landry and his colleagues find that the determinants of knowledge utilisation lie in the linkages between researchers and users (what they label the 'interaction model') instead of the nature of the research results ('science push'), the users' needs and context ('demand pull'), or the efforts made to disseminate results ('dissemination model'). The authors suggest that knowledge utilisation depends on various disorderly interactions occurring between researchers and users rather than on linear sequences beginning with the needs of the researchers or the needs of the users (Landry et al., 2001). From a different perspective, an international team of researchers placed the interactions between researchers and stakeholders at the core of a methodology to identify and trace research impacts (Molas-Gallart and Tang, 2011; Spaapen and van Drooge, 2011). This research found evidence of the indirect and interactive processes of application occurring across scientific disciplines including the social sciences and the humanities.

It has been further argued that the interaction between researcher and non-academic stakeholders is not only crucial for the use of research results but for social research itself (van Langenhove, 2011). From this perspective, researchers who deal with human issues should aim, first and foremost, at introducing change at the local level and therefore they should favour research that is local in scope. Only at later stages there would be room for more indirect contributions through the active dissemination of knowledge; for this reason, the argument continues, the sciences that deal with human affairs need to put more focus on knowledge-brokering and participatory approaches (van Langenhove, 2011). Other authors have placed interaction with society at the centre of what a social scientist should do; researchers dealing with human issues should aim their research at praxis, and therefore social links and interactions become central to the conduct of science (rather than just the 'dissemination' of its results) (Flyvbjerg, 2001, 2005). If this is the case, it follows that the analysis of such links should be a core element of research evaluation.

Yet, when evaluation focuses on the interactions between academics and those individuals and groups that the evaluator expects to be among the non-academic users or beneficiaries of the academic's work, the evaluation will seek to address the processes through which research and its results become socially valuable. A focus on processes requires the use of research methods capable of providing a fine-grained understanding of the ways in which research generates public value, and

the contextual factors that facilitate or hinder such processes. This work is very time consuming (and therefore expensive), and yields a type of result (awareness of the importance of specific, context-dependent factors) that is useful to improve practice, to adapt policies, and to tweak implementation routines, but can only provide contributory help to make decisions on the distribution of resources. Further, it cannot be used in evaluations carried out for control purposes: to learn in detail how researchers work and interact with non-academic partners it is necessary to obtain information about collaborative links that are often informal (Schiller, 2010) and may not even be part of the researchers' formal responsibilities. The evaluator needs to get close to the evaluands and work with them, establishing a level of trust that could not be achieved if the goals of the evaluation were to be controlling or distributive.

## *Towards an approach for the evaluation of A&H research: Understanding 'value' in formative evaluation*

Addressing the processes through which research and its results become socially valuable, aligns itself naturally with the 'improvement' function of evaluation and with the formative approaches described above. Evaluation practice can support the social engagement of academia, working with researchers to identify instances of interaction and valorisation and, in the process, helping them appreciate such activities and justifying their value (Molas-Gallart and Tang, 2011). Further, in a formative evaluation approach, the analysis of the social value of research will only provide one form of evidence and will contribute to framing the arguments accompanying policy definition and implementation.

In this process, questions about the nature of the effects observed, and in particular about whether a specific effect should be considered 'valuable', do not need to be solved *a priori*. It is not necessary for the evaluator to decide beforehand, whether and how a particular effect should be assessed as positive in itself or compared with other effects. The HERAVALUE project[5] has argued that humanities research contributes to social development, and such social development includes aspects like freeing the individual from the quotidian, or helping people develop more fulfilled lives. Not only are these objectives incommensurable, but social development is a political question (Benneworth, 2015); it follows that an analysis of the valorisation of A&H research needs to be explicit about the social development model to which such research may be contributing.

This is not the same, however, as making a preliminary decision about what form of social development is preferable before engaging in an evaluation exercise. As notions of social development, and therefore of social value, can be disputed, an evaluation could do nothing more than contribute to a process of policy development based on argument and discussion. As different publics will disagree on what is considered to be social progress, the value of A&H research cannot be measured in a way that could be consensually considered as objective or fair.

This means that the role of evaluation and impact assessment is not to provide final answers but to contribute to a process of policy development based on discussion and argument (Majone, 1989). Therefore, although evaluation needs to be explicit about the concept or concepts of public value on which a specific assessment is based, these concepts need not and should not be considered unique and beyond debate.

Not only does A&H research evaluation need to address how users and beneficiaries engage in research, but also requires an understanding of how they conceptualise value. Yet, such conceptualisation and engagement will vary across research areas, projects and application contexts. For instance, the way in which research in psychology is likely to generate social value and the nature of such value are different from the form of value generated by historical research. The understanding of these differences, and of how users and beneficiaries conceptualise value in different relevant contexts, needs to become part of any evaluation research.

This further strengthens the view of evaluation we are proposing here: evaluation needs to engage with researchers and 'stakeholders', it has to be aware of the context where knowledge is generated and applied, and it needs to focus on the processes that generate and are generated by the interactions between researchers and beneficiaries. Learning about the types of knowledge and other outputs that A&H research produces, the way they are disseminated and applied, and the contexts that enable such application, will clarify the nature of foreseeable or actual impacts and provide a perspective that can then be combined with others in a 'multiple perspective' evaluation framework (Kuhlmann, 1998).

Within this approach the 'expectation of an exact measurement of ''objective impacts'' of a policy in the sense of immutable truth is neither possible nor wise' (Kuhlmann, 1998: 138). Indeed, many research evaluation experts have long been advocating a formative approach contributing to policy argument and negotiation. Some funding bodies are starting to support this view. The UK Arts & Humanities Research Council has argued in favour of formative evaluation, noting that 'too much emphasis has been given to the evaluation of impacts and not enough to the assessment of the process' by which A&H generates socially valuable outcomes (Arts & Humanities Research Council, 2013).

These are, however, difficult goals to achieve in practice, mainly because they cut against the grain of current trends in evaluation and research management. Controlling and distributive functions are dominant in the current police drive to establish evaluation as an 'accountability' tool. Formative approaches emphasizing debate and progressive adaptation are confronted with rigid administrative systems unable to incorporate a formative evaluation approach (Silvani et al., 2005) and focusing on the controlling function of evaluation (Molas-Gallart, 2012). To incorporate formative approaches to the evaluation of A&H research will often require profound changes in research management, going well beyond the boundaries of evaluation practice.

## Acknowledgements

The author thanks two anonymous referees and Dr Paul Benneworth for the very helpful comments received on previous versions of this paper.

This paper builds upon the talk 'Research Assessment and the Evaluation of Public Value', given to the 2012 HERAVALUE final conference 'Measuring the Public Value of Arts and Humanities Research: From theoretical challenges and practical problems towards a policy, public and university consensus', held on 25–26 th October 2012 at the Dublin Institute of Technology.

## Notes

1. Interestingly, however, A&H researchers were more involved in research relevant for non-commercial organisations.
2. Two reviews of impact assessments and evaluations present interesting summaries of the different approaches taken by the UK Economic and Social Research Council and the Arts and Humanities Research Council. See Luiz de Campos (2010); Marjanovic et al. (2009).
3. Studies of research impact in the humanities are often done together with other social sciences. As argued above, the different ways of splitting the study of different aspects of human activities are often guided by practical, rather than epistemological, considerations.
4. For a concise but interesting review see Landry et al. (2001). These authors consider interaction between knowledge producers and users as an independent variable explaining the utilisation of social science research among policy-makers, but the study focuses on the specific context of policy use, and the processes of knowledge utilisation they consider remain cumulative (progressing from transmission to application). Other forms of knowledge co-production between policy-makers and social scientists are not addressed. See also Landry et al. (2003).
5. http://www.utwente.nl/mb/cheps/research/current_projects/heravalue/.

## References

Abreu M, Grinevich V, Hughes A, et al. (2009) *Knowledge Exchange between Academics and the Business, Public and Third Sectors*. Cambridge: Centre for Business Research.

Arts & Humanities Research Council (2012) *The Imapct of AHRC Research 2011/12*. Swindon: Arts & Humanities Research Council.

Arts & Humanities Research Council (2013) *Workshop taking on the Challenge*. Available at: http://www.ahrc.ac.uk/Funded-Research/Funded-themes-and-programmes/Cultural-Value-Project/Pages/Workshop-taking-on-the-challenge.aspx (accessed 3 August 2013).

Barker K (2007) The UK Research Assessment Exercise: The evolution of a national research evaluation system. *Research Evaluation* 16(1): 3–12.

Benneworth P (in press) Tracing how arts and humanities research translates, circulates, and consolidates in society. How have scholars been reacting to diverse impact and public value agendas? *Arts and Humanities in Higher Education* 14.

Benneworth P and Jongbloed BW (2010) Who matters to universities? A stakeholder perspective on humanities, arts and social sciences valorisation. *Higher Education* 59: 567–588.

Bozeman B (2007) *Public Values and Public Interest: Counter-balancing Economic Individualism*. Washington, DC: Georgetown University Press.

Bozeman B, Rogers J, Roessner D, et al. (1999) *The Research Value Mapping Project. Qualitative-Quantitative Case Studies of Research Projects Funded by the Office of Basic Energy Sciences*. Washington, DC: Office of Basic Energy Sciences. Department of Energy.

Bozeman B and Sarewitz D (2011) Public value mapping and science policy evaluation. *Minerva* 49: 1–23.

Cassity E and Ang I (2006) Humanities-industry partnerships and the 'Knowledge Society': The Australian experience. *Minerva* 44: 47–63.

Chapman L (1979) *Your Disobedient Servant*. Harmondsworth: Penguin Books.

Curie M (1921) The Discovery of Radium: Address by Madame M. Curie at Vassar College, 14 May 1921, Vassar College.

Dahler-Larsen P (2012) *The Evaluation Society* (trans. Sampson S). Stanford: Stanford University Press.

European Commission (2005) *Assessing the Social and Environmental Impact of European Research* (No. EUR 21702). Luxembourg: Office for Official Publications of the European Communities.

Flyvbjerg B (2001) *Making Social Science Matter. Why Social Inquiry fails and How it can Succeed Again*. Cambridge: Cambridge University Press.

Flyvbjerg B (2005) Social science that matters. *Foresight Europe* 38(October/March): 38–42.

Guston DH (1992) The demise of the social contract for science: Misconduct in science and the nonmodern world. Program in Science, Technology, and Society, Massachusetts Institute of Technology.

Hughes A, Kitson M, Probert J, et al. (2011) *Hidden Connections. Knowledge Exchange between the Arts and Humanities and the Private, Public and Third Sectors*. Swindon: Arts and Humanities Research Council and Centre for Business Research.

Kline R (1995) Constructing 'Technology' as 'Applied Science': Public Rhetoric and Engineers in the United States, 1880-1945. *ISIS* 86(2): 194–221.

Kuhlmann S (1998) Moderation of policy-making? Science and technology policy evaluation beyond impact measurement—The case of Germany. *Evaluation* 4(2): 130–148.

Landry R, Amara N and Lamari M (2001) Utilization of social science research knowledge in Canada. *Research Policy* 30(2): 333–349.

Landry R, Lamari M and Amara N (2003) The extent and determinants of the utilization of university research in government agencies. *Public Administration Review* 63(2): 192–205.

Lepori B and Reale E (2012) S&T indicators as a tool for formative evaluation of research programs. *Evaluation* 18(4): 451–465.

Levitt R, Celia C, Diepeveen S, et al. (2010) *Assessing the Impact of Arts and Humanities Research at the University of Cambridge*. Cambridge: RAND Corporation.

Luiz de Campos A (2010) *Economic Impact Assessment within the Research Councils. Report to RCUK Strategy Unit and Performance Evaluation Group*. Brighton: CENTRIM.

Majone G (1989) *Evidence, Argument, and Persuasion in the Policy Process*. New Haven: Yale University Press.

Marjanovic S, Hanney S and Wooding S (2009) *A Historical Reflection on Research Evaluation Studies, Their Current Themes and Challenges*. Santa Monica: RAND Corporation.

Martin B (2011) The research excellence framework and the impact agenda: Are we creating a Frankenstein Monster? *Research Evaluation* 20(3): 247–254.

Martin BR (2004) The changing social contract for science and the evolution of the university. In: Geuna A, Salter AJ, Steinmueller WE (eds) *Science and Innovation: Rethinking the Rationales for Funding and Governance*. Edward Elgar: Cheltenham, pp.7–29.

McGuire RH (2008) *Archaeology as Political Action*. Berkeley: University of California Press.

Molas-Gallart J (2012) Research governance and the role of evaluation: A comparative study. *American Journal of Evaluation* 33(4): 577–592.

Molas-Gallart J and Tang P (2011) Tracing 'Productive Interactions' to identify social impacts: An example for the social sciences. *Research Evaluation* 20(3): 219–226.

Nussbaum MC (2012) *Not for Profit: Why Democracy Needs the Humanities*. Princeton: Princeton University Press.

Olmos-Peñuela J, Molas-Gallart J and Castro-Martínez E (2013) Informal collaborations between social sciences and humanities researchers and non-academic partners. *Science and Public Policy*. doi: 10.1093/scipol/sct075.

Pawson R and Tilley N (1997) *Realistic Evaluation*. London: Sage Publications.

Pielke Jr R (2012) Basic research as a political symbol. *Minerva* 50(3): 339–361.

Radosevich R and Kassicieh S (1994) Participant roles in public-sector technology commercialization: Introduction. In: Kassicieh S, Radosevich R (eds) *From Lab to Market. Commercialization of Public Sector Technology*. New York: Plenum Press, pp.125–135.

Ravetz J (1988) A new social contract for science. *Bulletin of Science, Technology & Society* 8(1): 20–30.

Rip A (2003) Societal challenges for R&D evaluation. In: Shapira P, Kuhlmann S (eds) *Learning from Science and Technology Evaluation. Experiences from the United States and Europe*. Cheltenham: Edward Elgar, pp.32–53.

Schiller D (2010) Institutions and practice in cross-sector research collaboration: Conceptual considerations with empirical illustrations from the German Science Sector. *Science and Public Policy* 38(2): 109–121.

Scriven M (1967) The methodology of evaluation. In: Tyler RW, Gagné RM, Scriven M (eds) *Perspectives of Curriculum Evaluation*, Vol. 1. Chicago: Rand McNally, pp.39–83.

Silvani A, Sirilli G and Tuzi F (2005) R&D evaluation in Italy: More needs to be done. *Research Evaluation* 14(3): 207–215.

Spaapen J and Dijstelbloem H (2005) *Evaluating Research in Context*. Amsterdam: Consultative Committee of Sector Councils for Research and Development (COS).

Spaapen J and van Drooge L (2011) Productive interactions in the assessment of social impact of research. *Research Evaluation* 20(3): 211–218.

Stirling A (2008) Opening up and closing down: Power participation and pluralism in the social appraisal of technology. *Science and Human Values* 32(2): 262–294.

van Langenhove L (2011) Social sciences and policy impact: The case for a participatory approach. In: Papanagnou G (ed.) *Social Science and Policy Challenges. Democracy, Values and Capacities*. Paris: OECD, pp.95–112.

Weiss CB (1980) Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilisation* 1(3): 381–404.

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

## Author biography

**Jordi Molas-Gallart** is an economist with more than 20 years of experience as an analyst of science, technology and innovation policies. He is Research Professor and Deputy Director at INGENIO, a research institute of the Spanish National Research Council (CSIC) and the Polytechnic University of Valencia. Before joining INGENIO, Jordi worked for 13 years at SPRU, University of Sussex as Research Fellow and Senior Research Fellow. His research interests include science and technology policy evaluation and impact assessment, and university–industry relations. He has led and contributed to many evaluation studies for a variety of clients, including the UK Economic and Social Research Council, the European Commission, INSERM, CSIC, Queen Mary College, the Russell Group of Universities, and several Spanish regional governments. He has been a member of the European Commission 'Lisbon Expert Group' for the follow-up of the research aspects of the revised Lisbon strategy. He is the author of one book and of more than 80 articles, book chapters, monographs and reports, and co-editor of *Research Evaluation*.