

Report 1. State of the art in the scientific, policy and social impact of SSH research and its evaluation

Project acronym: **IMPACT-EV**

Project title: **Evaluating the impact and outcomes of European SSH research**

Grant Agreement number: 613202

Coordinator: Universitat de Barcelona

Funding Scheme: Collaborative Project

Due date of deliverable: Month 6

Actual submission date: Month 8

Start date of the project: 01/01/2014

Project duration: 48 months

Work package: 1

Task(s): 1 Deliverable: D1.1

Dissemination level: Public

Lead beneficiary for this deliverable: Consiglio Nazionale delle Ricerche. Emanuela Reale

Authors:

Emanuela Reale, Emilia Primeri (CNR-CERIS)

Ramon Flecha, Marta Soler, Esther Oliver, Lúdia Puigvert (UB-CREA), Teresa Sordé (UAB)

Andràs Schubert, Sàndor Soòs, Judith Mosoni-Fried (MAGYAR TUDOMANYOS AK)

Kubra Canhilal, Benedetto Lepori (USI)

Dragana Avramov (PSPC)

Paul Holm, Charles Travis (TCD), Charles Larkin (UCardiff)

Andrea Scharnhorst, Arjan Hogenaar, René Van Horik (KNAW-DANS)

Claire Donovan (BU)



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration (FP7/2014-2017) under grant agreement n° 613202.

Table of contents

1. Introduction	2
2. Methods	3
3. What do we know about the Scientific impact of SSH research?	6
3.1. Advances in measuring and assessing scientific impact	6
3.2. Scientific impact of SSH research.....	19
4. What do we know about the Political impact of SSH research?	20
4.1. Advances in Political impact assessment	20
4.2. Political impact of SSH Research	24
5. What do we know about the Social impact of SSH research?.....	27
5.1. Social impact assessment	28
5.2. Social impact of SSH research	35
6. What do we know about how SSH research has contributed to the ERA?	44
6.1. SSH research on Strengthening the ERA	44
6.2 Evaluation of SSH research impact on strengthening the ERA.....	46
7. Conclusions.....	48
8. References.....	51
ANNEX I.....	65
ANNEX II.....	71

1. Introduction

The project IMPACT-EV aims to develop a permanent system for selecting, monitoring, evaluating and comparing the impact and outcomes of European social science and humanities (SSH) research, taking into account the latest quantitative and qualitative evaluation tools, identifying new ways to implement them and exploring new standards and indicators to complement existing impact assessment processes. This report presents a review of the existing scientific literature on research evaluation procedures and the changing structure of scientific, social and political impacts of SSH research. This is a starting point for further developments and will be complemented by a comparative analysis of a number of selected SSH evaluation systems and an impact assessment of SSH European projects. Beyond that, the project will explore in depth the scientific, social, and political impacts, and the contributions of SSH research to the European Research Area (ERA), in several identified examples. It will integrate this knowledge, thus defining comprehensive indicators and processes for evaluation.

The goal of IMPACT-EV has special relevance in the current context where we face increased accountability of public funding agencies and an increase in the volume of scientific outputs worldwide. The demands come also from below, from social sectors that contribute to the discussions and arguments concerning the values and goals of science and politics. Scientific practice is of more public interest than ever due to its sometimes unexpected consequences, but also because research politics is conditioned by risks and social pressure, and also shaped by social movements and sub-politics (Beck, 1992). In this context, a dialogic process of research including end-users and different stakeholders is increasingly seen as a requirement for our society. Identifying scientific contributions and research methodologies that bridge the existing gaps between science and society can help to address this challenge. These changes are affecting traditional research programme evaluation not only in the field of Life and Natural Sciences but also in Social Sciences and Humanities. Traditional research evaluation schemes are not disappearing but rather being reconsidered and reformed (Arnold, 2004). Therefore, IMPACT-EV faces a momentous task mapping the impact of SSH research and to provide tools to measure this impact and to promote this impact in its diverse forms.

This report is structured in six parts. The first part summarises the methods that we have used in our literature review and analysis. The following sections focus on each of the four types of impact that are covered by the IMPACT-EV project: scientific, political, social, and strengthening the ERA. Each section is divided into two subsections, the first part focusing on what the scientific literature says about these different types of impact assessments, and the second identifying examples of SSH research that has achieved these different types of impact. These examples are not part of an exhaustive search and analysis. The purpose of the report is to show that when reviewing the literature of SSH research projects it is possible to find research projects that have scientific, political and social impact. This initial exploration is of the starting point for the wider IMPACT-EV project.

2. Methods

The present report is based on an extensive literature review and documental analysis of research outcomes.

Data collection

In order to facilitate this report, a comprehensive scientific literature review has been conducted. The consortium has reviewed the following sources:

- a) Books, reports, working papers, etc.
- b) Guidelines for applicants and evaluators, with searches of EC databases of funded projects.
- c) Scientific articles: the literature search has been conducted mainly using the Web of Science (WoS) and SCOPUS databases
- d) CORDIS database: exploration of EU FP6 and EU FP7 projects (period 2006 to 2012).
- e) The FP7 Flash-it project¹ was given priority as a source for relevant research reports.
- f) Web consultation: for instance, Science Europe Association and other investigation centres and institutes from around the world.
- g) Grey literature from relevant evaluation institutions has also been reviewed.

A snowball strategy has been followed in order to identify further sources. For instance, in those cases where specific projects were mentioned, the search for information has been extended to include project reports or other available online data.

The review of these sources has been conducted in the following disciplines agreed upon by the Consortium: Sociology and social-economic geography; Educational Sciences and Media and Communication; Humanities; Life Sciences; Economics and Business; Law; Political Science; Psychology, among others.

To conduct the search, numerous combinations of keywords were used in order to detect the impact of research, and to assess the influence of systems that evaluate the impact of SSH research. These combinations included terms such as: research, research system, evaluation, assessment, monitoring, indicators, impact, social impact, scientific impact, political impact, positive, successful, projects, among many others.

The analysis covers the period 2006-2012, that of EU FP6 and EU FP7, although some references to important pieces of literature published before 2006 are presented in the text when they were useful to understanding the evolution of the concept of, and approaches to, evaluating the impact of SSH research.

¹ "Flash-it - Facilitating Access to Socio-economic Research through Information and Communications Technologies" is a project funded under EU FP7 (Grant Agreement: 290431), that seeks, among others to standardize, analyse, synthesize and disseminate the results of investigations in the SSH area by means of creating a network, and the technological tools necessary to facilitate this.

Data Analysis

The researchers in the consortium shared a common analytical framework to analyse documents, which included exclusionary and transformative dimensions according to the approach of the communicative methodology of research (Gómez, Puigvert & Flecha, 2011). The exclusionary dimension is that which identifies the elements and factors that lead to a lack of research impact (scientific, political and social). The transformative dimension identifies actions that have promoted the success of the impact of the research.

The analysis has been performed in a cooperative manner by the members of the Consortium, through feedback and direct contribution to its elaboration. It was developed by filling dedicated grids for the analysis of research impact and system impact, aimed at discovering the elements that it was necessary to outline for the purposes of the IMPACT-EV project.²

We now present a few key figures about the content of the work. The literature surveyed includes 172 Journals covering the whole range of SSH, including also several journals from other interesting fields; 272 grids, each related to one piece of literature, have been surveyed in total.³ The literature is not concentrated in specific journals, with the exception of *Research Evaluation*, where more than 40 articles of this review are included. Other journals where we find a large concentration of papers are *American Psychologist* (10) and *Scientometrics* (7).

Articles in journals are the most important output: more than 90% of the pieces analysed belong to this type of publication. This result is partly due to the methods adopted for the review, partly depends on the type of argument presented, and is also due to the fact that the policy interest in the issue raised has only been relatively recent. This is also confirmed by the fact that 233 out of 272 outputs have been published from 2009 to 2012.

Nonetheless, the review also includes some “hidden” pieces of literature, in the form of FP6 or FP7 project documents. In this case, the reader is invited to explore the projects in order to discover the different outputs that are related to the impact assessment.

Figure 1 presents the distribution of the reviewed outputs by discipline; the distribution reflects the criteria adopted by the IMPACT-EV project for the literature review. All the fields are well represented in the literature; thus, the interest in the assessment of research impact is an issue whose importance is well perceived in all SSH areas, although the field of psychology has a number of papers surveyed that accounts for about 1/3 of the total number.

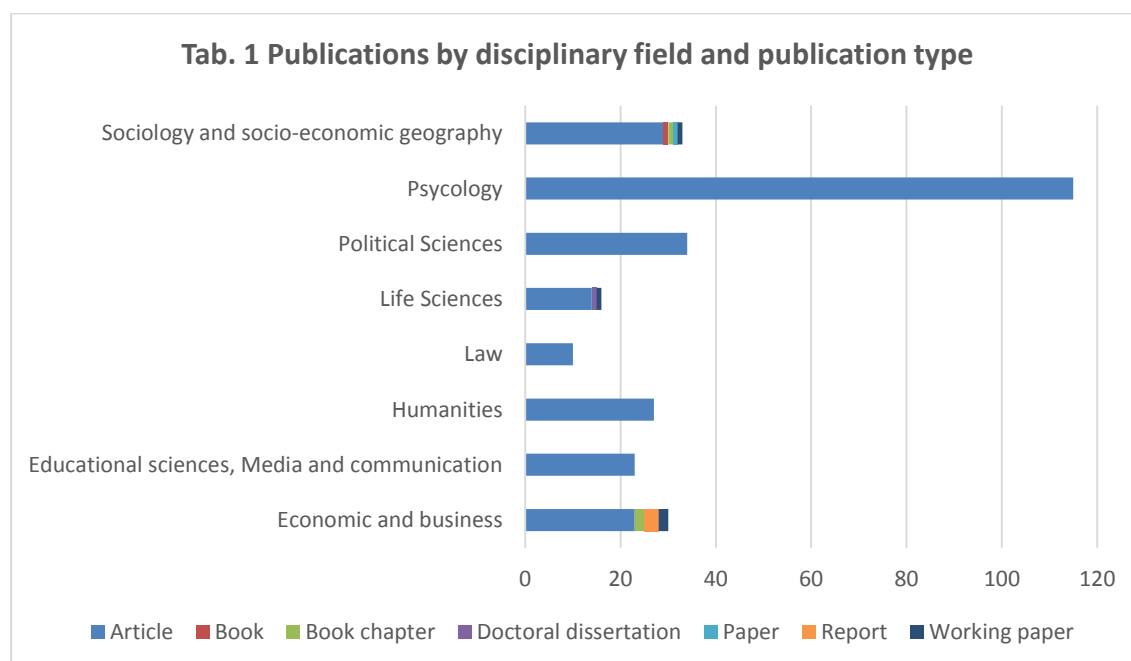
This is due to the fact that the approach followed for the field of psychology has been different from that used for the other fields; in this case the work followed a ‘computer-aided literature review’ technique, which is useful when you have to explore both a large number of papers and topics that might cover different disciplinary fields. Annex 1 presents one example of this technique, which outlines clusters of themes addressed within the field, and the corresponding keyword profiles. This investigation allows us to understand the level of

² IMPACT-EV, WP1 Guidelines, available for internal use on the project website <http://www.impact-ev.eu>

³ The grids are part of the documentation of the IMPACT-EV Project, and are available for internal use on the project website <http://www.impact-ev.eu>

interdisciplinarity of the fields, and the type of problems that have more impact-oriented activities, and can be useful when targeted research within the Europa 2020 strategy is concerned. Research programmes that want to address societal and political challenges are generally designed as multidisciplinary efforts; the assessment of their impact would need to be grounded on a literature review, which extensively explore several different disciplinary areas.

The exercise is based on a relatively large scale Web of Science record in the area of Psychology, where the habit of publishing in WoS journals is largely shared within the scholars' community. Other SSH areas do not have the same characteristics, thus the application of the mentioned technique is actually limited. Moreover, as far as the content of this Report is concerned, the high number of psychology papers does not imply that the field has more prominent results on impact assessment than others, rather that the issue of impact is often mentioned in the papers.



Limitations

The main limitation of this review is related to the ability to trace research outputs that are neither publicly available in repositories nor have been cited by other publications thus allowing them to be identified. This is mainly the case for books and book chapters in the area of Humanities and Law.

A second limitation, strongly linked to the former, is that literature written in national languages has been included only where those pieces of work have received citations in relevant non-national (i.e. largely English-language) outputs, thus reducing the chance of inclusion.

Despite these shortcomings, the review includes a wide range of journals, including those related to the humanities, and so we estimate that these limitations do not impact on the reliability and completeness of the picture provided.

3. What do we know about the Scientific impact of SSH research?

3.1. Advances in measuring and assessing scientific impact

History and development of scientometrics

The scientific impact of research has been on the agenda of science policy, society at large and the sciences themselves at least since the emergence of Big Science (Price 1963, p. 86) after WWII. The growth of science, the need to monitor (public) spending, and more recently the complete transfer to a knowledge-based economy for most of the highly industrialized countries led to the growth of a specialised scientific discipline in SSH, namely bibliometrics, scientometrics and informetrics (De Bellis, 2009). In this regard, one of the factors that triggered the growth of this field was an innovation in the information services, journal indexing (Garfield, 1962, 1979). For the first time, alongside journal abstracting services, one could search the reference lists of journal articles. Eugene Garfield's Institute for Scientific Information in Philadelphia, USA whose services are now continued under the trademark Web of Knowledge by Thomson Reuters (<http://thomsonreuters.com/thomson-reuters-web-of-science/>) was for many decades the focal point in the discussion about "How to measure the impact of science?". Testimony of those debates – which also from the beginning included the Social Sciences, Humanities and Arts, can be found in the Essays of an Information Scientist (Garfield, 1962-1993) – a 15 volume book series whose full texts are searchable online (<http://www.garfield.library.upenn.edu/essays.html>). Among those debates, the Impact Factor and other citation-based methods stand out (Garfield, 2006) because of its prominent use for the research assessment.

Impact of scientific research is often understood as impact on science, and impact on science has been traditionally understood as being measured in terms of the Impact factor. Since the very beginning there has been an extended debate about to what extent "impact" measured in terms of "being cited" says something about the "quality" or "importance" of a journal paper. Science history is full of stories about unexpected inventions and that innovation and new ideas almost always emerge at the boundaries of scientific fields, often carried out by outsiders or rather eccentric, atypical researchers (Joerges & Shinn, 2001). Connected to these processes, Merton (1968) reflected about the so-called *Matthew effect of science*, based on the observation that some very well-know persons in academia seem to receive more credit for their work than others (Merton, 1968). What seems to be an injustice on the individual level, is functional for the level of academia as a system. Such *light towers*, meaning individuals who act as placeholders for a certain idea, method, or direction, allow a faster and more effective navigation through science (Merton, 1968).

For citations and the journal impact factor in particular one can show that there is a difference between the expectation of being cited when publishing in a specific scientific journal with a certain impact factor and the actual citation rate that each individual article receives. For each journal one finds losing and winning papers and actors, compared with the average citation rate and the journal impact factor. Aggregated on the level of countries

one can even observe a systematic deviation leading to a small group of countries' papers which are more cited than expected, and a majority of countries' 'losing' citations when compared with the expectation (Bonitz, Bruckner & Scharnhorst, 1997). Observations of this kind add another critical dimension to the use of the impact factor and related measures for scientific impact (Glanzel & Moed, 2002). Still, one-dimensional measures such as the impact factor – one number, easy to handle and to apply – are preferred by many bodies dealing with evaluations. Therefore, in scientometrics there is a profound debate about the validity of different indicators, and the continuous development of new indicators. One example is SJR (Scimago Journal Rank), a citation impact index taking into account a journal's prestige where the authors' articles are cited (González-Pereira, Guerrero-Bote & Moya-Arnégón, 2010).

Before we review further the changes in output-based indicators, let us make one remark about another class of indicators. For decades scientometrics has focussed on indicators based on the output of scholarly communication – publications. One reason for this bias can be found in the availability of standardised databases, such as the Science Citation Index (SCI). One should, though, not forget that the primary goal for the SCI was to improve information seeking processes, and that its use for evaluation and assessment is a later addition. But next to metrics based on publications, quantitative science studies or scientometrics also use the system of Science and Technology Indicators, among which equally important, so-called input indicators are monitored. Examples are the number of PhD students, the number of R&D staff, and R&D expenditure (Godin, 2004).

When the systematic study of the sciences started, for instance marked by the foundation of the Society for Social Studies of Science in the US, quantitative and qualitative studies of the academic system were more entangled with each other, than can be observed today (Elkana, 1978). Branching off in different ways to analyse the science system is one of the stumbling blocks in the current impact debate, which is often conducted from the rather narrow viewpoint of quantitative measurements.

In general, one can state that parallel with changes in scholarly communication (Habermas, 1985) we observe the emergence of proposals for new indicators (Borgman 1990; 2007). Not all of these emerged out of the field of scientometrics, but to date the scientometrics community has effectively engaged in all these debates.

One way to order these developments in scientometrics is according to a timeline of first appearance:

- The use of the web in scholarly communication: Webindicators, webometrics, cybermetrics (see Scharnhorst, Wouters & van den Besselaar., 2006)
- Web 2.0 – user generated content and the emergence of altmetrics (see Priem, Taraborelli, Groth, & Neylon, 2010; Bornmann 2014)
- Semantic web – automatically generated impact stories (see <https://impactstory.org/>)

Use of the web in scholarly communication

One pioneer in the area of (gold) open access which also explores altmetrics is the Public Library of Science (PLOS). PLOS explores tools to track the post-publication reception of any research (Fenner, 2014). This process has also been promoted by scholars in the field of Social Sciences and the Humanities generating Open Access initiatives such as the Public Knowledge Project (PKP) (MacGregor, Stranack & Willinsky, 2014). In addition, some

scientists, such as the PLoS Medicine Editors (2006), stated that a journal's impact factor provides limited information on the ways a given article is being read and discussed inside and outside the scientific field as well as about its political impact.

In the field of SSH the Impact Factor only takes into account journal article citations. It does not capture article citations in books, and neither does it capture book citations in scientific articles. This has been a long-standing issue in traditional citation studies (Garfield, 1980; Leydesdorff, Hammarfelt & Salah, 2011). The different nature of scholarly communication in some areas of SSH, relying on books instead of journals, has eventually led to the emergence of a Book citation index and first experiments with this (Torres-Salinas, Rodríguez-Sánchez, Robinson-García, Fdez-Valdivia & García, 2013). Some scholars from the field of Social Sciences have recommended the use of alternative statistics, such as those derived from Google Scholar in order to capture citations that appear in both articles and books (Jacobs, 2011). Analysis of specific journals such as Jacobs' study of *Gender and Society* suggest that some journals have a far more influential impact than that displayed by ISI journal impact factor (Jacobs, 2009). Articles from the journal *Gender and Society* were more cited in books, dissertations, book chapters, proceedings and other reports. Although these types of statistics are not comprehensive tools, when they complement each other, they can provide a closer approximation of the scientific impact of a research project. Beyond the differences in scholarly communication and the difference between expected and observed citation rates, there has been also an argument that all those global information spaces (e.g. Web of Science or Scopus) might be insufficient because of some bias in the coverage. A more precise and almost complete coverage of output might lead to quite different insights in the performance and the impact of SSH research (Sivertsen & Larsen, 2012). Additionally, the different cultures of scholarly communication, and particularly the role of changing information practices in the humanities which Bulger and others (2011) analysed, have an influence on measuring scientific impact.

New approaches in scientometrics

In recent years one can observe a trend both in scientometric analysis as well as in scientific impact studies towards calculating the impact of individual authors. For many years – also due to availability of bibliographic databases – measurement of scientific impact focused on publications (Scharnhorst & Garfield, 2010). Increasingly we now observe a tendency to focus on authors as the unit of analysis (Wouters & Costas, 2012). A new indicator that has gained a lot of attention is the h-index, proposed by Hirsch 2005. “The h-index is a measure of a combination of productivity and citation impact. It is calculated by ordering the number of publications by a particular researcher on the basis of the total number of citations they have received. For example, someone who has an h-index of 40 has published at least 40 articles that have each been cited at least 40 times. Moreover, the remaining articles have not been cited more than 40 times each. The higher the h-index the better”.⁴ Hirsch proposed the h-index through which the impact of a researcher could be calculated without using time limitation – such as two years in the case of the Garfield-Sher Impact factor or the three years in the case of the SJR (Hirsch, 2005). Low-cited articles do not affect h-type index and h-type based indexes negatively. Some authors highlighted the fact that remaining time limitation is a better indicator for SSH where citation dynamics are more extended in time than in the Natural and Life Sciences (Jacobs, 2011). At the same time, it also takes into account book citations. Several improvements to the h-index have been developed in recent

⁴ Quotation from the blog post from Paul Wouters “Still using the Hirsch index? Don't!", web resource <http://citationculture.wordpress.com/2011/09/26/still-using-the-hirsch-index-dont/>

years. For instance, Leo Egghe proposed a new index that maintained the properties of *h*-index but taking also into account the citation of an author's highest cited papers, the *g*-index (Egghe, 2006). Braun, Glänzel and Schubert (2006) proposed to take Hirsch-type indexes as a useful complement to journal impact factors and for evaluating research scientific impact. However, as can be read in the blog post cited above, there are also critical remarks about the *h*-index. Moreover, impact studies on the level of individual authors need to be processed with special care.⁵

Other scholars, despite recognising the advances made still argue that citation counting through mechanisms such as the Impact Factor or *h*-type indexes is useful, but not sufficient. It can take years to know the real scientific impact of research especially in the Social Sciences and Humanities (Priem et al., 2010). Thus, the speed of ALTmetrics and similar tools can create real-time inputs about how an article or a research report is being used, bookmarked, shared and cited on the Web, and the analysis of these trends could be useful in creating standards adapted to the specificities of SSH. These webometric indicators, such as article usage data (HTML views and PDF downloads), should also be taken into account in the research evaluation process.

In the last five years, with the consolidation of web 2.0 and social and scientific media networks, new tools are required to observe the specific scientific impact that a piece of research has on its inner- and outer- scientific networks. Recent technological developments promoted by the open access and ALTmetrics intellectual and scientific movements are making visible the number of times that a research article is being viewed, downloaded, bookmarked, shared, commented upon, inserted into a researcher Blog, or tweeted. ALTmetrics tools and software (such as Article Level Metrics or Total Impact) have been promoted especially by scholars and publishing groups from the Life Sciences, such as the Public Library of Science (PLOS). The most recent tools are showing great advances that could be helpful for evaluating and monitoring the scientific impact of SSH research projects. For instance, researchers and funding agencies can measure the impact and outreach not only of an individual researcher, but also of a research project, because the unit of analysis is at the article level rather than that of the individual scholar. Article(s)' performance can be benchmarked against others, and funding agencies can conduct custom searches that account for research impact or gauge the value of any article with post-publication peer review and discussions. Funding agencies and scholars can also track tagging and cut/paste activities from their scientific articles, thus they can observe the frequencies and places in which the research has been used (including media reports, online newspapers and other sources). At the same time, several authors have stated the need to examine the role of science blogs as a tool in disseminating research outcomes published in Open Access journals or in journals behind paywalls. These research blogs serve as informal post-publication peer-review boards where peers (scholars from the same or other disciplines) as well as citizens can gain access to the knowledge, comments and open debates being held. Shema, Bar-Ilan and Thelwall (2012) observed that Life Sciences blogs were the most common and wide-spread type, while the Social Sciences and Humanities were the least represented in their sample.

The use of rankings in assessing scientific performance

Rankings are a further tool not explicitly devoted to assessing research impact, but used to some extent to assess research excellence of organizations. In this area, there are two key

⁵ see <http://citationculture.wordpress.com/2013/07/29/bibliometrics-of-individual-researchers/>

initiatives flagged in the literature. The first ranking mentioned by the literature is the first 'Economists Top 40', published in *Economisch Statistische Berichten* (ESB) in 1980 by two eminent Dutch economists, Tom Wansbeek and Arie Kapteyn. This grassroots ranking of economic researchers in the Netherlands, based on quantity and quality of research outputs, had the ambition of improving research output in international journals and raising standards of research performance on a national level. As a consequence of this ranking, Dutch economists felt they were engaged in a ranking competition, and the results were clear: from 1980, when the Top 40 started, only 74 economists in the Netherlands had produced at least one article in the source journals of the Top 40. This number increased to 305 by 1987, partially through an extension of the list of source journals. At the individual level, in the early 1980s, one 25-page long article in a journal with the lowest weight was enough to be among the Top 40 most highly ranked economists; in 1994–1998, however, more than 7 'standard' articles were needed for a Top 40 ranking, which increased to more than 8 standard articles in 1995–1999. The increase in the production and quality of the articles is clear (Nederhof, 2008).

The second ranking considered here is that of the academic impact of world national research institutes, a project initiated by the Chinese Academy of Sciences and conducted from 2006 to 2009. This project adopted Soft System Methodology (SSM) and 3E theory (efficacy, efficiency, and effectiveness) to design its theoretical framework and indicator system. Its indicator system included two independent systems of Efficacy (E1) and Effectiveness (E3), and a comprehensive system combining the two. In this project, the indicator system further introduced a hierarchical structure. This structure aimed to create a robust ranking, so that it did not only reflect the importance of different scientific research qualities, but also exhibited research outputs at various levels to balance the data randomness of different level indicators (Xu, Li, Meng, Liu & Mingers., 2013). To sum up, the first ranking proved that it could increase the scientific production of research in Economics, whereas the second ranking could lead to similar competitiveness between different national research institutes, at the same time that it enhanced the impact of the research within these national institutes.

More recently the U-Multirank tool for higher education institutions uses a multidimensional and multi-level user-driven approach, aimed at ranking institutions' performance across a wide range of higher education missions. It provides HEIs performance profiles at two levels: for the institution as a whole, and at the level of different disciplinary fields, enabling us to compare the institutions in terms of the activities they are engaged in, improving the comparative method used by similar cases (Van Vught & Ziegele, 2012). As an alternative ranking method, the use of web presence has been proposed by the Cybermetrics Lab (Spanish National Research Council, CSIC). The so-called Ranking Web or Webometrics is the largest academic ranking of Higher Education Institutions. Since 2004 and every six months an independent, objective, free, open scientific exercise is performed by the Lab for the providing reliable, multidimensional, updated and useful information about the performance of universities from all over the world based on their web presence and impact (Aguillo, Bar-Ilan, Levene & Ortega, 2010; Aguillo, Granadino, Ortega & Prieto, 2006).

The following table summarises key issues regarding measurement of scientific impact.

Data	What is measured crucially depends on the available data and data collections. Databases, as well as any machine readable information on the web, will influence what will be measured and reported.
------	--

Metrics	Next to data, indicators derive from metrics and statistics, and often are built on implicit assumptions about the size of the data ensemble and independence or dependence of events. For the application of such metrics insights into their mathematical foundation is needed.
Unit of analysis	Scientific impact has been measured on the level of countries, institutions, and individuals, often relying on journals or groups of journals indexed in citation databases. The error margin of indicators increases when moving from large to small ensembles of data points.
Scholarly culture	Scientific communication functions differently in different fields. They have different norms, including norms of publishing, different venues for output and different regimes of peer review. In scientometrics this has been captured in the debate about field-normalization of indicators. What remains is to pay attention to this effect, in particular when comparing performance across fields and for interdisciplinary work.
Science as an adaptive system	A measurement of impact is always an intervention into the dynamics of a social system, this also holds for academia. If a certain metric will be implemented one can be sure that the adaptive agents of the social system of science will react to this. An example is the increase of publications parallel with a decrease of content per publication as a result of publication number indicators.

Existing debates on how to develop the evaluation of scientific impact

Debates about the optimum method of assessing the scientific impact (or quality) of SSH research are framed by the suitability, or otherwise, of adopting evaluation techniques used by the so-called 'hard' sciences, most notably bibliometric indicators. The discussion ranges from the idea that SSH disciplines are less scientifically developed, and that as they mature the existing metrics will become a better fit, to the view that the scientific impact of SSH research cannot be captured by blunt metrics and can only truly be assessed by peer review (Donovan, 2007a). Most discussion occupies a middle ground, and also seeks alternative bibliometric techniques more suited to the production and consumption of SSH research. The use of Web of Science data, and the Social Sciences Citation Index, dominates the academic literature, as does the prospect of developing alternative 'non-standard' metrics. This literature tends to focus on testing the efficacy of various national research evaluation exercises, or applying novel measures in national or local contexts.

There is much discussion surrounding the issue of whether metrics can replace peer review in the assessment of the scientific impact of SSH research. Butler and McAllister (Butler, L. & McAllister, 2009, 2011) believed this to be the case, and attempted to model the outcomes of the 2001 UK Research Assessment Exercise (RAE) for political science using data that was available to the assessors (external earnings, percentage of staff submitted, student data) but swapping the peer review of four nominated publications per researcher with citation data. This was an innovative study as it included a novel (or 'non-standard') bibliometric technique that mined the Web of Science database to also capture citations made by indexed journal papers to books and chapters. They believed that this was 'far more

comprehensive than standard citation measures' and 'defensible for a discipline such as political science, where only a quarter of submitted [RAE] outputs were articles in ISI journals' (Butler & McAllister, 2009, p. 7). They concluded that for political science, their modelling produced results that closely matched the peer review results, and so the RAE panel process could be replaced by metrics.⁶

Butler and McAllister's 2009 paper was part of a journal symposium, and its methodology was subject to several criticisms. Donovan argued that there were major flaws with the inclusion and analysis of citations to books and chapters. Data was included only for the first listed authors of books and chapters (even where second authors returned this output in their department's RAE submission), wrong editions of texts were counted, and the citation window extended at least five years beyond the RAE assessment period, so their modelling did not compare like with like (Butler & McAllister, 2009, p. 77). Russell argued that 'Given the gaps in the CI [citation index] system – and the extra five years of citation data used (notwithstanding their improvement on the ISI data) it is still a considerable leap of faith to assume that metrics have captured something that the review missed, rather than vice versa' (Russell, 2009, p. 68).

As a separate undertaking, McKay attempted to use metrics to model the outcome of the 2008 RAE for social work, social policy and administration, but in contrast to Butler and McAllister concluded that 'objective data may be useful in assessing the quality of research environments, but cannot replicate the human decisions made in determining the quality of research outputs' (McKay, 2012, p. 527).

Several papers focus on the efficacy of metrics-only evaluation systems applied to SSH research at the national level. For example, Schneider (2009) described the 'Norwegian model', a publications-based funding formula, which has been used annually since 2006 to distribute around 2% funding of basic research to the Norwegian HE sector. The exercise is based on a national research documentation system, which is used to provide annual counts of all types of publication covering all research fields (e.g. papers in journals indexed and not indexed by Web of Science, books, and book chapters). All Norwegian publications are registered, and the bibliometric data is validated and then stored in a standardised format, so that 'Norway currently has one of the richest bibliometric databases used for bibliometric purposes' (Schneider, 2009, p. 370). In order to encourage and reward high quality publications, higher weightings are given to the most 'prestigious' publications following field-specific publication norms. Schneider is optimistic about the treatment of SSH research, as the database includes 'complete annual research publication data for the social sciences and humanities', and so allows 'comparison of institutions based on all their research activity, and likewise direct comparison among different fields of research and among such institutions' (Schneider, 2009, 373-374). He maintained that 'based on the 4 years of experience with the model in Norway, it is reasonable to conclude that it is indeed possible to include all scholarly publications within all scientific fields in a bibliometric indicator' (Schneider, 2009, p. 374).

⁶ Butler and McAllister extended their analysis to include chemistry departments where 98% of publications submitted to the 2001 RAE were indexed in the Web of Science and where standard bibliometric analysis was not thought to be a problem, thus highlighting that 'no single model will apply across science and non-science disciplines. Any metrics approach to performance evaluation has to use a discipline-specific suite of indicators' (2011, 30). They therefore noted 'strong differences between HASS disciplines (represented in our analysis by political science) and STEM disciplines (represented by chemistry)' (2011, 55), but did caution that their novel citation technique was 'significantly more expensive and time-consuming than the traditional approach used for STEM disciplines' (2011, 55-56).

Other papers are confident about the possibility of using bibliometric resources in the future to assess the excellence of SSH research. Viola et al. (2010) provided a report focusing on the efforts and approaches made in order to evaluate scientific excellence in the specific EU Key Action for the social sciences and humanities (SSH) – ‘Improving the Human Potential’ (IHP) of the 5th Framework Programme (FP5). The analysis revealed important findings concerning publication behaviour in SSH: book publications are less dominant for SSH than anticipated, books and chapters were estimated to account for 25% of all publications. This may be the result of rapidly changing funding mechanisms and promotion decisions which are increasingly based on ‘scientific impact’ – available for journal articles but not other publication means. Also the citation analysis of this study suggests that this ‘ripening’ period is not too different from other fields of science. A three to four year citation window will impact the overall citation score of an article – but a longer period can negatively impact the scores per article as well as per author. In this respect the study seems to confirm that the longest possible citation window is therefore not the best option. In this study, the longest possible window was seven years and it was applied to publications from 2002. The authors consider that bibliometric analysis could be improved if the authors of FP research outputs, in order to identify journal articles (as well as other publication means) more systematically, were contractually bound to acknowledge the specific Key Action funding in their publications. The analysis could also be improved if curricula vitae were taken into account as CVs contain a full list of publications, and also provide information on authors' age and affiliation(s).

New metrics and mixed data

In the context of federally funded research in the USA, Largent and Lane (2012) outlined STAR METRICS, an ex post (and potentially an ex ante) metrics-based evaluation system developed by 19 federal science agencies and research institutions. It is a data platform that aggregates data on research investments from federal agencies and federally funded research institutions. There are two sets of analyses which (1) focus on scientific workforce data from research institutions' payroll and accounting systems; and (2) focus on the development of ‘open automated data infrastructure and tools that will enable the documentation and analysis of the inputs, outputs, and outcomes resulting from federal investments in science, in collaboration with science agencies and research institutions, without increasing administrative burden’ (Largent & Lane, 2012, p. 433). A prototype STAR METRICS R&D Dashboard describes NSF and National Institutes of Health R&D research investments by topic and geographic area; and describes various outputs associated with those investments in terms of publications, patents, and patent applications (Largent & Lane, 2012, p. 435). There are also data with the potential to relate to wider social or political impact, described in Section 4.1 below.

Also in the context of the USA, Haak et al. (2012) described eSPA (US National Institutes of Health electronic Scientific Portfolio Assistant), a web-based analytics system that links several scientific databases that provide data for research managers in terms of scientific productivity, quality, and dissemination⁷. They, however, cautioned that ‘No single indicator

⁷ The indicators used are: Productivity – publication count (number of published articles resulting from a project), data source: Medline; Quality – funding (annual and total US\$ amounts of funding received by a project): data source: IMPAC II; Quality – times cited (number of other articles that have cited a given project's published articles): data source: Web of Science; Quality – journal impact factor: data source: Journal Citation Report; Dissemination of knowledge – author count (number of named authors for published articles – indicates degree to which an investigator collaborates with others): data source: Medline; Dissemination of knowledge – bibliography

will provide a definitive assessment, nor can any single collection of indicators be equally applicable in evaluating all federal agencies, or, for that matter, all NIH institutes. Ideally, a reporting system would allow for the use of indicators specific to an area of research or germane to a funding agency's mission' (Haak et al., 2012, p. 473).

Kellow (2012) outlined the national 2011 Excellence in Research for Australia (ERA) exercise, and focused on the assessment of political science departments. Social science disciplines assessed the scientific impact or quality of publications using a combination of ERA journal rankings, bibliometric measures and (with the exception of psychology) peer review. Kellow concluded that 'The ERA largely rewarded large budgets and failed to adjust for size. Its results therefore inevitably conflate size with quality to some extent' (Kellow, 2012, p. 576). He argued that the relative ranking of political science departments was 'broadly credible'. However, using the example of astronomy, he maintained that the 'relativities between disciplines ... are highly suspect' (Kellow, 2012, p. 579). The result was that the scientific impact or quality of political science research was undervalued relative to other disciplines, and that 'this risks affecting the perceptions of policymakers and impacting funding for the discipline as a whole' (Kellow, 2012, p. 567). We should, however, note that ERA remains an experimental exercise, and that research funding to universities in Australia remains tied to the Institutional Grants Scheme, a metrics-based funding formula comprised of data on research income, postgraduate students, and publication productivity (journal papers, books, chapters and conference proceedings). In the Australian context, Butler analysed ISI data (Butler, 2002; 2003; 2004) and found that linking research funding to publication productivity, in the case of journal publications, led to a rapid increase in the number of publications, the highest number of which were in lower quality journals.

Pontille and Torny (2010) turn to focus on three different initiatives that produced journal rankings in the social sciences and humanities. They compared and contrasted the key features, and reception of, draft rankings for Excellence for Research in Australia (ERA) produced by the Australian Research Council (ARC); the 'initial lists' of the European Reference Index for the Humanities (ERIH) produced by the European Science Foundation (ESF); and a list created by the French Agency for Evaluation Research and Higher Education (AERES). They argued that journal rankings are viewed as an 'alternative tool' to ISI data, and are 'raise[d] ... to the level of an instrument of assessment which is adapted to SSH and which is an intermediary between peer review and diverse metrics' (Pontille & Torny, 2010, p. 348).⁸ They were concerned, however, that while these alternative journal rankings signalled the end of judging the quality of a journal paper by its ISI journal impact factor, the new lists could nonetheless be similarly reified. Also, when considering the wider social impact of SSH research, 'social actors, and public actors in particular, are not in the slightest bit interested in the publication outlet, being more directly concerned by the content of the articles and its transferability, the local and adapted nature of which is of greater importance'. It followed that 'other measures ... will need to be developed, such as the

count (number of other articles cited in bibliography of published article – can provide an indication of domain of inter-connectedness): data source – Web of Science. There are also indicators that relate to political impact, described in Section 4.1 below.

⁸ The three journal rankings sought to assess the scientific impact or quality of SSH research while including journals not indexed in ISI databases. Pontille and Torny observed that there were variations between the meanings (and thus implications) of the three rankings, most notably that a 'C' rating for AERES meant a publication had no value, for ERA this indicated a lower quality publication, and for ERIH no hierarchy was implied and the 'C' rating indicated 'European added value' (2010, 352-353). They argued that the key benefit of journal rankings were that the centrality of ISI data could be bypassed, and that this opened the way for further ranking exercises that could include other important SSH publication outputs (e.g. chapters and books).

“impact criteria” envisioned by the defunct RQF⁹ ... or, more generally, S&T indicators’ (Pontille & Torny, 2010, p. 357).

Alternative data sources for SSH assessment

One attraction of using metrics to assess the scientific impact or quality of SSH research is the idea that the process can save time and money in comparison with a peer review exercise. In this respect, the prospect of remote data collection is very appealing. Lepori & Probst (2009) used a novel data gathering and analysis technique to map a heterogeneous social science field (communication studies) in a culturally, socio-political and linguistically diverse country (Switzerland). They gathered data remotely via CVs from departmental websites including: research interests, disciplinary background, geographical background, and publications. CV publication data was supplemented by Web of Science data. They argued that this approach was ‘an example of how one can combine and exploit easily accessible information to get a map of a scientific field where bibliometric techniques cannot be readily used’ (Lepori & Probst, 2009, p. 132). Lepori and Probst reached four key conclusions: (1) that ‘it is to a large extent possible to retrieve this information directly from institutional websites, without requesting CVs from people or resorting to other sources’; (2) that combining data sources to analyse publication activities (CVs, Web of Science data) allows data triangulation, thus ‘overcoming some of the limitations of [online CVs as a] data source’; (3) while some online CVs had low levels of detail, this did not impede basic data analysis; and (4) ‘at least for social sciences and humanities, it makes more sense to map a field by combining different points of view, namely from the institutional definition of the field, people and their education, publication activities and, finally, expert evaluation and the results of national evaluations, than just to rely on bibliometric analysis’ (Lepori & Probst, 2009, pp. 132-133). They concluded that this approach was best suited to small countries, or where a field of research was relatively small within a country, because ‘fine mapping using very detailed information is possible’ (Lepori & Probst, 2009, p. 133).

Another example of a metrics-only approach was the assessment of the value of universities’ publication databases for evaluation purposes (Reale et al., 2011), building positioning indicators to describe different profiles of university research activities, rather than their competitive position along the single dimensions of scientific production and academic reputation. Project results support evidence that institutional databases are social constructs, able to show a representation of the research performance of the universities, which is strongly affected by the interests of the different communities, influencing their development and evolution. Databases can also be valuable sources, when used in combination with international ones and with other information sources, to put together a broad picture of academic institutions and their scientific efforts and impact.

As we have already seen, an underlying concern in the literature on assessing the scientific impact or quality of SSH research is the extent to which metrics should be balanced with peer review, if indeed at all (see Donovan 2007a; 2007c). Butler and McAllister (2009) introduced the notion of ‘light touch’ peer review, although as Donovan (Donovan, 2009, p. 79) noted, what this constitutes is not clearly stated. She presents two definitions; (1) ‘One

⁹ The Research Quality Framework (RQF) was a national research evaluation exercise developed in Australia during 2005, and included a narrative case study approach to assessing the social, economic, environmental and cultural impact of research, which formed the basis of assessing research impact in the UK’s Research Excellence Framework. Before it could be implemented, the RQF was replaced by Excellence for Research in Australia (see Donovan, 2008).

option is to present expert panels with the results of a traditional peer-review exercise and a separate metrics-based assessment, where the task of the panel is to consider both sets of data in order to form a final quality judgement'; (2) 'Another is for metrics to be the dominant form of information which panels of peers then consider ... and actual review of "outputs" will only take place in rare cases where there are no appropriate quality metrics.' She cautions that within the bibliometrics community, best practice is to adopt the former approach, whilst the latter is frowned upon.

EURECIA developed a novel methodology for the study of the impact of research funding schemes on knowledge and its social conditions, and applied this to investigate the impact (effects) of the ERC and its funding schemes on science (EURECIA, 2012). The study constituted a departure from more traditional approaches in two ways: a) by interrogating the relationship between research funding and the science system rather than the economy and society at large; and b) by broadening the 'impact' question to include not only intended effects as read through the objectives but also other possibilities. To this end, bibliometric analyses have been integrated with data coming from documentary analysis and interviews in order to deepen understanding of the type of change the ERC funding produced.

Effects of metrics on research and on SSH research

Several authors pointed out negative consequences of metrics on the scientific quality of the research outputs. Frolich (2011) focused on the effects of the Norwegian performance-based funding (PBF) and quality reform (QR). Based on 2,000 responses to a survey of academics and teachers at Norwegian higher education institutions, she found that, 'According to the faculty members, the implementation of the QR has not had a significant impact on research Ninety per cent of faculty members report that the reform has not changed their publishing behaviour' (Frolich, 2011, p. 849).

Linková and Stöckelová (2012) described a complex points system operated by the Czech Republic's Research, Development and Innovation Council.¹⁰ Books were included but received 40 points relative to 500 points for a paper in Nature, Science, or the Proceedings of the National Academy of Sciences of the USA. They were negative about the impact of this scheme on the social sciences, and cited the report of an International Audit of Research, Development and Innovations in the Czech Republic (2011) which states that the research assessment system 'pays most attention to the wrong things Distorts behaviour, reduces the stability of the research system and hampers its performance.' Linková and Stöckelová concluded that 'No other system of performance-based research funding allocation known to us and still in use is equally radical in its exclusive focus on the past and its level of standardisation across different types of institutions and disciplines' (Linková & Stöckelová, 2012, p. 619).

An interesting strand within the SSH literature is the reflexive consideration of the potential effects of metrics or evaluation systems upon the future shape of SSH research. For example, there is a concern that research evaluation systems (and metrics in particular) are inherently biased as these have been constructed with the sciences in mind, and with SSH

¹⁰ Journal impact: $10 - 305 (10 + 295 \times \text{factor})$, where $\text{factor} = (1 - N) (1 + (N/0.057))$, where N is the standardised ranking of the journal for a given discipline; Journal impact: 500 (paper in Nature, Science, Proceedings of the National Academy of Sciences of the USA); Reviewed journal: 12/11/10/44; Book: 40; Patent: 500 (EU, US, Japanese patent); Patent: 40/200 (Czech or other national patent granted/used); Technology: 100 (includes applied technology, breeding variety, etc.) (Linková and Stöckelová, 2012: 264).

as an afterthought. Ouimet, Bedard and Gelineau (2011) applied various bibliometric indicators¹¹ to test if 'the h-index and some of its derivatives [are] discriminatory when applied to rank social scientists with different epistemological beliefs and methodological preferences' (Ouimet et al., 2011, p. 93). The study was set in Canada, and used the Publish or Perish (PoP) software, based on Google Scholar data, because it had greater coverage of non-English-speaking regions, and also captured citations to books: 'Indeed, using Web of Science or Scopus would have led to a dramatic underestimation of the true scientific productivity of the social scientists in our database, as many of them publish in French and produce book or book chapters that are rarely indexed in these databases' (Ouimet et al., 2011, p. 105). The study found that the h-index (and its derivatives) are biased against particular epistemological or methodological approaches so that 'on average, a quantitativist [sic] who is more prone towards positivism will have a larger h-index than a qualitativist [sic] who is more prone towards constructivism' (Ouimet et al., 2011, pp. 101-102) and that 'faculty members in Psychology ... tend to outperform social scientists in Anthropology, Sociology, Social Work and Political Science' (Ouimet et al., 2011, p. 99). Ouimet et al. stated that their study 'has demonstrated that fairly recent mainstream bibliometric indices such as the h-index discriminate against the analytical approaches employed by faculty members (reflexive, quantitative, qualitative, mixed), thus relativizing the relevance of using such indices to compare faculty members with different methodological preferences' (Ouimet et al., 2011, p. 104). They concluded that research funders, scientific managers and policymakers 'cannot use these indices to compare researchers from different academic disciplines' (Ouimet et al., 2011, p. 104).

Similarly, Donovan (2007b) conducted a forensic bibliometric examination of an ISI-based ranking of Australasian political science departments, supposedly sympathetic to the regional character of political science (multidisciplinary, historical, qualitative, humanistic) and found that its outcomes were skewed by data that favoured quantitative psychology and American research preoccupations. She found that quantitative social science research was more visible in ISI databases (due to more science-like journal-based publication practices and larger potential citing audiences), and that the use of this common metric led to an imagined hierarchy of science that may influence funding, hiring and promotion decisions (Donovan, 2007b, pp. 671-673). She concluded that standard ISI-based rankings reinforce a quantitative bias, and 'this apparently neutral metric is infused with judgements about the value of various types of social science, which privilege the quantitative, and either overlook or dismiss the interpretative' and in this respect 'political scientists are in danger of becoming the captives of badly chosen indicators' (Donovan, 2007b: p. 670). In this light, unquestioning acceptance of standard indicators may act to reshape social science away from regional traditions and concerns. In her critique of Butler and McAllister, Donovan (2009) concluded that 'Metrics have the potential to become technologies of governance ... and we should remain aware that while metrics may simplify assessment, they may also simplify the scope and aspirations of political science through privileging "positivistic" knowledge to the exclusion of interpretive knowledge' (Donovan, 2009, p. 80).

These concerns are also relevant to EU countries that are 'peripheral' in terms of a regional focus, and where social science includes not only quantitative research but also historical and qualitative traditions. Gantman (2012) took such regional considerations further and studied how the scientific productivity of 150 countries was affected by economic, linguistic,

¹¹ h-index; m-quotient; g-index; Individual h-index; Age-weighted citation rate; e-index; Contemporary h-index (ac).

and political factors.¹² He found that 'Scientific productivity on [*sic*] the social sciences obeys a different logic'. Like the 'exact sciences', financial resources were a central to countries' scientific productivity. However, unlike the exact sciences 'the government's degree of authoritarianism has a negative and significant effect on scientific productivity' (2012, p. 980). He also found that 'Linguistic imperialism is not a myth in the social sciences. Once other relevant factors are controlled for, the variable English as official language has a positive and significant effect on scientific productivity if publishing in international journals is important for the members of the scientific community, it is clear that social scientists from non-English speaking countries are at a disadvantage in this regard' (Gantman, 2012, p. 980).

Rafols et al. (2012) provide quantitative evidence on how the use of journal rankings can disadvantage interdisciplinary research in research evaluations. Using publication and citation data, they compare the degree of interdisciplinarity and the research performance of a number of Innovation Studies units with that of leading Business and Management schools in the UK. The study shows that: (i) Innovation Studies (IS) units are consistently more interdisciplinary in their research than Business and Management Schools (BMS); (ii) the top journals in the Association of Business Schools' rankings span a less diverse set of disciplines than lower-ranked journals; (iii) this results in a more favourable assessment of the performance of Business and Management schools, which are more disciplinary-focused. Results suggest that ABS journal rankings favour research within the dominant disciplines of BMS (mainly business, management, economics and finance) and disadvantage interdisciplinary IS units. Given the close correlation between RAE grades and assessments based on journal ranks in previous RAEs, this effect is large enough to have a substantial negative impact on the funding of IS units. The policy implications of these results are discussed in the light of studies on the consequences of biases in assessments. For example, research suggests that British economics departments have narrowed their recruitment to favour 'main-stream' economists, thus reducing the cognitive diversity of the research system's ecology. This may lead to intellectual impoverishment in the medium or long term.

However, to conclude this discussion, Lewis and Ross (2011) studied the perceived effects among 274 academics of the Australian Institutional Grants Scheme (IGS) / Excellence for Research in Australia (ERA); the New Zealand Performance-Based Research Fund (PBRF); and the UK's Research Assessment Exercise (RAE) / Research Excellence Framework (REF). The study was informed by the idea that simple metrics are perceived as being more cost-effective than peer review processes, but that metrics are less accurate when applied to SSH. Given their knowledge of the above literature, the authors were surprised to find that 'there was no clear pattern that indicated that science academics see the system as more beneficial than the humanities and social science academics, or as having more positive impacts' (Lewis & Ross, 2011, p. 393).

¹² The author used the SCOPUS database to analyse scientific productivity; and alongside science fields, included the social science categories of 'Economics, Econometrics, and Finance' with 563 journals (59 not published in English); and 'Sociology and Political Science' with 214 journals (33 not published in English).

3.2. Scientific impact of SSH research

The aim of this section is to highlight some examples of existing reports that analyze the scientific impact of SSH research projects. Although they are very few, some reports have carried out a detailed analysis of the scientific impact that Framework Programme projects had. For example, this is the case of the *Evaluation of the Impact of Framework Programme supported Social Sciences and Humanities Research* (Technopolis, 2010). This ex-post evaluation report shows how diverse research outcomes from SSH research projects are (45% grey literature, 28% journal articles, 19% book chapters and 8% books) and analyses their citation impact. When addressing citation, this report introduces new sources. Until that time, citation data was mainly retrieved from the Web of Science. But this report uses for the first time other sources to analyse citation data such as Google Scholar and creates an *h*-index for authors who have published results from the EU funded projects. Data shown in this report is more comprehensive for SSH than traditional citation analysis because it also captures citations to books, book chapters and grey literature (reports and working papers). In this sense, this report showed relevant data that can be used as benchmarks for assessment as the citation average for books as result of FP5 projects, being 35. In regard to journal article citation, the analysis also gives relevant information to take into account for gender. For example, it demonstrated that women were more cited in journal articles than men (5.2 versus 4.6 times) (Technopolis, 2010).

Among those outstanding SSH research projects that have an extended record of scientific publications we can find either R&D research projects or Networks of Excellence. For example, the network of excellence “RecWoWe - Reconciling Work and Welfare in Europe” (2009-2011) network gathers 15 academic books, 11 articles in peer-reviewed journals and 5 reports or working papers¹³ Another Network of Excellence with similar impact is “EQUALSOC – Economic Change, Quality of Life and Social Cohesion” (2006-2010).¹⁴ Its records are, with a high number of research outcomes being highly cited. In this case, team members published 15 academic books, 135 book chapters, 282 in peer-reviewed journals, and 7 working papers. Looking at outstanding R&D projects funded under EU FP7 some examples are: “MULTILINKS – How demographic changes shape intergenerational solidarity, well-being and social integration: A multilinks framework” (2009-2011) that published 38 articles in peer-reviewed journals¹⁵ and “MAFE: Migration between Africa and Europe” (2008 - 2012) having 19 articles in peer-reviewed journals, 10 book chapters and 5 edited volumes, and 34 working papers.¹⁶ These are only some from among many examples of SSH research projects that are already accounting for an important scientific impact.

¹³ Information retrieved from RecWoWe website on 29 June, 2014 <http://recwowe.vitamib.com/publications-1/books-and-issues/books-and-issues>

¹⁴ http://flash-it.eu/publications/doc_download/482-equalsoc-publications

¹⁵ <http://www.multilinks-project.eu/publications/>

¹⁶ <http://www.mafeproject.com/>

4. What do we know about the Political impact of SSH research?

4.1. Advances in Political impact assessment

In recent years, discussions about how Social Science research can have political impacts are present not only in social and academic forums but also as part of the political research agenda (Meagher, Lyall & Nutley, 2008; Lemay & Sá, 2012). In the literature, we find that the debate on how to assess the political impact of a research project draws on the possibility of identifying attribution (how could we identify that political impact - creation of a new norm or policy or its reform - is related to the findings of a research project) and time (when this political impact takes place). The literature in the field of research assessment has tended to avoid causal inference between research and policy change, arguing that policymakers and researchers tend to live in different worlds.

Relationships between science and policy

A body of literature has been dedicated to the study of the relationship between research and politics. On the one hand, Boaz and Ashby (2003) have pointed out the need for changes in traditional research assessment through creating mechanisms more able to identify how research generates findings that can be reported usefully in informing politics and practice. On the other hand, we also need to know how policy-makers use evidence from social sciences in their practice to address social problems. According to Sanderson (2009) better contexts, beyond instrumental rationality, can be built and are needed for enhancing an appropriate process for policy making.

For understanding the processes and actors behind successful policymaking that use evidence from scientific research, some authors have emphasised the need to explore the existing “productive interactions” between researchers and different actors – such as policymakers, stakeholders, social movements, etc. (Spaapen & van Drooge, 2011; Molas-Gallart & Tang, 2011; de Jong et al., 2014). One of the main questions behind this work is how research findings are taken up and used by these actors in order to create social innovation and impact. These authors argue that is difficult to attribute political impact to a specific piece of research because such impact would depend not only on researcher strategies but, crucially, on how other social agents adopt research outcomes. Hence this is why the SIAMPI project¹⁷ uses this approach based on interactions.

¹⁷ SIAMPI. Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society. FP7 project (Grant Agreement: 230330).

Another method that can be used to identify the political impact of research was developed by HERG at Brunel University. The Payback Framework gathers the policy benefits from undertaking research (Donovan & Hanney, 2011). Traditionally, it identifies how the use of research findings influences policy/decision making processes specifically in the Health sector, although the Payback Framework has also been successfully applied to SSH research (Klautzer et al., 2011; Levitt et al., 2010). Similarly, the role of different stakeholders in research has been the object of study of many studies aiming to describe the most effective processes in translating evidence into political impact while taking into account occasional or more structured partnerships between stakeholders in the scientific research process (Wehrens, Beeker & Bal, 2012; de Jong et al., 2011).

However, scholars have emphasised the need to have more cases to test new methods for evaluating the political impact of research and to move the field on (Brown, 2012; Boaz, Fitzpatrick & Shaw, 2009). But in order to carry out this task more evidence is needed on how the research process enhances and increases political impact. In relation to this debate and the need to advance in the identification of these social processes, questions related to attribution and the time lag of the research's political impact are also connected and relevant. Penfield, Baker, Scoble and Wykes (2014) argue that these are two of the main challenges associated with evaluating research impact.

Literature highlights that political impact is difficult to attribute to a specific research project if researchers do not participate in political assessment (Rymer, 2011). They could do this by producing evidence briefings based on systematic reviews (Chambers et al. 2012) or by taking part in advisory committees on legal practice and policy (de Jong et al., 2011).

In this light, during recent years, the number of problem-oriented or policy-oriented research calls has been growing in Europe, thus defining the nature of SSH research projects funded under these programmes. When assessing the political impact of this type of research, many indicators do not sufficiently measure policy-relevant effects because traditional indicators do not take into account how research is contributing to ameliorate the problems that societies face or how evidence is used by policymakers. By focusing their research on Denmark and the United Kingdom, Ernø-Kjølhede and Hansson (2011) conceptualised this type of policy oriented research as Mode 2 research and highlighted the need to build new indicators – Mode 2 indicators - for better monitoring research impact.¹⁸

In the context of the US, an example of indicators of political impact being developed include, within the National Institutes of Health eSPA system, proxies of the political impact of research in the form of 'Direct policy impact – news reports (reporting of biomedical research in mass media)' with data derived from the news sources Science News Daily and Medline Plus (Haak et al., 2012).

Participation and public engagement

An important theme in the literature on the political impact of SSH research is the need for active reciprocal engagement with end-users throughout the research process to ensure the maximum uptake and implementation of research findings. The political science literature

¹⁸ Mode 2 refers to the conceptualisation of modes of knowledge production (Mode 1 and Mode 2) elaborated by Gibbons et al. (1994). Mode 1 is the traditional academic approach to knowledge production (i.e curiosity-driven, discipline-based), and a new emerging Mode 2 is characterised by transnational collaborations, transdisciplinarity, and problem-oriented approaches.

has tended to focus on the lack of engagement between political scientists and political actors. For example, in the UK context, Donovan and Larkin considered the relationship between political science and 'practical politics', and why political scientists were failing to inform public policy and provide research that was 'readily usable by practitioners, be they policymakers from government or the civil service, or those seeking to influence, challenge or simply understand policy from the backbenches, NGOs, unions, trade associations, business or the media' (Donovan & Larkin, 2006, p. 11). They identified professional values within political science that dismissed applied work and prioritised 'increasing abstractness'; while policymakers and practitioners tended to focus on problem solving. They noted that proposed solutions tended to focus on a lack of policymakers' awareness of relevant research, and academics' ignorance of how findings might be fed into the policy process. But although several initiatives had been launched to address this gap, they argued that a distance remained between social scientists and practitioners because the underlying assumption was that 'the only thing preventing greater use of social science by practitioners is ignorance: the product is fine and only practitioner awareness needs addressing' (Donovan & Larkin, 2006, p. 16). However, they concluded that the abstract focus of the discipline works against producing useable knowledge for practitioners, and that 'there are few incentives from within academia or under the [Research Assessment Exercise] to change this.'

Dibb and Quinn (2010) later wrote that in the UK there remained a 'growing tension around the "double hurdle" of scholarly quality and relevance'. Although the then new Research Excellence Framework had introduced the assessment of the wider impact of research, 'academics are questioning the advisability of prioritising impact over peer review, and the potential consequences for career progression' (Dibb & Quinn, 2010, p. 327).

Co-production of research and research impact

However, while the political science literature has tended focus on barriers to engagement and political impact, the literature more closely allied to public administration has focused on how the co-production of research between academics and policymakers can facilitate research impact. Several papers focused on this topic in a 2010 special edition of *Public Money and Management*. Duijn et al. (2010) focused on the co-production of research between academics and policymakers, particularly in terms of negotiating complex governance processes. They believed that 'If public managers and policy-makers become more reflective and researchers more action-oriented, they can meet in joint enquiry' (Duijn et al., 2010, p. 228), and so championed the idea of a 'community of inquiry' located 'in the middle between science and practice' and where social scientists and practitioners can 'co-produce knowledge to cope with practical challenges' (Duijn et al., 2010, pp. 230-232). Antonacopoulou similarly believed that taking steps to actively engage in research 'as a mode of collaboration where scholars and business executives as co-researchers can develop a closer understanding of the subject matter, is more likely to overcome the perennial problem of translation' (Antonacopoulou, 2010, p. 225). Martin presented five strategies for achieving 'more engaged and engaging scholarship' based on research design, evidence gathering, data analysis and the dissemination of research findings (Martin, 2010, p. 211). This took the form of a sliding scale of the involvement of practitioners in the research process as: (1) informants; (2) recipients; (3) endorsers; (4) commissioners; and (5) co-researchers. He argued that practitioners may be involved in one or more of these elements, and the more engaged they are in each of these processes, the more likely it is that the research will meet their needs and have impact. O'Hare and colleagues highlighted the negotiated context of co-produced research, and introduced the idea of academics and

practitioners working together as 'critical friends' so that they may 'negotiate clear independence' (O'Hare, Coaffee & Hawkesworth 2010, p. 246).

While the above papers draw attention to the positive and negative aspects of engaging in co-produced research, they remain theoretical. Lovbrand (2011) studied co-production in the context of a large FP6 project "ADAM - Adaptation and Migration Strategies: Supporting European Climate Policy", in order to examine 'how knowledge-making practices are incorporated into European climate policy-making, and more importantly, how EU climate policy has shaped the funding, making and interpretation of useful European climate policy research' (Lovbrand, 2011, p. 226). In the context of this case study, Lovbrand remained sceptical about the impact of co-produced research. Although ADAM 'aimed to generate reflection about the suitability of existing climate policy goals ... the interaction with the EC offered little scope for thinking beyond policies already formulated by EU officials'. In this respect, 'many scholars of science and society hope that research co-produced with non-scientists will stimulate reflection and debate, the ADAM story tells us that the politics of useful science requires further attention' (Lovbrand, 2011, p. 226). Lovbrand found that on the one hand, when economic modellers changed their approach to fit the policy agenda, 'they received positive attention from the policy community and were more effective in putting their knowledge to work' (Lovbrand, 2011, p. 235), yet on the other hand, 'in contrast to positive stories of co-production found in the science studies literature ... the ADAM story told here paints a less rosy picture' of a 'co-production process that "closed down", rather than "opened up", the interpretation of feasible and desirable climate policy goals in the post-Kyoto era' (Lovbrand, 2011, p. 234).

On a more positive note, Stoker (2010) discussed the role of experiments (randomised control trials) in providing evidence to affect the direction of government policy, and reviewed various barriers to research utilisation. He alerts us to the fact that 'the dynamic of [research utilisation] rarely follows the logic of a simple problem-solving sequence where the problem is identified, a gap in knowledge is specified, the gap is filled by experimental research, and afterward the policy is enacted in the light of this information' (Stoker, 2010, p. 54). He therefore offers recommendations for researchers wishing to effectively translate experiments into policy: (1) develop 'people' skills; (2) understand the role that politics plays in driving the policy process; (3) accumulate evidence and wait for the right moment to use it; (4) be aware that research is a political act in itself, and that researchers need to be self-aware and self-critical; (5) build lasting relationships with policymakers to gain trust.

While Stoker dealt with theory, Cotterill and Richardson (2010) assessed the benefits of co-produced research with local government as a research partner, using randomised control trials (RCTs). They reported on a series of local level experiments co-produced with policymakers and public service providers, including: (1) to evaluate the promotion of household recycling participation in 6,580 households. Half the households were visited by canvassers trained to promote and encourage recycling; and (2) to evaluate the impact of school-based education on environmental attitudes and behaviour of 715 primary school students and their families in 27 primary schools. Evidence of social impact was: (1) measured by observing dustbin set-out rates over the 3 week period. Recycling participation rose by 5 per cent immediately after canvassing, but after 3 months had reduced to 2 per cent; and, (2) pre- and post- surveys by students in class and at home; the results showed no significant difference. In terms of co-production, Cotterill and Richardson concluded that 'partners want to have equal say over the research methodology, and negotiations cover both the research and the intervention', and that because the 'collective nature of civic interventions can militate against individual randomization' researchers can encounter

'ethical and moral objections from principled public service practitioners' (Cotterill & Richardson, 2010, p. 161).

Armstrong and Alsop (2010) write from the perspective of research management from within the UK's Economic and Social Research Council, arguing that both ex ante and ex post evaluation criteria need to rest on a conceptual model that includes 'the crucial role of co-production in achieving impact' because the ESRC's impact evaluation work has demonstrated that 'sustained involvement of [non-academic research] users is one of the most important determinants of policy impact' (Armstrong & Alsop, 2010, pp. 209-210). They argue that engaging with potential research users at a very early stage of the research process is the most important action that can lead to research findings being adopted (i.e. waiting to disseminate research findings to policymakers is too late). The effective co-production of research entails that research users should be 'involved throughout the research process, from agenda-setting, through design, fieldwork and communication of outcomes' (Armstrong & Alsop, 2010, p. 209). They identified 'certain key factors that are vital for impact generation, relating the process of generating impact, the context in which research messages are delivered and the content of the research. These include:

- Established relationships and networks with user communities.
- Involving users at all stages of research – not as subjects, but as project partners.
- Well-planned (and properly resourced) user engagement and knowledge-exchange strategies.
- Portfolios of research activity that build reputations with research users.
- Good infrastructure and management support.
- Where appropriate, the involvement of intermediaries and knowledge brokers as translators, amplifiers, and network providers.' (Armstrong & Alsop, 2010, pp. 209-210).

4.2. Political impact of SSH Research

The scientific literature explored and analysed in this second section shows how research projects having relevant impact on policy do not tend to highlight such impact in the research outputs they publish. Some policy impact can actually be found in reports of the project or other works that can acknowledge it but this is not a common practice. The same is true of laws and policy texts which rarely acknowledge SSH research that provided the evidence for reform.

However, a closer examination allows us to identify several examples of different SSH research projects that in collaboration with stakeholders have contributed in its origins, to the development of laws, policies and action plans in different levels and areas. Some of the examples that were found in the scientific literature are related to the creation or reform of the following policy levels: City-level policies; State level laws or policies; European law, regulations and recommendations; Promotion of political discussions, which influence processes of policy making at any level; Civil society action plans/ programmes/ policies.

A first example is the Boston Gun Project (Massachusetts, US), sponsored by the National Institute of Justice, and directed by David M. Kennedy, Anthony A. Braga, and Anne M. Piehl of Harvard University's John F. Kennedy School of Government. This project, began in early

1995¹⁹ and established an innovative partnership between researchers and practitioners to assess the Boston's youth homicide problem and design an intervention to have a substantial near-term impact (Kennedy, Piehl, & Braga, 1996). The fruit of this research project was the implementation of a Citywide Strategic Plan (city policy), in 1996, which had impressive results: the intervention accounted for a 63% drop in the number of Boston youth homicides, a 32% decrease in shots fired, and a 25 percent decline in gun assaults (Butts & Roman, 2011). At least, partly as a result of this success, the National Institute of Justice within the US Department of Justice funded projects in ten additional cities to replicate the Boston Gun Project, such as Los Angeles, Oakland and various cities in New Jersey.

In the United Kingdom the Economic and Social Research Council gave a special award to the research of Clifford Stott because of its impact on public policy. Stott, Adang, Livingstone and Schreiber (2007) observed collective behaviour patterns among football fans and protest crowds. They concluded that their work had policy implications especially with policing tactics and strategies used to prevent riots and conflicts by using specific tactics of low profile policing taking into account dialogue, consensus and not creating a strong antagonistic identity that could be spread among the crowds and generate or recreate conflicts. Stott assessed the design and implementation of police use of force strategy for the 2004 UEFA European Football Championships and the low levels of violent confrontation were due to this form of policing. At the same time, proposals based on this approach were accepted as amendments in the 2005 and the 2010 editions of the European Union's Handbook on International Police Cooperation and Measures to Prevent and Control Violence and Disturbances in Connection with Football Matches with an International Dimension. His research also was included in the recommendations of the UK Home Office Public Order Unit to manage football fans travelling to the UK. He also helped to design, create, train and implement the first UK Police Liaison Teams in 2011.

Impact on policy for social exclusion

Although interactions and partnerships of researchers with practitioners and policymakers have been widely studied, less attention has been placed on studying how these partnerships include the voices of the most vulnerable end-users throughout the whole research process (Gatt, 2011).

Indeed, in the area of social exclusion/inclusion an important impetus for changing legislation came from the joint work of scientists and organisations working with disadvantaged groups. By way of example, the right to housing is enshrined in many constitutions in Europe, but in most cases the right to access housing, like the right to work, are not enforceable rights. In France, for example, the right to housing was a recognised (non enforceable) social right as it is enshrined in the preamble to the Constitution since 27 October 1946. The first systematic study of the legal context of housing exclusion in Europe was published in 1996 (Avramov, 1996). Another ten years of research and lobbying by organisations working with homeless people in France contributed to the bill being passed of 5 March 2007, introducing the enforceable right to housing (Loison, 2007). Not only has a constitutional right become an enforceable right thanks to stakeholders' involvement, but researchers bringing knowledge of research methodology and stakeholders' grass-root organisations with strong lobbying capacities jointly carry out also monitoring of implementation.

¹⁹ <http://www.hks.harvard.edu/programs/criminaljustice/research-publications/gangs,-guns,-urban-violence/operation-ceasefire-boston-gun-project>, last checked: 19/06/14.

Impact on policies toward societal challenges

The FP6 “Integration of Female Immigrants in Labour Market and Society FEMIPOL” (CIS8-CT-2004-022666) concluded that there is need to adopt a rights-based approach to migrant workers and separate residence rights from employment rights. This proposal endorsed by the FEMIPOL project was part of the UK’s Trade Union Congress’ (TUC) approach to migrant workers, and it was suggested that giving all workers employment rights, including the regular and irregular, is needed in order to counteract the current power of employers to exploit vulnerable migrant workers.

A certain link may be observed between the impact of research and the publication of a report on abortion by a governmental scientific institute in Belgium. Namely, in 1972, the Population and Family Study Centre (CBGS), in cooperation with researchers from different universities, produced a voluminous multidisciplinary monograph on abortion (Cliquet & Thiery, 1972). Abortion was illegal and largely a taboo topic in research quarters at the time. This publication was not a commercial success. Nevertheless, it appears that this book had an important influence in policy quarters: it facilitated the political discussions of this otherwise taboo issue. Furthermore, the study was used for the policy proposal in the State Committee for the Ethical Problems regarding contraception and abortion in 1975. In 1990, a new law legalising induced abortion was passed in the Belgian parliament. Is there a causal relation between the publication of the CBGS book on abortion and the law allowing abortion under certain conditions? The CBGS research and publication of results certainly played a role in the political decision making process, but there can be no doubt that also without this particular research and scientific publication, the Belgian legislation would have been adapted.

Another sensitive societal challenge was tackled in Belgian legislation without an obvious link between research, on euthanasia and the euthanasia law in Belgium. Around 1985, a governmental scientific institute CBGS was not given permission to undertake research on euthanasia. In 2002, the law on euthanasia was passed in Belgium. Would CBGS research on euthanasia, if it had taken place, facilitated or accelerated the policy-making process and adoption of legislation, just as it did for contraception and abortion? Even without population surveys the law was changed. Of course, many scientific publications on euthanasia existed in other countries and there was, in particular, the experience and practice of euthanasia in the Netherlands and Switzerland.

The contribution of research within European Framework Programmes

European Framework Programmes provide examples of political impact. One of the European projects, funded under FP6 that had an impact at the political level is “CoPECL - Joint network on European private law”. The objectives of this project integrated by various groups was to deliver a proposal for the “Common Frame of Reference” (CFR) for European contract law as described both in the Commission’s Action Plan (COM [2003] 68 final) and the Commission’s Communication on “European Contract Law and the Revision of the Acquis: The Way Forward” (COM (2004) 651 Final) of 11 October 2004. As a result of the work conducted by the network, the European Parliament stated in its Resolution of 23 March 2006 that the project of a “Common Frame of Reference” is 'by far the most important initiative underway in the field of civil law'. Moreover, the results of the CoPECL Network, in

particular the DCFR- Draft Common Frame of Reference-, have attracted extraordinary attention from the EU Institutions and Member States²⁰.

An FP7 funded project having similar policy impact was “EUROJUSTIS - Scientific Indicators of Confidence in Justice Tools for Policy Assessment”, led by Mike Hough. This project aimed to provide EU, national and local institutions with new indicators for assessing public confidence in policing and in justice. The results of this project were widely disseminated. Hough, Jackson, Bradford, Myhill and Quinton (2010) stated that there were two routes for policing and criminal justice systems to act, the correct one generating repressive and not inclusive actions, negatively affecting the trust and legitimacy citizens give to their institutions; and a correct route by being more inclusive and taking into account the voices of the marginal group and having a police and a judiciary system that is fair. By having these indicators and this information on trust and legitimacy, policy makers can act and generate more effective policies for policing and against crime that can reduce the costs of the system and increase trust and the legitimacy citizens give to those institutions. Hough was member of different assessing bodies and this research as well as others projects he conducted were taken into account in different policy documents such as the revision of the *London Policy and Crime Plan* (2013).

The Science and Society Programme (SaS) under FP6 and the Science in Society Programme (SiS) under FP7 addressed governance from many perspectives, such as encouraging dialogue between scientists and other members of the public, improving the use of science in policymaking, promoting an adherence to ethical standards, and developing better ways for the results of research to be accessed by all. The SiS Programme also supports specific research activities such as the connection between science, democracy and law, as well as governance issues linked to advanced participatory societies and to the ERA's integration. The movement from good governance to more elaborated concepts of democratic, participative, sustainable and responsible governance, is reflected in the aims, objectives, activities, outcomes and political impact of many projects funded.

5. What do we know about the Social impact of SSH research?

In contrast to the abundant scientific literature about the impact of a wide variety of technological, economic and health factors, there are few investigations focused on the direct or indirect impact of SSH research with regards to social or societal phenomena. Yet, this gap stands also in sharp contrast to the widespread interest, practice and documentation about the impact of research on academia and scientific knowledge (e.g. Leduc, 1994; Penfield et al., 2014). Nevertheless, in recent years, a substantial increase in efforts aiming at measuring or creating tools for identifying such impact has been carried out by experts in the field as well as by research funding agencies. These efforts include ‘ex ante assessments’ or ‘ex post assessments’ of the potential or achieved social impact of SSH research. See, for instance, the recent literature review of the societal impact of research (in general) by Lutz Bornmann (2013) (see in this context also Gibbons et al, 1994; Newby, 1994; Buxton et al., 2000; Hessels & Van Lente, 2010; Holbrook & Frodeman, 2010; de Jong et al 2011; United States Government Accountability Office, 2012).

²⁰ http://cordis.europa.eu/result/report/rcn/51949_es.html

5.1. Social impact assessment

In the review of the specialised literature, we have identified three key topics discussed - conceptualising social impact, evaluation methodologies, and indicators. In what follows, a summary of the main findings is presented.

Conceptualizing social impact in SSH research

The concept of social impact has been defined as “the process of assessing or estimating, in advance, the social consequences that are likely to follow from specific policy actions or project development, particularly in the context of appropriate national, state, or provincial environmental policy legislation, and not of (SSH) research” (Burdge & Vanclay, 1995, p. 31). However, a number of different terminologies and conceptualisations reflect also a lack of generally accepted conceptual and instrumental frameworks (eg, van der Meulen & Rip, 2000, p. 11; Bornmann, 2013, p. 220).

First of all, there is clearly no general consensus on what is to be understood by the word ‘social’ in the ‘social impact’ literature. There is apparently a broad diversity in the conceptualisation and definition of this domain. In some publications, for instance those of the European Commission, a very broad spectrum of social impact areas are listed as part of the work programme: human rights, social cohesion, economic cohesion, employment, human capital formation, public health and safety, social protection and social services, liveable communities, culture, consumer interests, security, governance, international cooperation, role of SMEs, lessons learnt and success stories (European Commission, 2005, 2011, p. 12ff). Hemling et al. (2011) include a long list of various social domains to be considered in social impact investigations: employment and labour markets; standards and rights related to job quality; social inclusion and protection of particular groups; gender equality, equality treatment and opportunities, non-discrimination; individuals, private and family life, personal data; governance, participation, good administration, access to justice, media and ethics; public health and safety; crime, terrorism and security; access to and effects on social protection, health and educational systems; culture; social impacts in third countries. At the other extreme of the social impact spectrum, the social impact domain is strongly limited to a few items pertaining to the living conditions of people: welfare, wellbeing and quality of life, customs and habits of life (consumption, work, sexuality, sports and food) (e.g. Godin & Doré, 2005).

Here it is important to note the blurring of the divide between what is defined as political impact and social impact, because these categories overlap. In this Report we maintain the distinction for analytical purposes, considering the overlapping when conclusions have to be outlined.

In the context of the development of Australia’s Research Quality Framework (RQF), definitions of research impact were co-produced with the research community, and were defined as ‘adding to the social, economic, natural, and cultural capital of the nation’ (Donovan, 2008, p. 54):

- *Social Benefit*. Improving quality of life; stimulating new approaches to social issues; changes in community attitudes, and influence upon developments or questions in society at large; informed public debate and improved policy-making; enhancing the knowledge and

understanding of the nation; improved equity; and improvements in health, safety and security.

- *Economic Benefit.* Improved productivity; adding to economic growth and wealth creation; enhancing the skills base; increased employment; reduced costs; increased innovation capability and global competitiveness; improvements in service delivery; and un-quantified economic returns resulting from social and public policy adjustments.
- *Environmental Benefit.* Improvements in environment and lifestyle; reduced waste and pollution; improved management of natural resources; reduced consumption of fossil fuels; uptake of recycling techniques; reduced environmental risk; preservation initiatives; conservation of biodiversity; enhancement of ecosystem services; improved plant and animal varieties; and adaptation to climate change.
- *Cultural Benefit.* Supporting greater understanding of where we have come from, and who and what we are as a nation and society; understanding how we relate to other societies and cultures; stimulating creativity within the community; contributing to cultural preservation and enrichment; and bringing new ideas and new modes of experience to the nation.

Other approaches to conceptualizing social impact are closer to economic impact. For example, the STARMETRICS initiative in the United States,²¹ where Weinberg et al. (2014) identified the effects of science funding on short-term economic activity. They concluded that scientific activity has economic impact on society by identifying the number of people directly employed in the research as well as by knowing products and goods purchased by scientific institutions. While it could be relevant to connect EU DG Research and Innovation datasets with Higher Education institutions creating jobs with EU funds; our analysis of social impact is more focused on the external jobs SSH research can generate through actions based on SSH research findings.

The diversity in the conceptualisation and definition of the social impact domain is also partly connected with the different meanings given to the terms 'social' and 'societal'. Sometimes both terms are often used interchangeably (Bornmann, 2013, p. 218). When they are distinguished, the concept social (impact) is sometimes limited to effects or benefits on traits or behaviours at the individual level (e.g. Godin & Doré, 2005); the concept societal impact, in contrast, refers to broader community based phenomena such as demographic change, human rights, social cohesion, economic cohesion, employment, human capital formation, public health and safety, social protection and social services, etc. (e.g. European Commission, 2011; Technopolis, 2009).

With regard to the domain of social/societal effects of research, Mostert et al. (2010) divide a somewhat broader classification into societal products (outputs), societal use (societal references), and societal benefits (changes in society). However, social/societal effects of (social) research may not only be of a positive nature ('benefits'), they may also be of a negative nature, and have disadvantageous consequences. Moreover, the normative nature of the evaluation – benefit or disadvantage – depends on the value judgment adopted (Brewer, 2011).

²¹ <https://www.starmetrics.nih.gov>

In understanding social impact as the use that society can make of research outputs, several projects have analysed the channels or processes that these outputs follow before they are “used” by society.

The FP7 project “SIAMPI - Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society” (2007-2013), proposed a shift in research evaluation from the existing social impact framework to that of ‘productive interactions’, while also including the contributions of researchers and stakeholders alike. In this context, research evaluation is understood more as learning tools as for generating greater future research impact.

In this sense, as Flecha indicates²² there exists some confusion between the concepts of Scientific impact, Dissemination, Knowledge transfer and Social impact, which is important to clarify:

- Scientific impact: refers to scientific and academic publication of research results including citations to publications.
- Dissemination: is when – in addition to having been published scientifically they are made known, both to the scientific community, as well as to policymakers, stakeholders and citizens in general (i.e. press, social media, networks, etc).
- Knowledge transfer: when the published and disseminated results are taken up by policymakers and/or social actors as the basis for their policies and/or actions regardless of whether they have made social improvements or not.
- Social impact: is when the published, disseminated results, which have been transferred, lead to an improvement in relation to the stated goals of European society. In the case of Horizon 2020 these refer to the objectives of the EU2020 Strategy.

Evaluation Methodologies

In regard to approaches and methodologies for the evaluation of social impact, we find both *ex-ante* evaluation of research projects concerning possible social impacts, and *ex-post* evaluation monitoring the impact of research that has been undertaken (Holbrook & Frodeman, 2011; Potì & Cerulli, 2011; Social Sciences and Humanities Scientific Committees, 2013; Bornmann, 2013). Furthermore, two major groups of methods are distinguished: qualitative (including peer review, case studies and surveys) and quantitative (development and use of statistical indicators and - in certain fields - also advanced mathematical models such as econometric models).²³

Notwithstanding the fact that peer review of societal impacts is often perceived as being more complex than reviewing scientific quality, and has been met with resistance within the scientific community, Holbrook and Frodeman (Holbrook & Frodeman, 2011, p. 240) argue that “there is little evidence to suggest that peer review is any less effective at *ex ante* assessments of societal impact than it is at *ex ante* assessments of scientific, technical, or intellectual merit”. On the other hand, van Raan (van Raan, 2000, p. 82) is of the view that

²² Flecha, R. “Buenas prácticas: Cómo lograr impacto en los proyectos del Reto 6.” [Good practices: How to achieve impact in the projects of the Challenge 6]. In *H2020: Informative session about Challenge 6*, ed. FECYT and European Commission. Bilbao: Universidad de Deusto (24 January 2014).

²³ In this report we classify peer review, case studies and surveys as qualitative methods; this does not imply that these methods do not use data and metrics, rather they develop original activities of data collection, which are likely to be treated either using qualitative techniques or quantitative statistical treatments.

peer reviewing for social impact purposes should be complemented by other methods, including recording stakeholders' views and bibliometric mapping (see below, section on quantitative methods). There are, of course, several other criticisms regarding peer review (e.g. its subjective and contingent nature, its time consuming procedure and high costs) (see Pontille & Torny, 2010, p. 347; van den Besselaar & Leydesdorff, 2009). But overall, peer review remains one of the major cornerstones of a comprehensive and integrated impact assessment process, especially in the domain of SSH impact assessments (Barker, 2007; Ernø-Kjølhede & Hansson, 2011).

Although not strictly purely qualitative, *case studies and surveys* are useful tools to grasp the social impact of research. Several authors have discussed the advantages and disadvantages of case studies (e.g. Martin, 2011; Bornmann, 2013). While some authors point out criticisms of lack of objectivity and quantification, and its labour-intensive nature, the unique advantage of qualitative approaches consists in their ability to deal with complexity, especially for a phenomenon such as societal impact. Case studies are particularly valuable in the assessment of interactions between scientists and stakeholders (Bornmann, 2013, p. 226). In the case of surveys, although they usually combine the application of qualitative (e.g. questionnaire construction) and quantitative procedures (e.g. statistical analyses of survey results), surveys are particularly valuable in measuring changes of opinion and the user perceptions of various groups of stakeholders.

Indicators for the evaluation of social impact and mixed methods

Studies have identified large numbers of indicators relevant to particular societal phenomena, which are usually domain specific (Martin, 2011; European Commission, 2010; Lepori et al., 2011). However, many authors complain about the inadequacy of these impact indicators (eg Barré, 2010; Ernø-Kjølhede & Hansson, 2011, p. 141; Spaapen & van Drooge, 2011, p. 211; (Gregersen, et al., 2009, p. 152) and the lack of internationally comparable databases (Lepori et al., 2007; Edler & Flanagan, 2011). Therefore, efforts have been made to develop more sophisticated indicators, which are more suitable than the simple basic indicators (e.g. van Raan, 1998; 2000; Pontille & Torny, 2010; Wagner & Leydesdorff, 2012).

A study on indicators for community research evaluation (Braun et al., 2009) showed that S&T indicators could make substantial contributions to the improvement of evaluation and monitoring practices. An indicator-based framework was proposed. It built on six indicator domains, covering the areas of programme management; FP participation and funding, scientific results, technological results, economic and social impacts and structuring of the ERA. To apply the indicators which are identified in these domains efficiently, important prerequisites are: (1) a good understanding of the role and added value of indicators in the evaluation process, (2) the integration of indicator production and application in the evaluation process, (3) a sound methodological background and (4) the availability of high-quality data and necessary support structures.

There is a general consensus in the social/societal impact literature that neither qualitative nor quantitative methods alone suffice to meet the social impact assessment goals related to scientific research. David Roessner (2000) calls even the choice of quantitative versus qualitative measures in research evaluation a false one, and he adds, "especially for evaluators isolated from the real world". Hence, in recent years, a shift in evaluation methods can be observed, moving from the application of simple to combined methods (e.g. Penfield et al., 2014; Donovan, 2007, pp. 592-593).

Best practice is considered to combine or integrate narratives with relevant qualitative and complementary quantitative indicators, in order to seize the multidimensional and contextual nature of complex societal phenomena (e.g. Spaapen & Sylvain, 1993; Gabolde, 1998; Evaluating Research in Context (ERiC), 2010; Schmoch, et al. 2010; de Jong et al., 2011; Donovan, 2011; Penfield et al., 2014). In recent years several specific combined or integrated social impact tools have been developed, e.g. the Payback framework (Buxton and Hanney, 1994; 1996; 1998; Wooding et al, 2007; Donovan & Hanney, 2011; Henshall, 2011; Penfield et al. ,2014; Klautzer et al., 2011), the AGORA model (Barré, 2001), and the SIAMPI approach (SIAMPI, 2011; Molas-Gallart & Tang, 2011; Spaapen & van Drooge, 2011; Penfield et al., 2014).

As has been pointed out above, there are several authors dealing with the methodological problems of involving various groups of *stakeholders* in the process of social impact analysis (e.g. Spaapen et al. 2007; de Jong et al., 2011; Bornmann, 2013). Authors argue that scientists alone should not conduct the qualitative assessment of societal impact only, because they often appear to have difficulties in dealing with assessing the societal impact of research. Interactions with non-academic stakeholders are not only important for transferring knowledge between science and society, but carefully selected stakeholders can be valuable in contributing to evaluating societal impact.

In addition, the scientific literature on social impact assessment of research (and in particular SSH research) addresses methodological difficulties in this endeavour. Several authors have listed and discussed such problems (e.g. Leduc, 1994; Martin, 2007; de Jong et al., 2011; Molas-Gallart & Tang, 2011; Spaapen & van Drooge, 2011; Bornmann, 2013). Others drew the attention to the fact that it is apparently more difficult to measure social impacts of (SSH) research than its effects on other domains such as environment, health, economy, or science itself (Spaapen & van Drooge, 2011, p. 211).

Martin (2007), for instance, distinguishes, four major difficulties: the causality problem, the attribution problem, the evaluation time scale problem, and the internationality problem. Bornmann (2013) mentions, in addition, several difficulties that pertain to the scientific process or practice itself (the peer review system; the specificity of the research institute) or the nature of the (SSH) research impact on the social/societal process (desirable or undesirable effects; the multidimensional social/societal impact of research).

Main challenges of impact assessment

Besides the decision of using qualitative or quantitative methods or a combination of the two, there are several challenges in research impact assessment (Proneos et al., 2009). Here we present some problems related to both social and political research impact assessment, which can also be largely extended to the assessment of research impact of other fields of science:

Causality and attribution: Due to the multiplicity and complexity of social phenomena and their causes, it is often not possible to identify the exact portion of the impact that is attributable to the research or other co-varying factors (e.g. van der Meulen & Rip, 2000; Nightingale & Scott, 2007; Gray et al., 2009, p. 139; de Jong et al., 2011; Spaapen & van Drooge, 2011; Bornmann, 2013; Penfield et al., 2014). Measuring difficulties are common due to the fact that basic data about causes and effects may not have been collected or may no longer be available (Penfield et al., 2014), and 'tracking backward' from an impact to its causes 'is extremely resource-intensive and an inadequate methodology when the objective is to analyse the impacts of a specific

research programme or project' (Molas-Gallart & Tang, 2011, p. 225). These, among other reasons, make it difficult to ascertain the particular impacts of a given piece of research.

Internationalisation: the scientific enterprise is, par excellence, an intrinsically international activity with very weak, if any, borders in matters of dissemination and influence. Hence, attribution is very difficult to identify, even in the presence of relevant national research (Martin, 2007). As mentioned above, it is particularly difficult to design suitable indicators for the assessment of internationalisation of science policy (see Flanagan, 2011; Reale et al. 2012; Reale et al., 2013).

Evaluation time scale: there is often a substantial time lapse between the publication of scientific results and policy measures that stem therefrom, let alone the social or behavioural changes that may be related to or be a consequence of those results (e.g. Buxton, 2011; Morris *et al.*, 2011; ERiC, 2010; Ruegg & Feller, 2003; Van der Meulen & Rip, 2000, p. 13; Spaapen & van Drooge, 2011). Hence, short-term policy or social impact evaluations may miss the often more important distant impacts of research findings, particularly because impact assessment bodies selectively use only short-term evaluation criteria (Buxton, 2011; Arnold, 2012; Bornmann, 2013). It may take many years before knowledge is applied or becomes observable and measurable (De Jong *et al.*, 2011, p. 62; Buxton, 2011, p. 260; Molas-Gallart & Tang, 2011, p. 224).

Problem-focused research: a preliminary facilitator or mediator in obtaining social/societal impact from (SSH) research is that the research projects should either be specifically problem-focused or should include a work package, panel or element that focuses on the post-research societal or policy oriented application of the findings. This is not self-evident, because scientists are often reluctant to get involved in other than purely scientific preoccupations such as policy implications or social engineering (e.g. Holbrook & Frodeman, 2011, p. 244; Dib & Quinn, 2010, p. 327). Under the influence of the Framework Programmes of the European Commission, and also at the national level, there are increased expectations for scientific research to include a societal dimension. Research is thus evaluated *ex ante* or *ex post* on the basis of its impact intentions or merits (e.g. European Commission, 2005, p. 11).

Dissemination/communication: In order to enhance social/societal impacts, research projects need to clearly state the outreach they expect to occur and importantly, how they will to achieve this. To this end, they are asked to include a well-designed dissemination strategy that can contribute to ensure that the findings are reaching policy makers, relevant professionals and other stakeholders, as well as the population as a whole (see e.g. Rigby, 2002; Newby, 1994; Solesbury, 1994; Stipich, 1994).

Positive vs. negative impact: research may impact in positive or in negative ways, and whether an impact is positive or negative can change over time and at different points of assessment. Also, a good impact for one group may be a bad impact for another (Molas-Gallart & Tang, 2011, p. 219), and so evaluating social impact is an inherently political task as various viewpoints and interests must be considered.

The economic returns of research

Impact assessment is grounded on the concepts of impact that one decides to adopt. The definition of impact on other societal subsystems is often in line with both general policy impact studies and the political interest in the societal impact of research. Hence, the notion of impact in this context is associated mainly with the economic and/or social effects that science may have, or it has had on society and the economy.

Next to social impact, increased importance is being attached to the extent to which research results achieve expected economic returns, understood as market benefits. A key contribution on this topic was from Salter and Martin (2001). This article critically reviewed the literature on the economic benefits of publicly funded basic research. In that literature, three main methodological approaches were adopted. Econometric studies are subject to certain methodological limitations, but they suggest that the economic benefits are very substantial. From the literature based on surveys and on case studies, it is clear that the benefits from public investment in basic research can take a variety of forms. The relative importance of these different forms of benefit apparently varies with scientific field, technology and industrial sector. Consequently, no simple model of the economic benefits from basic research is possible. The authors critically reconsidered the rationale for government funding of basic research, arguing that the traditional 'market failure' justification needs to be extended to take account of these different forms of benefit from basic research.

From the perspective of the IMPACT-EV project, one interesting piece is the Rudd's paper (2011), which argues that "research in the humanities, arts, and social sciences (HASS) tends to have impacts that enhance quality of life (QOL) but that are not amenable to pricing in established markets. If the economic value of 'non-market' research impacts is ignored when making the business case for HASS research, society will under-invest in it" (Rudd, 2011, p. 127). The results were that 21% of the sample, placed a strong priority on QOL indicators relating to people, community, and culture, and an estimation of the money they were willing to pay for these outcomes was made. Therefore, non-market valuation techniques can certainly help HASS researchers demonstrate the economic benefits of their research impacts more fully than is currently the case. However, it is not sufficient to rely only upon this type of evaluation in order to secure funding, because STEM research impacts might also incorporate non-market human health and environmental QOL impacts in their cost-benefit calculations, and the outcome would be uncertain for the SSH in that case.

Another key item is Foresight that has evolved as a distinct prospective analytical tool: it considers alternative futures of various S&T fields or socio-economic systems by bringing together the perspectives of various stakeholder groups, and thus assists the decision-making processes at different levels. However, in order to avoid over-promising — and subsequent disappointment — about what foresight can deliver, the potential contributions to decision-making processes by foresight planning should be clearly understood. Havas et al. (2010) for instance put foresight into the broader context of policy-making processes, with a particular emphasis on innovation policy. They described the evolution of different policy rationales since the 1960s, developed a framework to classify the impacts of various types of prospective analyses, and reviewed the evaluation results of several national foresight programmes by using this framework. On that basis, future directions of how foresight might evolve were considered to spur discussions. Moreover, the authors pointed out that it is crucial to prove the impact of foresight on decision-making. This impact depends on relevance to major issues faced by society, but also its timing and the quality of its 'products' — reports and recommendations — are crucial. Only substantive, carefully formulated

proposals can grab the attention of opinion-leaders and decision-makers and are thus likely to be implemented.

Relying on efficiency analysis, the study of Jiménez-Sáez et al. (2011) evaluates to what extent policymakers have been able to promote the establishment of consolidated and comprehensive research groups to contribute to the implementation of a successful innovation system for the Spanish food technology sector, oriented to the production of knowledge based on an application model. Using data envelopment analysis (DEA) techniques that allow calculation of a generalised version of the traditional distance function model for productive efficiency, pervasive levels of inefficiency and a typology of different research strategies are found. Among these, established groups do not play the pre-eminent benchmarking role; rather, partially oriented, specialised and "shooting star" groups are the most common patterns. These results correspond with an infant innovation system, where the fostering of higher levels of efficiency and the promotion of the desired research patterns are ongoing. This study demonstrates the deviations that exist between the innovation system objectives and the mechanisms (incentive schemes) established to assess research group participation.

5.2. Social impact of SSH research

Based on a preliminary examination of the secondary literature as well as directly analyzing primary sources, a selection of SSH research projects with social impact are presented. The thorough review conducted (articles, working papers, reports and other sources) has identified some examples of social impact. Our analysis is better understood as exploratory in nature rather than being comprehensive. In this report, social impact is attributed to a specific research project when there is evidence that its outcomes have contributed to the accomplishment of the Europe 2020 targets on employment, education and poverty and social exclusion.²⁴ In so doing, this section is divided into three parts according to the three EU 2020 targets. Within each of these there are different examples of the varied ways SSH research is contributing to improve the everyday life of hundreds of people from all around the world. The challenges of identifying social impact have been already discussed in the previous section and should be taken into consideration here. It is necessary to acknowledge the limitations of the present review and the need to continue with this work in the next three years of the IMPACT-EV project.

SSH research contributions to the EU 2020 target for employment

The Europe 2020 Strategy aims to attain an employment rate of 75% of people aged between 25 and 64. The most used indicator for this target is "Employment rate aged group 20-64" (Eurostat, 2014).²⁵ In our analysis, we have identified three different ways through which SSH research contributes to this target:

- (a) By creating new jobs or companies;
- (b) By improving the employability; and, finally,
- (c) By identifying actions that generate new jobs.

²⁴ Targets on *Climate change* and on *R&D* were not included in the analysis because of the difficulties to find evidence on these specific targets.

²⁵ EUROSTAT (2014). Employment rate by sex, age group 20-64 (t2020_10). Indicator Profile (ESMS). Retrieved on June 25, 2014 from [here](#).

While the two first consist in creating new interventions as a product of the close collaboration between SSH researchers and stakeholders, the third is based on describing the main features of actions that have been proven to be effective in doing so.

a) Creating new jobs and companies

The archaeological research and excavations conducted in Atapuerca (Spain) which led to the discovery of the *homo antecessor* is an example of how research can generate new jobs and companies. This project began in the 70s with the discovery of the first archaeological remains, since then, the research conducted by the The Catalan Institute of Human Paleoecology and Social Evolution (IPHES) have achieved a great international scientific impact, being a model in its field. As an example, the paper published in *Science* (Carbonell et al., 1995) had received 230 citations in Web of Science and 241 in Scopus on May 2014. Beyond this, the research has produced 105 papers in JCR publications in ISI from 1995 to 2004, of which 38 belong to IPHES. This scientific value has led to relevant economic and social developments in this northern rural region of Spain.

Between 1995 and 1998, four new catering and hospitality companies were created in the town of Atapuerca, also, three SME had been created around the activities of the sites (Granja Escuela Arlanzón, Paleorama and ArqueOcio). In 1999 the Atapuerca' Foundation was instituted in a Public-Private-Partnership model. Additionally, the research has supported the creation of two non-profit organizations (ACAHIA and Paleorama) which provide the public visits to the sites and the management of the Archaeological Park of Atapuerca. Furthermore, the project collaborates with NGOs such as Aspanias Burgos, a charity organization that works with disabled people.

In 2001 UNESCO recognized the deposits of the Sierra de Atapuerca as World Heritage site because it represents an exceptional reserve of data, the scientific study of which provides priceless information about the appearance and the way of life of these remote human ancestors (UNESCO, 2001). As a result of the interventions or actions derived from the project's findings, 1130 direct stable jobs have been created and maintained since 2010 through the creation of the Museum of Human Evolution in Atapuerca besides to other foundations and businesses (Moreno Lara & Fernández, 2001).

Thus, besides these direct developments, jobs have been created as an indirect result of the project's findings. This is the case, for instance, for the number of jobs generated in the tourism sector because of the significant increase of the region's visitors, attracted by the museum as well as other sites around the area. The promotion of the Museum of Human Evolution has generated an overall economic impact of over 53 million Euros. In addition, most of 470,982 visitors to the sites officially accounted for by Regional Administration between 1995 and 2004, were catered for or took accommodation in the towns around the sites (Saíz Martín, 2011). Other consequences of the social impact that the research have had in the area is the positive effect for the revitalization of the affected localities. In this regard, there has been a positive change at the Atapuerca population with a growth of over 10% in the last five years, those data are reinforced when it is stated that there has been a rejuvenation of the average age of the population, which stands at 42 years old and in their work, because 15% is dedicated to tourism-related jobs (Saíz Martín, 2011).

b) Improving employability

Besides creating new jobs, there are other research projects that generate actions that contribute to attaining the EU 2020 target by improving citizens' employability. The FP5 "WORKALO - Creation of new occupational patterns for cultural minorities: The Gypsy Case"

is an example of a research project that has generated effective partnerships between researchers and other stakeholders in improving employability. The *WORKALO* project (2001-2004) aimed to create new occupational patterns for cultural minorities, specifically for the Roma. Findings highlighted the need to change the old paradigm in which many vocational training programmes were oriented to marginalized groups and specifically towards the Roma just to respond to legal duties but not to insert them to the labour market. Traditionally, these courses were not having results in the employability of the Roma because they were based on deficit thinking theories rather than taking into account the competences and skills that Roma already had. In many cases, courses could provide some specific skills to the attendees in one specialized profession (such as hairdressing), but they did not find a job because of ethnic discrimination. The project analysed the skills that Roma people already had, most of them required by the labour market in the information society. In this sense, based on *WORKALO* results and as result of specific partnerships between researchers and Romani associations specific vocational training courses were designed that could connect Roma women and migrants needs with finding a job (Sordé-Martí, Munté, Contreras, & Prieto-Flores, 2012). This intervention aimed to overcome all the above-mentioned barriers that Roma face when getting access to the labour market. It consisted of a training course for becoming canteen monitors. As the contribution of Romani women is very relevant in the school canteens with Roma children, their labour insertion was very successful in comparison to other courses, with an 80% success rate, whereas these types of programme usually do not achieve more than a 20% rate for labour market inclusion (Sordé-Martí, Serradell, Puigvert, & Munté, 2014).

c) Identifying actions that generate new jobs

A major range of SSH research which has produced key knowledge in identifying innovations and strategies of diverse stakeholders needs to be considered here. Two examples are provided under this category.

The first example of SSH research that has contributed to shedding light on cases of job creation is the Food deserts project led by Neil Wrigley and funded by the ESRC. Food deserts are deprived urban areas with poor access to retail stores affecting thus access to healthy food among their residents. The project analysed the effects that the urban renewal of these food deserts had on food consumption and other factors. Specifically, in 2000, they conducted a case study in one of the most deprived areas of Leeds looking at the impact that the arrival of TESCO (one the largest food retailers in the UK) had and ten smaller independently run shops in the area. This project counted on the participation of the Local Authorities and one of the agreements was to give priority in employment to long-term unemployed residents of the area. Thus, the impact of this renewal was the creation of 320 new jobs positions from which 230 were occupied by unemployed local residents (Wrigley, Warm & Margetts, 2003). This research thus showed the impact it had on the employment of the residents of this area. Besides generating new jobs, qualitative research demonstrated that residents are having better access to healthy food and that the consumption of fruit and vegetables per day increased by half. Before the renewal project took place residents had to travel long distances to carry groceries or became dependent on others to buy healthy food. That was one of the reasons why they ended up buying high calorie and processed food. Before the project took place, 70% of the residents consumed less fruit and vegetables than the national average (Desjardins, 2010). This project, which gathered a multidisciplinary team of geographers, public health nutritionists and city and regional planners, has been awarded for its outstanding impact (Wrigley, Lowe & Guy, 2002) as the initial study in Leeds has been replicated in other deprived neighbourhoods, providing evidence of the social benefits of this urban renewal procedure.

The FP7 project “WILCO - Welfare innovation at the local level in favour of cohesion” is the second example. It analysed seventy-seven cases of social innovations from 20 cities and covering ten countries in the EU. The innovations have been developed in order to respond to social problems present in those cities during the most recent economic crises. As it is stated, "the majority of social innovations are new service arrangements, making a difference in terms of organizations, processes and types of service offers" (WILCO, 3rd press release). WILCO'S objective was to study and disseminate these cases in order to facilitate their successful transfer and implementation to other places (Evers, Ewert, & Brandsen, 2014). All these cases were innovative but not all were successful in promoting social inclusion and employment. Three examples stood out. First, the Neighbourhood Mothers (Berlin, Germany) consisted in involving one hundred migrant women in participating as “District Mothers” reaching out other mothers in schools, playgrounds, mosques, shops or markets and giving them information on health, language or social services for promoting their social inclusion. These mothers received training connected to being a social assistant in the future, and a small salary. The local youth welfare office in Friedrichshain-Kreuzberg has already created five part-time positions for neighbourhood mothers who additionally passed professional training to become a social assistant for intercultural family care (Evers, Ewert & Brandsen, 2014, p.13). Second, the *Neighbourhood Mothers Catering* (Amsterdam, the Netherlands) case relied on the capacity-building of migrant women and their expertise in cooking. It consisted of creating a cooperative of migrant mothers to start a business in the catering sector and employing people who were at risk of poverty. Among the women participating in the project, some did this as a regular job, others to get involved to and to receive an additional income, and others saw it more as a way of making a bit of extra money and they have managed to become self-employed with guidance (for instance, concerning food safety and/or administrative issues) (Evers, Ewert & Brandsen, 2014, p.216). The last case, *Her Second Chance* (Varazdin, Croatia), aimed to improve the socioeconomic conditions of disadvantaged women through their empowerment, training and development of business skills. Over time, 63 beneficiaries successfully acquired new professions that made them more competitive in the labour market, for example, 15 women become certified nurses (Evers, Ewert & Brandsen, 2014, p.54).

SSH research contributions to the EU 2020 target for education

The main two targets in the EU2020 strategy with regards to education are to reduce the rates of early school leavers (reducing school drop-out rates below 10%) and to raise the percentage of Higher Education graduates (at least, 40% of 30-34 year-olds completing third level education). In what follows, three different ways through which SSH research projects have contributed to these goals are presented:

- a) By reducing school dropout rates;
- b) By raising the percentage of higher education graduates, and
- c) By identifying actions that lead to improvements in educational outcomes.

a) Reducing school dropout rates

Among many, two projects are highlighted here: the University of Cambridge-based “Thinking together” and the US government based Head Start Programme.

The Thinking Together research lead by Neil Mercer and colleagues at the University of Cambridge has analysed in the last decades group work, interactions and talk in many different classroom settings. This internationally recognised research team in psychology has been developing and testing an intervention programme based on a dialogic approach for fostering collaborative activity and exploratory talk (Kershner, Warwick, Mercer, & Kleine Staarman, 2012). The outcomes of experimental research on the effects of such an

approach has shown that the targeted children who participated in the intervention programme, not only improved the quality of joint reasoning as a group, but they also improved their individual scores, compared with children in control classes (Kershner et al., 2012). The Thinking Together research has widely influenced practice, informed the UK's National Curriculum guidance, and has generated projects in several other countries.

Similarly, the Head Start Project is a programme of the United States Department of Health and Human Services that provides comprehensive early childhood education, health, nutrition, and parent involvement services to low-income children and their families. Since 1964, it is the largest federally funded early childhood education programme in the United States. The Head Start Project is based on the theory of the psychologist Jerome Bruner who has shown that the human mind actively constructs experience, rather than functioning as a passive blank slate, and his insights into child development paved the way for preschool programmes such as Head Start. By 2005 more than 21 million children from low-income families had taken advantage of both from the educational programme and the social, health and nutritional services it provides (Barnett & Hustedt, 2005). The project differs from other programme for children in its aspiration to improve not only the lives of children but transforming communities by giving all their members new chances to be involved in the nurturing and education of children. Head Start became the sole opportunity for parents at risk of poverty to participate in institutional change at the local level while paving the way to offer child-centred services as well (Ellsworth, 1998). Various studies provide evidence on short-term and long-term benefits. Regarding short-term benefits, according to the study of Abbot-Shim et al (Abbott-Shim, Lambert, & McCarty, 2003), results of Head participants in receptive vocabulary and phonemic awareness were salient compared to non-participants. Besides, the former together with their parents also had more positive outcomes in health than the latter. Concerning long-term benefits, meta-analyses of longitudinal studies (Gorey, 2001; Nelson, Westhues, & MacLeod, 2003) and causal modelling (W S Barnett, Young, & Schweinhart, 1998) shed light on the permanent character of the results attained. A high-quality pre-kindergarten education has been linked to improvements in school readiness, requiring skilled teachers to provide safe, caring and supportive environments and to facilitating active engagement and learning across diverse domains (Lloyd & Modlin, 2012). In this regard, professional development has become important in terms of improving the interactions and instructional practices of early childhood teachers (Lloyd & Modlin, 2012). The programme represents an example of how a psychological theory can lead to great advancements in education.

b) Raising the percentage of higher education graduates

There are some examples of interventions that have been carried out by scientists that help raise graduation rates in Higher Education Institutions. Walton and Cohen (2011) designed an intervention that has proven to improve the graduation rates and health status among African-American college students. The researchers identified in the scientific literature that the members of socially stigmatized groups (such as African-Americans) may have more problems with their social belonging in mainstream institutions like school and work (Walton & Cohen, 2011). Moreover, this problem with the sense of belonging might undermine their school and health outcomes, particularly, in transition periods, such as the transition to college. Building on their own previous research (Walton, & Cohen, 2003, 2007; Yeager, D. S. & Walton, 2011), they designed the intervention based on allowing minority students to develop a non-threatening explanation for feelings of non-belonging to school. More specifically, it consisted in exposing students (White and African-American) with “a narrative that framed social adversity in school as shared and short-lived” (Walton & Cohen, 2011, p. 1448) by exposing the outcomes to other students; encouraging students to write an essay describing the way in which their experiences resonate with those from the narratives, and

finally, turn that essay into a videotaped speech becoming a media resource to be showed to prospective students in order to facilitate their transitions to college.

The results of this intervention were very positive for the African-American students who participated in the intervention (for White students it had very little impact), and in their last year of college reduced the gap between their academic outcomes and those of their White classmates by 79%. Moreover, the African-American students that participated in the intervention reported less visits to the doctor (60% of untreated African Americans had seen a doctor recently, only 28% of treated African Americans had) and also the gap in self-reported health was eliminated in the treatment condition (Walton & Cohen, 2011).

c) Identifying actions that lead to improvements in educational outcomes

Researchers from the Abdul Latif Jameel Poverty Action Lab (J-PAL) at the Economics Department at the Massachusetts Institute of Technology developed a study funded by the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD) and the DIME initiative at the World Bank (Banerjee, Banerji, Duflo, Glennerster, & Khemani, 2010) which analysed the effectiveness of three different participatory programmes undertaken by the most prominent educational NGO in India, called Pratham and aimed at universal primary education in India. They made a randomized trial evaluation with a baseline survey in 280 villages in Juanpur district in the state of Uttar Pradesh (UP). Each intervention was implemented in 65 villages, randomly selected out of 280 villages; a fourth group of 85 villages formed the control group. The first and the second interventions were providing information on the channel of interventions available to villagers, and helping citizens gather information on the status of education in their villages, in order to achieve greater involvement of parents in the school systems, or to private responses (e.g. tutoring, volunteering). The third intervention is the one that has been proven the most successful of all three, and it consisted in applying the "Read India" programme developed by the Pratham organization. Thus, this intervention involved offering reading camps for children run by volunteers that had been previously trained by the Pratham' organization staff. The results of this last intervention show the impact that it has been achieved: children in the villages that received intervention 3 are 1.7 percent more likely to read at least letters, 1.8 percent more likely to read words or paragraphs, and 1.7 percent more likely to read stories (Banerjee et al, 2010, p. 24).

In the last decade, several scholars have dedicated research efforts to analyse the positive impacts of particular educational and community based projects, for instance, the Harlem Children's Zone (HCZ), a non-profit organization for poverty-stricken children and families living in Harlem, providing free support in the form of parenting workshops, a pre-school programme, three public charter schools, and child-oriented health programmes for thousands of children and families. Among these scholars, Wilson (Wilson, 2010), for example, describes how the HCZ is providing evidence about how schools in economically marginalized neighbourhoods can create opportunities for their students to reach College and fight against early school leaving. Other scholars, such as Dobbie and Fryer (Dobbie & Fryer, 2011), show how these schools are combining a high-quality education with wider community actions. The evidence they provided demonstrated how HCZ schools are closing the achievement gap in Maths (97% of third graders in HCZ Promise Academy scored at or above grade level) and in Language. Finally, the public benefits estimated for each student that leaves school and graduates in the HCZ is more than \$250,000.

SSH research contributions to the EU 2020 target for social exclusion and poverty

The target set by the EU2020 strategy is to reduce by 20 million the number of people who are living at risk of poverty and social exclusion. Poverty is not only linked to income but also comprises several other dimensions such as health, dwelling, employment and education (Alkire, Roche & Sumner, 2013; Sen, 1999). In this section, we present SSH research projects that provide evidence on the impact they have on these dimensions, thus reducing poverty and social exclusion of the most excluded communities. In this sense, based on the present exploratory analysis, SSH research has contributed to combating poverty in two different ways:

- (a) By informing interventions that contribute to reducing poverty, and,
- (b) By identifying programmes, actions or interventions that contribute to reducing poverty.

The FP6 project “INCLUD-ED -Strategies for Inclusion and Social Cohesion in Europe from Education” analysed educational strategies that contribute to overcoming inequalities and promoting social cohesion and educational strategies that generate social exclusion, particularly focusing on vulnerable and marginalized groups. From 2006 to 2011, researchers from 15 European universities and research institutions, representatives of vulnerable groups, teachers, educators and other professionals, family members and policy makers worked together in order to identify the Successful Educational Actions (SEAs) that contribute to overcoming school failure and early school leaving, as well overcoming the risk of suffering exclusion in other areas. These SEAs have already informed several EU official communications and resolutions (European Parliament resolution of 2 April 2009 on educating the children of migrants; European Parliament resolution of 9 March 2011 on the EU strategy on Roma inclusion) and have been adopted by regional and local government policies, as well as by civil society organizations. The implementation of the SEAs has led to great social improvements in highly deprived areas. For instance, in one of the schools participating in the project, it led to the reduction of absenteeism from 30% to 10% in one year period and occurred only occasionally in a two year period (Aubert, 2011). This European research detected that the implementation of SEAs led to an important decrease in absenteeism rates in schools or high schools located in disadvantaged areas.

The INCLUD-ED project also collected scientific evidence of other benefits of implementing SEAs, for instance, it has also generated a statistically significant increase in instrumental learning for students from disadvantaged contexts. Therefore, in a case studied in Spain, after implementing Interactive Groups (an SEA), the percentage of 4th grade students (primary ed.) with low mathematics achievement in the school decreased by 47% between 2009 and 2011, and those with high levels increased by 18%. Similarly in another Spanish case, the results of 4th grade students (primary ed.) in maths improved from 2008 to 2010, where they moved from score 1 to 3 (out of 5), which is close to the regional average (Valls & Kyriakides, 2013). Additionally, another case analysed by INCLUD-ED in Malta, shows that after implementing an “after school club”, in 2011, students improved in maths and language. Their scores show improvement especially in English (67%) and maths (82%) (Gatt & Armeni, 2012). In this club, parents and children stayed at school together after hours with a professional educator where they were involved together in activities intended to reinforce the contents taught in the classroom. As a result, students’ reading and writing skills have improved, it was also found to be especially important for pupils from disadvantaged backgrounds.

Long-term unemployment goes usually hand in hand with poverty. Very few SSH research projects provide data on the long-term jobs for the most vulnerable population derived from the interventions scholars have developed together with stakeholders (public administration, civic organizations and others) based on research findings. One specific research project that shed light on this endeavour is the FP6 project *INCLUD-ED*. The project was grounded in education but also analysed the contributions education can make on other relevant fields of social inclusion such as Health, Housing and Employment. The research team first identified the most successful actions in generating sustainable and flexible employment. They analysed the case of Mondragon cooperatives in Spain, the seventh largest industrial business group of the country. Flecha and Santa Cruz (2011) identified six Successful Cooperativist Actions that can be transferred to different contexts. Scholars from the INCLUD-ED project participated in the creation of a cooperative in the poorest neighbourhoods in Spain. Scholars initiated a dialogic process with the citizens of the neighbourhood and, after a training session, the citizens (mostly of Roma descent) decided to start a cooperative based on the most successful cooperativist actions identified (Padros, Garcia, de Mello, & Molina, 2011). This cooperative contributed to creating jobs for long-term unemployed people. Since 2011, “the cooperative currently provides stable employment for 11 people, all of whom had previously been in a situation of social exclusion. The cooperative is facing today’s crisis successfully in one of the poorest *barrios* in Southern Europe. It has improved living conditions of many families in the *barrio*. Furthermore, 80 people have been hired in seasonal jobs in the fields and other agricultural sectors” (Flecha, & Soler, 2014).

(b) By identifying programmes, actions or interventions that contribute to reducing poverty

There is a set of SSH research projects that tend to identify actions that have social impact on the ground, describing the main features of these interventions, and providing data on their effectiveness in promoting social improvements. Three examples are provided in what follows as an illustration.

There are several SSH research projects that have identified programmes that contribute to ameliorating poverty, one example being those programmes aimed at improving financial literacy skills. Research has already shown the positive association between financial knowledge and the betterment of household financial decision-making. For instance, Lusardi and Mitchell (2007) found that those who were financially literate when they were young are more likely to plan for retirement and less likely to face poverty in the future. Stango and Zinman (Stango & Zinman, 2007) also demonstrated that those who are unable to correctly calculate interest rates out of a stream of payments end up borrowing more and accumulating lower amounts of wealth. Other scholars found that those who severely underestimated the power of interest compounding are more likely to end up with excessive amounts of debt and raising delinquency rates (Lusardi & Tufano, 2009; Elliehausen, Christopher Lundquist, & Staten, 2007; Hirad & Zorn, 2001). This wealth of research on financial literacy has informed initiatives aimed at improving these types of skills, such as Individual Development Accounts (IDAs), which are matched savings accounts that help people with modest means to save towards the purchase of a lifelong asset, such as a home. One of the features of IDAs is that they require general financial education; participants are offered several hours and sessions of financial education. Financial education proves effective and is associated with a sizable increase in savings among IDA participants (Schreiner & Sherraden, 2007). Mills et al. (2008) demonstrated in the first experimental evidence study on IDA how people who have been participating in IDA literacy training increased their homeownership and reduced non-retirement financial assets.

A second example is found among those research projects that have analysed the *Bolsa Familia* programme (Brazil) as an effective political intervention in the reduction of poverty.

The basis for the programme draws from the *Bolsa Escola* programme, initiated in 1995 to fight against poverty, and which is still ongoing. It consists of a basic income for the poorest families that is designed to keep their children in education and to provide them with access to vaccines. The aim of this Conditional Cash Transfer programme was to keep children in education by preventing families from sending their children to work before finishing compulsory education. The programme was extended to the whole Brazil as part of the *Fame Cero* programme of the Lula Government (2003-2011), and the targets of the programme were extended beyond education and social improvements in all areas were reported. Among the studies conducted, Shei et al. (Shei, Costa, Reis, & Ko, 2014) conducted quantitative research focused on health improvements, comprising a household survey of approximately 567 families, including beneficiaries and non-beneficiaries of the programme. They found that the programme significantly improved the utilization of preventive health care services by poor beneficiaries' families. On their side, De Lima Amaral and Do Prado Monteiro (2013) also used a quantitative methodology, to analyse the school drop out of children around the country between the years 2004 and 2005, and absenteeism. This research concluded that the school dropout rate is lower among children from beneficiary families, highlighting that the programme is globally responsible for a 57.2% reduction in the dropout rate. Besides health and education access, De Brauw, A et al. (2014) focused on the programme effects on gender equality. They interviewed 15,426 households, which were part of a study promoted by the Federal Ministry of Social Development of Brazil and the findings reveal the overall positive effects of *Bolsa Familia*, with a clear reinforcement of women's power in decision making processes at the household level.

The last example is found in Jones and Lansdverk's (Jones & Lansdverk, 2006) study which analysed a programme called "The Academy", a long-term placement for discharged foster youths that have no other placement options and no chance of returning to their biological families. The Academy was created in 2001 in response to the findings of previous studies and the experience of professionals, that showed alarming numbers of youths leaving foster care without suitable education or life skills, which prevented them from being independent as adults. First established in San Diego and further expanded in other cities and states, it provided a stable home and a comprehensive educational programme. Jones and Lansdverk's research was carried out between 2001 and 2006 with a sequential cohort design, allowing them to monitor the first three graduating classes. They analysed student academic records and school documentation, conducted interviews to assess the behavioural, emotional and social competences and problems affecting the youths, school performance and post-academy functioning. Among their results, Jones and Lansdverk found that criminality and victimization behaviours were not identified among the foster youths which contrasts with data identified in other studies. Instead, engagement is a major trend in the community among families and peers. Finally, 28% of the youths interviewed six months after leaving the Academy reported attending college, half of them being in a four-year college. This figure is significantly higher than the one indicated by Courtney et al (Courtney, Dworsky, Keller, Havilicek, & Bost, 2005) (only 7% of attendance among former foster youths).

6. What do we know about how SSH research has contributed to the ERA?

6.1. SSH research on Strengthening the ERA

The 2020 Vision for the European Research Area encapsulates six building blocks: a single labour market for researchers; excellent research institutions and universities; world-class research infrastructures; joint programming in research; effective knowledge sharing; opening of the ERA to the world (European Commission, 2009). In this section we focus on literature more related to praxes for research evaluation systems, developed by different actors and related to national and transnational policy levels. In this section, the scope of analysis is limited to some examples of research projects addressing the analysis of the main features of the European Research Area. A more extensive analysis will be covered in later work during the IMPACT-EV project.

In recent decades, we have witnessed relevant changes in the structure of SSH research praxis with a variety of implications for research evaluation systems. In recent years, for example, multi-university collaboration has been identified as the fastest growing type of co-authorship worldwide. This type of collaboration has been proven to produce the highest-impact papers when top tier universities are present (Jones, Wuchty & Uzzi, 2008). At the same time, it challenges traditional solo-authored and disciplinary modes of scientific production within SSH. This paradigm shift in research practice is a recent reality that should be taken into account by research evaluation systems and methods. Other studies have also shown how co-authorship or multi-authorship increased in recent decades especially in scientific publications indexed in the *Science Citation Index* (Glänzel & Schubert, 2004). An analysis of FP5 and FP6 Integrated Projects and Networks of Excellence on Aerospace, Energy and Environment, ICT and Life Sciences concluded that ERA funding and activities (such as FP Collaborative projects, ERA-NETs and ERA-NETs plus and many others) boosted international collaboration as well as the global impact of European centres of excellence that would otherwise not have had the same visibility. At the same time, the majority of these actions are reported to reduce duplication and fragmentation of activities by joining research calls and promoting these collaborative and transnational research teams (Harrap & Boden, 2012).

The use of scientific and other networks and the Web has increased among all sciences. Network analysis on research impact and information flows has thus become a relevant cutting edge issue (Barabási et al., 2002). For instance, the SISOB project - an ongoing FP7 Science and Society action - focuses on how to develop measurements of production and distribution networks as well as how to relate these measurements to outcomes (SISOB, 2011-2013). At the same time, the most recent developments in network analysis among academics tell us how researchers can improve their productivity (in grant seeking for example) by being embedded in networks with more experienced researchers. Gender also plays a key role, and institutions should bear in mind, for instance, that women have more difficulties than men in obtaining research grants (Rawlings & McFarland, 2011). This analysis shows us that researchers observe and learn from their contacts and more experienced peers, and therefore institutional initiatives focusing on the promotion of these ties and behaviours do not only affect scientific productivity. In addition, these networks and

mentoring activities should also be promoted for building effective relationships among scholars who work to further developing methods to improve SSH research impact on society and politics.

These patterns in “top-down” funding initiatives can vary greatly among sciences. For instance, in Physics, collaboration and co-authorship tends to be greater than other disciplines (Mattson, Laget, Nilsson Vinderfjärd & Sundberg, 2010). It also depends on the thematic priorities (Heller-Schuh et al., 2011). For the IMPACT-EV project, more in-depth knowledge about co-authorship or sole authorship dynamics- within the different fields of SSH and across time - will be relevant for carrying out *ex-ante* and *ex-post* evaluation processes and for adopting realistic evaluation standards. For example, there are studies that have shown differences between the Social Sciences and the Humanities in publication trends. We have evidence for the last decade that shows how the annual number of articles, especially in English, is increasing more in the Social Sciences than in the Humanities. This trend is related to how research evaluation schemes have provided funding rewards linked to research outcomes published in Web of Science indexed journals (Engels, Ossenblok & Spruyt, 2012).

When referring to interdisciplinarity, research has indicated the growth of “top-down” and “bottom-up” initiatives in promoting interdisciplinary research teams over recent decades. Many interdisciplinary projects have produced limited gains in blurring disciplinary boundaries and in creating transdisciplinary research (Jacobs & Frickel, 2009). This situation could be influenced by the lack of good methods for identifying how interdisciplinarity affects productivity (Anzai, Kusama, Kodama & Sengoku, 2012). Scholars highlight that there are Success Stories worldwide that can provide useful information on interdisciplinarity and transdisciplinarity patterns, and generating new fields of research. In this sense, there is a need to conduct further research on what types of interdisciplinary or transdisciplinary research centres and/or projects are the most dynamic and enduring (Jacobs & Frickel, 2009).

Important efforts have been made to assess the impact of researchers' mobility and career. Ivancheva and Gourova (2011) focused on researchers' mobility, and presented some results from their project 'European Career of Researchers' (E*CARE), funded by the FP7 Programme. The aim of the project was to identify opportunities and barriers for international mobility and career development for researchers from eight European countries (Austria, Bulgaria, Czech Republic, Cyprus, Greece, Hungary, Slovakia and Switzerland) and to 'assess the general working environment, the forms and quality of services for mobile researchers and the degree of available international support; to reveal the main characteristics of international mobility of European scientists and its major impacts and consequences for them', and so provide advice to EU policymakers (Ivancheva & Gourova, 2011, p. 187). Rather than focus on the idea of the 'brain drain', they sought to focus on what might 'create conditions for brain gain and brain circulation in Europe, effectively removing obstacles to international mobility' (Ivancheva & Gourova, 2011, p. 189).

Two questionnaires were circulated: (1) to researchers (PhD students, postdoctoral researchers, experienced researchers, university lecturers, etc.), and there were 869 respondents; (2) to stakeholders (representatives of industry, research organisations, non-governmental organisations, public bodies, etc.), with 313 respondents. Ivancheva and Gourova found that the greatest barriers to researchers' mobility were: research funding, housing and accommodation, bureaucracy within the host institution, employment services and employment legislation (Ivancheva & Gourova, 2011, p. 196). Their key recommendations included: (1) that specific EU initiatives intended to support the mobility of

European researchers, such as the EURAXESS Jobs Portal and EURAXESS Service Centre, were not well known within the research community, and so their visibility needed to be raised along with awareness of their potential to support mobility and career development; and (2) EU policies for researcher mobility needed to be backed up by Member State initiatives to be effective, including national action plans with clear success indicators (Ivancheva & Gourova, 2011, p. 197).

More recently the POCARIM project (Mapping Population Careers, Mobilities and Impacts of Advanced Research Degree Graduates in the Social Sciences and Humanities) developed under FP7 investigates the career paths, employment patterns and contributions of doctoral graduates in the Humanities and Social Sciences in Europe. The project sought to identify the population of Humanities and Social Science researchers, its core characteristics, and trends in employment and mobility, identifying the diversity of post-doctoral career paths of Humanities and Social Science graduates, and assessing the contribution of this group of researchers to Europe's knowledge-based economy and society. Results outlined both factors facilitating and impeding PhD holders in these fields to have an impact.

The issue of mobility and career prospects will be extensively covered in WP7 of the IMPACT-EV Project, which will review in detail initiatives under the People Programme, reviews, and impact assessments.

6.2 Evaluation of SSH research impact on strengthening the ERA

There are many studies focusing on the features, inefficiencies and proposals for **improving SSH research evaluation systems**. Some analyses at the **transnational** level highlight that one of the challenges for the European Union is to have a common, sustainable and feasible research evaluation system for the Social Sciences and Humanities. This system should take into account the specificities of, and variety across, disciplines and national funding agencies; but it should also take into account clear impact indicators that could go beyond scientific dissemination to broader social and political impacts. In this sense, some projects have identified difficulties that the majority of SSH research projects face with regard to social and political impact. According to the FLASH-IT project (APRE, 2011-2013), the current deficient dissemination process for research results and political recommendations from SSH research projects in the Framework Programme should be replaced by better connecting of these results with political recommendations, especially those related to Europe 2020 targets. One possibility is to connect researchers and academics with different types of decision-makers (European, national and regional), civil society organisations, journalists and other stakeholders. Another FP7 project, METRIS, identifies current emerging trends within SSH that challenge research evaluation (Viola, 2011-2013). In their recommendations, METRIS researchers highlight the need for a better understanding of the European Research Area, its organisation and its funding impact especially in SSH (by taking into account different meanings of excellence between disciplines and national funding research agencies). This project also recommends that the European Commission “develop monitoring capacities for reliable statistics on public and private funding of the SSH” as well as fostering new publishing dynamics (such as the extension of Open Access scientific journals and repositories) which can improve research dissemination (European Commission, 2009). At the same time, while SSH research in all observed countries underwent an *ex-ante* processes of evaluation and selection based on peer review, there are more countries that do not perform *ex-post* evaluations (especially new Member States) than those that do (Roman & Peter, 2012). At the national level, regarding the *ex-ante* processes of research evaluation the work of Michele Lamont and her colleagues (2009) has been

relevant in identifying the existing dynamics, features and culture behind peer review for fellowships and research grants for academics in the United States, Canada, France, the United Kingdom and Finland. This study analysed the context of evaluation, definition of standards and boundaries of what is considered excellent within academia. There is no consensus however about the definition of excellence among SSH scholars. Lamont and her colleagues identified how panellists engage in a deliberative process of decision-making about the research grant proposals submitted. Panellists may take into consideration the different set of political priorities recommended by the research funding agency and, while they are not forced to take them into consideration, they often discuss the political implications of the proposals presented. Panellists within SSH also attribute different meanings to originality besides excellence, but not usually conditioned by disciplines (Guetzkow, Lamont, & Mallard, 2004). In addition, the comparative research on peer evaluation systems in the Social Sciences and Humanities in different countries such as the UK, US, France, and Finland shows how evaluation rules vary across countries and systems. Different typologies of research systems could be made. In this sense, Mallard and Lamont labelled the United States system as the professional system of evaluation because it relies only on peer evaluation and fairness; the UK system as the managerial system because it relies on evaluation from the government and it is opposed to scholars and to control; and the French system as the post-corporatist system because it relies on hiring decisions, on who is submitting the proposal and the position he/she holds (Lamont & Mallard, 2005). Thus, the evaluative tool or technique imposed by the research-funding agency would and should condition behaviour, not the values and beliefs of panellists (Lamont & Huutoniemi, 2011).

Regarding other national contexts, the Netherlands Research Council for the Economic and Social Sciences (NaGW-NWO), for instance, applies a mixed method process of research grant selection by combining peer review, past performance measures based on citations, and panel review. Scholars observed that final decisions on the best proposals were made on the basis of the external peer-review process and other contingent decision elements rather than on scientometric indicators of past performance (Van den Besselaar & Leydesdorff, 2009). In contrast, some authors in the field of scientometrics have proposed some tools to evaluate research impact in comparison with other national or regional contexts. Leydesdorff (2013), for example, proposed the Integrated Impact factor (I3) for evaluating research impact applied to institutions as well as to journals. It is a non-parametric indicator and a tool comparable to the one developed by the NSF (National Science Foundation) in the United States in six rank items.

In this context of constant changes in the field, some governments have developed some changes in their evaluation system policies because if they relied exclusively on peer review evaluations without having clear guidelines, the selected projects would not address political and social agenda aims. In order to address this fact, the Canadian government, for example, conducted research to improve guidelines for reviewers' panellists and for all the participating agents in the SSH research evaluation process (Lamont & Mallard, 2005). In this respect an interesting development is the creation of transnational evaluation systems such as EC DGs, Research Council, European Science Foundation, in the context of ERA, and more recently the H2020 Vademecum, Section on Proposal Submission and Evaluation (European Commission, 2014).

One interesting document is The Standard Evaluation Protocol (SEP), which describes the methods used to assess research conducted at Dutch universities and NWO and Academy institutes every six years, as well as the aims of such assessments. The focus in the SEP has changed from the productivity of researchers to the quality of their output. Besides, relevance to society has become more important as an assessment criterion. As a spin-off of

recent discussions on data manipulation in the Netherlands, research integrity has also become an important issue in this new SEP.

As a result, the SEP 2015-2021 consists of three major evaluation criteria: research quality, relevance to society, and viability. In this way a 'publish or perish' culture becomes less dominant. Relevance to society is measured by evaluation of the research output, the re-use of the output (software; datasets; publications and so on), and of demonstrable marks of recognition by peers and societal groups.

Several other examples could be provided about the changing methods of research evaluation systems, where the improvement of integration towards the ERA is a key item in this context, although the mentioned practices were not specifically devoted to SSH research. Nonetheless, the analysis of these national practices is outside the scope of this Report.²⁶

7. Conclusions

This Report presents the state of the art in scientific, political and social impact of SSH research, and its evaluation, by the way of an extensive literature review of research impact. Information has been retrieved from journals with the highest impact factor and indexed in databases of the various fields and disciplines (i.e. WoS, SCOPUS), the most relevant books and research reports, and other grey literature from relevant evaluation institutions. Furthermore, scientific discussions, contributions, approaches and new tools have been taken into account. The main focus of the literature search covered from 2006 to 2012. Several contributions related to periods before 2006 have been included as well, based on their relevance.

Scientific impact

The literature on scientific impact is dominated by interest in bibliometric analyses, indicators and tools focusing on publications and other research outputs (mainly patents). In the mid-80s, changes in scholarly communication emerged, and practices informed by open access principles gained more attention from the epistemic communities and from research institutions, new indicators were developed, based on the web (webindicators, webometrics, cybermetrics, altmetrics), as well as new approaches in scientometrics (e.g. h-Index and g-index). The most recent tools have the potential for measuring the impact and outreach of articles, which allow benchmarking their impact not only with citations, but also tracking the different uses of papers (cut/paste activities, citations in media reports, online newspapers, peer review discussions, blogs, etc.).

Generally speaking, changes in scholarly communication are likely to transform and improve our capability to understand the scientific impact of research outputs, going beyond the simple paper publication. The diffusion of open access practices are supposed to further reinforce the capability to monitor scientific impact.

²⁶ A more systematic analysis of the national evaluation systems, and how they deal with the evaluation of the impact of SSH research, is developed within the WP2 of the IMPACT-EV Project.

As for SSH, despite efforts to use WoS data and scientometric techniques, and the transformation of some disciplinary fields, the analysis of the literature found many shortcomings with the methods and solutions proposed, which largely agree that assessments based on bibliometric resources generally underestimate the value of the SSH research outputs. Alternative metrics, methods and data sources have been explored in order to understand their potential for scientific impact assessment; negative consequences for the quality of research due to the extensive use of bibliometrics concerns scholars and countries where qualitative research traditions are prevalent. All in all, bibliometrics is largely considered no more than one resource among many others for scientific impact assessment, which can provide better results when used in combination with other metrics and information sources than when it is used as sole tool. This evidence raises the problem of choosing the most suitable mixed method for impact assessment.

Political impact

The assessment of the political impact of research is an important topic, which gained a momentum in Europe, especially in investigating the relationship between science and policy, and how to improve the impact of the results of research on the policy process. Its specific features relate to the fact that it deals with transformations produced in policy development and the policy process (motivations and rationales, policy design, policy implementation, policy assessment).

Participation and public engagement of researchers and stakeholders in policy making is considered one key element to strengthening the impact of research on decision-making; also the co-production of research between academics and policy makers has been assessed as a promising practice, which is likely to lead to greater political impact.

These results also apply in the case of SSH research; moreover, plenty of evidence emerged in the literature on the contribution and impact of SSH research on policies related to social problems like exclusion, gender discrimination, and other relevant social challenges. To this end, the contribution of the Science and/in Society actions, funding research projects dealing with governance and public engagement, have been of crucial importance.

Social impact

The attention that social impact has received is increasing in recent times. Research efforts have overcome the deterministic linear model that foresees the impact as a natural destination of all research results soon after they are delivered; they also support a more precise conceptual distinction of impact from other activities, like dissemination or knowledge transfer, etc.

The present report provides several examples of SSH research producing social impact, including areas that are key components of the EU2020 strategy (e.g. employment, education, social exclusion, poverty, researchers' career and mobility, knowledge-based society). There are several ways by which SSH research with impact has been identified. For example, when researchers generate interventions based on research findings and provide evidence on its social improvements, or when researchers identify actions that are having a positive impact on society and analyse their features to create possibilities for transferability.

Strengthening of the ERA

SSH research also contributes to the structure of the ERA, with a view to strengthening the integration of research activities, projects, programmes and government agendas. This

contribution is documented by several pieces of literature discussing changes in the modes of knowledge production, and the new relevance of transnational and transdisciplinary approaches in research projects, and collaboration in research outputs (from sole-authored to co-authored research, from disciplinary oriented to interdisciplinary oriented research).

Change has also affected SSH research evaluation systems at the transnational level and the national level, to overcome inefficiencies and limitations in understanding the type of impacts that research outputs actually have achieved.

Summing up, despite SSH being so differentiated that no one single evaluation method is suitable to assess the impact of the research produced within the different disciplinary fields, the analysis of the literature showed that:

- Both new metrics and mixed methods of evaluation are promising tools for research impact assessment;
- Evidence of actual impacts of SSH research for solving societal challenges, and for improving policy-making, is available. Thus, the transformative dimension of impact is quite well documented, and can be taken into account for further developments in the IMPACT-EV project;
- The literature analysis documented the European Framework Programme's prominent role in sustaining and promoting research projects whose results allow the production of important knowledge advances on the topic of research impact, including SSH disciplinary fields. Research results also illustrated important outputs coming from EU Framework Programmes for promoting integration toward the building of the ERA;
- The exclusionary dimension, which identifies the elements and factors that lead to a lack of impact (scientific, political and social) of the research, is not well outlined in the literature. One possible explanation might be that scholars generally consider those outputs to be publishable where successful positive achievements are represented; another explanation might be the fact that the research aims to show unintended and eventually negative consequences of existing assessment approaches rather than searching for explanations of the lack of impact. Thus, the exclusionary dimension is a topic deserving special attention in the development of IMPACT-EV analyses.

The literature review has demonstrated that there is plenty of room for further improvements in methods, techniques, metrics, and methodologies to better grasp the impact of research in SSH fields. The IMPACT-EV project will contribute towards filling this gap.

8. References

- Abbott-Shim, M., Lambert, R., & McCarty, F. (2003). A comparison of school readiness outcomes for children randomly assigned to a Head Start programme and the programme's wait list. *Journal of Education for Students Placed at Risk*, 8(2), 191-214.
- Aguillo, I.F., Bar-Ilan, J., Levene, M. & Ortega, J.L. (2010). Comparing university rankings. *Scientometrics*, 85, 243-256.
- Aguillo, I. F., Granadino, B., Ortega, J. L., & Prieto, J. A. (2006). Scientific research activity and communication measured with cybermetric indicators. *Journal of the American Society of Information Science and Technology*, 57(10), 1296-1302.
- Alkire, S., Roche, J.M., & Sumner, A. (2013). Where Do the World's Multidimensionally Poor People Live? OPHI Working Paper 61.
- Antonacopoulou, E. P. (2010). Beyond Co-production: Practice-relevant scholarship as a foundation for delivering impact through powerful ideas. *Public Money and Management*, 30(4), 219-226.
- Anzai, T., Kusama, R., Kodama, H., & Sengoku, S. (2012). Holistic observation and monitoring of the impact of interdisciplinary academic research projects: an empirical assessment in Japan. *Technovation*, 32, 345-357. doi: 10.1016/j.technovation.2011.12.003
- APRE (Agenzia per la promozione Della Ricerca Europea). (2011 – 2013). *FLASH-IT Facilitating access to socio-economic research through information and communications technology*. 7th Framework Programme. Activity 8. Horizontal actions. European Commission.
- Armstrong, F., & Alsop, A. (2010). Debate: Co-production can contribute to research impact in the social sciences. *Public Money and Management*, 30(4), 208-210.
- Arnold, E. (2004). Evaluating research and innovation policy: a systems world needs systems evaluations. *Research Evaluation*, 13(1), 3-17. doi:10.3152/147154404781776509
- Arnold, E. (2012). Understanding long-term impacts of R&D funding: The EU framework programme. *Research Evaluation*, 21(5), 332-343.
- Aubert, A. (2011). Moving beyond social exclusion through dialogue. *International Studies in Sociology of Education*, 21(1), 63–75. doi:10.1080/09620214.2011.543854.
- Avramov, D. (1996). *Homelessness in the European Union: Social and Legal Context of Housing Exclusion in the 1990s*. Brussels: FEANTSA.
- Banerjee, A. V, Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of Participatory Programmes: Evidence from a Randomized Evaluation in Education in India. *American Economic Journal: Economic Policy*, 2(1), 1–30. doi:10.1257/pol.2.1.1
- Barabási, A.L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, 311, 590-614.
- Barker, K. (2007). The UK Research Assessment Exercise: the evolution of a national research evaluation system. *Research Evaluation*, 16(1), 3-12.

- Barnett, W. S., Young, J. W., & Schweinhart, L. J. (1998). How preschool education influences long-term cognitive development and school success: A causal model. In W. S. Barnett and S. S. Boocock (Eds.), *Early care and education for children in poverty: Promises, programmes, and long-term results* (pp. 167–184). SUNY Press.
- Barnett, W. S., & Hustedt, J. T. (2005). Head Start 's Lasting Benefits. *Infants and Young Children*, 18(1), 16–24.
- Barré, R. (2010). Towards socially robust S&T indicators: indicators as debatable devices, enabling collective learning. *Research Evaluation*, 19(3), 227-231.
- Beck, U. (1992). *Risk Society. Towards a New Modernity*. London: SAGE.
- Boaz, A., Fitzpatrick, S., & Shaw, B. (2009). Assessing the impact of research on policy: a literature review. *Science and Public Policy*, 36 (4), 255-270.
- Boaz, A., & Ashby, D. (2003). *Fit for purpose ? Assessing research quality for evidence based policy and practice*, ESRC UK Centre for Evidence Based Policy and Practice, Working Paper 11.
- Bonitz, M., Bruckner, E., & Scharnhorst, A. (1997). Characteristics and impact of the Matthew effect for countries. *Scientometrics*, 40(3), 407-422.
- Borgman, C. L. (Ed) (1990). *Scholarly communication and bibliometrics*. Newbury Park: Sage.
- Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, Mass: MIT Press.
- Bornmann, L. (2013), What is societal impact of research and how can it be assessed? a literature survey. *Journal of the American Society for Information Science and Technology*, 64, 2: 217-233.
- Bornmann, L. (2014). *Measuring the broader impact of research: The potential of altmetrics*. Digital Libraries; Physics and Society; Applications. Retrieved from <http://arxiv.org/abs/1406.7091>
- Braun M., Leopori B., Reale E., Slipersaeter S., Kaloudis A., Filliatreau G., & Laredo P. (2009). *Tools and Indicators for Community Research Evaluation and Monitoring*. Final Report, EC.
- Braun, T., Glnzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, 69(1), 169-173.
- Brewer, J.D. (2011). The impact of impact. *Research Evaluation*, 20(3), 255-256.
- Brown, C. (2012). The 'policy-preferences model': a new perspective on how researchers can facilitate the take-up of evidence by educational policy makers, *Evidence and policy*, 4 (8), 455-472.
- Bulger, M.E., Meyer, E.T., De la Flor, G., Terras, M., Wyatt, S., Jirotko, M., Eccles, K., & McCarthy Madsen, C. (2011). *Reinventing Research? Information Practices in the Humanities. A Research Information Network Report*. Available at SSRN: <http://ssrn.com/abstract=1859267> or <http://dx.doi.org/10.2139/ssrn.1859267>
- Burdge, R., & Vanclay, F. (1995). Social impact assessment. In *Environmental and Social impact Assessment*, Vanclay, F., and Bronstein, D. (Eds.). Chichester: Wiley.
- Butler, L. (2002). A List of Published Papers in No Measure of Value. *Nature*, 419(6910), 877.
- Butler, L. (2003). Explaining Australia's Increased Share of ISI publications – The effects of a funding formula based on publication counts. *Research Policy*, 32(1): 143-155.

- Butler, L. (2004). *What Happens when Funding is Linked to Publication Counts?* in H.F. Moed et al. (Eds.), *Handbook of Quantitative Science and Technology Research*, Springer, Netherlands.
- Butler, L. & McAllister, I. (2009). Metrics or Peer Review? Evaluating the 2001 UK Research Assessment Exercise in Political Science. *Political Studies Review*, 7, 3-17.
- Butler, L. & McAllister, I. (2011). Evaluating University Research Performance Using Metrics. *European Political Science*, 7(1), 30-17.
- Butts, J.A., Roman, J.K., & Lynn-Whaley, J. (2011). *Varieties of Juvenile Court. Nonspecialized Courts, Teen Courts, Drug Courts, and Mental Health Courts*. In *Juvenile Justice: Advancing Research, Policy, and Practice*. Francine T. Sherman and Francine H. Jacobs (Eds.). Hoboken, NJ: John Wiley and Sons.
- Buxton, M., & Hanney, S. (1994). *Assessing Payback from Department of Health Research and Development: Preliminary Report. Volume 1: The Main Report*. HERG Research Report, NO.19. Uxbridge: HERG, Brunel University.
- Buxton, M., & Hanney, S. (1996). How Can Payback from Health Services Research Be Assessed? *Journal of Health Services Research and Policy*, 1(1), 35-43.
- Buxton, M., & Hanney, S. (1998). Evaluating the NHS research and development programme: will the programme give value for money? *Journal of the Royal Society of Medicine*, 91 Supl. 3 (35), 2-6.
- Buxton, M., Hanney, S., Packwood, T., Roberts, S. & Youll, P. (2000). Assessing benefits from Department of Health and National Health Service Research and Development, *Public money and Management*, 20(4), 29-34.
- Buxton, M. (2011). The payback of 'Payback': challenges in assessing research impact. *Research Evaluation*, 20(3), 259–260. doi:10.3152/095820211X13118583635837
- Carbonell, E., Bermúdez de Castro, J. M., Arsuaga, J. L., Díez, J. C., Rosas, A., Cuenca-Bescós, G., Rodríguez, X. P. (1995). Lower Pleistocene Hominids and Artifacts from Atapuerca-TD6 (Spain). *Science*, 269, 826–832.
- Chambers, D., Grant, R., Warren, E., Pearson, S.A., & Wilson, P. (2012). Use of evidence from systematic reviews to inform commissioning decisions: a case study. *Policy and Evidence*, 8 (2), 141-148.
- Cotterill, S. & Richardson, L. (2010). Expanding the Use of Experiments on Civic Behaviour: Experiments with local government as a research partner. *The Annals of the American Academy of Political and Social Science*, 628, 148-164.
- Courtney, M., Dworsky, A., Keller, R., Havlicek, R., & Bost, N. (2005). *Midwest evaluation of the adult functioning of former foster youth: outcomes at age 21*. Chicago: Chapin Hall Center for Children.
- De Bellis N. (2009). *Bibliometrics and citation analysis: From the Science citation index to cybermetrics*. Lanham, Md: Scarecrow Press.
- De Brauw, A., Gilligan, D.O., Hoddinott, J.F., & Shalini, R. (2014). The impact of Bolsa Família on women's decision-making power. *World Development*, 59, 487-504.
- De Jong, S., van Arensbergen, P., Daemen, F., van der Meulen, B., & van den Besselaar, P. (2011). Evaluation of research in context: an approach and two cases. *Research Evaluation*, 20(1), 61-72.
- De Jong, S., Barker, K., Cox, D., Sveinsdottir, T., & Van den Besselaar, P. (2014). Understanding societal impact through productive interactions: ICT research as a case. *Research Evaluation*, 23(2), 89–102.

- De Lima Amaral, E.F., & Do Prado Monteiro, V. (2013). An evaluation of the impact of the educational conditions of Brazil's Bolsa Familia Programme (2005 and 2009), *DADOS: Revista de Ciências Sociais*, 56(3), 531.
- Desjardins, E. (2010). *The Urban Food Desert: Spatial Inequality or opportunity for change?* In A. B. Palmer (Ed.), *Imagining Sustainable Food Systems. Theory and Practice* (pp. 87–112). Ashgate Publishing Limited.
- Dibb, S. & Quinn, L. (2010). Debate: Research impact or career progression?. *Public Money and Management*, 30(6), 326-328.
- Dobbie, W., and Fryer, R.G. (2011). Are High Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 3(3), 158-87. doi: 10.1257/app.3.3.158
- Donovan, C. & Larkin, P. (2006). The Problem of Political Science and Practical Politics. *Political Studies*, 26(1), 11-17.
- Donovan, C. (2007a). Guest Editor's Introduction: Future Pathways for Science Policy and Research Assessment: Metrics vs peer review, quality vs impact. *Science and Public Policy*, 34(8), 538-542.
- Donovan, C. (2007b). The Hidden Perils of Citation Counting for Australasian Political Science. *Australian Journal of Political Science*, 42(4), 665-678.
- Donovan, C. (2007c). The Qualitative Future of Research Evaluation. *Science and Public Policy*, 34(8), 585-597.
- Donovan, C. (2008). The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research. *New Directions for Evaluation*, 118, 47-60.
- Donovan, C. (2009). Gradgrinding the Social Sciences: The politics of metrics of political science. *Political Studies Review*, 7, 73-83.
- Donovan, C. (2011). State of the Art in Assessing Research Impact: Introduction to a special issue. *Research Evaluation*, 20(3), 175-179.
- Donovan, C. & Hanney, S. (2011). The 'Payback Framework' explained. *Research Evaluation*, 20(3), 181-183.
- Duijn, M., Rijnveld, M. & van Hulst, M. (2010). Meeting in the Middle: Joining reflection and action in complex public sector projects. *Public Money and Management*, 30(4), 227-233.
- Edler, J., & Flanagan, K. (2011). Indicator needs for the internationalisation of science policies. *Research Evaluation*, 20(1), 7-17.
- Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics*, 69(1), 131-152, doi:10.1007/s11192-006-0144
- Elkana, Y. (Ed). (1978). *Toward a metric of science: The advent of science indicators*. New York: Wiley.
- Elliehausen, G., Lundquist, E. C. & Staten, M. E. (2007). The Impact of Credit Counseling on Subsequent Borrower Behavior. *Journal of Consumer Affairs*, 41(1), 1–28. doi:10.1111/j.1745-6606.2006.00066.x
- Ellsworth, J. & Ames, L. J. (Ed.). (1998). *Critical Perspectives on Project Head Start: Revisioning the Hope and Challenge*. Albany, NY: SUNY Press.
- Engels, T.C.E., Ossenblok, T.L.B., & Spruyt, E.H.J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000-2009. *Scientometrics*, iFirst article. doi:10.1007/s11192-012-0680-2

- Ernø-Kjølhede, E., & Hansson, F. (2011). Measuring research performance during a changing relationship between science and society. *Research Evaluation*, 20(2), 130-142. doi: 10.3152/095820211X12941371876544
- EURECIA (2012). *Understanding and assessing the impact and outcome of the ERC and its funding schemes*. EC, available at: <http://www.erc.europa.eu>
- European Commission (2005). *Assessing the Social and Environmental Impacts of European Research*. Brussels: Directorate-General for Research, European Communities.
- European Commission (2009). *The European Research Area Partnership*. Brussels: Directorate-General for Research, European Communities.
- European Commission (2014). *H2020 Vademecum: Section on Proposal Submission and Evaluation*. Version adopted by the Commission on 30 April 2014, C/2014/2746.
- Evaluating Research in Context (ERiC). (2010), *Evaluating the societal relevance of academic research: A guide*. Delft, The Netherlands: Delft University of Technology.
- Evers, A., Ewert, B., & Brandsen, T. (2014). *Social Innovations for Social Cohesion. Transnational patterns and approaches from 20 European cities*. (pp. 1–420). Liege: EMES European Research Network.
- Fenner, M. (2014). *Altmetrics and Other Novel Measures for Scientific Impact*. In Bartling, S. and Friesike, S. *Opening Science*. (pp. 179-189). Springer: Dordrecht.
- Flecha, R. & Santa Cruz, I. (2011) Cooperation for economic success: the Mondragon case. *Analyse and Kritik*, 1, 157-170.
- Flecha, R., & Soler, M. (2014). Communicative Methodology: Successful Actions and Dialogic Politics. *Current Sociology*, 62 (2), 232-242.
- Frolich, N. (2011). Multi-layered Accountability. Performance-based funding of universities. *Public Administration*, 89(3), 840-859.
- Gabolde, J. (1998). New challenges for indicators in science and technology policy-making: a European view. *Research Evaluation*, 7(2), 99-104.
- Gantman, E. R. (2012). Economic, Linguistic and Political Factors in the Scientific Productivity of Countries'. *Scientometrics*, 93(3), 967-985.
- Garfield, E. (1962-1993). *Essays of an information scientist*. Vol. 1-15. Philadelphia: ISI Press.
- Garfield, E. (1979). *Citation indexing - its theory and application in science, technology, and humanities*. New York: Wiley.
- Garfield, E. (1980). Is Information Retrieval in the Arts and Humanities Inherently Different from That in Science? *The Library Quarterly*, 50, 40–57. Retrieved from <http://www.jstor.org/stable/4307182>.
- Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA: The Journal of the American Medical Association*, 295(1), 90–3. doi:10.1001/jama.295.1.90
- Gatt, S., & Armeni, L. S. (2012). The Community as a Resource to Educate Primary School Children. *Literacy Information and Computer Education Journal*, 3(2), 560–567.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge. The dynamics of Science and Research in Contemporary Societies*. SAGE Publication.
- Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171–193. doi:10.1023/A:1014848323806

- Glänzel, W., & Schubert, A. (2004). *Analysing Scientific Networks through Co-authorship*. In H. Moed; W. Glänzel; U. Schmoch (Eds.). *Handbook of Quantitative Science and Technology Research* (pp. 257-276). Amsterdam: Kluwer.
- Godin, B. (2004). *Measurement and Statistics on Science and Technology: 1930 to the Present*. London: Routledge.
- Godin, B., & C. Dore.(2005). Measuring the impacts of science; beyond the economic dimension, INRS Urbanisation, Culture et Sociult. Paper presented at the HIST Lecture, Helsinki Institute for Science and Technology Studies, Helsinki, Finland.
- Gómez, A., Puigvert, L., & Flecha, R. (2011). Critical Communicative Methodology: Informing Real Social Transformation Through Research. *Qualitative Inquiry*, 17(3), 235-245. doi: 10.1177/1077800410397802
- González-Pereira, B., Guerrero-Bote, V.P., & Moya-Arnégón, F. (2010). A new approach to the metric of journals' scientific prestige: the SJR indicator. *Journal of Informetrics*, 4(3), 379-391. doi: 10.1016/j.joi.2010.03.002
- Gorey, K. M. (2001). Early childhood education : A meta-analytic affirmation of the Short- and Long-Term Benefits of Educational Opportunity. *School Psychology Quarterly*, 16(1), 9–30.
- Gray, A., Broadbent, J. & Lavender, M. (2009). Editorial: PMM, the impact of research on practice, and the emerging Research Excellence Framework. *Public Money and Management*, 29(3): 139-140.
- Gregersen, B., Linde L. T., & Rasmussen, J. G. (2009). Linking between Danish universities and society. *Science and Public Policy*, 36(2), 151-156.
- Guetzkow, J., Lamont, M., & Mallard, G. (2004). What is Originality in the Humanities and the Social Sciences. *American Sociological Review*, 69 (2), 190-212. doi: 10.1177/000312240406900203.
- Haak, L., Ferriss, L., Wright, W., Pollard, K., Barden, M. E., Probus, K., Tartakovsky, M. A., Hackett, M. & Charles J. (2012). The electronic Scientific Portfolio Assistant: Integrating scientific knowledge databases to support programme impact assessment. *Science and Public Policy*, 39(4): 464-475.
- Habermas, J. (1985). *The Theory of Communicative Action: Reason and the rationalization of society*. Boston: Beacon Press.
- Harrap, N. & Boden, M. (2012). *ERA-NETs and the realisation of ERA: Increasing coordination and reducing fragmentation*. NETWATCH Policy Brief Series No. 1. doi: 10.2791/9503.
- Hasan, D. (2001). Between Kismet and Karma: Charting the impact of a HEI and cultural sector collaboration. *South Asian Popular Culture*, 9(01), 103-114.
- Havas, A., Schartinger, D. & Weber, M. (2010). The impact of foresight on innovation policy-making: recent experiences and future perspectives. *Research Evaluation*, 19 (2), 91-104.
- Heller-Schuh, B., Barber, M., Henriques, L., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G., & Weber, M. (2011). *Analysis of Networks in European Framework Programmes (1984-2006)*. European Commission Joint Research Centre. Institute for Prospective Technological Studies. doi: 10.2791/54167
- Helming, K., et. al. (2011). Ex ante impact assessment of policies affecting land use, Part B: application of the analytical framework. *Ecology and Society*, 16(1), 29.

- Henshall, C. (2011). The impact of Payback research: developing and using evidence in policy. *Research Evaluation*, 20 (3): 257-258.
- Hessels, L. K., & Lente, H. (2010). The mixed blessing of Mode 2 knowledge production. *Science, Technology and Innovation Studies*, 6(1), 65-69.
- Hirad, A., & Zorn, P. (2001). *A Little Knowledge Is a Good Thing: Empirical Evidence of the Effectiveness of Pre-Purchase Homeownership Counseling*. Working paper, Joint Center for Housing Studies, University of Harvard.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569-16572. doi: 10.1073/pnas.0507655102
- Holbrook, J.B., & Frodeman, R. (2011). Peer review and the ex ante assessment of societal impacts. *Research Evaluation*, 20(3), 239–246.
- Horizon 2020 Work Programmes 2014/15, Opportunities for Researchers from the Socio-economic Sciences and Humanities (SSH) Analysis of SSH-relevant Topics.
- INCLUD-ED Consortium (2009). *Actions for success in schools in Europe*. Brussels: European Commission.
- Ivancheva, L., & Gourova, E. (2011). Challenges for Career and Mobility of Researchers in Europe. *Science and Public Policy*, 38(3), 185-198.
- Jacobs, J.A., & Frickel, S. (2009). Interdisciplinarity: A Critical Assessment. *Annual Review of Sociology*. 35, 43-65. doi:10.1146/annurev-soc-070308-115954
- Jacobs, J.A. (2009). Where Credit is Due: Assessing the Visibility of Articles Published in Gender and Society with Google Scholar. *Gender and Society*, 23, 817-832. doi: 10.1177/0891243209351029
- Jacobs, J. A. (2011). Journal Rankings in Sociology: Using the H Index with Google Scholar. Population Studies Center, University of Pennsylvania, PSC Working Paper Series, PSC 11-05.
- Jiménez-Saez, F., Zabala-Iturriagagoitia, J. M., Zofío, J. & Castro-Martriag, E. (2011). Evaluating research efficiency within National R&D Programmes. *Research Policy*, 40, 230–241.
- Joerges, B., & Shinn, T. (Eds.) (2001). *Instrumentation: Between science, state, and industry*. Dordrecht: Kluwer Academic Publishers.
- Jones, B.F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: shifting impact, geography, and stratification in science. *Science*, 322(5905), 1259-1262. doi: 10.1126/science.1158357
- Jones, L., & Lansdverk, J. (2006). Residential education: Examining a new approach for improving outcomes for foster youth. *Children and Youth Services Review*, 28(10), 1152–1168. doi:10.1016/j.childyouth.2005.10.015
- Kennedy, D. M., Piehl, A. M., & Braga, A. A. (1996). Youth Violence in Boston: Gun Markets, Serious Youth Offenders, and A Use-Reduction Strategy. *Law and Contemporary Problems*, 59(1), 147-196.
- Kellow, A. (2012). Assessing Political Science Quality: 'Excellence in Research for Australia. *European Political Science*, 11(4), 567-580.
- Kershner, R., Warwick, P., Mercer, N., & Kleine Staarman, J. (2012). Primary children's management of themselves and others in collaborative group work: "Sometimes it takes patience". *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, 42(2), 201–216.

- Klautzer, L., Hanney, S., Nason, E., Rubin, J., Grant, J. & Wooding, S. (2011). *Assessing Policy and Practice Impacts of Social Science Research: The application of the Payback Framework to assess the Future of Work programme*. *Research Evaluation*, 20(3): 201-209.
- Lamont, M., & Mallard, G. (2005). *Peer Evaluation in the Social Sciences and Humanities Compared: The United States, the United Kingdom, France*. Ottawa: the Social Sciences and Humanities Research Council of Canada.
- Lamont, M. (2009). *How Professors Think. Inside the Curious World of Academic Judgement*. Cambridge, MA: Harvard University Press.
- Lamont, M., & Huutoniemi, K. (2011). *Opening the black box of evaluation: How quality is recognized by peer review panels*. In Camic, C., Gross, N. and Lamont, M. (eds.), *Social Knowledge in the Making* (pp. 209-232). Chicago: University of Chicago Press.
- Largent, M. A., & Lane, J. I. (2012). STAR METRICS and the Science of Science Policy. *Review of Policy Research*, 29(3): 431-438.
- Leduc, L. (1994). Evaluation in the social sciences: the strategic context. *Research Evaluation*, 4 (1), 2-5.
- Lemay, M.A., & Sá, C. (2012). Complexity sciences: towards an alternative approach to understanding the use of academic research. *Evidence and Policy*, 8(4), 473-494.
- Lepori, B., van den Besselaar, P., Dinges, M., van der Meulen, B., Pot de B., Reale, E., Slipersaeter, S., & Theves, J. (2007). Indicators for comparative analysis of public project funding: concepts, implementation and evaluation. *Research Evaluation*, 16(4), 243-255.
- Lepori, B., & Probst, C (2009). Using Curricula Vitae for Mapping Scientific Fields: A small-scale experience for Swiss communication sciences. *Research Evaluation*, 18(2), 125-134.
- Lepori, B., Reale, E., & Tijssen, R. (2011). Designing indicators for policy decisions: challenges, tensions and good practices: introduction to a special issue. *Research Evaluation*, 20 (1), 3-5.
- Levitt, R., Claire, C., Diepeveen, S., Chonaill, S.N., Rabinovich, L., & Tiessen, J. (2010). *Assessing the Impact of Arts and Humanities Research at the University of Cambridge*. Cambridge: RAND Europe.
- Lewis, J. M., & Ross, S. (2011). Research Funding Systems in Australia, New Zealand and the UK: Policy setting and perceived effects. *Policy and Politics*, 39(3): 379-398.
- Leydesdorff, L., Hammarfelt, B., & Salah, A. (2011). The structure of the Arts and Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the American Society for Information Science and Technology*, 62, 2414–2426. doi:10.1002/asi.21636
- Leydesdorff, L. (2013). An evaluation of impacts in “Nanoscience and nanotechnology”: steps towards standards for citation analysis. *Scientometrics*, 94, 35-55. doi:10.1007/s11192-012-0750-5
- Linková, M., & Stöckelová, T. (2012). Public Accountability and the Politicization of Science: The peculiar journey of Czech research assessment. *Science and Public Policy*, 39(5): 618-629.
- Lloyd, C. M., & Modlin, E. L. (2012). *Coaching as a Key Component in Teachers' Professional Development Improving Classroom Practices in Head Start Settings*. OPRE Report.

- Loison, M. (2007). The Implementation of an Enforceable Right to Housing in France. *European Journal of Homelessness*, 1, 185-197.
- Lovbrand, E. (2011). Co-producing European Climate Science and Policy: a cautionary note on the making of useful knowledge. *Science and Public Policy*, 38(3): 225-236.
- Lusardi, A., & Mitchell, S.O. (2007). Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education. *Business Economics*, 42(1), 35-44.
- Lusardi, A., & Tufano, P. (2009). *Debt Literacy, Financial Experiences, and Overindebtedness*. NBER Working Paper n. 14808.
- MacGregor, J., Stranack, K., & Willinsky, J. (2014). *The Public Knowledge Project: Open Source Tools for Open Acces to Scholarly Communication*. In Bartling, S. and Friesike, S. *Opening Science*. (pp. 165-175). Springer: Dortdech.
- Martin, B.R. (2007). *Assessing the impact of basic research on society and the economy*. Paper presented at the Rethinking the impact of basic research on society and the economy (WF-EST International Conference, 11 May 2007), Vienna, Austria.
- Martin, B.R. (2011). The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Evaluation*, 20 (3), 247–254.
- Martin, S. (2010). Co-production of Social Research: Strategies for engaged scholarship. *Public Money and Management*, 30(4): 211-218.
- Mattson, P., Laget, P., Vinderfjgies, A.N., Sundberg, C.J. (2010). What do European research collaboration networks in life sciences look like? *Research Evaluation*, 19(5), 373-384. doi: 10.3152/095820210X12809191250924
- McKay, S. (2012). Social Policy Excellence – Peer Review or Metrics? Analyzing the 2008 Research Assessment Exercise in Social Work and Social Policy and Administration. *Social Policy and Administration*, 46(5), 526-543.
- Meagher, L., Lyall, C., & Nutley, S. (2008). Flows of knowledge, expertise and influence: a method for assessing policy and practice impacts from social science research. *Research Evaluation*, 17 (3), 163-173.
- Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*, 159, 56–63. doi:10.1126/science.159.3810.56
- Mills, G., Gale, W.G., Patterson, R., Engelhardt, G.V., Eriksen, M.D. & Apostolov, E. (2008). Effects of individual development accounts on asset purchases and saving behavior: Evidence from a controlled experiment, *Journal of Public Economics*, 92 (5–6), 1509–1530. doi: 10.1016/j.jpubeco.2007.09.014
- Molas-Gallart, J., & Tang, P. (2011). Tracing “productive interactions” to identify social impacts: An example from the Social Sciences. *Research Evaluation*, 20(3), 219–226.
- Moreno Lara, V., & Fernández, M. E. (2001). La difusión de la investigación en la sierra de Atapuerca. Actividades y reflexiones desde el aula arqueológica “Emiliano Aguirre”. *REVISTA ÍBER. Didáctica de Las Ciencias Sociales, Geografía E Historia*, 29, 27–36.
- Morris, Z. S., Wooding S. & Grant J. (2011). The Answer is 17 Years, What is the Question: Understanding time lags in translational research. *Journal of the Royal Society of Medicine*, 104, 510-520.
- Mostert, S., Ellenbroek, S., Meijer, I., van Ark, G. & Klasen, E. (2010). Societal output and use of research performed by health research groups. *Health Research Policy and Systems*, 8(1), 30.

- Nederhof, A. J. (2008). Policy impact of bibliometric rankings of research performance of departments and individuals in economics. *Scientometrics*, 74(1), 163–174. doi: 10.1007/s11192-008-0109-0
- Nelson, G., Westhues, A., & MacLeod, J. (2003). A meta-analysis of longitudinal research on preschool prevention programmes for children (Article 31). *Prevention and Treatment*, 6(1).
- Newby, H. (1994). The challenge for social science: a new role in public policy-making. *Research Evaluation*, 4 (1), 6-11.
- Nightingale, P., & Scott, A. (2007). Peer review and the relevance gap: ten suggestions for policy-makers. *Science and Public Policy*, 34(8), 543-553.
- O'Hare, P., Coaffee, J., & Hawkesworth, M. (2010). Managing Sensitive Relations in Co-produced Panning Research. *Public Money and Management*, 30(4), 243-250.
- Ouimet, M., Bedard, P. O., & Gelineau, F. (2011). Are the h-index and Some of it Alternatives Discriminatory of Epistemological beliefs and Methodological Preferences of Faculty Members? The case of social scientists in Quebec. *Scientometrics*, 88(1), 91-106.
- Padros, M., Garcia, R., de Mello, R., & Molina, S. (2011). Contrasting Scientific Knowledge With Knowledge From the Lifeworld: The Dialogic Inclusion Contract. *Qualitative Inquiry*, 17(3), 304–312. doi:10.1177/1077800410397809
- Penfield, T., Baker, M.J., Scoble, R., & Wykes, M.C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation*, 23 (1), 21-32.
- PLoS Medicine Editors (2006). The Impact Factor Game. *PloS Med*, 3(6), e291. doi:10.1371/journal.pmed.0030291
- Pontille, D., & Torny, D. (2010). The controversial policies of journal ratings: evaluating social sciences and humanities. *Research Evaluation*, 19(5), 347-360.
- Potì, B., & Cerulli G. (2011). Evaluation of firm R&D and innovation support: new indicators and the ex-ante prediction of ex-post additionality. *Research Evaluation*, 20(1), 19-29.
- Price, D. J. S. (1963). *Little science, big science*. New York: Columbia Univ. Press.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: A manifesto*. Retrieved from: <http://altmetrics.org/manifesto/>.
- Rafols, I., Leydesdorff, L. & O'Harea, A. (2012), How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business and Management. *Research Policy* 41, 7, 1262–1282.
- Rawlings, C.M. & McFarland, D.A. (2011). Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. *Social Science Research*, 40(3), 1001-1017. doi:10.1016/j.ssresearch.2010.10.002
- Reale, E., De Filippo, D., Gomez, I., Lepori, B., Potì, B., Primeri, E., Probst, C., & Sanz Casado E. (2011). New uses of the institutional databases of Universities: indicators of research activity, *Research Evaluation*, 20(1), 47-60.
- Reale E., Inzelt A., Lepori B., & van den Besselaar P. (2012). The social construction of indicators for evaluation: Internationalization of Funding Agencies. *Research Evaluation*, 21, 245-256.
- Reale et al. (2013). EC-JOREP. Investments in Joint and Open R&D programmes and analysis of their economic impact. Final Report. DG Research and Innovation. Contract No. RTD/DirC/C3/2010/SI2.561034.

- Rigby, R. (2002). Social Science Research Impact – Mapping and Measurement. ISSC Workshop, BBAW. Berlin, 14-15 March 2002. <http://www.stage-research.net/SPSG/paris-issc/BerlinRigby.html>
- Roessner, D. (2000). Quantitative and qualitative methods and measures in the evaluation of research. *Research Evaluation*, 9 (2), 125-132.
- Roman, L., & Peter, V. (2012). *The situation of SSH relevant evaluations in Europe and abroad. Social Sciences and Humanities' Monitoring. METRIS Policy Brief. NR. 2-2012.* Brussels: European Commission.
- Rudd, M.A. (2011). An Exploratory Analysis of Societal Preferences for Research-Driven Quality of Life Improvements in Canada. *Social Indicators Research*. 101 (1), 127-153.
- Ruegg, R., & Feller, I. (2003). A Toolkit for Evaluating Public R&D Investment Models, Methods, and Findings from ATP's First Decade. Diane Pub Co.
- Russell, A. (2009). Retaining the Peers: How peer review triumphs over league tables and faulty accounting in the assessment of political science research. *Political Studies Review*, 7(1), 63-72.
- Rymer, L. (2011). *Measuring the impact of research –the context for metric development.* Australia: Group of Eight.
- Saíz Martín, E. (2011). Documentos que acompañan al PAU de la Sierra de Atapuerca (pp. 1–24). Burgos. Retrieved from <http://www.jcyl.es/jcyl/patrimoniocultural/atapuerca/DocumentosQueAcompanan.pdf>
- Salter, A. J., & Martin B. R. (2001). The economic benefits of publicly funded basic research: a critical review. *Research Policy*, 30(3), 509-532.
- Sanderson, I. (2009). Intelligent policy making for a complex world: pragmatism, evidence and learning. *Political Studies*, 57 (4), 699-719.
- Scharnhorst A., Wouters P., & van den Besselaar P. (Eds.) (2006). What does the Web represent? From virtual ethnography to web indicators, Special issue of *Cybermetrics* 10(1), Paper 0 (See URL <http://www.cindoc.csic.es/cybermetrics/vol10iss1.html>)
- Scharnhorst, A., & Garfield, E. (2010). Tracing scientific influence. *Journal of Dynamics of Socio-Economic Systems*, 2(1), 1–33.
- Schmoch, U.; Schubert, T.; Jansen, D.; Heidler, R. & von Ger, R. (2010). How to use indicators to measure scientific performance: a balanced approach. *Research Evaluation*, 19(1), 2-18.
- Schneider, J. W. (2009). An Outline of the Bibliometrics Indicator Used for Performance Based Funding of Research Institutions in Norway. *European Political Science*, 8(3), 364-378.
- Schreiner, M., & Sherraden, M. (2007). *Can The Poor Save? Saving and Asset Building in Individual Development Accounts.* New Jersey: Transaction Publishers.
- Sen, A. (1999). *Development as Freedom.* Oxford University Press.
- Shei, A.; Costa, F.; Reis, M. G. & Ko, A. I. (2014). The impact of Brazil's Bolsa Família conditional cash transfer programme on children's health care utilization and health outcomes. *BMC International Health and Human Rights*, 14(1), 10. doi:10.1186/1472-698X-14-10
- Shema H, Bar-Ilan J, & Thelwall M. (2012). Research Blogs and the Discussion of Scholarly Information. *PLoS ONE* 7(5): e35869. doi:10.1371/journal.pone.0035869

- SIAMPI (2011). Social Impact Assessment Methods for Research and Funding Instruments Through the Study of Productive Interactions (Final report on social impacts of research). Retrieved from [http://www. siampi.edu/](http://www.siampi.edu/)
- SISOB. An Observatorium for Science in Society based in Social Models. Collaborative project. (2011-2013), call FP7-SCIENCE-IN-SOCIETY-2010-1, as a Collaborative Project under the 7th Framework Programme, Grant agreement no.: 266588.
- Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics*, 91(2), 567–575. doi:10.1007/s11192-011-0615-3
- Smith, J. & Duckett, N. (2011). *Toolkit: Early Intervention programmes to prevent Youth Homelessness: some examples from the UK*. Retrieved June, 3, 2014, from http://www.depauluk.org/_uploads/documents/projects/tools-for-service-delivery/early-intervention-toolkit-family-mediation.pdf
- Solesbury, W. (1994). Knowledge transfer: the medium and the message. *Research Evaluation*, 4 (1), 55-58.
- Sordé-Martí, T., Munté, A., Contreras, A., & Prieto-Flores, O. (2012). Immigrant and Native Romani Women in Spain: Building Alliances and Developing Shared Strategies. *Journal of Ethnic and Migration Studies*, 38(8), 1233–1249.
- Sordé-Martí, T., Serradell, O., Puigvert, L., & Munté, A. (2014). Solidarity networks that challenge racialized discourses: The case of Romani immigrant women in Spain. *European Journal of Women's Studies*, 21(1), 87–102.
- Spaapen, J., & Sylvain, C. (1993). Assessing the value of research for society. *Research Evaluation*, 3(2), 117-126.
- Spaapen, J., Dijkstra, H., & Wamelink, F. (2007). *Evaluating Research in Context: a Method for Comprehensive Assessment*. The Hague: COS.
- Spaapen, J., & van Drooge, L. (2011). Introducing eproductive interactions” in social impact assessment. *Research Evaluation*, 20(3), 211–218.
- Stango, V., & Zinman, J. (2007). Fuzzy Math and Red Ink: When the Opportunity Cost of Consumption is Not What It Seems. Working Paper, Dartmouth College.
- Steele, C. (2011). *Whistling Vivaldi: How Stereotypes Affect Us and What We can Do (Issues of Our Time)*. New York: Norton and Company.
- Stipich, N. (1994). Improving social knowledge transfer: new approaches and structures. *Research Evaluation*, 4(1), 59-64.
- Stoker, G. (2010). Translating Experiments into Policy. *The Annals of the American Academy of Political and Social Science*, 628, 47-58.
- Technopolis (2009). *Impact Europese Kaderprogramma's in Nederland*. Technopolis Group.
- Torres-Salinas, D., Rodríguez-Sánchez, R., Robinson-García, N., Fdez-Valdivia, J. & J.A. García. (2013). Mapping citation patterns of book chapters in the Book Citation Index. *Journal of Informetrics*, 7(2), 412–424. doi:10.1016/j.joi.2013.01.004
- UNESCO. (2001). *CONVENTION CONCERNING THE PROTECTION OF THE WORLD CULTURAL AND NATURAL HERITAGE* (pp. 1–155). Cairns. Retrieved from <http://whc.unesco.org/archive/2000/whc-00-conf204-21e.pdf>
- United States Government Accountability Office (2012). *Designing evaluations*. Washington, DC: United States Government Accountability Office.

- Valls, R., & Kyriakides, L. (2013). The power of Interactive Groups: how diversity of adults volunteering in classroom groups can promote inclusion and success for children of vulnerable minority ethnic populations. *Cambridge Journal of Education*, 43(1), 17–33.
- Van den Besselaar, P. & Leydesdorff, L. (2009). Past performance, peer review, and project selection: A case study in the social and behavioral sciences. *Research Evaluation*, 18(4), 273-288.
- Van der Meulen, B. & Rip, A. (2000). Evaluation of societal quality of public actor research in the Netherlands. *Research Evaluation*, 8(1), 11-25.
- van Raan, A.F.J. (1998). Assessment of social sciences: The use of advanced bibliometric methods as a necessary complement of peer review. *Research Evaluation*, 7(1), 2-6.
- van Raan, A.F.J. (2000). R&D evaluation at the beginning of the new century. *Research Evaluation*, 9(2), 81-86.
- Van Vught, F.A., & Ziegele, F. (eds, 2012). *Multidimensional Ranking: The Design and Development of U-Multirank*. Springer, Science+Business Media
- Viola, P.; Rivera Leon, L.; Cadiou, Y., & Doussineau, M. (2010). *Evaluation of the Impact of Framework Programme supported Social Sciences and Humanities Research, A bibliometric approach*. EC Directorate-General for Research Science, Economy and Society.
- Viola, P. (2011 – 2013). *METRIS II Monitoring European Trends in Social Sciences and Humanities. Framework service contract*. 7th Framework Programme. European Commission.
- Wagner, C.S., & L. Leydesdorff (2012). An Integrated Impact Indicator: A new definition of 'Impact' with policy relevance. *Research Evaluation*, 21(3), 183-188.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39, 456–467.
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, 92, 82–96.
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, 331(6023), 1447–51. doi:10.1126/science.1198364
- Wehrens, R., Bekker, M., & Bal, R. (2012). Dutch Academic Collaborative Centres for Public Health: development through time – issues, dilemmas and coping strategies. *Evidence and Policy*, 8(2), 149-170.
- Weinberg, B.A., Owen-Smith, J., Rosen, R.F., Schwarz, L., McFadden Allen, A., Wess, R.E., & Lane, J. (2014). Science Funding and Short-Term Economic Activity. *Science*, 344 (6179), 41-43.
- Wilson, W. J. (2010). Why Both Social Structure and Culture Matter in a Holistic Analysis of Inner-City Poverty. *The ANNALS of the American Academy of Political and Social Science*, 629, 200-219. doi: 10.1177/0002716209357403
- Wooding, S.; Nason, E.; Klautzer, L.; Rubin, J.; Hanney, S. & Grant, J. (2007). Policy and practice impacts of research funded by the Economic and Social Research Council. A case study of the Future of Work programme, approach and analysis. Rand Corporation. TR-435-ESRC.
- Wouters, P. (1999). *The citation culture*. PhD Thesis Amsterdam (available at <http://garfield.library.upenn.edu/wouters/wouters.pdf>)

- Wouters, P., & Costas, R. (2012). *Users , narcissism and control – tracking the impact of scholarly publications in the 21 st century*. (p. 50 pages). Amsterdam. Retrieved from http://www.surffoundation.nl/en/publicaties/Pages/Users_narcissism_control.aspx
- Wrigley, N., Lowe, M., & Guy, C. (2002). Urban regeneration, social inclusion and large store development: the Seacroft development in context. *Urban Studies*, 39, 2101–2114.
- Wrigley, N., Warm, D., & Margetts, B. (2003). Deprivation, diet, and food-retail access: findings from the Leeds of the Future of Work. *Environment and Planning A*, 35(1), 151–188. doi:10.1068/a35150
- Yeager, D. S. & Walton, G. M. (2011). (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81, 267–301.
- Xu, F.; Li, X. X.; Meng W.; Liu, W. B. & Mingers, J. (2013). Ranking academic impact of world national research institutes—by the Chinese Academy of Sciences. *Research Evaluation*, 22 pp. 337–350.

ANNEX I

A COMPUTER-AIDED LITERATURE REVIEW FOR WP1

Introduction

in order to increase the efficiency/coverage/depth of our literature review, as a paradigmatic methodology borrowed, partially, from text mining for modelling research fronts (in scientometrics), we have applied a computer-aided methodology for literature review, consisting of the following two stages:

(1) based on a moderately "large-scale" WoS record collected on the topic (impact dimensions of psychology, in our case), the documents in the collection are being **clustered** based on their **textual descriptors** (title- and keywords), each cluster representing a more-or-less distinctive discourse related to the topic.

(2) In the second step, these clusters – based on their key or **characteristic documents** and **keyword profile**-- are being described along the "analysis grid".

The method implies that it is groups of documents (instead of individual pubs) that are being described "gridwise". The gain of the method is twofold: (1) in this way, a more systematic and more exhaustive view of the state-of-the-art could be obtained, (2) the description of the literature could be more concise and "processable".

Materials and methods

The corpus covering the theme "social impact of psychology" has been obtained from the Web of Science, based on the following search query (SU: Subject Area, TS: topic, TI: title):

SU = Psychology AND (TS=policy OR TI=policy)
Timespan: 2009-2013. Indexes: SCI-EXPANDED, SSCI, AandHCI.

This search has resulted in **4230 docs**. For surveying the most recent literature, we have narrowed this set to cover the last 3–4 years (Timespan: 2011–2013). As a result, about three thousand documents (**n=2711**) have been included in the final sample.

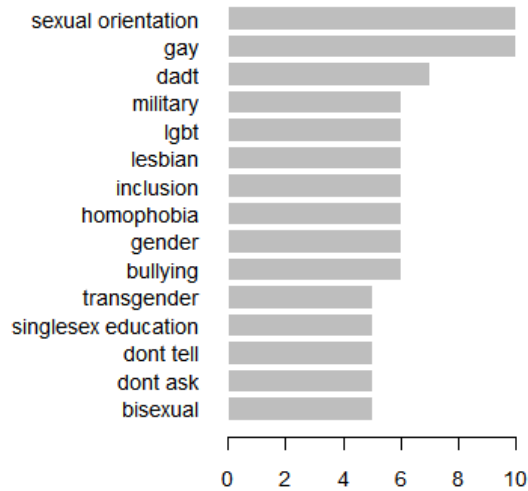
In the next step, sample documents have been clustered based on their thematic similarity. (Indicators used for clustering: keywords, title words and keywords extracted from references [titles of references]; a standard method using tf-idf-like weighting scheme and the cosine similarity measure has been applied for establishing relations among docs; the clustering was based on the community-detection of the document similarity graph).

For each cluster (1) a **keyword profile** – frequency distribution of keywords, with a cutoff leaving the most frequent, most characteristic concepts – has been devised. (2) For each cluster, a set of "**key documents**" and a **cluster description** have been provided (**about 100 docs**) (3) Key documents have been represented via **grids**, along with the reference to the corresponding cluster (description and profile).

Results

In the following, the cluster profiles and descriptions for the most characteristic clusters of the sample are provided.

Keyword profile



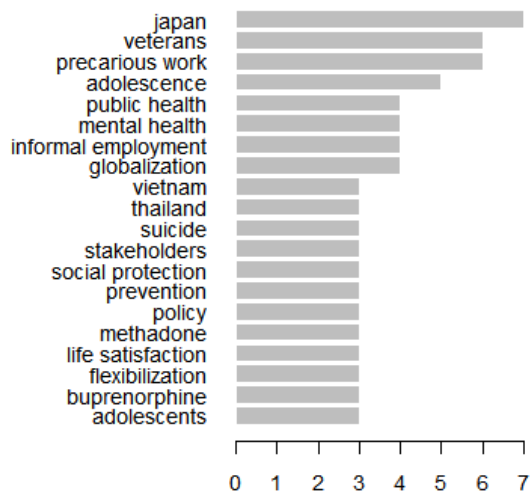
Cluster ID and Description

Cluster #1

Description:

Understanding and contributing to the societal problems of the issues of sexual orientation, especially in relation to institutional contexts (military), from a behavioural perspective

Cluster #2

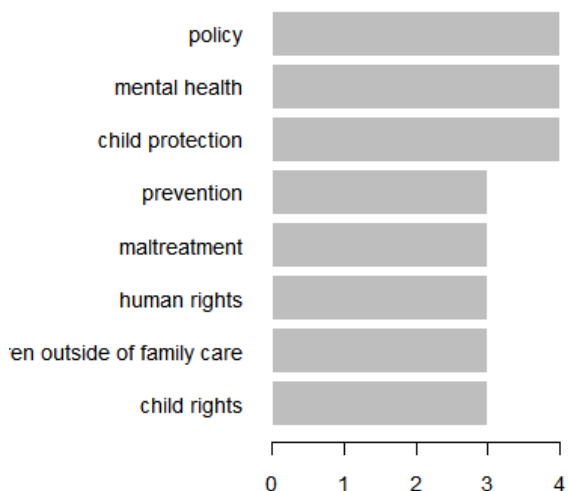


Description:

Understanding the psychology and contributing to the societal problems of modern and emergent forms of employment and globalization, related to mental health issues, with a special focus on Asian countries

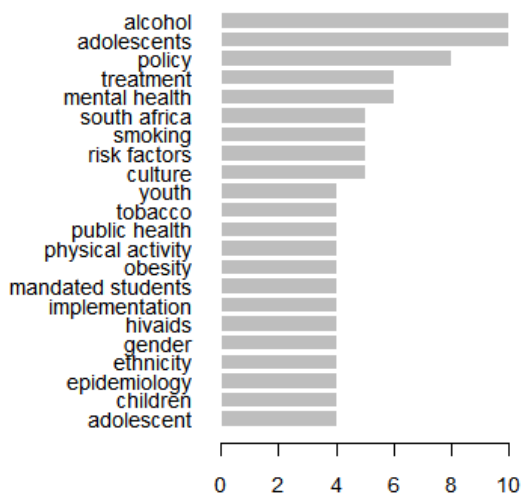
Keyword profile	Cluster ID and Description
-----------------	----------------------------

Cluster #3



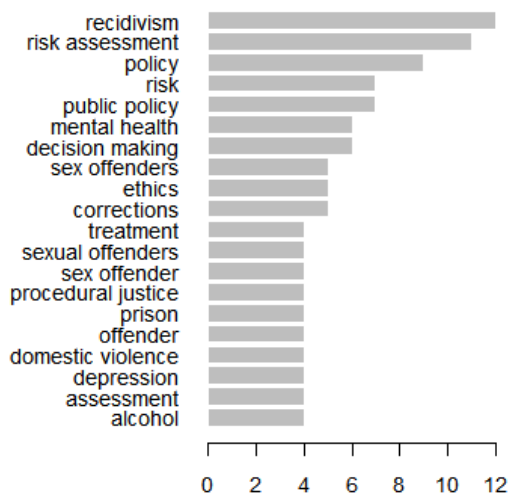
Description:
 Understanding the psychological relations between and contributing to the societal problems of child protection, family care and child/human rights

Cluster #6



Description:
 Understanding the psychology and contributing to the societal problems of adolescent health in relation to habitual/cultural risks factors (especially alcohol consumption and smoking).

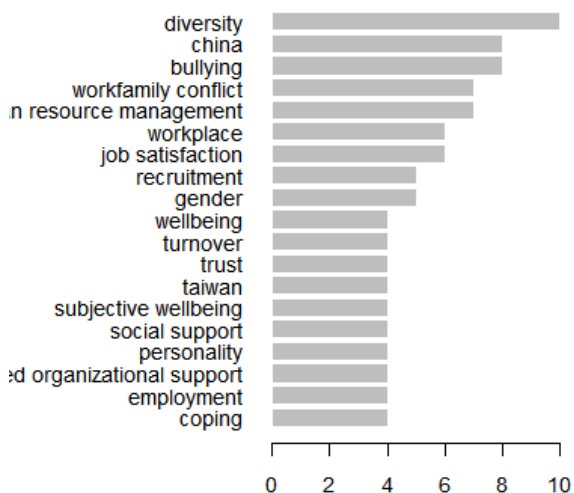
Cluster #7



Description:

Understanding the psychology and contributing to the policy/risk assessment aimed at recidivism, with a special focus on sex offenders and the related legislation

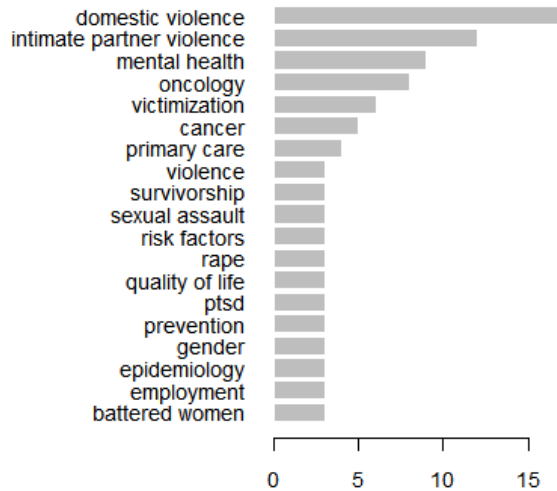
Cluster #8



Description:

Understanding the psychology of and contributing to the societal problems/policies of human resource management in relation to recent challenges: diversity issues, work–family conflict, bullying phenomena, gender issues; special focus is being put on Asian countries (China, Taiwan).

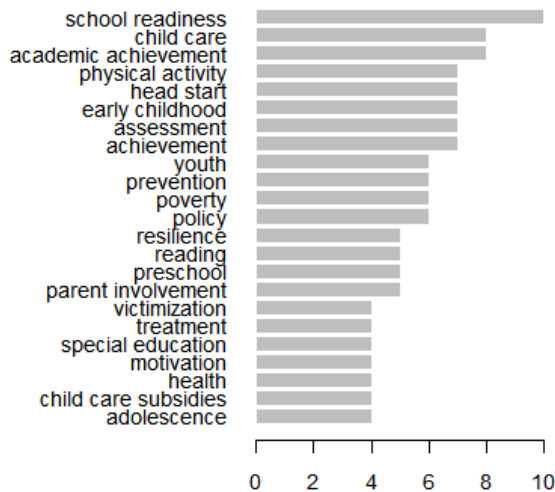
Cluster #9



Description:

Understanding the psychology of and contributing to the societal problems of domestic violence phenomena in relation to sexual violence and mental health in general

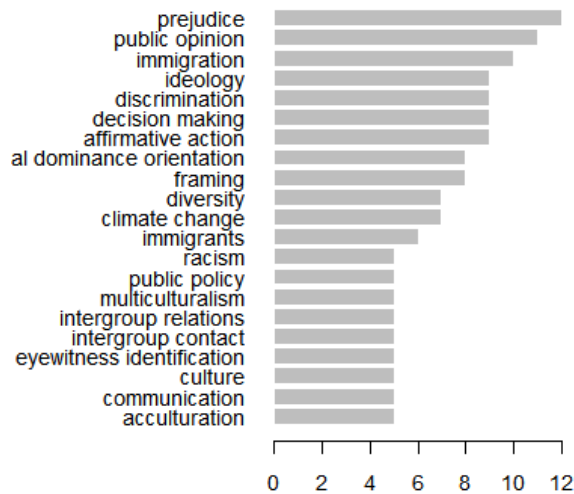
Cluster #11



Description:

Understanding the psychology of and contributing to the policies/assessment concerning early child development in relation to academic achievement, programmes/initiatives (USA Head Start), poverty, and parent involvement.

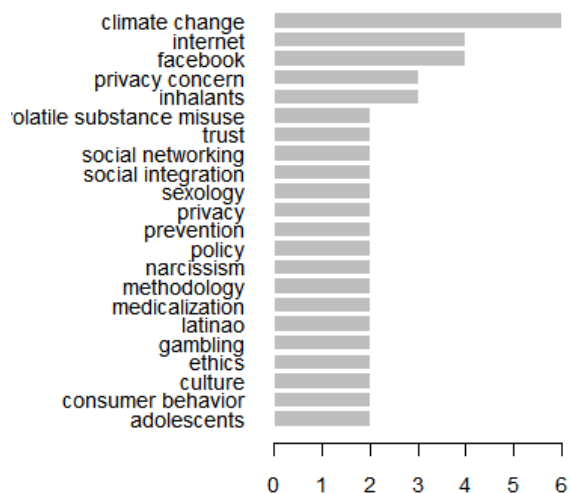
Cluster #15



Description:

Understanding the psychology of and contributing to the policies concerning public attitudes towards immigration and racism phenomena in relation to ideology, discrimination, diversity, framing (communication), intergroup and cultural issues.

Cluster #15



Description:

Understanding the psychology of and contributing to the policies concerning attitudes towards global problems: climate change and internet-related phenomena, the latter mainly in relation to social media and privacy issues

ANNEX II

List of Social Impact Indicators

This is a tentative and non-exhaustive list of social impact indicators based on the social impacts identified and explained in Section 5.2. Following the list, the table summarises the evidence identified for each of the social impact indicators.

SOCIAL IMPACT (EU2020 TARGET): 75% of 20-65 year olds employed

Social Indicator: Employment rate - age group 20-64 (Europe 2020 Indicators – Eurostat)

Social Impact Indicators:

- Number of job positions created
- Increase in tourism rate
- Number of companies created
- Number of cooperatives created among women and people in vulnerable groups
- Number of people from vulnerable groups (e.g. migrants, minorities, women) who have participated in vocational training courses
- Number of buildings renewed in low SES areas

SOCIAL IMPACT (EU2020 TARGET) - Reducing school drop-out rates below 10%

Social indicator: Early leavers from education and training (% of population aged 18-24) (Europe 2020 indicators – Eurostat)

Social Impact Indicators:

- Increase in children's reading ability skills
- Linguistic competence of children from low income families in early childhood education
- Rate of students from disadvantaged neighbourhoods who score at or above the national level in maths and language.

SOCIAL IMPACT (EU2020 TARGET) - at least 40% of 30-34-year-olds completing third level education

Social indicator: Tertiary educational attainment (% of population aged 30-34) (Europe 2020 indicators – Eurostat)

Social Impact Indicators:

- Reduction in the achievement gap between minority university students and non-minority students.

SOCIAL IMPACT (EU2020 Target): at least 20 million fewer people in or at risk of poverty and social exclusion

Social indicator: People at risk of poverty or social exclusion (Europe 2020 Indicators – Eurostat)

Social Impact Indicators:

- Percentage of students from low SES areas passing the official tests in instrumental learning (reading, maths).
- Decrease in the rates of absenteeism of students in the schools or high schools in disadvantaged contexts
- School completion rates of youth at-risk
- Increase in the use of preventive health care services by poor families
- Criminality and victimization rates of youth at-risk
- Increased financial knowledge among vulnerable groups and people with low income.
- Number of cooperatives created in low SES areas.

Social Impact Evidences

SOCIAL IMPACT (EU2020 TARGET): 75% of 20-65 year olds employed

Social Indicator: Employment rate - age group 20-64 (Europe 2020 Indicators – Eurostat)

Social Indicator	Impact	Evidence from SSH research
Number of job positions created		<p><i>Atapuerca. Spain</i> The findings of Atapuerca contributed to directly create job positions. As a result of the interventions or actions derived from the projects' findings 1130 direct stable jobs have been created and maintained since 2010 due to the creation of the Museum of Human Evolution in Atapuerca as well as to other foundations and businesses.</p> <p>Moreno Lara, V. & M. E. Fernández (2001). La difusión de la investigación en la sierra de Atapuerca. <i>Actividades y reflexiones desde el aula arqueológica "Emiliano Aguirre"</i>. Iber. <i>Didáctica de las Ciencias Sociales, Geografía e Historia</i>, 29, 27-36.</p>
Increase in tourism rate ²⁷		<p><i>Atapuerca. Spain</i> Jobs were created as an indirect result of the project's findings, in the tourism sector.</p> <p>In the three year period 2010-2013, the Atapuerca System – formed by the Museum of Evolution, two visitor reception centres, the archeological site, and the archeological park received 1.14 million tourists. The percentage of foreign visitors was 13% in 2013, more than twice the numbers recorded in 2010 (6%). In these three years, the Museum received 744,000 visitors, 166.500 visited the park, and 230,500 the archeological site.</p> <p>Casado, M. (2013). Las visitas de extranjeros se duplican en el Museo desde 2010. <i>El Correo de Burgos</i>, 17 de julio de 2013. "El sistema Atapuerca ha</p>

²⁷ Previous research has already shown that this particular achievement correlates with the target: Kocak, O. (2007). E-commerce opportunities in the tourism sector and it's employment effect. *Changes in Social and Business Environment, Proceedings. 2nd International Conference on Changes in Social and Business Environment*, Panevezys, LITHUANIA, pp. 120-127. Eslami, S., Farahani, A.M., & Asadi, H. (2013). The effects of development of sport tourism on the employment: a review of related research. *International Journal of Sport Studies*, 3(1), 105-110. British Tourist Authority. (1982). *Employment in tourism*. United Kingdom.

	<p>recibido 1,14 millones de turistas en estos tres años". http://www.elcorreodeburgos.com/noticias/2013-07-17/las-visitas-de-extranjeros-se-duplican-en-el-museo-desde-2010</p> <p>Additionally, businesses in the service sector have been created: the Abba Hotel and Atapuerca Conference Center.</p>
<p>Number of companies created²⁸</p>	<p><i>Atapuerca. Spain</i></p> <p>Between 1995 and 1998, four catering and hospitality companies were created in the town of Atapuerca. Also, three SME had been created around the activities of the sites:</p> <ul style="list-style-type: none"> • Granja Escuela Arlanzón, SL (1999). Focuses on environmental education, hospitality, transformation and production of energy, education in peace and solidarity, and leisure activities. • Paleorama, SL (2001). Develops courses and cultural activities, entertainment and leisure, and conducts research projects on archaeological and historical heritage. • ArqueOcio, SL (2001-2007). Designed and realized activities related to the dissemination of archeological, paleontological and natural heritage. <p>Additionally, the research has supported the creation of two non-profit organizations (ACAHIA and Paleorama) which provide public visits to the sites and the management of the Archaeological Park of Atapuerca</p>
<p>Number of cooperatives created among women and people of vulnerable groups²⁹</p>	<p><i>Neighbourhood Mothers Catering. WILCO. Netherlands</i></p> <p>The WILCO project identified social innovation projects in Europe. One of these projects consisted of the creation of a catering cooperative of migrant mothers. The project was based on their expertise in cooking, and employed people who were at risk of poverty, either as a regular job or to earn an additional income. The research found that, with appropriated guidance, women could become self-employed.</p> <p>Everts, A., Ewert, B., and Brandsen, T. (2014). <i>Social Innovations for Social Cohesion. Transnational patterns and approaches from 20 European cities</i>. Liege: EMES European Research Network.</p>
<p>Number of people from vulnerable groups (e.g.</p>	<p><i>Neighbourhood Mothers. WILCO. Germany</i></p> <p>The WILCO project identified social innovation projects in Europe.</p>

²⁸ Previous research has already shown that this particular achievement correlates with the target: van Praag, C. M., & Versloot, PH. (2007). What is the value of entrepreneurship? A review of recent research. *Small Business Economics*, 29 (4), 351-382.

Hart, M., & Hanvey, E. (1995). Job generation and new and small firms - some evidence from the late 1980s, *Small Business Economics*, 7 (2), 97-109.

²⁹ Previous research has already shown that this particular achievement correlates with the target: Sanyang, S. E., & Huang WenChi. (2008). Green cooperatives: a strategic approach developing women's entrepreneurship in the Asian and Pacific region. *World Journal of Agricultural Sciences*, 4(6), 674-683. International Labour Organization. (2004). Employment Creation Potential of the Ethnic Minority Cooperatives in Viet Nam. ILO Office in Vietnam.

<p>migrants, minorities, women) who have participated in vocational training courses³⁰</p>	<p>One project consisted of one hundred migrant women receiving training to become social assistants, and provided a small salary. These mothers reached out to other mothers to inform them about health, language, and social services. The local youth welfare office created positions for neighbourhood mothers as social assistants.</p> <p>Everts, A., Ewert, B., & Brandsen, T. (2014). <i>Social Innovations for Social Cohesion. Transnational patterns and approaches from 20 European cities</i>. Liege: EMES European Research Network.</p> <p><i>Her Second Chance. WILCO. Croatia</i> The WILCO project identified social innovation projects in Europe. One project consisted in improving the socio-economic conditions of disadvantaged women through their empowerment, training and development of business skills. 63 beneficiaries successfully acquired new professions that made them more competitive in the labour market, for example, 15 women become certified nurses.</p> <p>Everts, A., Ewert, B., & Brandsen, T. (2014). <i>Social Innovations for Social Cohesion. Transnational patterns and approaches from 20 European cities</i>. Liege: EMES European Research Network.</p> <p><i>WORKALO. Spain</i> WORKALO demonstrated the need to create successful vocational training courses for Romani women. Based on this evidence, and as result of specific partnerships between researchers and Romani associations, vocational training courses were designed that could connect Roma women and migrants needs with finding a job. This intervention consisted of a training course to become canteens monitors, and achieved an 80% success rate in labour market inclusion, whereas the usual rate does not exceed 20%.</p> <p>Sordé, T., Munté, A., Contreras, A., & Prieto-Flores, Ò. (2013). Immigrant and Native Romani Women in Spain: Building Alliances and Developing Shared Strategies, <i>Journal of Ethnic and Migration Studies</i>, 38(8), 1233-1249.</p> <p>Sordé, T., Serradell, O., Puigvert, L., & Munté, A. (2013). Solidarity networks that challenge racialized discourses: The case of Romani immigrant women in Spain. <i>European Journal of Women Studies</i>, 20(4).</p>
<p>Number of buildings renewed in low SES areas³¹</p>	<p><i>Employment creation in Food deserts. United Kingdom</i> The impact of urban renewal in a food desert was the creation of 320 new job positions of which 230 were occupied by unemployed local residents.</p>

³⁰ Previous research has already shown that this particular achievement correlates with the target: Santa Cruz, I., Siles, G., & Vrečer, N. (2011). Invest for the Long Term or Attend to Immediate Needs? Schools and the Employment of Less Educated Youths and Adults. *European Journal of Education*, 46(2), 197-208.

³¹ Previous research has already shown that this particular achievement correlates with the target: Wong, F.K.W., Hui, E.C.M., Wong, J.T.Y., & Wan, J.K.M. (2010). The impact of urban renewal to the labour force in Hong Kong. *Facilities*, 28, (13), 611-640. Kawaguchi, N. (2013). Development of a social and solidarity economy and urban regeneration in Roubaix, France. *Japanese Journal of Human Geography*, 65(4), 40-56.

	Wrigley, N., Warm, D., & Margetts, B. (2003). Deprivation, diet, and food-retail access: findings from the Leeds "food deserts" study. <i>Environment and planning A</i> , 35,151-188.
--	--

SOCIAL IMPACT (EU2020 TARGET) - Reducing school drop-out rates below 10%

Social indicator: Early leavers from education and training (% of population aged 18-24) (Europe 2020 indicators – Eurostat)

Social Impact Indicator	Evidence from SSH research
Increase in the children's reading ability skills ³²	<p><i>India</i></p> <p>The participatory programme aiming at universal primary education in India, "Read India", developed by the Pratham organization had impact on improving the reading skills of the children. It offered reading camps for children run by volunteers that had been previously trained by the Pratham organization staff. As a result, children in the villages that received intervention 3 are 1.7 percent more likely to read at least letters, 1.8 percent more likely to read words or paragraph, and 1.7 percent more likely to read stories.</p> <p>Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India. <i>Economic Policy</i>, 2 (1), 1–30.</p>
Linguistic competence of children from low income families in early childhood education ³³	<p><i>Head Start (USA)</i></p> <p>It is an early childhood education programme aimed at improving the lives of children and also those of whole communities. Based on the theory of the psychologist Jerome Bruner, it provides comprehensive early childhood education, health, nutrition, and parent involvement services to low-income children and their families. By 2005 more than 21 million children from low-income families had taken advantage of both the educational programme and the social, health and nutritional services it provides. Both short-term and long-term benefits have been identified. Short-term benefits include salient results in receptive vocabulary and phonemic awareness, and positive outcomes in the health of children and their parents. Concerning long-term benefits, a permanent results have been identified. A high-quality pre-kindergarten education has been linked to improvements in school</p>

³² Previous research has already shown that this particular achievement correlates with the target: Maughan, B., Hagell, A., Rutter, M., Yule, W. (1994). Poor Readers in Secondary-School. *Reading and Writing*, 6(2), 125-150. Boland, T. (1993). The Importance of Being Literate - Reading Development in Primary-School and its Consequences for the School Career in Secondary-Education. *European Journal of Psychology of Education*, 8, 289-305.

³³ Previous research has already shown that this particular achievement correlates with the target: Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Childhood Language Skills and Adult Literacy: A 29-Year Follow-up Study. *Pediatrics*, 125(3), e459 -e466 (doi: 10.1542/peds.2008-2111). Temple, J., Reynolds, A. and Miedel, W. (2000). "Can Early Intervention Prevent High School Dropout?" *Urban Affairs*. March, 35(1), pp. 31-56.

	<p>readiness requiring skilled teachers to provide safe, caring and supportive environments, and in facilitating active engagement and learning across diverse domains</p> <p>Abbott-Shim, M., Lambert, R. & McCarty, F. (2003). A Comparison of School Readiness Outcomes for Children Randomly Assigned to a Head Start Programme and the Programme's Wait List. <i>Journal of Education for Students Placed at Risk (JESPAR)</i>, 8(2), 191-214.</p> <p>Barnett, S. & Hustedt, J. (2005). Head Start's Lasting Benefits. <i>Infants and Young Children</i>, 18(1), 16–24.</p> <p>Lloyd, C. & Modlin, E. (2012). <i>Coaching as a Key Component in Teachers' Professional Development: Improving Classroom Practices in Head Start Settings</i>. OPRE Report 2012-4, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.</p>
<p>Rate of students from disadvantaged neighbourhoods who score at or above the national level in maths and language³⁴</p>	<p><i>Harlem Children's Zone. NYC (USA)</i></p> <p>It is a non-profit organisation for poverty-stricken children and families living in Harlem. HCZ is providing evidence on how highly racialized schools in economically marginalised neighbourhoods are creating opportunities for their students to reach College and fight against early school leaving. These schools combine a high-quality education with wider community actions. HCZ schools are closing the achievement gap in Maths (97% of third graders in HCZ Promise Academy scored at or above grade level) and in Language.</p> <p>Dobbie, W., & Fryer, R.G. (2011). Are High Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone. <i>American Economic Journal: Applied Economics</i>, 3(3), 158-87.</p> <p>Wilson, W. J. (2010). Why Both Social Structure and Culture Matter in a Holistic Analysis of Inner-City Poverty. <i>The ANNALS of the American Academy of Political and Social Science</i>, 629, 200-219.</p>

SOCIAL IMPACT (EU2020 TARGET) - at least 40% of 30-34-year-olds completing third level education

Social indicator: Tertiary educational attainment (% of population aged 30-34) (Europe 2020 indicators – Eurostat)

Social Impact Indicator	Evidence from SSH research
Reduction of the achievement gap between	An intervention to reinforce the sense of belonging of university students in their first year, by lessen their psychological

³⁴ Previous research has already shown that this particular achievement correlates with the target: Korhonen, J., Linnanmaki, K., & Aunio, P.(2014). Learning difficulties, academic well-being and educational dropout: A person-centred approach. *Learning and Individual Differences*, 31, 1-10.

<p>minority university students and non-minority students³⁵</p>	<p>perceptions of threat at campus, found that the African American students who participated in the intervention reduced the gap with their European-American classmates by 79%, and improved their self reported health and well-being while reducing their reported number of doctor visits.</p> <p>Walton, G.M. & Cohen, G.L. (2011). A Brief Social-Belonging Intervention Improves Academic and Health outcomes of Minority Students. <i>Science</i>, 311, 1447-1451.</p>
--	---

SOCIAL IMPACT (EU2020 Target): at least 20 million fewer people in or at risk of poverty and social exclusion

Social indicator: People at risk of poverty or social exclusion (Europe 2020 Indicators – Eurostat)

Social Impact Indicator	Evidence from SSH research
<p>Percentage of students from low SES areas passing official tests in instrumental learning (reading, maths)³⁶</p>	<p><i>INCLUD-ED. (FP6)</i></p> <p><u>Case 1</u> After implementing “interactive groups” the percentage of 4th grade students (primary ed.) with low mathematics achievement in the school decreased by 47% between 2009 and 2011, and those with high levels increased by 18%.</p> <p><u>Case 2</u> After implementing “interactive groups”, the results of 4th grade students (primary ed.) in maths improved from 2008 to 2010. They moved from score 1 to 3 (out of 5), which is close to the regional average.</p> <p><u>Case 3</u> After implementing the “after school club”, in 2011, students improved in maths and language. Their scores showed an improvement especially in English (67%) and maths (82%).</p> <p>Valls, R. & Kyriakides, L. (2013). The power of Interactive Groups: How Diversity of Adults Volunteering in classroom groups can promote inclusion and success for children of vulnerable minority ethnic populations. <i>Cambridge Journal of Education</i>, 43(1), 17-33.</p> <p>Gatt, S. & Armeni, L.S. (2012). The Community as a Resource to Educate Primary School Children. <i>Literacy Information and Computer Education Journal (LICEJ)</i>, 3(2), 560-567.</p>

³⁵ Previous research has already shown that this particular achievement correlates with the target: Whittaker, J.A., & Montgomery, B.L. (2012). Cultivating Diversity and Competency in STEM: Challenges and Remedies for Removing Virtual Barriers to Constructing Diverse Higher Education Communities of Success. *Journal of undergraduate neuroscience education*, 11(1), 44-51. Swail, W.S. (2003). Retaining Minority Students in Higher Education: A Framework for Success. ASHE-ERIC Higher Education Report. Jossey-Bass Higher and Adult Education Series. San Francisco: Jossey-Bass.

³⁶ Previous research has already shown that this particular achievement correlates with the target: Behrman, J.R. (2011). How much might human capital policies affect earnings inequalities and poverty? *Estudios de Economía*, 38(1), 9-41.

<p>Decrease in the rates of absenteeism of students in schools or high schools in disadvantaged areas³⁷</p>	<p><i>INCLUD-ED (FP6)</i> In a school located in a deprived area, the implementation of actions such as interactive groups and extending learning time led to the reduction of absenteeism from 30% to 10% in one year period, and occur only occasionally in a two year period.</p> <p>Aubert, A. (2011). Moving beyond social exclusion through dialogue. <i>International Studies in Sociology of Education</i>, 21(1), 63 – 75.</p>
<p>School completion rates of youth at risk³⁸</p>	<p><i>The Academy, Residential education (USA)</i> It is a placement option for discharged foster youths that have no other placement options. They are also provided with education. School completion and college attendance rates are higher than the rates reported in most other studies. 28% of the youths interviewed six months after leaving the Academy reported attending college, half of them being in a four-year college.</p> <p>Jones, L., & Lansdverk, J. (2006). Residential education: Examining a new approach for improving outcomes for foster youth. <i>Children and Youth Services Review</i>, 28(10), 1152-1168.</p> <p><i>Bolsa Familia (Brasil)</i> This programme consists of providing a basic income for poor families, designed to keep their children in education and provide them with vaccinations. The aim is to avoid families sending their children to work before finishing compulsory education. Research has shown that the school dropout rate is lower among children from beneficiary families, highlighting that the programme is globally responsible for a 57.2% reduction in the dropout rate.</p> <p>De Lima Amaral, E.F. & Do Prado Monteiro, V. (2013). An evaluation of the impact of the educational conditions of Brazil's Bolsa Familia Programme (2005 and 2009). <i>DADOS: Revista de Ciências Sociais</i>, 56(3), 531.</p>
<p>Increase in the use of preventive health care services by poor families³⁹</p>	<p><i>Bolsa Familia (Brasil)</i> This programme consists of providing a basic income for poor families, designed to keep their children in education and provide them with vaccinations. The aim is to avoid families</p>

³⁷ Previous research has already shown that this particular achievement correlates with the target: Adow, A. K., Buor, D., Tagoe-Darko, E., & Kyei, P. O. (2013). Longterm poverty reduction through boreholes provision in rural communities - the quality education platform: practical insights from the Atebubu and Afram Plains Districts of Ghana. *Online Journal of Social Sciences Research*, 2(1), 16-26.

³⁸ Previous research has already shown that this particular achievement correlates with the target: Maddox, B. (2010). Marginal returns: re-thinking mobility and educational benefit in contexts of chronic poverty. *Compare: A Journal of Comparative and International Education*, 40(2), 213-222.

³⁹ Previous research has already shown that this particular achievement correlates with the target: Sin, D.D., Svenson, L.W., Cowie, R. L., & Paul Man, S. F. (2003). Can Universal Access to Health Care Eliminate Health Inequities Between Children of Poor and Nonpoor Families?: A Case Study of Childhood Asthma in Alberta, 124(1), 51-56. Strauss, Z., & Horsten, D. (2013). A human rights-based approach to poverty reduction: the role of the right of access to medicine as an element of the right of access to health care. *Potchefstroomse Elektroniese Regsblad*, 16 (3).

	<p>sending their children to work before finishing compulsory education. The programme has increased the preventive use of health care services by poor families.</p> <p>Shei, A., Costa, F., Reis, M.G., & Ko, A.I. (2014). The Impact of Brazil's Bolsa Familia Conditional Cash Transfer Programme on Children's Health Care Utilization and Health Outcomes. <i>BMC International Health and Human Rights</i>, 14, 10.</p>
<p>Criminality and victimization rates of youth at risk⁴⁰</p>	<p><i>The Academy, Residential education</i> It is a placement option for discharged foster youths that have no other placement options. They are also provided with education. No criminality or victimization behaviours were identified, the opposite to other studies focused on this population.</p> <p>Jones, L., & Lansdverk, J. (2006). Residential education: Examining a new approach for improving outcomes for foster youth. <i>Children and Youth Services Review</i>, 28(10), 1152-1168.</p>
<p>Increased financial knowledge among vulnerable groups and people with low incomes⁴¹</p>	<p><i>Financial Literacy</i> Financial literacy improves financial knowledge, which is positively associated with household financial decision-making.</p> <p>Stango, V., & Zinman, J. (2007). <i>Fuzzy math and red ink: When the opportunity cost of consumption is not what it seems</i>. Working Paper, Dartmouth College.</p> <p>Hilgert, M., Hogarth J., & Beverly, S. (2003), Household Financial Management: The Connection between Knowledge and Behavior. <i>Federal Reserve Bulletin</i>, 309-322.</p>
<p>Number of cooperatives created in low SES areas⁴²</p>	<p><i>INCLUD-ED (FP6)</i> The creation of a competitive cooperative in the poorest neighbourhood of Albacete contributed to creating jobs for people at risk of exclusion. This cooperative contributed to creating jobs for long-term unemployed people. Since 2011, the cooperative provides stable employment for 11 people, and 80 people have been hired in seasonal jobs in the fields and other agricultural sectors. The cooperative has improved</p>

⁴⁰ Previous research has already shown that this particular achievement correlates with the target: Gaetz, S. (2004). Safe Streets for Whom? Homeless Youth, Social Exclusion, and Criminal Victimization. *Canadian Journal of Criminology and Criminal Justice*, 46, 423-455.

⁴¹ Previous research has already shown that this particular achievement correlates with the target: Bernheim, B. D. & Garrett, D. M. (2003). "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households," *Journal of Public Economics*, 87(7-8), 1487-1519.

Schreiner, M., & Sherraden, M.W. (2007). *Can the Poor Save?: Saving and Asset Building in Individual Development Accounts*. Transaction Publishers. Financial education reduces debt levels: "This study provides evidence that the receipt of one-on-one credit counseling is associated with improvement in borrower credit profiles over an extended period." Elliehausen, G., Lundquist, E.V., & Staten, M.E. (2007). The Impact of Credit Counseling on Subsequent Borrower Behavior. *Journal of Consumer Affairs*, 41 (1) 1-28.

⁴² Previous research has already shown that this particular achievement correlates with the target: Burgués, A.; Martín, S.; Santa Cruz, I. (2013). La relación entre cooperativas transformadoras y desigualdades sociales en los territorios. *Scripta Nova. Revista Electrónica De Geografía Y Ciencias Sociales*, XVII, 427 (4).

	<p>living conditions of many families in the neighbourhood.</p> <p>Flecha, R. & Santa Cruz, I. (2011) Cooperation for economic success: the Mondragon case. <i>Analyse and Kritik</i>, 1, 157-170.</p> <p>Padros, M., Garcia, R., de Mello, R., & Molina, S. (2011). Contrasting Scientific Knowledge With Knowledge From the Lifeworld: The Dialogic Inclusion Contract. <i>Qualitative Inquiry</i>, 17(3), 304–312.</p> <p>Flecha, R., & Soler, M. (2014). Communicative Methodology: Successful Actions and Dialogic Politics. <i>Current Sociology</i>, 62, 232–242.</p>
--	--