47

# fteval JOURNAL for Research and Technology Policy Evaluation

ISSUE 47 | MAY 2019

CONFERENCE
"Impact of Research and Innovation Policy at the Crossroads of Policy Design, Implementation and Evaluation"
VIENNA 2018

## Impact of R&I Policy at the Crossroads of
POLICY DESIGN
IMPLEMENTATION
EVALUATION

PIERRE-BENOIT JOLY, MIREILLE MATT AND DOUGLAS K. R. ROBINSON

RESEARCH IMPACT ASSESSMENT: FROM EX POST TO REAL-TIME ASSESSMENT

NELLY BRUNO AND MARTINA KADUNC

IMPACT PATHWAYS: TRACKING AND COMMUNICATING THE IMPACT OF THE EUROPEAN FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION

JULIEN RAVET, BAPTISTE BOITIER, MARCO GRANCAGNOLO, PIERRE LE MOUËL, LIVIU STIRBAT, PAUL ZAGAMÉ

THE SHAPE OF THINGS TO COME: EX-ANTE ASSESSMENT OF THE ECONOMIC IMPACT OF HORIZON EUROPE

## FEDERAL MINISTRY OF EDUCATION, SCIENCE AND RESEARCH
Minoritenplatz 5, 1014 Vienna
**Mag.ª Irene Danler**
**E:** irene.danler@bmbwf.gv.at
**Mag.ª Simone Mesner**
**E:** simone.mesner@bmbwf.gv.at

## FEDERAL MINISTRY FOR DIGITAL AND ECONOMIC AFFAIRS
Stubenring 1, 1014 Vienna
**Mag.ª Sabine Pohoryles-Drexel**
**E:** sabine.pohoryles-drexel@bmwfw.gv.at

## FEDERAL MINISTRY FOR TRANSPORT, INNOVATION AND TECHNOLOGY
Radetzkystraße 2, 1030 Vienna
**Dr. Rupert Pichler**
**E:** rupert.pichler@bmvit.gv.at
**Dr. Mario Steyer**
**E:** mario.steyer@bmvit.gv.at

## ACR – AUSTRIAN COOPERATIVE RESEARCH
Sensengasse 1, 1010 Vienna
**Dr.in Sonja Sheikh**
**E:** sheikh@acr.ac.at

## AUSTRIAN COUNCIL FOR RESEARCH AND TECHNOLOGY DEVELOPMENT
Pestalozzigasse 4/DG 1, 1010 Vienna
**Dr. Johannes Gadner**
**E:** j.gadner@rat-fte.at

## AIT – AUSTRIAN INSTITUTE OF TECHNOLOGY
Giefinggasse 4, 1210 Vienna
**Mag. Michael Dinges**
**E:** michael.dinges@ait.ac.at
**Mag.ª Barbara Heller-Schuh, MAS**
**E:** barbara.heller-schuh@ait.ac.at

## AQ AUSTRIA – AGENCY FOR QUALITY ASSURANCE AND ACCREDITATION AUSTRIA
Franz-Klein-Gasse 5, 1190 Vienna
**Dr.in Elisabeth Froschauer-Neuhauser**
**E:** elisabeth.froschauer-neuhauser@aq.ac.at
**Dr.in Eva Maria Freiberger**
**E:** eva.maria.freiberger@aq.ac.at

## AWS – AUSTRIA WIRTSCHAFTSSERVICE GMBH
Walcherstraße 11A, 1020 Vienna
**Mag.ª Marlis Baurecht**
**E:** M.Baurecht@aws.at
**Mag. Norbert Knoll**
**E:** n.knoll@awsg.at

## CDG – CHRISTIAN DOPPLER RESEARCH ASSOCIATION
Boltzmanngasse 20, 1090 Vienna
**DIin Mag.ª Brigitte Müller**
**E:** brigitte.mueller@cdg.ac.at

## CONVELOP – COOPERATIVE KNOWLEDGE DESIGN GMBH
Bürgergasse 8-10/I, 8010 Graz
**DIin Drin. Karin Grasenick**
**E:** karin.grasenick@convelop.at

Erdbergstraße 82/4, 1030 Wien
**Mag. Thomas Jud**
**E:** thomas.jud@convelop.at

## FFG – AUSTRIAN RESEARCH PROMOTION AGENCY
Haus der Forschung,
Sensengasse 1, 1090 Vienna
**DIin Dr.in Sabine Mayer**
**E:** sabine.mayer@ffg.at
**Mag. Leonhard Jörg**
**E:** leonhard.joerg@ffg.at

## FWF – AUSTRIAN SCIENCE FUND
Haus der Forschung
Sensengasse 1, 1090 Vienna
**Dr. Falk Reckling**
**E:** falk.reckling@fwf.ac.at
**Dr. Thomas Völker**
**E:** thomas.voelker@fwf.ac.at

## IHS – INSTITUTE FOR ADVANCED STUDIES
Josefstädter Straße 39, 1080 Vienna
**Dr.in Angela Wroblewski**
**E:** wroblews@ihs.ac.at
**Mag. Richard Sellner**
**E:** richard.sellner@ihs.ac.at

## INDUSTRIEWISSENSCHAFTLICHES INSTITUT – IWI
Mittersteig 10, 1050 Wien
**FH-Hon.Prof. Dr. Dr. Herwig W. Schneider**
**E:** schneider@iwi.ac.at
**Mag. Philipp Brunner**
**E:** brunner@iwi.ac.at

## JOANNEUM RESEARCH FORSCHUNGSGESELLSCHAFT MBH
Haus der Forschung
Sensengasse 1, 1090 Vienna
**Mag. Wolfgang Polt**
**E:** wolfgang.polt@joanneum.at
**Mag. Jürgen Streicher**
**E:** juergen.streicher@joanneum.at

## KMU FORSCHUNG AUSTRIA – AUSTRIAN INSTITUTE FOR SME RESEARCH
Gusshausstraße 8, 1040 Vienna
**Mag.ª Iris Fischl**
**E:** i.fischl@kmuforschung.ac.at
**Mag. Peter Kaufmann**
**E:** p.kaufmann@kmuforschung.ac.at

## LUDWIG BOLTZMANN GESELLSCHAFT
Nußdorfer Straße 64, 1090 Vienna
**Mag. Patrick Lehner**
**E:** patrick.lehner@lbg.ac.at

## OEAW – AUSTRIAN ACADEMY OF SCIENCE
Dr. Ignaz Seipel-Platz 2, 1010 Vienna
**Nikolaus Göth, MSc**
**E:** nikolaus.goeth@oeaw.ac.at

## TECHNOLPOLIS GROUP AUSTRIA
Rudolfsplatz 12/11, 1010 Vienna
**Mag.ª Katharina Warta**
**E:** warta@technopolis-group.com

## VIENNA BUSINESS AGENCY. A SERVICE OFFERED BY THE CITY OF VIENNA.
Mariahilfer Straße 20, 1070 Vienna
**Robert Mayer-Unterholzer**
**E:** mayer-unterholzner@wirtschaftsagentur.at

## WIFO – AUSTRIAN INSTITUTE OF ECONOMIC RESEARCH
Arsenal, Objekt 20
Postfach 91, 1103 Vienna
**Dr. Jürgen Janger**
**E:** Juergen.Janger@wifo.ac.at
**Dr. Andreas Reinstaller**
**E:** andreas.reinstaller@wifo.ac.at

## WPZ RESEARCH GMBH
Mariahilfer Straße 115/16, 1060 Vienna
**Dr.in Brigitte Ecker**
**E:** brigitte.ecker@wpz-research.com

## WWTF – VIENNA SCIENCE AND TECHNOLOGY FUND
Schlickgasse 3/12, 1090 Vienna
**Dr. Michael Stampfer**
**E:** michael.stampfer@wwtf.at
**Dr. Michael Strassnig**
**E:** michael.strassnig@wwtf.at

## ZSI – CENTRE FOR SOCIAL INNOVATION
Linke Wienzeile 246, 1150 Vienna
**Dr. Klaus Schuch**
**E:** schuch@zsi.at

# CONTENT

# EDITORIAL

## RATIONALE AND LESSONS (TO BE) LEARNED FROM THE AUSTRIAN PRESIDENCY CONFERENCE ON 'IMPACT OF RESEARCH AND INNOVATION POLICY AT THE CROSSROADS OF POLICY DESIGN, IMPLEMENTATION AND EVALUATION'

KLAUS SCHUCH, CENTRE FOR SOCIAL INNOVATION (ZSI) & AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION (FTEVAL)

DOI: 10.22163/fteval.2019.358

## DEAR READERS!

Most impact evaluations of R&I policy interventions focus either on scientific-technical effects or on economic effects. For this purpose, suitable indicators, data bases and methods have been created and continuously developed in recent decades. However, the comprehensibility and assessment of social and societal effects of R&I policy interventions has only recently gained new attention. One reason for this is the orientation of R&I policy towards major societal challenges ('new mission-oriented R&I policy'[1]). The European Commission's "Horizon Europe", the 9th European Research and Innovation Framework Program, explicitly provides within the second pillar of the next Framework Program specific R&D missions still to be selected. For these missions as well as for the global challenges postulated in Horizon Europe - as in Horizon 2020 - the social impact dimension is highly relevant as it explicitly addresses the goals set by society (for example, the United Nations Sustainable Development Goals / SDGs[2]). In order to better track and measure the impact dimensions of Horizon Europe, an expert report[3] was presented immediately after the publication of the European Commission's proposal for Horizon Europe in July 2018, which distinguishes between the following three impact dimensions: (1) scientific impact, (2) societal impact and (3) economic impact. Already the year before, an ERAC ad-hoc working group submitted a report[4] that also argues for different dimensions of impact, but focuses on measuring the impact of European framework programs at national level.

In anticipation of developments at the European level, the Austrian Federal Ministry of Transport, Innovation and Technology suggested in 2017 to hold an international conference for the measurement of mission-oriented R&I interventions within the framework of the Austrian Council Presidency. The Austrian Platform for Research and Technology Policy Evaluation (fteval) was commissioned with this task and organised the conference in November 2018 together with the Manchester Institute of Innovation Research and the Institut Francilien Recherche, Innovation et Société from Paris. The starting point for the conference was that, first and foremost, not only the European, but also national R&I policies are required to make a contribution to society and to document the corresponding effects, and second, that the new impact agenda has an impact on the whole policy cycle, including policy-making, policy implementation and policy evaluation.

Both the presidency event and the expert report of the European Commission have chosen the approach of impact pathways to further discuss the measurement of the three different dimensions of impact in order to emphasize the design and process character of effect creation and effect development. In particular, the impact pathways for measuring societal effects are challenging. These are confronted with basic definitional problems. While "social impact" in the EU context is understood as a generic term (e.g. in the case of the Better Regulation Toolbox of the European Commission), which implies effects on society, politics,

1 See Gassler et al. (2006).
2 https://sustainabledevelopment.un.org/?menu=1300
3 See Van den Besselaar et al. (2018).
4 See ERAC (2017).

environment, economy and other dimensions, "societal impact" is understood as more specifically. Also, the approaches and models commonly used in the scientific literature to establish social impact of R&I policies refer to a variety of issues, including policy implications, and lack clear demarcations[5]. So far, existing assessments of the social impact of R&I policy interventions are often only contextual and specific as well as qualitative and anecdotal in nature.

In addition to the theoretical problems of demarcation, there are serious deficits with regard to the indicators for assessing societal effects of R&I policies as well as a lack of systematically collected, quality-assured data. Moreover, there is often a falsely equation of social impact with dissemination or transfer, to which most of the so-called alternative metrics (altmetrics) focus. Particular challenges for the development of appropriate indicators to measure societal impact include

1. that the time taken to achieve the actual impact on society is longer than the achievement of concrete results;
2. that the assignment of social changes is more difficult than the assignment of scientific references or economic attributes;
3. that the availability and comparability of data to track social and political impacts of R&I interventions is severely limited.

According to the literature review[6] in the European Commission's expert report, specific and commonly used indicators for measuring social impact are almost non-existent, or if so, often only as suggestions without systematic application[7]. It is therefore hardly surprising that most agencies and evaluation projects do not consider the social (or societal) impact of R&I. In a few cases, societal impact in ex-ante evaluations is sometimes cited as a criterion to consider, but without specific indicators.

The Austrian Council Presidency Conference 'Impact of Research and Innovation Policy at the Crossroads of Policy Design, Implementation and Evaluation' has therefore addressed the question of how impacts along the three dimensions of impact mentioned above (scientific, economic and social) can be better understood, designed and measured by a favorable R&I policy. The conference structured the topic impact evaluation into four blocks:

1. The essence of impact-oriented R&I policy
2. Design, implementation and support measures for an impact-oriented R&I policy
3. Novel concepts, tools and methods for assessing social impact of R&I policies and
4. Effects of impact evaluations on policy learning

These topics were addressed in five key-note presentations, four panel discussions, seven specific paper sessions featuring 40 ex-ante selected papers, three workshops, a case study on impact measurement at the French National Agricultural Research Institute and a poster session, in which 11 posters were presented.

296 experts from 39 countries and all continents have registered for the conference. Of these, 255 actually attended the conference. 131 of the accredited persons can be assigned to the research area, 73 came from agencies, 70 from politics, 13 from intermediary institutions including research infrastructures, 8 from the business enterprise sector and one from the press. 42% of the participants came from Austria. Larger contingents came from the category „International Institutions" (n = 24), especially from the European Commission, but also from OECD, EUREKA and COST, which made the European dimension of the event visible. 21 of the accredited persons came from Germany; 17 from the UK; 10 from France and Norway; 9 from Belgium and Spain and 8 from the Netherlands. With the exception of Malta, Slovakia and Slovenia, all EU countries were represented. More accredited persons from non-EU countries came from Iceland, Norway, Russia, Switzerland, Ukraine, and from Australia, Brazil, Chile, Iran, Japan, Nepal, South Africa and the United States.

Feedback on the conference was consistently positive. 93% said that the organization was very good or good; 97% would recommend the conference.

Overall, the conference could contribute to the following immediate results:

- The level of knowledge about impact evaluations in the R&I area has been widely consolidated.
- Recent experiments to promote effective policies and measures have been put forward for discussion.
- Methodological experiences to better assess the social impact of mission-oriented R&I policy have been extensively shared.
- Indicators to measure progress on key pathways or actual impact in the short, medium and long term were presented and reflected.
- An increased use of more comprehensive impact assessment approaches in the field of R&I policy was suggested.
- The need to use unique identifiers and better data bases was discussed.
- Awareness was raised for the use and impact of big data approaches and artificial intelligence for text mining, automated data collection, and automated data analysis.
- The need for clear expectation management was recognized.
- It has been widely acknowledged that for impact assessment, both research organizations and agencies themselves should set up appropriate procedural arrangements to support societal impacts and to document them.

These conference proceedings collect 21 papers and 11 posters presented and discussed during the conference.

I am very grateful to all authors who contributed to these conference proceedings and to the success of the conference!

Yours sincerely

Klaus Schuch

Executive Manager of the Austrian Platform for Research and Technology Policy Evaluation

5 See Brewer (2011) und (2013); Flecha (2018); Raua et al. (2018); Reale et al. (2017).
6 See Van den Besselaar (2018).
7 See Barré (2010); Reale et al. (2017).

# LITERATURE


**Barre, R.** (2010): Towards Socially Robust S&T Indicators: Indicators as Debatable Devices, Enabling Collective Learning', Research Evaluation, 19(3), 227–31.

**Brewer, J.** (2011): The impact of impact. Research Evaluation; 20(3), 255–256, https://doi.org/10.3152/095820211X12941371876869.

**ERAC** (2017): Final Report of the ERAC Ad-hoc Working Group on Measuring the Impact of EU Framework Programmes for Research and Innovation at National Level, ERAC 1206/17.

**Flecha, R.** (2018): A Note on Key Concepts for Social Impact of Research Assessment. https://archive.org/details/ANoteOnKeyConceptsForSocialImpactOfResearchAssessment.

**Gassler, H., Polt, W., und Rammer, C.** (2006): Schwerpunktsetzungen in der Forschungs- und Technologiepolitik – eine Analyse der Paradigmenwechsel seit 1945, Österreichische Zeitschrift für Politikwissenschaft (ÖZP), 35(1), 7-23.

**Raua, H., Gogginsb, G. und Fahyb, F.** (2018): From invisibility to impact: Recognising the scientific and societal relevance of interdisciplinary sustainability research, Research Policy, 47, 266-276.

**Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P. und Van Horik, R.** (2017): A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research, Research Evaluation, January, 1–11.

**Van den Besselaar, P., Flecha, R. und Radauer, A.** (2018): Monitoring the Impact of EU Framework Programmes, Expert Report, Publications Office of the European Union: Luxembourg.


# DEBATING IMPACT AND MISSION-ORIENTATION OF R&I POLICIES

WOLFGANG POLT, KLAUS SCHUCH, MATTHIAS WEBER, ELKE DALL, MAXIMILIAN UNGER AND NELA SALAMON

DOI: 10.22163/fteval.2019.322

This paper summarizes the main findings from a survey[1] carried out at the occasion of the conference 'RTI Policy in Service of Society: Impact at the Crossroads of Policy Design, Implementation and Evaluation'. This Austrian Presidency of the EU Council conference was organised on behalf of the Austrian Federal Ministry of Transport, Innovation and Technology by the Austrian Platform for Research and Technology Policy Evaluation together with Manchester Institute of Innovation Research and IFRIS – Institut Francilien Recherche, Innovation et Société, in Vienna in November 2018. It was devoted to the challenge of generating, understanding and assessing impact, in particular societal impact, through R&I policy. It discussed new rationales and new demands for R&I policy in service of society, reflected challenges in R&I policy-making triggered by these rationales and demands, and scrutinised what is expected and delivered from different policy intelligence approaches, in particular impact assessment and evaluation.

A part of the conference dealt with developing an understanding of mission-oriented policies (MOPs). The respective results are in the main focus of this paper (based on a survey which was carried out during the conference). The focus on mission-oriented policies emerged against the background of current discussion about the relevance, the pros and cons and the challenges for implementation of such approaches both in the context of the EU as well as on the national level.

By mission-oriented research, technology and innovation policy we understand *"initiatives [which] typically are ambitious, exploratory and ground-breaking in nature, often cross-disciplinary, targeting a concrete problem/challenge, with a large impact and a well-defined timeframe. More specifically, they have a clearly defined (societal or technological) goal with preferably qualified and/or quantified targets and progress monitored along predefined milestones. Directionality and intentionality of these initiatives is what differentiates them from other types of initiatives, such as systemic or challenge-oriented policies"* (JIIP, 2018a, p4). MOPs were suggested as a focusing device to bridge the gap between societal challenges and specific R&I projects (Lamy et al., 2017). With the recently published programmatic paper on mission-orientation in European R&I policy (Mazzucato, 2018), the rationales for a mission-oriented approach have been visibly spelled out as a trigger of further political debate and public consultation. This debate is backed up further by the recommendations from other expert groups (ESIR, 2017; RISE, 2018), two major analytical studies on the empirical evidence on mission-oriented policies (JIIP et al., 2018a and 2018b) and foresight activities (Weber et al., 2018).

While not being the sole topic of the conference, MOPs were addressed in several key-notes by Engelbert Beyer [Federal Ministry of Education and Research, Germany], Mireille Matt [INRA], Goran Marklund [Vinnova] and Matthias Weber [AIT], as well as in dedicated sessions (e.g. sessions on 'Policy designs for impact generation', 'Pathways to impact of R&I Policies'), workshops (e.g. on 'The new mission orientation' and on 'The assessment of societal impact of R&I policy') and plenary debates (e.g. Plenary 1 on *'Designing and supporting mission oriented research policy')*. In addition, at several points of the conference, the audience was encouraged by the moderator to participate in the live survey via the mentioned tool mentimeter. This survey also covered some general questions concerning impact assessments more broadly. The use of the mentimeter tool was regarded as suitable means to elicit some first views on a concept that has only recently re-emerged in policy debates, and on which there are currently no systematic studies available on the expectations that different stakeholder groups attach to it.

296 experts from 39 countries and all continents had registered for the conference. Of these, 255 actually attended the conference. 41.9% of the accredited participants came from Austria. Larger contingents came from the category 'international institutions' (8.1%), especially from the European Commission, but also from the OECD, EUREKA and COST, which made the European dimension of the event visible. 7.1% of the accredited persons came from Germany; 5.7% from the UK; 3.4% each from France and Norway; 3% each from Belgium and Spain and 2.7% from the Netherlands. With the exception of Malta, Slovakia and Slovenia, all EU countries were represented. Other accredited persons from non-EU countries came from Iceland, Norway, Russia, Switzerland, Ukraine, as well as Australia, Brazil, Chile, Iran, Japan, Nepal, South Africa and the USA. 42% of the participants were women and 58% men.

131 of the accredited persons can be assigned to academic research and evaluation. 73 came from agencies, 70 from policy, 13 from intermediary institutions including research infrastructures, 8 from the business enterprise sector and one from the press. For the following analysis they were grouped into 'researchers/evaluators', 'policymakers/agency' and 'other'. The latter group consists of experts from intermediary organisations, the business enterprise sector and media.

As such, they represented a highly qualified audience to discuss the topic. Overall, 242 participants chose to log in the online survey at one or the other point of this two-day event. Generally, we observed a balance between the participants that identified as "policy maker / agency" (42% in one of the survey questions) and "researcher / evaluator" (47%, with the rest identifying as "other").

While the first three questions were asked during the panel, it has to be noted that the MOP related questions (Q7 – Q10) were asked in

1 Using mentimeter ™ technology: https://www.mentimeter.com/

a separate parallel session where attendance was considerably lower (around 60 to 70 persons). Less than half of the participants chose to express their opinions. This in itself might be seen as an indication of the lack of information on, understanding of or interest in the concept and a pointer to the need for further, in-depth discussion. It also needs to be mentioned that, given the overall focus of the conference on the impact of R&I, the researchers attending the conference primarily came from applied and policy research rather than from basic research.

The statements addressed and analysed in this article were:

**Q1:** We are able to measure the social impact of R&I policy (n= 120 or 47.1% of the conference participants)

**Q2:** We are able to attribute R&I Impacts to specific policies (n=120)

**Q3:** We are able to radically change our funding system (n= 119)

**Q4:** What do you think is most important for missions to succeed (n=103)

**Q7:** Missions should be an important part of STI policy in the future (n= 27 or 10.6% of the conference participants)

**Q8:** For implementation of missions, you need substantially new approaches to governance (n= 27)

**Q9**: Missions can be more easily implemented on the national than at the international/EU level (n= 26)

**Q10**: Missions should be more narrowly defined in order to be successful (n= 27)

In the case of the questions 1-3 and 7-10 (see list above), participants were asked to agree/disagree with different statements. These were answered by a Likert scale item, whereby the Likert scale was a number between 1 and 5; 1 standing for "strongly disagree" and 5 for "strongly agree". Note that due to the large difference among the response rates between question groups 1-3 and 7-10, any induction based on comparison of observations among these groups would be misleading.

The main results in our perspective were:

- When assessing the ability to measure social impact of R&I policy (Q1; see figure 1) - a question that is also very important in the context of MOP[2] - one can observe a considerable amount of scepticism (the median values for all groups of respondents ranging from 2 to 3 (= average and below). What is remarkable though is the difference between the groups, with researchers/evaluators being considerably more up-beat about these capabilities than policy makers / agencies or others.
- A slightly more (though again not very) optimistic picture emerges in the assessment of the possibility to attribute R&I impacts to specific policies (Q2), with the median hovering around 3 for both policy makers / agencies as well as researchers / evaluators. This was rather surprising when considering that attribution questions are in general more difficult to answer than impact questions. The impact needs to be identified first, before it can be attributed to the influence of specific policies.
- To a somewhat greater extent, both policy makers and researchers alike believe in the ability to radically change the R&I funding instruments (Q3), although again the overall assessment in these respects is only average, and it spreads across the full specturm from strong agreement to strong disagreement.Hence, the overall estimation with respect to our abilities, both in terms

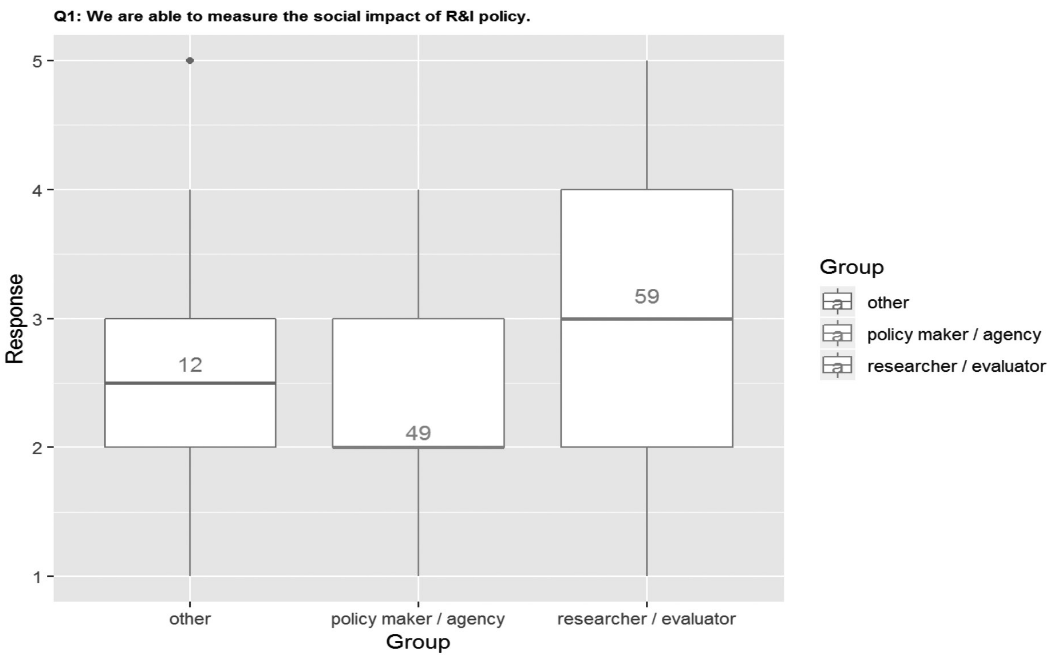



**Fig. 1:** Response to the statement *"We are able to measure the social impact of R&I policy"* by target groups
**Note:** The "heavy"" line in the box-plot is the median and the ends of the box are the first and third quartile (25th and 75th percentile respectively). The extent of the whiskers are the most extreme values still within 1.5 times the box itself (by default). Values beyond the extent of the whiskers are considered to be outliers and are depicted as circles.

2 See Polt / Weber (2014)

of analytic capabilities as well as in terms of abilities to radically change policies might be labelled as a kind of 'sober realism'.

- When it comes to the *questions specifically addressing MOPs*, it has to be kept in mind that response rates were considerably lower than for the general questions. Against this caveat, it can be said that a substantial majority of those answering the question (Q7) supported the view that MOP should play an important part of STI policy in the future [see Fig. 2]. While the median value of this assessment did not differ between policy makers and researchers, the latter were slightly more enthusiastic about this policy approach when taking into account the positive / negative spreads of the answers.
- Also, there is a general recognition that for the implementation of missions, a substantially new approach to governance would be needed (Q8; see Fig. 3).

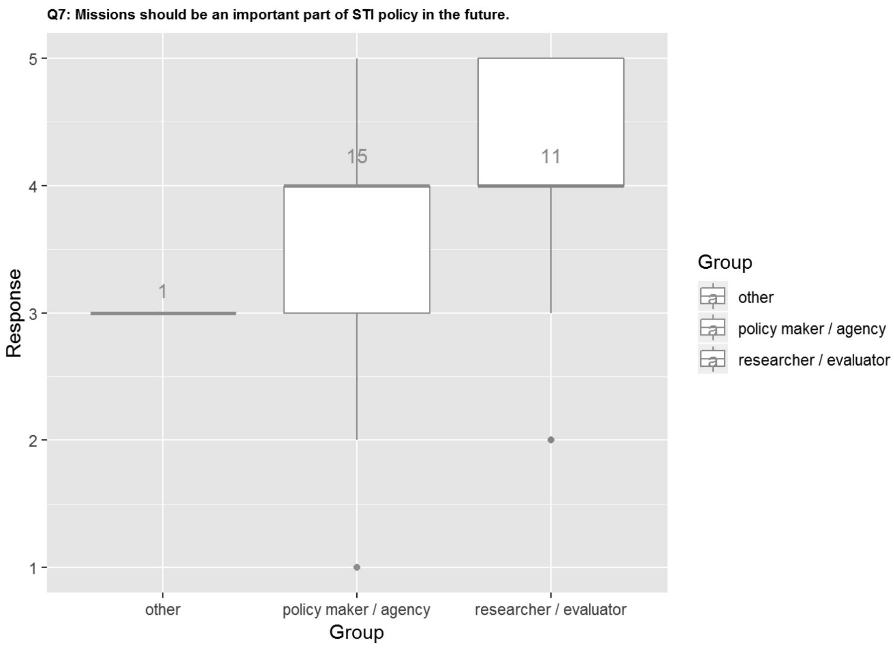



**Fig. 2:** Response to the statement 'Missions should be an important part of STI policy in the future' by target groups

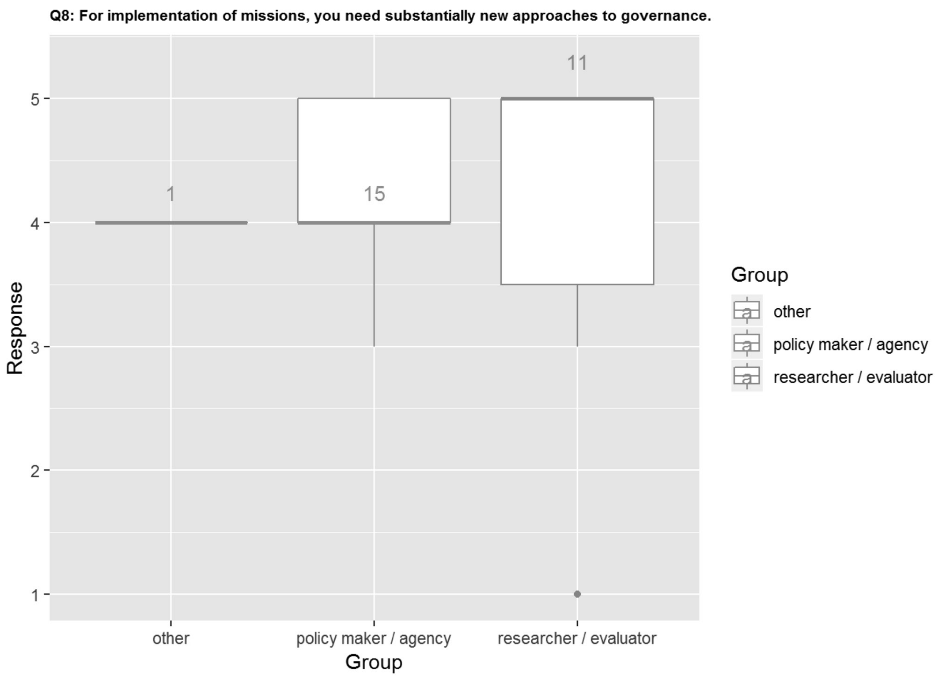



**Fig. 3:** Response to the statement *'For the implementation of missions you need substantially new approaches to governance'* by target groups

- Major differences in the perceptions of MOP emerge when actors responded to the questions whether MOP could be more easily implemented at the national than at the EU level (Q9): While policy makers predominantly perceived the national level as less suitable entry point (median=2), researchers strongly saw the national level as the one to prefer (median=4, see Fig 4.). This picture might be explained by the strong recent emphasis on MOPs in the conceptual debates in the European Commission, while on the national level, policy debates only very recently have also centred on this issue. On the other hand, researchers, from their experiences with the empirical material might be led by the observation that most MOPs currently in place are in fact carried out at the national level and hence their perception might be a 'positivist' one. Moreover, the granularity of missions may vary considerably: some missions can well be addressed at the level of even smaller EU member countries, but others (and probably the better known examples) require the bundling of capacities of several European countries to have a chance to be addressed successfully.
- Likewise, the perception whether a MOP should be more broadly or more narrowly defined in order to be successful (Q10) was markedly different between policy makers and researchers. The former being much more in favour of a more narrow definition (median=4, range from 5 to 3, see Fig. 5), while the latter seemingly leaning towards a broader concept of MOP (median=3, range towards 2). Here, policy makers seem to show some hesitation with respect to broader and hence more managerially more challenging MOPs, which is coherent with the answering patters vis-a-vis the questions on the implementation challenges. In line with this argument, the answers also seem to reflect a different understanding of policy makers/agencies and researchers/evaluators when referring to "success" in addressing a mission. For the former, running a good R&I programme relevant to a mission may well be a success, whereas the latter may see this from a longer-term perspective of triggering change in society and economy.
- When asked, which factors are most important for a mission to succeed (Q4, see Fig. 6), the 'engagement of national and regional stakeholders' ranked first, followed by 'the development of capacities for pro-active, flexible management' and the 'portfolio of instruments'. Of lesser importance was seen the 'measurement and impacts by goals and milestones'. This perception is in line with the one seeing MOP as a challenging task of aligning the actions of a considerable number of actors associated to a mission and the corresponding management challenges. This ranking broadly coincides with the one of the importance of challenges (again stakeholder engagement being seen as the most important challenge) and the capacity development of management on second place. Interestingly, though, the 'portfolio of instruments' was seen as a major challenge only by a minority – maybe reflecting the fact that the respondents mostly came from countries with well-developed tool boxes of STI policy instruments.

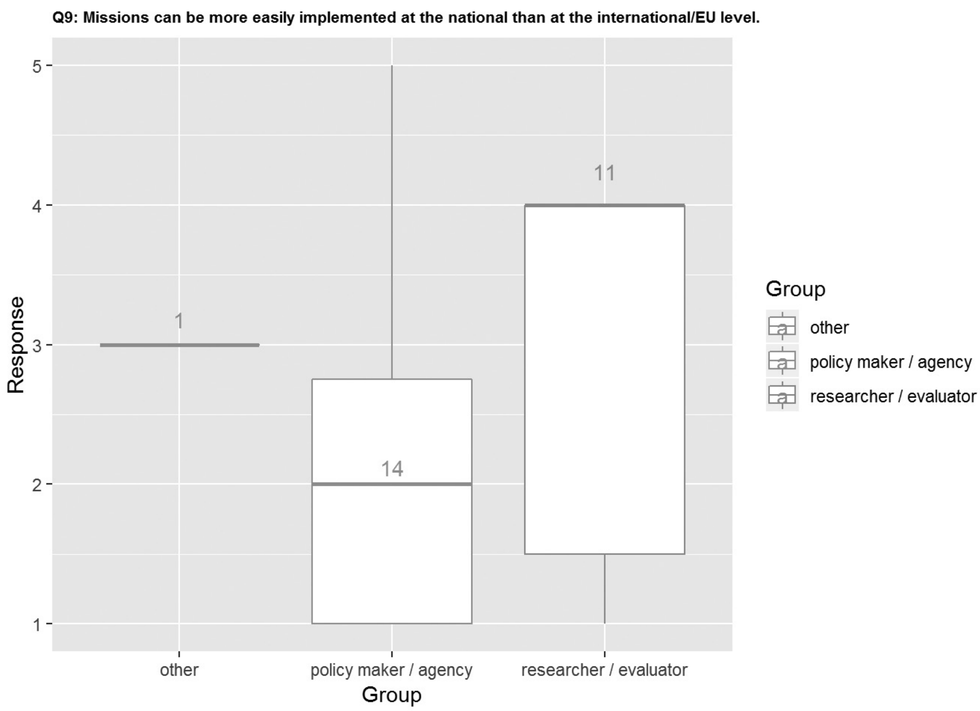



**Fig. 4:** Response to the statement *'Missions can be more easily implemented on the national than at the international/EU level'* by target groups

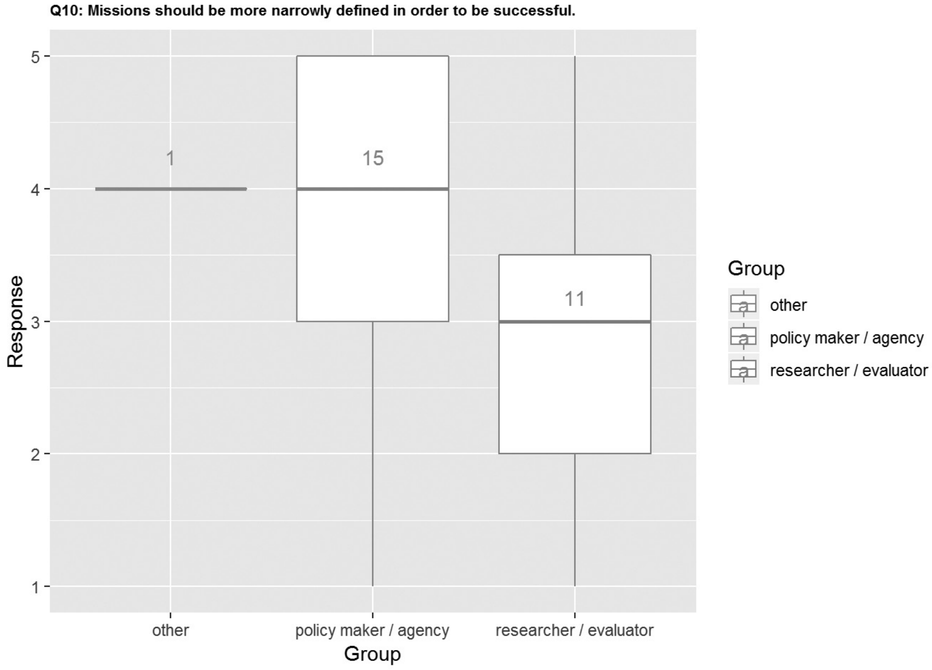
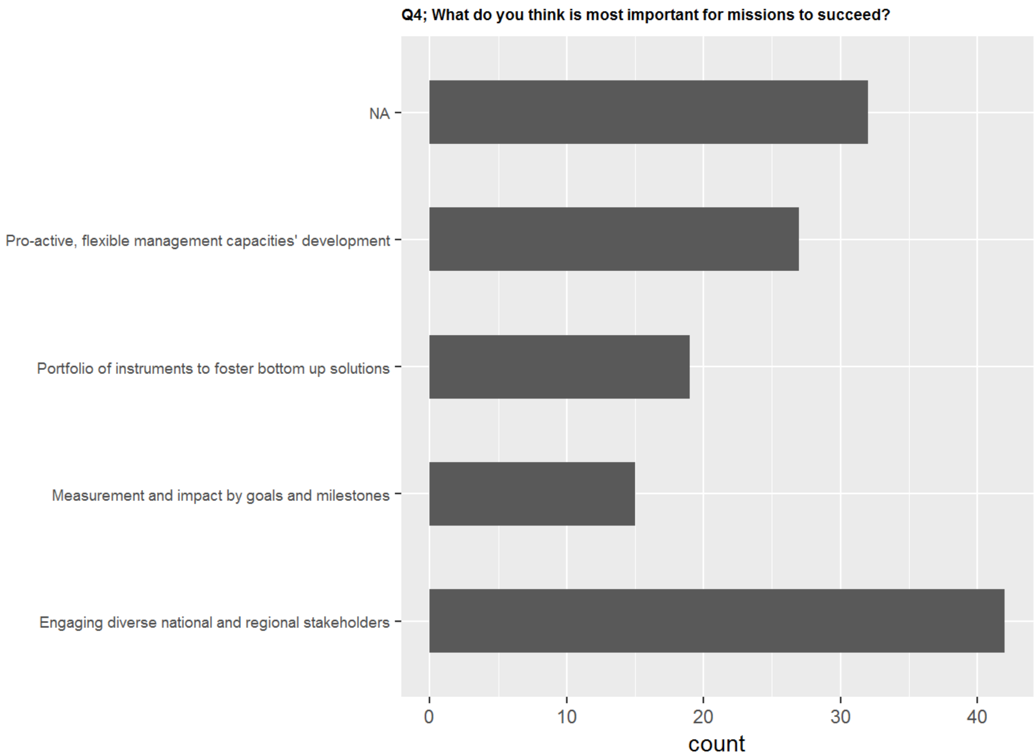



**Fig. 5:** Response to the statement *'Missions should be more narrowly defined in order to be successful'* by target groups




Code: NA = no answer

**Fig. 6:** Response to the question *"What do you think is most important for missions to succeed"*

To sum up: the survey might shed some light on the current state of debate on MOPs, especially on the differences in perceptions between actor groups: its implementation is seen as challenging and would have to be accompanied with the development of substantial new management capabilities and probably a quite radical change in policy orientation. There seems to be some hesitation (especially on the side of the policy makers) whether such a change can be achieved and the respective capabilities could really be developed. By analysing the answers to the open question about "perceived challenges", it seems that at the stage of discussion we are, the definition and selection of missions is perceived as the main concern. This major concern is closely followed by issues addressing the governance of MOP, centring on the issue of necessary political support. The participants also addressed the challenge of coordination and communication with the main stakeholders and the resistance that might be encountered. Subcritical funding of the missions and over-ambition are other potential critical issues mentioned.

On the positive side, most respondents would see and welcome an increased role of MOP in STI policy. Apparently, there is still need for an intense debate about MOP for which the near future will already provide quite some opportunities.

# REFERENCES

**ESIR** (2017). Towards a Mission-Oriented Research and Innovation Policy in the European Union: An ESIR Memorandum: Executive Summary. European Commission, Brussels.

**JIIP et al.** (2018a). Mission-Oriented Research and Innovation: Inventory and characterization of initiatives. Final Report. European Commission, Brussels.

**JIIP et al.** (2018b). Mission-Oriented Research and Innovation: Assessing the impact of a mission-oriented research and innovation approach. Final Report. European Commission, Brussels.

**Lamy, P., Brudermüller, M. et al.,** (2017). LAB – FAB – APP. Investing in the European future we want. Report of the independent High Level Group on maximising the impact of EU Research & Innovation Programmes. European Commission, Brussels.

**Kuittinen, H., Polt, W. and Weber, K.M.** (2018). Mission Europe? A revival of mission-oriented policy in the European Union: In: RFTE – Council for Research and Technology Development (Ed.): RE:THINKING EUROPE. Positions on Shaping an Idea. Vienna, September 2018, pp. 191-207

**Mazzucato, Mariana** (2018). "Mission-oriented research & innovation in the European Union: A problem-solving approach to fuel innovation-led growth." European Commission, Brussels.

**RISE** (2018): "Mission-Oriented Research and Innovation Policy: A RISE Perspective." European Commission, Brussels.

**Weber, M., Andreescu, L. et al.** (2018). Transitions at the Horizon: Perspectives for the European Union's future research- and innovation-related policies. Final Report of BOHEMIA project. European Commission, Brussels.

# AUTHORS

**WOLFGANG POLT, MAXIMILIAN UNGER**
*Joanneum Research POLICIES*

**KLAUS SCHUCH, ELKE DALL, NELA SALAMON**
*Zentrum für Soziale Innovation ZSI*

**MATTHIAS WEBER**
*Austrian Institute for Technology*

# EVALUATION OF THE TOP-LEVEL RESEARCH INITIATIVE

GÖRAN MELIN

DOI: 10.22163/fteval.2019.323

## INTRODUCTION

This paper presents an evaluation of an unusual research programme, as it is an example of a supra-national research programme, where five countries joined forces and created a common pot of funds, to be distributed to researchers in the five countries without respect to national origin, only research quality. Cooperation between at least three of the countries was however required in each of the funded projects.

The aim of the paper is to present the programme to a wider audience and especially its outcome with respect to impact of research. Of particular interest is an attempt to evaluate and measure societal impact and societal 'readiness' of the projects that were included in the programme. Towards the end, a discussion is held of what impact that can be expected when funding research of the kind at hand in this case.

## THE TOP-LEVEL RESEARCH INITIATIVE

In the autumn of 2008, the prime ministers of the five Nordic countries met and joined forces to create the largest joint Nordic research and innovation initiative to date: The Top-level Research Initiative (TRI). The budget of the programme was DKK400m over five years (~€50–55m).

The TRI addressed issues of climate, energy and the environment with the overarching idea to strengthen the Nordic competitive advantage in science and innovation in these areas. The initiative involved the very best agencies and institutions in the Nordic region, and some 200 researchers from universities and research institutions and 63 companies participated. Multi-disciplinary coordination was emphasised, including sciences and social sciences as well as business and industry.

Budgetary funding was divided among the Nordic countries in proportion to their GDP. On top of this, the Nordic Council of Ministers, as well as the Nordic organisations NordForsk, Nordic Energy Research and Nordic Innovation, all contributed to the financing of the TRI. The TRI was organised as a true common pot, with none of the financially contributing partners being guaranteed an equal share of the research grants.

The TRI was a result of an ambition in the Nordic Council to establish a Nordic research arena that would increase the level and ambition of collaboration among Nordic research as well as creating a basis for enhanced Nordic participation in EU framework programmes. Whereas it today can be concluded that several of the programmes of the TRI not only had the potential to enhance collaboration and to support innovation in key future technology areas, the TRI was tightly coupled with the climate change agenda. Although the aim of the TRI was threefold – climate, energy and the environment – accounts of achievements of the TRI seem to stress the fight against global warming.

The research was organised in 20 projects. Each project belonged to one of six sub-programmes with different profile. There were three types of projects: Nordic Centres of Excellence (NCoE) and Integrated Projects. NCoE were large centres for existing Nordic research communities with participants from at least three Nordic countries. The NCoE aim to increase and facilitate cooperation between excellent researchers, research groups or institutions in the Nordic countries to strengthen the communities and enhance the international profile in prioritised areas in the Nordic countries through joint research and researcher training, joint management and leadership, and shared infrastructure.

Integrated Projects were research projects involving research partners from the Nordic countries and more decidedly involving business partners. These projects focused on involving non-academic partners and thus facilitating ties to business and end-users. The IPs included industry partners and operated under four of the six sub-programmes: Energy Efficiency with Nanotechnology, Integration of large-scale wind power, Sustainable biofuels, and Carbon Capture and Storage (CCS). Last, there were a couple of projects that were labelled 'Studies'.

## DESIGN; METHODS

A set of evaluation questions guided the work of the evaluation. There are twelve evaluation questions, organised under four headings.

- Societal and scientific impact of the TRI
  1. How has the TRI contributed to societal and scientific impact?
  2. In what ways have the TRI funded projects reached out and influenced stakeholders outside the scientific community?
  3. In which areas have the TRI been most successful in reaching out?
- Nordic added value of the TRI
  4. In what ways have visibility and attractiveness of Nordic research increased in a European and global context?
  5. In what ways has the TRI facilitated appropriate division of work and specialisation between the Nordic countries?
  6. To what extent have the TRI projects been integrated and fed back into the national research systems?
  7. How has the efficient and flexible use of the Nordic resources been ensured?
- Societal readiness for innovation and research
  8. How are TRI funded projects distributed on the Societal Readiness Level scale?

- Applicability and utilisation of the innovation and research outcome
  9. In what ways have the activities supported by the TRI contributed to innovation?
  10. How has the TRI contributed to knowledge and innovation that serves the needs of business and society?
  11. How has the TRI contributed to increased international cooperation in research?
  12. How has the TRI contributed to strengthened Nordic international competitiveness?

The focus for the evaluation, thus, was on the results and impact the TRI had, and on the Nordic added value that the programme brought. The concept of Nordic added value is rather vague, but commonly agreed to exist where initiatives or activities are best and most efficiently carried out in a Nordic context rather than on national or EU level.

Data used in the evaluation were collected from a wide range of sources:

- Document studies
- Interviews with 33 individuals
- Self-assessments of funded projects
- eSurvey to project leaders and participants
- Case studies of seven projects
- Bibliometric analyses

# RESULTS WITH RESPECT TO IMPACT

The TRI was ambitious, and several important results and effects can be observed. This was already a conclusion from the final report of the ongoing evaluation, presented in 2014. Now, four years later, the following can be concluded:

- **THE RESEARCH FUNDED BY TRI IS GENERALLY OF HIGH OR VERY HIGH SCIENTIFIC QUALITY**

The TRI projects have produced scientific publications that are more cited in high-end journals than would be expected from a statistical viewpoint. The TRI projects produced more high-end scientific publications, and also a broader base of high-end publications. In recent years there continues to be a steady stream of publications.

- **THE NCOE HAVE HAD LARGER SCIENTIFIC IMPACT THAN THE INTEGRATED PROJECTS**

This is, of course, to be expected. The Integrated Projects have had more industry-oriented impact, and at the same time attracted a large portion of additional research funding. Some of the Integrated Projects have been able to produce profitable solutions – some of which have been commercialised.

- **THE TRI PROJECTS HAVE ATTRACTED A LARGE AMOUNT OF ADDITIONAL FUNDING**

The research carried out in TRI projects in total have attracted at least €73.5m in additional funding, or close to 150 per cent of the total budget of the programme. This shows that the research was of high quality, and also highly relevant. The additional funding stems to a large extent from national funding sources, but there are several examples of related research projects in EU consortia enabled through the TRI. Among several researchers, there is an impression that additional funding opportunities from Nordic institutions have been few.

- **THE TRI LARGELY FUNDED ALREADY ESTABLISHED RESEARCH AND RESEARCHERS**

This was the purpose from the start, and this also partly explains the much better-than-average scientific quality and citing rates.

- **THE TRI HAS CONTRIBUTED TO INCREASED AND SUSTAINABLE NORDIC RESEARCH COLLABORATION**

The NCoE have been able to more firmly consolidate their collaborations through additional funding from Nordic and EU sources. Due to the fact that the additional funding for Integrated Projects is secured mainly from national funding sources, there is less international cooperation, although some is enabled through EU funding. However, the sustainability of these collaborations depends on availability of continued and relevant funding opportunities. There are several examples of potential international collaborations between research organisations.

- **THERE ARE NO CLEAR DIFFERENCES IN IMPACT BETWEEN THE SUB-PROGRAMMES**

There are high-impact projects in all sub-programmes, as well as projects with less impact. That said, it is difficult to label projects with less impact, as this in several cases may still be too early to fully appraise.

- **TRI ENABLED RESEARCH COLLABORATIONS THAT WOULD OTHERWISE NOT HAVE BEEN POSSIBLE**

Although well-established researchers were funded, as noted above, they had not always collaborated before. Through TRI, existing cross-border research collaborations were strengthened, and in several cases included research partners (and in some cases companies) formerly not part of the network.

- **TRI ENABLED TO CARRY OUT PROJECTS THAT WOULD OTHERWISE NOT HAVE EXISTED**

Some of the project ideas would obviously have found other means of funding, but they would then in most cases have been more national in scope.

- **THE TRI CONTRIBUTED TO THE TRAINING OF AT LEAST 81 PHD STUDENTS**

This is an important contribution to national and Nordic strengths in these areas. The presupposed Nordic orientation of these researchers' continued professional careers also helps to create Nordic added value, as does the mobility of these individuals.

- **PARTICIPATION IN THE TRI PROJECTS CLEARLY CONTRIBUTED TO VALUE CREATION AND NEW CONTACTS AND PARTNERSHIPS FOR THE COMPANY PARTNERS**

For the companies, TRI contributed to increased R&D cooperation, mainly within the Nordic countries. There are examples where the Integrated Projects had an impact on the development and application of scientific methods for participating companies as well as external companies.

- **THERE ARE SOME EXAMPLES OF CLEAR SOCIETAL IMPACT**

The results from the NCoE are clearly useful for public policy actors, although there are yet few clear signs of direct policy impact. The results from Integrated Projects have had some influence on public actors, especially in Iceland regarding the country's potential for wind power and for CO2 storage. When discussing potential societal impact, the future and potential importance of the large number of PhDs co-funded by the programme, and who thus have received a Nordic perspective and grown a Nordic network, should be noted.

- **THE TRI CLEARLY CONTRIBUTED TO NORDIC ADDED VALUE**

All the points above indicate this direction.

# SOCIETAL READINESS FOR INNOVATION AND RESEARCH

Societal Readiness Levels (SRL) is a way of assessing the level of societal adaptation of, for instance, a particular social project, a technology, a product, a process, an intervention, or an innovation (whether social or technical) to be integrated into society.  If the societal readiness for the social or technical solution is expected to be low, suggestions for a realistic transition towards societal adaptation are required. The lower the societal readiness is, the better the plan must be for transition. These are the SRL levels:

- SRL 1 – identifying problem and identifying societal readiness
- SRL 2 – formulation of problem, proposed solution(s) and potential impact, expected societal readiness; identifying relevant stakeholders for the project
- SRL 3 – initial testing of proposed solution(s) together with relevant stakeholders
- SRL 4 – problem validated through pilot testing in relevant environment to substantiate proposed impact and societal readiness
- SRL 5 – proposed solution(s) validated, now by relevant stakeholders in the area
- SRL 6 – solution(s) demonstrated in relevant environment and in cooperation with relevant stakeholders to gain initial feedback on potential impact
- SRL 7 – refinement of project and/or solution and, if needed, retesting in relevant environment with relevant stakeholders
- SRL 8 – proposed solution(s) as well as a plan for societal adaptation complete and qualified
- SRL 9 – actual project solution(s) proven in relevant environment

The SRL scale is still not a broadly recognised concept, and it was indeed difficult for interviewees to assess projects according to this scale. We therefore chose to focus this assessment on the smaller number of projects that were selected as case studies, altogether seven projects. Also, the SRL scale has several features in common with the more broadly recognised Technology Readiness Level concept. The former, to some extent, mirrors the latter.

Technology Readiness Levels (TRL) are used to assess the maturity level of a particular technology. Each technology project is evaluated against the parameters for each technology level and is then assigned a TRL rating based on the projects progress. There are nine technology readiness levels. TRL 1 is the lowest and TRL 9 is the highest. The TRL levels are as follows:

- TRL 1 – basic principles observed
- TRL 2 – technology concept formulated
- TRL 3 – experimental proof of concept
- TRL 4 – technology validated in lab
- TRL 5 – technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 6 – technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies)
- TRL 7 – system prototype demonstration in operational environment
- TRL 8 – system complete and qualified
- TRL 9 – actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space)

Given the difficulties in assessing the projects on the SRL scale, we chose to map the projects on both levels. A comparison of how the projects perform on the TRL scale provides better possibilities to assess how they are positioned on the SRL scale. For the purpose of this paper, where the evaluated projects and their individual characteristics may be less relevant, the outcome of the mapping is shown as an example of how the SRL and TRL scales can be used side by side in order to investigate societal readiness. Figure 1 depicts the seven projects' positions on the SRL and TRL scales.

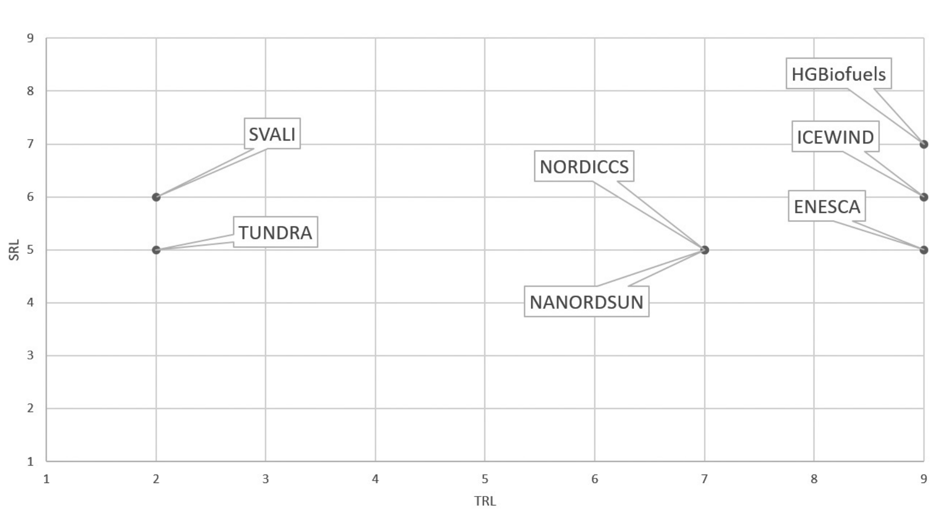



**Fig. 1:** The projects' positions on the SRL and TRL scales.
**Source:** Technopolis Group

The mapping of the projects on the two levels shows that six of the seven are at SRL levels 5 and 6. They are well past the initial stages of identifying and formulating the problem, and address validating proposed solution(s) by relevant stakeholders in the area or demonstrating these solutions in relevant environment and in cooperation with relevant stakeholders to gain initial feedback on potential impact. One project – HG Biofuels – is at the stage of refinement of project and/or solution. This suggests that these projects (the solutions) have come a relatively long way to be integrated into society.

The span is larger on the TRL scale. The Integrated Projects (projects with more industrial participation) are, not surprisingly, at higher levels. Some of them have actual systems proven in operational environments and products on the market. Two NCoE (projects of a solid academic character) – SVALI and TUNDRA – are, as is logical, at the other extreme of the TRL scale.

# DISCUSSION ABOUT IMPACT OF THE TRI

No doubt, the TRI was a successful programme. It delivered in accordance with the expectations – and in some cases more. The programme performed well on its overall objectives, in particular those addressing "the highest quality in research and innovation by combining the strongest Nordic communities" and "strengthen national research and innovation systems".

The programme's impact still ought to be put in some perspective. It was indeed the largest Nordic research programme in history, but the total budget was still modest compared to what the Nordic countries invest individually in research in these areas. Just to take one example, the Swedish Energy Agency has an annual budget of around SEK1.6b (close to €160m) for research and innovation in the field of energy for ecological sustainability, competitiveness and security of supply (Government bill 2016). The Agency's mandate and area of support reflect a much larger commitment and a much larger area than that of the TRI, but budget figure does give a perspective to the relative weight of national and Nordic investment in research and innovation, in related research areas. Results and impact that came out of the TRI need to be regarded in this perspective.

While the TRI was a unique effort with strong political backing, it was not enough to radically change the Nordic research landscape in its target research area. That would have required a long-term commitment with subsequent programmes or funding opportunities on Nordic level. It would probably also have required a closer alignment with national priorities and funding schemes. This is an insight that could be taken into account if launching similarly ambitious initiatives in the future.

Neither was the creation of Nordic collaboration as a platform for increased international cooperation within the EU and beyond a central aim for most of the projects and their participants. There is evidence from some projects that this was after all achieved, but it was not necessarily a key driving force or motivation for the researchers when applying for and carrying out research collaboration with funding from the TRI.

The TRI still had a clear Nordic added value. Through the TRI, real cross-border collaboration between researchers and some companies did take place, including networking of importance for PhDs and senior researchers, resulting in several examples of continued collaboration/contacts. It is likely that the training of the (at least) 81 PhDs will have long-term impact on joint Nordic research, 'marinated' in Nordic collaboration as they are.

# REFERENCES

**Arnold E and Eriksson M** (2009). Experience from the first Nordic Top Level Research Initiative. See e.g. http://www.toppforskningsinitiativet.org/filer/TRI_brochure_spread.pdf-en

**Government bill Sweden** (2016). Prop. 2016/17:66 Forskning och innovation på energiområdet för ekologisk hållbarhet, konkurrenskraft och försörjningstrygghet.

# AUTHOR

**GÖRAN MELIN**

*Technopolis Sweden, Faugert & Co Utvärdering*

Skeppargatan 27, 11452 Stockholm, Sweden

# FROM MEASURING IMPACTS TO MAPPING IMPACT SYSTEMS:

## LESSONS FOR THE IMPACT ASSESSMENT OF RESEARCH INFRASTRUCTURES FROM A STUDY ON THE EUROPEAN SOCIAL SURVEY ERIC[1]

PETER KOLARZ

DOI: 10.22163/fteval.2019.324

## ABSTRACT

Commonly agreed standards and methodologies for Research Infrastructure (RI) impact assessment continue to be elusive, despite the efforts of several expert groups across the globe. Against this backdrop, Technopolis conducted a major impact study of the European Social Survey ERIC in 2016/17. A mixed methods approach yielded a broad range of valuable findings on the academic, non-academic and teaching impacts of the ESS. The study also arrived at the notion of 'impact systems': sets of RI-, country-, and sector-specific framework conditions that simultaneously highlight impact pathways, help formulate avenues for future impact optimisation, and enable meaningful comparison and benchmarking between participating countries. This paper posits that mapping impact systems will be a valuable component to any future RI impact study and contributes to the on-going debates about RI impact assessment standards.

## BACKGROUND – IMPACT ASSESSMENT OF RESEARCH INFRASTRUCTURES (RIS)

Research Infrastructures (RIs) are focal points for continuous interaction between scientific, technological, socio-economic, political and policy development.[2] But operating RIs requires a growing share of public funding, and government and research funding institutions are increasingly interested in the the added value that RIs provide. Yet, it is difficult to quantify and understand returns on investments into RIs in conventional commercial terms. Investment in RIs brings a broad range of benefits that spread across wider society rather than serving merely the direct stakeholders (owners and users of RIs).

In 2014, The Global Science Forum (GSF) set up an expert group to examine potential priorities for RI policy that should be addressed at the global level. The GSF secretariat then carried out a review of existing reports and identified that a standard impact assessment framework is missing and there is no agreed model shared between funding agencies and/or RIs' organisations to measure socio-economic impact.[3] Other organisations, including most recently an ESFRI Strategic Working Group, are dealing with these concerns.

Currently, a heterogeneous set of methods is applied to capture the effects of RIs, most of which address standard economic impacts (direct effects) and to some extent economic multipliers. However, comprehensive and methodologically demanding studies are still rare. Core aspects of RI benefits, such as their impact on policy, human and social capital formation and innovation, are not extensively explored. Moreover, impact assessment will differ with scale (e.g. national mid-scale vs. large international facilities), type (e.g. different pathways and productive interactions for single-sited vs. distributed vs. virtual e-RI) or discipline (e.g. applied technical science vs. social sciences and humanities vs. environmental observation platforms).[4] A fully standardised set of performance indicators uniformly applicable to all RIs is unlikely ever to materialise: the breadth of different RIs (thematically, conceptually, structurally) does not appear to allow for such a level of standardisation in evaluation and impact assessment. However, a move towards more common frameworks (even if this does not extend to the point of standard indicators) would benefit the policy community, especially in terms of comparative endeavours to weigh up the value of various RIs.

1 Acknowledgements: Jelena Angelis, Annemieke Biesma, Victoria Blessing, Neil Brown, Anneloes de Ruiter, Katre Elias-Taal, Patrick Eparvier, Samuel Grange, Adam Krcal, Bea Mahieu, Ernestine Mbengue, Göran Melin, Kerli Müürisepp, Reda Nausedaite, Kalle Nielsen, Loic Perroud, Leonor Rivoire, Jasper Schipper, Paul Simmonds, Ingvild Storsul-Opdahl, Julia Synnelius, Appoline Terrier, Vincent Traag, Erik van Wijk (CWTS), Martin Wain, Elisabeth Zaparucha, Frank Zuijdam.

2 Griniece E., Reid A. and Angelis J. (2015) Evaluating and Monitoring the Socio-Economic Impact of Investment in Research Infrastructures, Technopolis Group

3 Moulin J. (2016) Workshop on Methodologies and Tools for assessing Socio-Economic Impact of Research Infrastructures, Global Science Forum (Paris, 3 November 2015)

4 Ibid.

# THE IMPACT STUDY OF THE EUROPEAN SOCIAL SURVEY (ESS) ERIC

The ESS is an international, comparative survey of social and political values and attitudes, which was launched in 2002 and is now in its 9th round of data collection. In 2013, it was given the status of a European Research Infrastructure Consortium (ERIC). In total, 24 countries (including 'guest' countries) participated in the eighth round of data collection. Since its inception, over 120,000 people have registered as ESS users. Around 64% of these are students, a further 27% can be classified as academics (research/ faculty/ PhD) and just under 10% come from other societal domains (e.g. policy, NGOs, businesses, private individuals).

The impact study of the ESS ERIC[5] was undertaken in 2016/17 as a work package of the Horizon 2020 project 'ESS-SUSTAIN', and was carried out by Technopolis, with bibliometric analysis by the Centre for Science and Technology Studies (CWTS). The study assessed the academic, non-academic and teaching impacts that have been achieved through the ESS, by all different user groups and in all ESS member and observer countries. It also assessed how these impacts came about ('pathways' to impact), identified best practice, and made recommendations to ensure the long-term sustainability of the ESS. This study presents one of the largest and most recent endeavours to assess the impact of a major pan-European RI. In the absence of an existing standard approach, we opted for a mixed methods approach, comprising:

- Desk research/ document review of existing evaluations and impact studies of the ESS and other related material (e.g. literature on the impact of other European research infrastructures)
- Analysis of ESS user data (supplied by the ESS data warehouse situated at NSD in Bergen, Norway)
- Observation/ attendance of events organised by the ESS or featuring presentation of ESS data (e.g. the 3rd ESS conference, Lausanne, July 2016)
- 100 interviews with internal stakeholders (National Coordinators, General Assembly members, members of other ESS advisory boards and committees) and external stakeholders (academic and non-academic ESS users)
- An online survey (n=2238) of active ESS users (users who logged in to the ESS data portal at least once in the 12 months leading up to the point of surveying)
- A short online survey of student users
- Analysis of publication information captured by the ESS in the 'ESS Bibliography'
- Publication and citation analysis of ESS-based publications listed in Web of Science (WoS)
- 36 case studies featuring detailed description of specific instances of ESS use and its academic, non-academic or teaching impact.

Each method step produced valuable information in its own right. However, there was an over-arching logic in the mixed-methods approach, in that it was critical for the study team to develop a detailed understanding of the benefits that the ESS brings to its users (including advantages over other survey resources). These benefits could be qualitatively assessed once the ESS user-base had been defined and mapped, and only after this step did the study assess what outputs, outcomes and impacts had been generated as a result of the benefits brought about by the ESS.

The study thus progressed from general assessments (size of the user base, reasons for using the ESS) to specific examples of impact. Additionally, the study sequence helped develop an understanding of the ESS, particularly through the consideration of 'benefits'. These are not uniform, but often differ country-by-country, and highlighted many unanticipated benefits (for example around ESS use for teaching), which in turn shaped the selection of output and impact indicators later in the study.

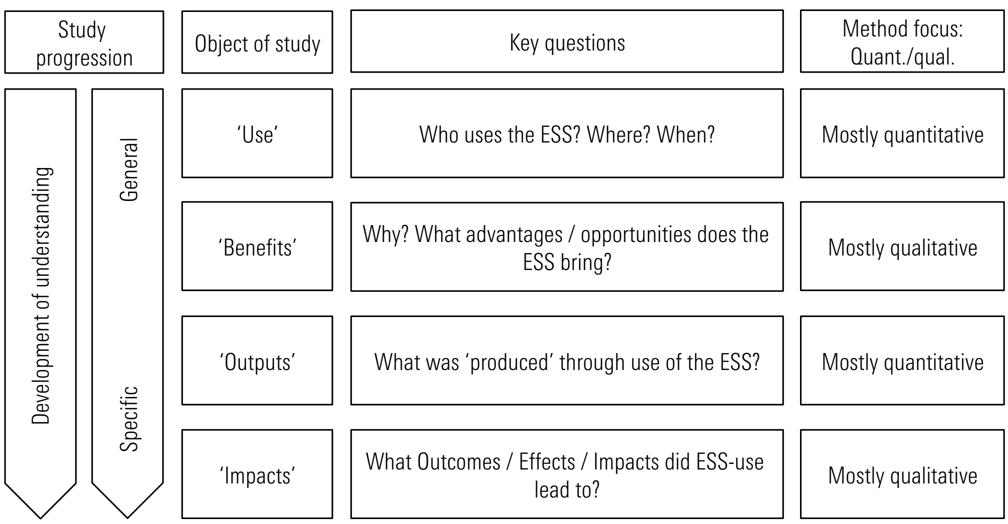



**Fig. 1**: Methods rationale for the ESS ERIC Impact Study

5 Kolarz P, Angelis J, Krcal A, Simmonds P, Traag V and Wain M (2017) Comparative Impact Study of the European Social Survey (ESS) ERIC. Technopolis Group. Available: http://www.europeansocialsurvey.org/findings/impact

# MEASURING IMPACTS – FINDINGS OF THE STUDY AND SUCCESS OF THE METHOD

The methodology was successful in that it enabled a comprehensive picture to be created of the use-intensity, the benefits, and the academic, non-academic and teaching impacts of the ESS ERIC. The study highlighted substantial differences between countries on a range of measures, and reached findings in both quantitative (e.g. user numbers, institutional concentration, output numbers, citation impacts) and qualitative terms (e.g. types of impact, types of value added, new fields and research questions enabled). Key identified impact 'highlights' include:

- There are over 2,700 known ESS-based outputs, including 1,373 journal articles. 817 ESS-based journal articles are listed in Web of Science (WoS). 22% of these fall into the top-10% most cited articles within their respective microfield (10% would be the expected average). Even at the level of individual institutions, ESS-based work almost always scores higher on citation metrics than is generally the case for each institution's WoS-listed publications in the social sciences overall (based on Leiden Rankings).
- Whilst high-quality and highly impactful research has been conducted in many different places, there are several institutions that form major 'hotspots' of ESS-based work, both in terms of high publication output and impact, and in terms of high student numbers learning statistical methods via the ESS. The Universities of Ghent, Leuven, Radboud Nijmegen, Tartu, LSE, NTNU, Cologne and Zurich are all examples of such clustering.
- Non-academic impacts appear in a wide range of different organisations, often in government ministries or agencies. Immigration and quality of life/wellbeing are fields where many non-academic impacts have occurred, but several other fields also feature non-academic impacts, including law enforcement, policing and justice, health inequalities, LGBT rights, children and family policy, and active ageing.
- Impacts identified include supporting policy creation or policy change, political agenda-setting, as well as influence on political and public debate more broadly. Additionally, the ESS often influences government monitoring: statistical agencies and other entities have in several cases drawn on the ESS, either by integrating certain ESS data into their own monitoring reports, or adopting various methodological standards practiced by the ESS.

Robustly assessing the impact of an RI doubtlessly has merits in itself: it ensures money is well spent and demonstrates areas of particular strength. However, the ESS impact study also moved beyond descriptive to analytical dimensions to arrive at recommendations, and considered 'impact pathways', i.e. how the observed impacts were achieved, as well as the drivers and barriers to impact.

A difficulty in this task is that the impacts of the ESS are so varied that a short typology of impact pathways is almost certainly non-exhaustive and prone to over-simplification. Moreover, substantial differences observed between individual member countries in terms of use-intensity, output, perceived benefit and types of impact constitute a further complicating factor. To generate a better framework to assist future impact optimisation of the ESS, the study posited the notion of 'impact systems'.

# THE USE OF IMPACT ASSESSMENT: MAPPING THE IMPACT SYSTEM OF THE ESS

Typically, research impact is thought of as a linear process. In the case of the ESS, a generic model might involve that a user first accesses ESS data and will then process it further. The ESS data might be immediately put to use as a teaching resource, replacing other data sets used in the past, leading to better teaching materials and more capable students (teaching impact). The ESS user might use the data to do further analysis and gain new knowledge, which is then published. The resulting outputs would be read, cited, and drawn on or responded to by other researchers, leading to changes in debates and academic perspectives (academic impact). Further, the new knowledge gained through the ESS may be disseminated (via published outputs or otherwise, via intermediaries or directly) to non-academic users. Research users then draw on the information, leading to debate input, policy or practice development (non-academic impact).

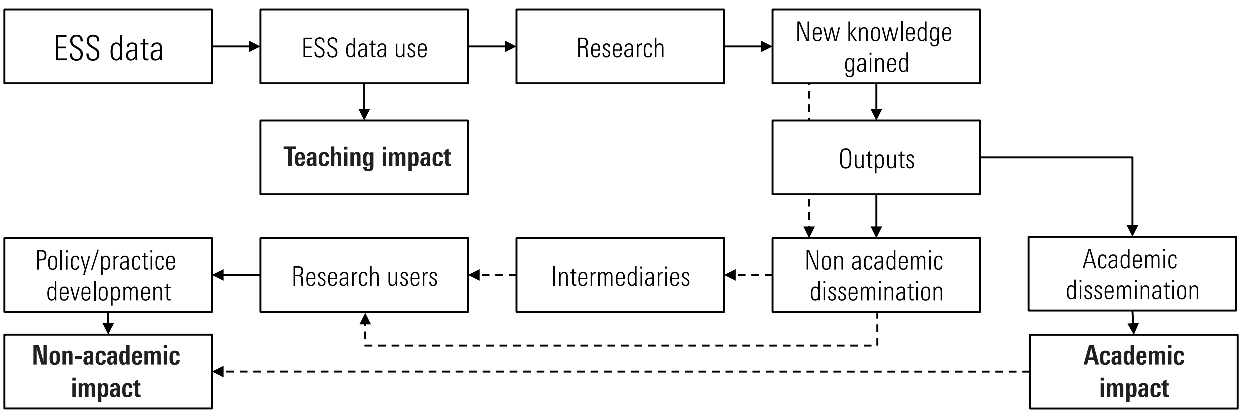



**Fig. 2:** The linear model of ESS impact

Linear models of this type have been envisaged in the past by organisations including the UK's Economic and Social Research Council.[6] Indeed, many impact case studies conducted as part of the ESS impact study follow variations of this generic formula.

However, a critical further finding of the study is that such linear 'stories' do not occur in isolation, and the likelihood of their incidence is dependent on context. The study identified a range of framework conditions that affect the extent to which people use the ESS in the first place, the purposes for which it can be of further use, and the overall ease with which knowledge transfer between academic and non-academic domains can take place.

Conditions of this type variously apply to the overall organisation and continuity of the ESS, the organisation and activities undertaken in terms of funding and at the level of national coordination, as well as more broadly at the level of overall academic, policy and knowledge transfer cultures in different countries. Research for the ESS impact study yielded a broad range of such framework conditions, notably including the following:

- Our survey results show that non-student ESS users most commonly first became aware of the ESS as students. When used in teaching, a generational effect therefore occurs, where student users move on to becoming academic or professional ESS users by virtue of existing familiarity in their subsequent academic or non-academic careers (should their remit permit this). However, the extent of ESS use for teaching purposes is also dependent on the availability of alternatives: some countries have many existing, high quality open access national datasets that students can use to learn, for instance, about statistical analysis and survey methods. Other countries have fewer alternatives, so the ESS becomes a more attractive option for teachers to use.
- To facilitate non-academic impact, a degree of 'translation' is often necessary. This can be in terms of simple data presentation (i.e. simplifying, visualising), so that ESS use in the news media becomes more feasible. ESS undertakes some such activities centrally, and National Coordination teams also make such efforts in some countries. Think Tanks, NGOs or other intermediaries may undertake further efforts of this kind, but different countries have different types and levels of proliferation of such organisations. In short, 'translation' may occur at central ESS level, or at country level by NC teams, or by organisations unconnected to the ESS.
- The notion of 'evidence based policymaking' differs between countries. Some have long-standing norms around making extensive use of survey data, others not so much. Moreover, in some countries direct use of data by ministries or government agencies is typical (and in some countries, sectoral ministries even part-fund the ESS with the intention of using ESS data for policy), whilst in others it is more common to contract academic experts to bring their knowledge into the relevant non-academic sphere in person. This affects the way in which policy impact is likely to occur.
- At the purely academic level, some countries have more pronounced traditions of quantitative methods in the social sciences, whilst others will place a far greater emphasis on qualitative and theoretical approaches, both in terms of research and teaching. Where the latter is the case, the ESS is likely to struggle much more to be used widely, especially when quantitative methods do not feature strongly on teaching curricula.
- Long term sustained funding of the ESS is an important condition for impacts to occur: without this, potential users have no guarantee of data availability in future, which presents difficulties for establishing the ESS as a go-to data source, or to use ESS data in policy monitoring activities. Likewise, many research questions or practical concerns require data from particular sets of countries to be available. Researchers or practitioners often wish to compare their country with other countries that are nearby, so inclusion of adjacent countries can be an important requirement. This is especially important in countries that are often 'grouped', e.g. the Baltics, the Visegrad group, the Eurozone, Scandinavia.
- In each country, some individuals may naturally gravitate to the ESS, but the national coordination team has an important role to play in terms of promotion: where promotion of the ESS is undertaken, user numbers grow, and so does the scope for impact. However, resources for promotion activities vary between countries, and over time.
- Different countries prioritise the transfer of knowledge from academia to practical fields in different ways, which in turn affects the extent and shape of that transfer. The UK's 'impact agenda' for instance ensures that the national research assessment system rewards cases of non-academic impact, providing an incentive to engage with non-academic domains. However, such impacts need to be based on excellent research, so outputs are an important part of the impact 'pathway'. Academics communicating ESS-based information without the presence of any particular outputs (for instance by providing a simple data training workshop to a non-academic organisation) may be more strongly incentivised in other systems.

Several other framework conditions were identified by the ESS impact study, mostly through qualitative engagement and often including highly country-specific institutions and norms. When put together, these conditions can be mapped into an impact system, which tracks the possible channels of ESS use and impact pathways and the likely intensity of use in different domains. It also helps account for why various channels of use and pathways to impact are more pronounced in some countries than in others. For example, the ESS is a particularly valuable teaching resource in countries that do not have existing social surveys as long-established teaching tools for statistical methods in universities. Whilst this is characteristic of smaller countries and those with weaker research systems, heavy ESS-use in teaching also entails a 'generational' effect, where student users become academic or professional users later in life; an effect likely to be especially strong in precisely those countries.

Figure 3 provides a generic overview of the ESS impact system. The various linkages (represented as arrows) may be stronger or weaker per country (even per topic or field), or affected by any number of contextual factors and framework conditions.

6 ESRC (2009) Taking Stock. A Summary of ESRC's Work to Evaluate the Impact of Research on Policy and Practice

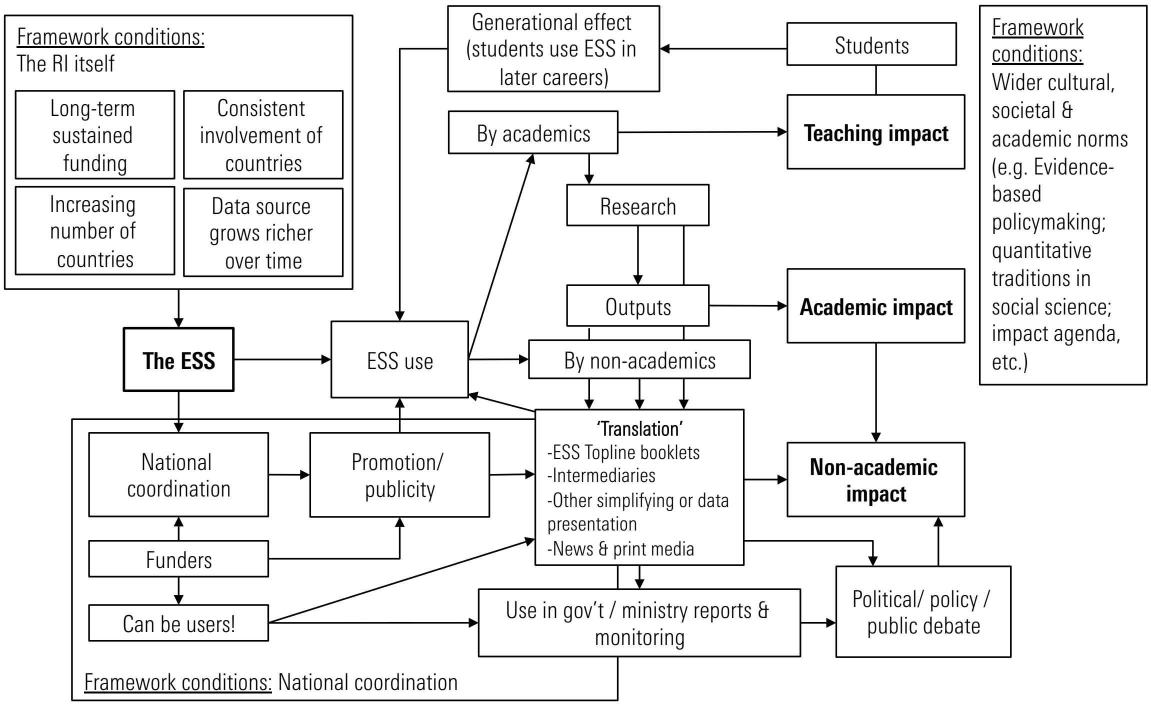



**Fig. 3:** The ESS impact system and framework conditions

Some aspects of the ESS impact system cannot readily be changed, or will only change slowly over time (e.g. an overall more qualitative or theory-driven social science tradition in a given country, where survey-based research is rare in the first place). These factors can help explain and contextualise lower levels of ESS-use or fewer clusters of ESS-based teaching or research activity. However, others can be affected, such as the consistent involvement of a country in the ESS, or the level of outreach and publicity conducted by the national coordinating team or the inclusion of major potential data users in the coordination (or funding) itself.

# KEY POINTS ON IMPACT SYSTEMS

The notion of impact systems contains two related implications for the impact assessment of Research Infrastructures (RIs): the importance of contextual understanding (system comprehension), which needs to form a critical part of any standardised approach to RI impact assessment and, secondly, the consequent importance of mixed methods, where system comprehension shapes indicator selection and informs scope for comparability – both between countries and, potentially, between different RIs.

The impact system can be mapped for an RI as a whole, as was done for the ESS ERIC. For each participating country (or indeed, for each relevant field of research or practice), particular system components and systemic strengths or weaknesses can be highlighted. In the first instance, this helps the formulation of recommendations for future optimisation. It also contextualises the measurable impacts and prevalence (or lack) of certain impact types, reducing the risk of meaningless and un-qualified comparison on impact indicators between different member countries or fields.

In the ESS impact study specifically, this approach helped generate several findings that affected the feasibility of comparative performance assessment between individual member countries. For example, low teaching use and few teaching impacts in certain countries are a reflection of existing, nationally long-established teaching resources, rather than a failure of those countries to appropriately harness the ESS for such purposes. Likewise, ESS use in countries with predominantly qualitative traditions in the social sciences cannot readily be compared with ESS use in countries where quantitative traditions dominate: in the former, the ESS must be assessed in terms of whether it has brought around cultural shifts, whilst in the latter an expectation of widespread, high-impact academic work is more appropriate. The same principle applies to countries with sophisticated vs. embryonic cultures of evidence-based policymaking.

Mapping impact systems likewise holds some promise for comparison between different pan-European RIs. The notion of impact systems foremost helps to highlight country-, topic- and RI-specific particularities and as such acts as a warning against benchmarking all member countries of an RI, or even different RIs, uniformly and with identical indicators. However, where benefits and framework conditions are similar, meaningful comparisons may become possible. Figure 4 shows how, in the hypothetical scenario of two RIs being assessed with the same system-oriented framework, some 'common ground' may be identified, allowing for comparisons on certain indicators to take place. The systemic perspective can highlight whether a certain indicator is relevant to both cases, and whether system components mean that the indicator can be interpreted in the same way for both or whether adjustments need to be made (e.g. if RI1 has a much larger user base than RI2 due to a broader thematic coverage).

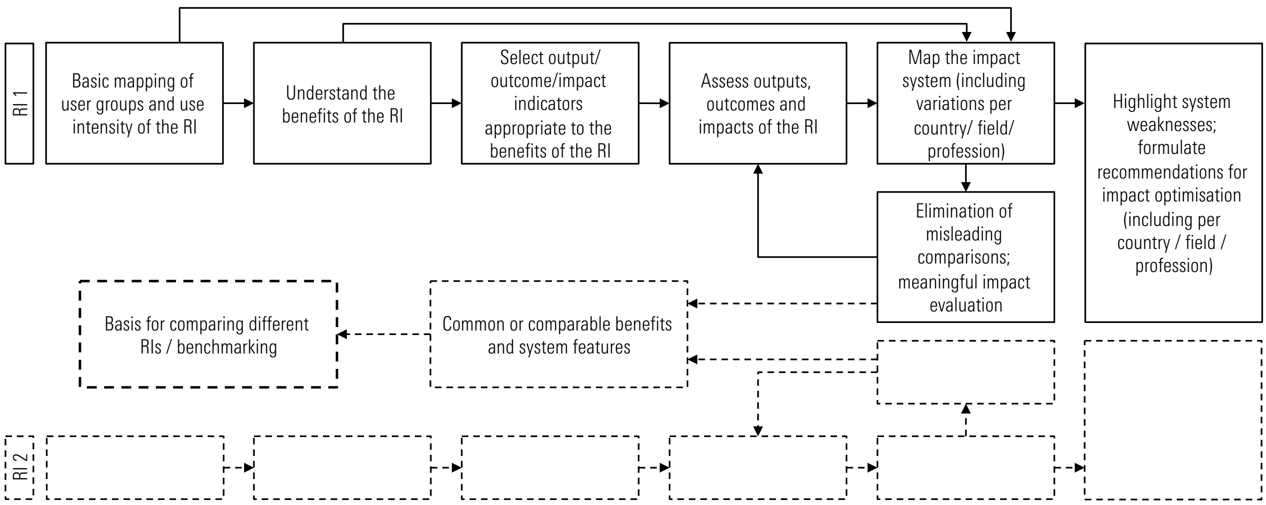



**Fig. 4:** Using Impact system mapping for RI impact assessment and comparison

Methodologically, the notion of impact systems also highlights the importance of mixed methods: quantifiable indicators are critical in order to demonstrate the value of RIs, and indeed to consider the comparative value of different RIs. However, to make meaningful judgements of this kind, The identification of output, outcome and impact indicators must be underpinned by qualitative investigation. Understanding the benefits of an RI (to all identifiable user groups) at the early stages of the impact assessment, and working in the later stages towards mapping the impact system has been shown in the ESS ERIC impact study as a helpful way of directing these qualitative method components. When impact systems have been understood, even modest impacts can be suitably highlighted if they are known to occur under adverse system conditions, while the scope for meaningful benchmarking in the pan-European RI-landscape is strengthened.

# SOURCES

**ESRC** (2009). Taking Stock. A Summary of ESRC's Work to Evaluate the Impact of Research on Policy and Practice.

**Griniece E., Reid A. and Angelis J.** (2015). Evaluating and Monitoring the Socio-Economic Impact of Investment in Research Infrastructures, Technopolis Group.

**Kolarz P, Angelis J, Krcal A, Simmonds P, Traag V and Wain M** (2017). Comparative Impact Study of the European Social Survey (ESS) ERIC. Technopolis Group. Available: http://www.europeansocialsurvey.org/findings/impact

**Moulin J.** (2016). Workshop on Methodologies and Tools for assessing Socio-Economic Impact of Research Infrastructures, Global Science Forum (Paris, 3 November 2015).

# AUTHOR

**PETER KOLARZ**
*technopolis /group/ United Kingdom*
3 Pavilion Buildings
UK-Brighton BN1 1EE
E: peter.kolarz@technopolis-group.com

# IMPLEMENTING S3 WITH CLUSTERS – AN INNOVATION MODEL FOR TRANSFORMATIVE ACTIVITIES

MICHAEL KELLER, IRIS REINGRUBER, MATEJA DERMASTIA, JACQUES BERSIER AND GERD MEIER ZU KÖCKER
DOI: 10.22163/fteval.2019.325

## ABSTRACT

The ongoing debates on updating Smart Specialisation Strategies (S3) in the European innovation policy framework mainly focus on practical implementation challenges. This paper draws on the specific experience from the Interreg Alpine Space project S3-4AlpClusters, which put the interplay between S3 and clusters at the core of its conceptual and practical study of S3-implementation across the Alpine Space. While overlaps between the two concepts are evident and cluster initiatives are acknowledged in the relevant literature as tools in the context of S3, practical implementation of S3 with cluster initiatives is found to be far from trivial and involves specific challenges. We therefore introduce an innovation model as a practical effort to better integrate cluster initiatives in the S3 process. The model is a systematic process for the regional and cross-regional identification and development of transformative activities (TA). Tools and methodologies for S3-implementation, such as S3-synergy diamonds, entrepreneurial discovery workshops (EDW) or action development workshops (ADW) are valuable individual contributions for future policy designs. Nevertheless, it is only by putting them into the context of a systematic innovation model, with a strong focus on transformative activities, and by making them the levers for cross-regional cooperation and a systematic involvement of cluster initiatives in regional economic development, that they become fully relevant for smart transformation processes leading to innovation within businesses, new value chains and jobs in innovative new areas.

## INTRODUCTION

In the context of regional economic development, there is an increasing interest to identify industrial transformation processes that lead to the emergence of new value chains and related industries. Such processes can provide competitive advantage for regions if they are timely identified and properly supported and represent huge potentials for regions to develop and ultimately to create jobs in innovative new fields. In its communication on Strengthening Innovation in Europe's Regions, the European Commission highlights that globalization requires regions to tackle the transformation of existing economic structures, *inter alia* by designing Smart Specialization Strategies (S3) and cluster policies (European Commission, 2017). This paper draws on the recent experience from the Interreg Alpine Space project S3-4AlpClusters[i], which put the interplay between S3 and clusters at the core of its conceptual and practical study of S3-implementation across the Alpine Space. While overlaps between the two concepts are evident and cluster initiatives are acknowledged in the relevant literature as tools in the context of S3 (see Ketels, 2013a), there is, to our knowledge, no comprehensive study on how clusters are currently involved in the practical development and implementation of S3. Moreover, practical implementation of S3 with cluster initiatives is found to be far from trivial and involves specific challenges. We therefore propose a novel focus on the interplay between S3 and clusters (Chapter I) and introduce an innovation model as a practical effort to better integrate cluster initiatives in the S3 process (Chapter II). The model is a systematic process for the regional and cross-regional identification and development of transformative activities (TA), which is currently implemented across the Alpine Space in the regions participating in the S3-4AlpClusters project[ii]. We provide insight into this practical experience to illustrate the proposed innovation model with examples (Chapter III) and conclude the paper with recommendations for current and future policy debates on S3-implementation.

## I. SMART SPECIALIZATION STRATEGIES AND CLUSTERS

### THE S3 FRAMEWORK

Smart Specialization Strategies (S3) play a crucial role in European regional development and innovation policy. Article 2(3) of the Common Provisions Regulation for the European Structural and Investment Funds (EU, 2013) defines S3 as intended *"to build competitive advantage by developing and matching research and innovation own strengths to business needs in order to address emerging opportunities and market developments in a coherent manner"* (p. 338). As a practical matter, S3 are of fundamental importance for the thematic objective of *"strengthening research, technological development and innovation"* within the common strategic framework of the European structural and investment funds (ESI Funds) (EU, 2013, pp. 347 ff.). As an *ex ante* conditionality for funds of the European Regional Development Fund (ERDF) in the 2014-2020 programming period (see EU, 2013, p. 438), they have become a common policy lever at national and regional levels within the European Union. While concrete implementation agendas for S3 strongly depend on regional and thematic contexts, some recognized basic principles guide the overall S3 process. The challenge at the heart of Smart Specialization Strategies (S3) approach is the need for regions to use their limited resources effectively to become and remain competitive in the global economy (see *inter alia* Foray *et al.,* 2009; Foray *et al*., 2012; Foray, 2015). Based on a principle of targeted spending (see *e.g.* Enos, 1995), regions need to achieve diversification by specializing on a limited number of prioritized economic activities to take advantage of knowledge spillovers and economies of scale and scope. Successful diversification is contingent on exploiting existing related variety (see Breschi *et al.*, 2003; Frenken *et al*., 2007; Boschma, 2017). In other words, regions should aim at tapping into opportunities for transformation to meet structural challenges by combining their existing capacities into unique innovative activities (smart specialization). Opportunities for transformation are critical in the S3 framework. Regional competitive advantage is created when opportunities for transformation are exploited by regions to combine their existing capacities into unique new domains (see Foray *et al,* 2012). As an ultimate goal, these activities in new domains of opportunities should translate into structural transformation within the economy in an "accumulative process that links the present and future strengths of a regional economy in a particular domain of activity and knowledge" (Foray and Goenaga, 2013, p.6).

Based on the finding that innovation requires prioritization and the provision of specific capacities and coordination devices (see *e.g.* Hausmann and Rodrik, 2006), Foray *et.al.* (2012) conclude that *"smart specialisation involves making choices, leading to priority setting and channelling resources towards investments with a potentially higher impact on the regional economy"* (p. 114). Specialization priorities are best identified through an entrepreneurial discovery process (see Coffano and Foray (2014). The bottom-up character of this approach is crucial. As noted by Foray and Goenaga (2013), *"Entrepreneurs […] are in the best position to discover the domains of R&D and innovation in which a region is likely to excel given its existing capabilities and productive assets"* (p.5). The term entrepreneurs is understood in a very broad sense and includes actors such as innovative firms, research leaders from academia, representatives of the regional innovation system or specialists from tech-transfer with knowledge of the scientific and technological domains covered in the region (see Foray *et al.*, 2012). Once identified, priorities need to be implemented. Foray and Goenaga (2013) note that "new options" for diversified regional systems and *"emergence and early growth of new activities, which are potentially rich in innovation and spillovers"* should be enabled through the generation of *"critical mass, critical networks [and] critical clusters"* (p.9). In this process of creating critical mass, connectivity is decisive. Cross-sectoral links are key drivers of specialized technological diversification. It has to be noted, that such links in related variety are not limited by regional borders. Cross-regional cooperation is a decisive element in the endeavour to generate critical mass in the presence of economies of scale and scope and indivisibilities in infrastructures and other assets. To quote Foray *et al.* (2012), *"match what you have with what the rest of the world has!"* (p.17).

As this short conceptual introduction hints at, there is obvious common ground between the principles underpinning S3 and the abundant literature on economic geography. Economies of agglomeration are widely acknowledged as a key driver of diversification and specialization processes (see Rosenthal and Strange (2004) and Cortright (2006) for a comprehensive review of the economies of agglomeration literature). The positive impact of agglomerations of related economic activity on regional innovation performance has been studied extensively (see *inter alia* Feldman and Audretsch, 1999; Porter, 2003; Feser *et al.,* 2008; Glaeser and Kerr, 2009; Delgado *et al.,* 2010 and 2014; Neffke *et al.,* 2011). More particularly, the work of Michael Porter (Porter, 1990; Porter, 2003; Porter, 2008, Ketels and Keller, 2015) established the concept of clusters and cluster initiatives as a cornerstone for regional innovation policies. Given these apparent conceptual overlaps, clusters are also acknowledged as tools in the context of S3 (see Ketels, 2013a). Nevertheless, there is, to our knowledge, no comprehensive study on how clusters are currently involved in the practical development and implementation of S3. We therefore propose a novel focus on the interplay between S3 and clusters.

## A BENEFICIAL INTERPLAY

As a *"geographical proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and externalities"* (Porter, 2011, p.215), clusters are of apparent interest in the development and implementation process of S3. More specifically, we understand clusters as groups of companies, mainly SMEs, and other actors (government, research and academic community, institutions for cooperation, financial institutions) co-locating within a geographic area, cooperating around a specialized niche, and establishing close linkage and working alliances to improve their competitiveness (see Ketels, 2011; Delgado *et al.,* 2012). A cluster initiative is the organized effort aiming at fostering the development of the cluster either by strengthening the potential of cluster actors or shaping relationships between them. They can be compared to regional networks and are usually organized by a cluster management (see Christensen *et al.,* 2012; Ketels, 2013b; Lindqvist *et al.,* 2003).

The interplay between S3 and clusters implies a two-way relationship with reciprocal benefits between the two concepts (Figure 1). The reliance on specific regional capacities in S3 emphasizes the importance of existing local resource concentrations. Cross-sectoral connectivity, inherent in the cluster concept, is a crucial determinant for the creation of critical mass for Transformative Activities (see Foray *et al.,* 2012). Moreover, clusters typically reunite the actors of the quadruple helix, crucial for cooperative leadership in an entrepreneurial discovery process. Strongly paralleling the definition of clusters, Foray (2015) concludes that preference in the process of developing and implementing S3 should be given to a *"mid-grained level of aggregation – the level at which activities group together a certain number of firms and partners that collectively explore and discover a new pathway to transformation"* (p.3). Finally, clusters are not limited to borders, but often stretched over several regions, which facilitates the cross-regional cooperation often beneficial for creating critical mass (see Foray, 2012). These considerations emphasize that cluster initiatives, as an organized form of the cluster concept, are ideal tools to use in the process of developing and implementing S3. On the other hand, clusters are also recognized as typical beneficiaries and direct recipients of S3-enhanced innovation. Indeed, *"generating a vibrant innovative cluster"* is considered *"a logical outcome"* of S3 (Foray, 2015, p.59). The whole process of establishing and collectively exploring new areas of opportunity*, "will possibly form the basis for [new] local resource concentration"* (Foray, 2015, p.15), by sparking entrepreneurship, spillovers and innovation at the cluster level. In this perspective, clusters are vehicles transmitting S3-enhanced innovation processes to the business level, ultimately contributing to establish new value chains and create jobs in innovative new fields. Translated to the policy level, this means that cluster policies benefit from being driven by S3 (see Foray, 2015, p.59), a view confirmed *inter alia* by Ketels (2013a) stressing that in relation to

S3, cluster policy becomes fully relevant at a later stage. In other words, meaningfully integrating clusters in the process of developing S3 opens

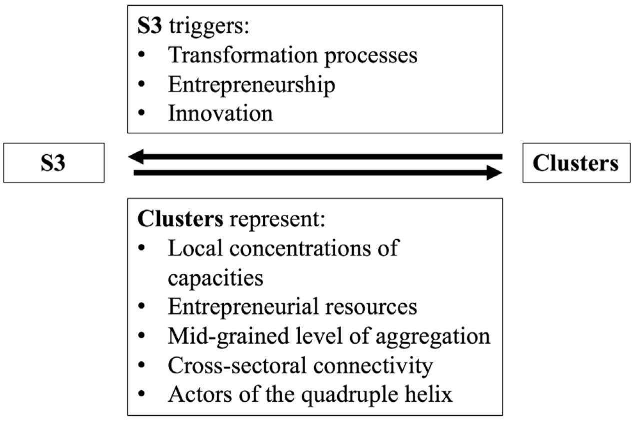



**Fig. 1:** Beneficial Interplay between S3 and Clusters
**Source:** Authors' elaboration.

up vast new perspectives for clusters in regional development policy.

As a practical consequence, the interplay between S3 and clusters represents a huge potential for implementation of S3. In a nutshell, the involvement of clusters into S3 helps to identify entrepreneurial resources and areas of strategic potential. Located at an intermediate level between individual firms and broad sectors, clusters typically reflect strong partnerships, vibrant communities and relevant connections between related businesses, suppliers and associated institutions. Clusters embrace all relevant actors of the innovation process and provide important information about needs, opportunities and ongoing transformations – all essential elements of S3. In addition, clusters are not limited to any border, but often stretched over several regions where they can facilitate the implementation of actions through interregional cooperation. In short, clusters are ideal vehicles to transmit S3-enhanced transformation processes to the business level and to give S3 real impact in terms of innovation within enterprises, new value chains and jobs in innovative new fields with high growth potential. Nevertheless, recent experiences from the Alpine Space, backed by studies from other regions (*e.g.* Nögel *et al.,* 2018), show that the potential of the interplay between S3 and clusters is not fully exploited in current S3-implementations efforts (see Meier zu Köcker *et al.*, 2017; Bersier and Keller, 2018).

## IMPLEMENTATION CHALLENGES

We have gathered and analysed extensive experience of real-world S3 development and implementation with clusters during the last two years within the Interreg Alpine Space project S3-4AlpClusters[i]). All 11 regions participating in the project have set up cluster initiatives and developed S3 or similar regional strategies[iii]). For all participating regions, we studied the role of clusters in the implementation process of S3 and compared it with experiences from outside the Alpine Space in a stress test approach based on an online survey of regional stakeholders, consisting *inter alia* of regional clusters and policymakers concerned with regional development and innovation policy (Meier zu Köcker and Dermastia, 2017). In addition, a thorough synergy analysis of regional S3 documents resulted in a report on strategic Alpine Space topics for interregional cooperation (Meier zu Köcker *et al.,* 2017). The analytical process was paralleled by strong interactions in several series of workshops with all regional stakeholders, including cluster managers, enterprises, SMEs, policymakers and academia (see Foray, 2017; Foray *et al*., 2018).

The real-world experience with cluster initiatives within the project provides strong evidence on how the interplay between S3 and clusters is currently being implemented at regional level. Overall, the results of the project activities confirm the relevance of the interplay between S3 and clusters. Clusters are well-acknowledged tools in the context of S3 and cluster-based regional development policy is recognized to yield good results. However, ways and extent to which clusters are involved in the development and implementation of S3 vary significantly between the studied regions and reveal untapped opportunities for cluster initiatives in the process (see Meier zu Köcker and Dermastia, 2017). Two elements in particular have been identified as critical:

a. Lack of focus on transformation
The role clusters can play to trigger real transformation processes in the transmission of S3 to the real-world business level remains insufficiently exploited because of a lack of focus on real transformation processes. The investigations revealed that the scope of priority areas defined in S3 tends to be very broad and driven by a focus on existing specialization, rather than opportunities for real transformation. If priorities are defined too broadly, connections, synergies, and spillovers will hardly happen and critical mass will not emerge. As a result, many regions tend to end up with similar broad priority areas and the intended diversification across regions is hampered (see Meier zu Köcker *et al.,* 2017). The practical experience with S3 development in the regions of the Alpine Space demonstrates that the identification of priorities and the generation of critical mass is far from trivial and requires appropriate processes and tools (see also Coffano and Foray, 2014; Nögel *et al.,* 2018). In a context of innumerable potential combinations of existing capacities and diffuse hopes of bonanza behind any new trend, the identification of transformation opportunities requires a solid base of evidence to guide the entrepreneurial discovery process. Sticking to broad priority areas, regions systematically neglect to focus on transformation processes in their S3 documents (Meier zu Köcker *et al.,* 2017).

b. Lack of need-based cross-regional cooperation
Clusters are crucially lacking tools for need-based interregional cooperation, which would enable them to contribute critical mass, connectivity and cross-sectoral links across regional borders. While the focus on related broad priority areas across Alpine Space regions impedes the identification of real transformation opportunities, it also represents an untapped potential and common ground to jointly tackle Alpine Space related challenges (ranging from issues such as economic globalization over demographic change to energy) through the development of cross-regional activities. Regrettably, the analysis conducted within the S3-4AlpClusters project[i]) revealed a quasi-total absence of cross-regional cooperation to exploit such synergy potentials within the Alpine Space. Indeed, the business environments and framework conditions for cross-regional cooperation tend to be weak, poorly aligned between regions

and completely lacking focus on need-based cooperation (see Meier zu Köcker and Dermastia, 2017 and Meier zu Köcker *et al.,* 2017). A need-based approach to cross-regional cooperation would be particularly vital for regions that are too small to implement structural transformation on their own. Tapping into external capacities and bundling regional competences would allow them to generate necessary critical mass, especially for resources confronted with economies of scope, scale and indivisibilities. Opportunities for transformation are often present at the intersection between different existing traditional industries. Regions lacking a strong and broad industrial base crucially depend on need-based cooperation to succeed in gaining sufficient critical mass to implement S3 (see Meier zu Köcker *et al.,* 2017).

The lessons learned from the S3-4AlpClusters project reveal clear challenges in current development and implementation of S3 in the Alpine Space (see Figure 2). The systematic identification of priorities is a complex exercise requiring new tools to support the entrepreneurial discovery process. The development of concrete actions is in many cases hampered by the lack of critical mass. Cross-regional cooperation based on complementary needs is critically missing from the given framework conditions. Given the huge potential of cross-regional cooperation and cluster-based processes, these challenges represent a clear call for action to enhance practical implementation of S3. Regions and their cluster initiatives need to be equipped with a systematic process for the development and implementation of S3 to boost their impact on businesses, new value chains and job growth in innovative new fields.

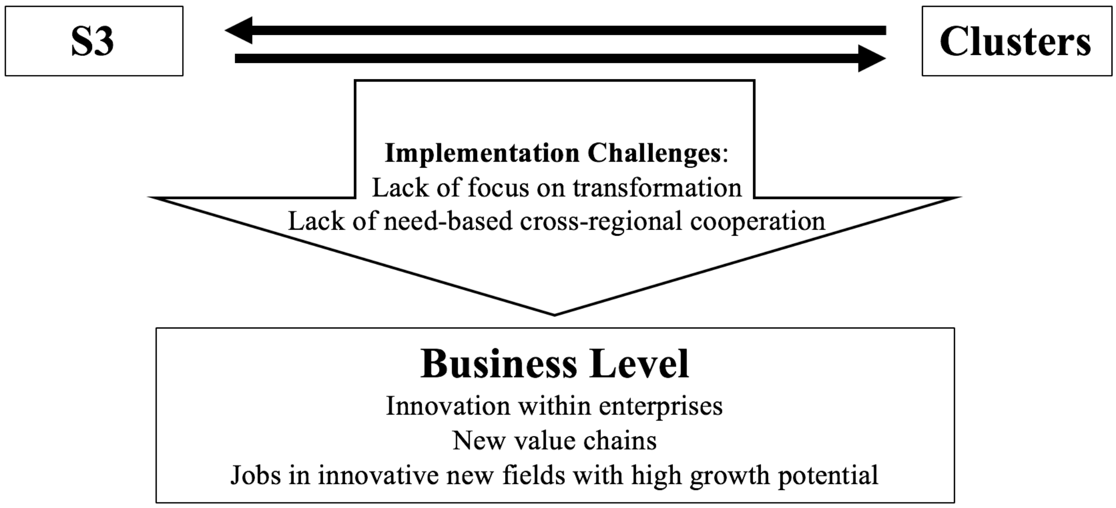



**Fig. 2:** Implementation Challenges
**Source:** Authors' elaboration.

# II. AN INNOVATION MODEL FOR TRANSFORMATIVE ACTIVITIES

## OBJECTIVES

To address the identified challenges, we introduce an Innovation Model as a systematic approach to implement S3 with clusters. The model has three core objectives:

1. Ensure a focus on transformative activities (TA)
2. Provide a process to implement S3 with cluster initiatives
3. Enable cross-regional cooperation

In a nutshell, the model offers a new perspective for cluster initiatives and regions to explore capacities and opportunities for transformation and to develop actions to create critical mass in innovative new fields both regionally and cross-regionally. The approach is a timely and innovative contribution because it directly addresses main obstacles in current S3 implementation (see Chapter I above).

## A NOVEL FOCUS ON TRANSFORMATIVE ACTIVITIES (TA)

The idea of transformative activities (TA) has been inherent in the concept of S3 since the latter was first formalized in 2009 by Foray *et al.* as a result of the reflections of the Knowledge for Growth Expert Group, established by the European Commissioner for Science and Research Janez Potocnik. Nevertheless, it has been the intense practical experience with S3-implementation in European regions (as evidenced *inter alia* in the Interreg Alpine Space project S3-4AlpClusters[ii]) that really put the spotlight on the importance to focus the S3 process on TA. Recently, the concept of TA has been more solidly grounded and is now recurrently referred to in the academic literature (see Foray *et al.* 2018; Foray 2018). Foray *et al.* (2018) note that *"S3 should be understood as a process aimed at transforming the economic structures of a region or any other geographical unit through the formation and development of new activities based on a combination of existing capacities on the one hand and opportunities for structural transformations on the other"* (p.3). The focus of S3 should not be on *"sectors but on modes of transformation of sectors or of establishing new ones".* The outcome of the S3 process should neither be *"an individual project nor a sector as a whole"*, but a transformative activity (TA), understood as a *"collection of innovation capacities and actions, that have been extracted from an existing structure or several structures, to which can be added extra-regional capacities and that is oriented towards a certain structural change"* (Foray *et al*. 2018, p. 1).

An example of what a focus on TA means in practice is provided by Foray (2017), documenting the experience from an entrepreneurial discovery workshop organized within the S3-4AlpClusters project (Milan, 30.05.2017). Existing policies in Lombardy currently support *"a bunch of great start-ups […] inventing new high-tech products and services with strong application potentials in the agrifood sector"* (p.98). Instead of prioritizing a high-tech sector as such, the idea of S3 suggests to seek opportunities for transformation at cross-sectoral intersections in a policy *"aiming at supporting the development of a* ***real transformative activity*** *[emphasis by Foray, 2017] which would likely drive structural changes – not only in the high tech but in the huge agrifood sector"* (p.98). In the case of Lombardy, a stringent transformative activity should focus on innovation capacities for high-tech innovations in agriculture and integrate a collection of concrete actions to *"support the absorption and adoption of new knowledge and technologies offered by [high tech] start- ups"* (Foray, 2018, p.13).

Viewed through this novel TA lens, S3 can be described as regional strategies aiming at transforming the economic structures of a region through the identification and development of transformative activities, based on a reflection about existing capacities on the one hand and opportunities for change on the other. Hence, regional implementation of S3 ultimately consists of two fundamental practical aspects: on the one hand the identification of the innovation capacities through which opportunities for structural change can be tackled, and on the other hand the definition of actions to develop these activities in a given region (Figure 3). The aim of the innovation model can thus be summarized as a process for the identification and development of transformative activities (TA), as defined in Box 1.

**BOX 1: TRANSFORMATIVE ACTIVITIES**

**TA: Transformative Activities** can be understood as a collection of ***innovation capacities*** and ***actions*** of a group of actors, derived from an innovative combination of existing structures, targeting related areas and having the potential to significantly transform existing industries.

**Source:** Authors' definition based on Foray et al., 2018.

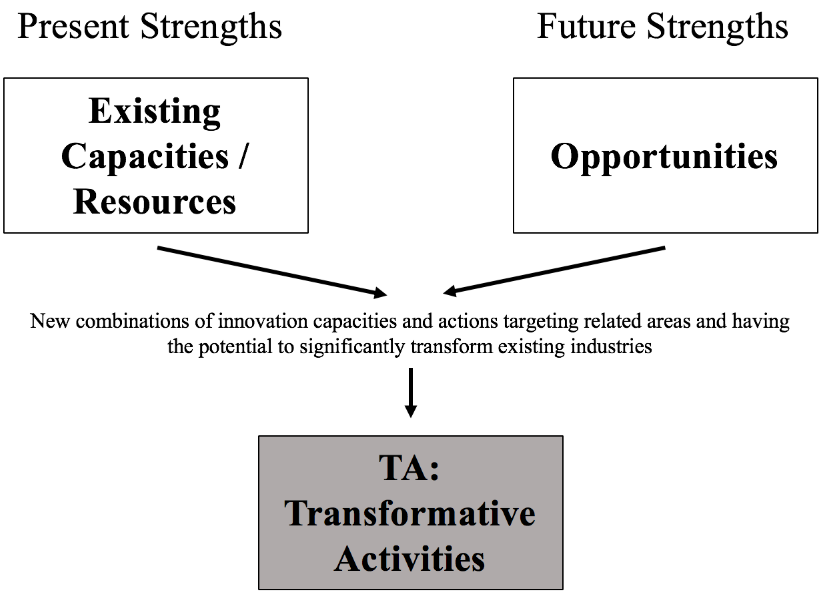



**Fig. 3:** Transformative Activities (TA) for Smart Specialization
**Source:** Authors' elaboration.

## A PROCESS FOR THE IDENTIFICATION AND DEVELOPMENT OF TRANSFORMATIVE ACTIVITIES

In order to operationalize the focus on transformative activities for cluster initiatives and cross-regional cooperation, we consolidate the fundamental questions of S3 development and implementation into a systematic process for the identification and development of TA (see Figure 4). Faced by global competition, regions need to distinguish themselves (diversification) in order to create competitive advantage. Limited resources compel them to specialize on a limited number of prioritized innovative activities, which should meet structural challenges and translate into structural transformation. Thus, the overall goal of S3 can be modelled as the successful regional or cross-regional development of TA, understood as a collection of related innovation capacities and actions with sufficient critical mass to lead to a structural transformation within the economy and the creation of new value chains and jobs in innovative new fields. To reach this goal, TA first need to be identified in an entrepreneurial discovery process based on a solid base of evidence. They then need to be developed into concrete actions whose implementation generates the necessary critical mass for structural transformation in the region. Generating critical mass presupposes to exploit cross-sectoral links (connectivity) and cross-regional cooperation. In order to evaluate the outcome of the process, the development of TA has to be monitored. The whole process should be a collective endeavor including all relevant actors of the innovation process. From identification to monitoring of TA, cluster initiatives are thus key players. They are located at a level of granularity between individual firms and broad sectors, reunite actors of the quadruple helix, reflect connectivity and are predestined to benefit directly from S3-enhanced innovation processes. Therefore, the model includes methodologies to involve cluster initiatives and enable cross-regional cooperation at each stage of the process (Figure 4).[ii)]

| Innovation Model | Process Description | **How to** involve **cluster initiatives** (CI) and enable **cross-regional cooperation**? |
|---|---|---|
| **Base of Evidence** ↓ | Generation of a **base of evidence** covering existing capacities and transformation opportunities. | • Stress Testing Role of CI in S3-implementation (Meier zu Köcker and Dermastia, 2017)<br>• S3-Synergy Diamonds (Meier zu Köcker *et al.*, 2017) |
| **Identification** ↓ | **Identification** of the set of ***innovation capacities*** necessary for the aspired transformation, based on a reflection about existing capacities on the one hand and opportunities represented by new technologies and challenges that can support and drive the process of structural transformation on the other. | • Regional and cross-regional Entrepreneurial Discovery Workshops (EDW) with CI (Foray *et al.*, 2018; Meier zu Köcker *et al.*, 2018) |
| **Development** ↓ | **Development** of the ***actions*** necessary to establish the transformative activities in a region – in particular through gaining critical mass. | • Regional and cross-regional Action Development Workshops (ADW) with CI (Foray *et al.*, 2018; Meier zu Köcker *et al.*, 2018) |
| **Implementation** ↓ | **Implementation**: execution of the developed actions, resulting in gain of critical mass in the defined set of innovative capacities. | • Best Practices for Services provided by CI (Antonioni *et al.*, 2018) |
| **TA: Transformative Activities** | **TA**: Transformative Activities can be understood as a collection of ***innovation capacities*** and ***actions*** of a group of actors derived from an innovative combination of existing structures, targeting related areas and having the potential to significantly transform existing industries. | |
| ↑ **Monitoring** | **Monitoring** the roles and contributions of cluster initiatives at the different stages of the process and evaluate its outcome | • Evaluation Toolbox for CI supporting a formative evaluation during the implementation of the process. |

**Fig. 4:** Innovation Model for the Identification and Development of TA and the Potential Role of Cluster Initiatives
**Source:** Authors' elaboration.

The process starts with the generation of a base of evidence based on qualitative and quantitative analytics. Solid information on existing capacities, clusters, entrepreneurial resources and opportunities for transformation is crucial to guide the subsequent entrepreneurial discovery process for the identification of transformative activities. An analysis of the current role of cluster initiatives in S3-implementation is a useful element of such a base of evidence to set the basis for a systematic involvement of cluster initiatives in the complete process. The experience from the S3-4AlpClusters project[i)] has shown that the stress test approach outlined by Meier zu Köcker and Dermastia (2017) is a valuable contribution to this effort (see section *Implementation Challenges* in Chapter I above). Foray *et al.* (2018) provide a comprehensive overview of the necessary data for a regional analysis to include in a useful base of evidence, notably *"employment per sector / industry, sectoral location quotients (LQ), sectoral productivity data, sectoral exportation data, sectoral innovation data, and regional cluster portfolios"* (p.5). Foray *et al.* (2018) further note that the entrepreneurial discovery process will benefit from a *"pre-determination of the covered field"* (p.7). A way to limit the covered field and disclose existing capacities and opportunities for transformation that are particularly contributory to evidence-based entrepreneurial discoveries is provided by Meier zu Köcker *et al.* (2018) by means of S3-synergy diamonds. Based on an analysis of existing S3 documents, regional priority areas are depicted as the cornerstones of the diamonds. Potential new combinations between priority areas form the axes and thus illustrate where relevant transformative activities can emerge from. The diamonds also disclose complementarities between regions with similar priority areas and thus contribute to facilitate need-based cross-regional cooperation in the subsequent process (see *e.g.* Figure 5 in Chapter III).

The generated evidence is used as an input for the identification and development of transformative activities (TA) in an entrepreneurial discovery and action development process. Per definition, TA consist of ***innovation capacities*** and ***actions*** of a group of actors derived from an innovative combination of existing structures, targeting related areas and having the potential to significantly transform existing industries (p.8 above, based on Foray *et al.*, 2018). In consequence, identification of TA means to identify, based on a reflection about existing capacities on the one hand and opportunities represented by new technologies and challenges that can support and drive the process of structural transformation on the other, a set of ***innovation capacities*** needed for the aspired transformation process. As noted previously, clusters represent local resource concentrations of specific regional capacities and provide, embracing the actors of the quadruple helix, important information about opportunities and ongoing transformations. Entrepreneurial discovery workshops (EDW) are acknowledged tools to involve cluster initiatives in the discovery process (see Coffano and Foray 2014). Foray *et al.* (2018) propose a workshop methodology for the identification of TA, which includes *"representatives of clusters with a comprehensive knowledge of the regional cluster-ecosystem"* as relevant actors (p.6). The methodology is designed to assess novel combinations of *"existing capacities and opportunities"*, to evaluate *"the relatedness of projects well located in this capacity/opportunity space"* and to prioritize and select a TA (or multiple thereof) *"consisting of a set of projects based on related innovation capacities"* (p.10). Meier zu Köcker *et al.* (2018) document how to implement EDW cross-regionally by using the S3-synergy diamonds as a basis for jointly identifying *"similar transformative activities which are of relevance to several regions"* (p.14) (see *e.g.* Figure 6 in Chapter III). In order to further develop the identified TA, the innovation capacities need to be completed with the ***actions*** necessary to enhance structural transformation in a region. As noted by Foray *et al.* (2018), *"developing and ultimately establishing a TA in a region requires building and gaining critical mass (capacity building)."* This can involve a broad range of actions, such as the "identification of missing critical inputs which need to be privately or publicly provided (specific training, research, infrastructure), the development of coordination devices (such as platforms or networks) to connect firms, suppliers, buyers, technology and research, the support of R&D projects or the inclusion of potential adopters of the innovation through training, integration of novel management practices or adoption of new technologies" (p.11). Again, cluster initiatives are key actors in such a process. Foray *et al*. (2012) assert the crucial importance of cross-sectoral connectivity, inherent in the cluster concept, for the creation of critical mass for transformative activities. Meier zu Köcker *et al.* (2018) lay out a methodology for action development workshops (ADW) aiming at developing action plans to create critical mass for TA both regionally and cross-regionally, if access to extra-regional capacities is needed (see *e.g.* Figure 6 in Chapter III).

Further down the process, the developed actions need to be executed regionally or cross-regionally (implementation). As noted above, cluster initiatives are ideal vehicles to transmit S3-enhanced transformation processes to the business level because they typically embrace all relevant actors of the innovation process and can facilitate the implementation of actions resulting in gain of critical mass in the defined set of innovation capacities (see Foray *et al.,* 2012; Foray, 2015). Since both the identified transformative activities and the concrete developed actions are unknown *ex ante* the way in which cluster initiatives can contribute to the implementation of transformative activities can take a multitude of different forms and concretizations. Best practices for cluster initiatives are abundantly available in the literature (see *e.g.* Lindqvist *et al., 2013).* More specifically, based on an analysis of innovation processes within cluster initiatives across the Alpine Space, Antonioni *et al.* (2018) provide a broad set of best practices of cluster services in support of different kinds of potential implementation actions, covering transversal fields such as education, technology, growth, research or collaboration. As noted by Foray *et al*. (2018), an entrepreneurial discovery and action development process typically involves *"success, failures and surprises"* and requires *"strong monitoring and flexibility mechanisms"* (p.3) (see also Coffano and Foray, 2014). Therefore, our innovation model for S3-implementation with cluster initiatives finally proposes to systematically monitor the roles and contributions of cluster initiatives at the different stages of the process and evaluate its outcome.

The systematic process for regional and cross-regional identification and development of TA, described in the present innovation model, is currently implemented across the Alpine Space in the regions participating in the S3-4AlpClusters project[ii]. In order to further illustrate the proposed process, we provide an insight into this practical experience in the next chapter.

# III. PILOT EXPERIENCE FROM THE ALPINE SPACE

Since its start in November 2016, the S3-4AlpClusters project[i] has been gathering experience with a broad range of issues related to practical S3-implementation. In particular, the project served as a testbed for the systematic identification and development of transformative activities (TA), as sketched out in the innovation model in the previous chapter (see Figure 4). 30 cluster initiatives from 11 regions of the Alpine Space are currently involved in these pilot activities. A solid base of evidence was produced for all regions. Synergies in regional S3 were identified and represented in 4 S3-synergy diamonds (Figure 5) targeting opportunities for transformative activities related to major challenges for the alpine macro-region, as outlined in the EU Strategy for the Alpine Region (EUSALP)[iv]:

1. Economic globalization that requires the alpine region to distinguish itself as competitive and innovative by developing a "knowledge and information" society
2. Demographic trends characterized particularly by the combined effects of ageing and new migration models
3. Climate / energy change and its foreseeable effects on the environment, biodiversity and on the living conditions of its inhabitants
4. The specific geographical position in Europe as a transit region and as an area with unique geographical and natural features, which will set the frame for all future developments, notably with respect to mobility (Meier zu Köcker *et al.*, 2017).

Drawing on the generated evidence, all regions identified and developed new TA in a series of entrepreneurial discovery (EDW) and action

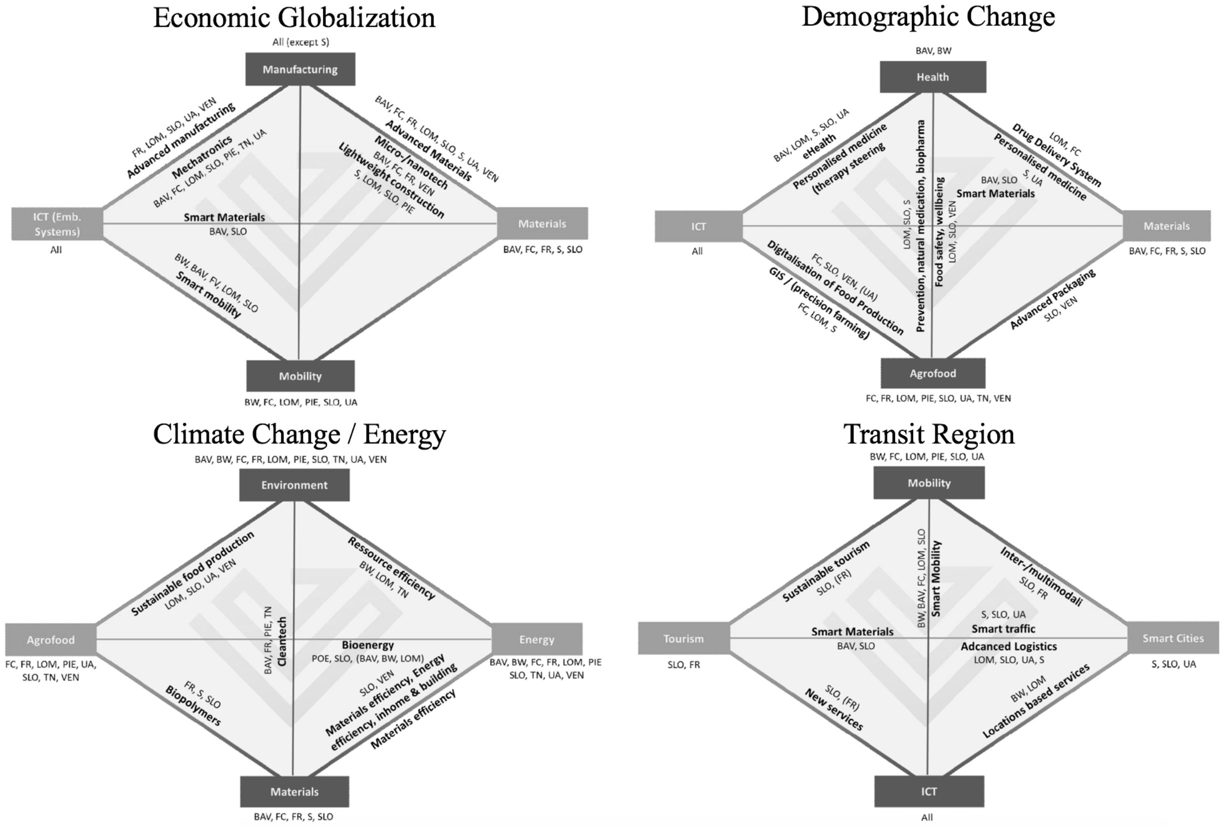



**Fig. 5:** S3-Synergy Diamonds for the Alpine Space Macro-Region
**Note:** BAV: Bavaria, BW: Baden-Württemberg, FC: Franche-Comté, FR: Canton of Fribourg, LOM: Lombardy, PIE: Piedmont, S: Salzburg, SLO: Slovenia, TN: Trentino, UA: Upper Austria, VEN: Veneto
**Source:** Meier zu Köcker et al., 2017.

development workshops (ADW), each involving 10 to 30 regional key actors including cluster initiatives, firms, policymakers and representatives of the regional innovation system (Bersier and Keller, 2018). The participants of the EDW assessed existing capacities and opportunities for transformation and prioritized a set of related innovation capacities to constitute a TA. Action plans were then developed in a series of ADW to complete the TA with the concrete actions necessary to gain critical mass in the identified innovation capacities and ultimately establish the TA in the concerned regions.

The character of the EDW and ADW and the applied methodologies varied among the different pilot activities and were shaped by specific regional demands. All workshops had in common however, that they followed the general process of the innovation model with a strong focus on TA and an active involvement of cluster initiatives. In two instances, the pilot activities were carried out cross-regionally. First, Upper Austria collaborated with Veneto on the development of safety, quality and food traceability along the food value chain. Second, Upper Austria also engaged in a cross-regional process of EDW and ADW to identify and develop TA together with Bavaria, Franche-Comté and Slovenia. The identification and creation of a common understanding on the TA to be further developed into concrete cross-regional actions and need-based cross-regional cooperation is a complex exercise. The use of S3-synergy diamonds (Meier zu Köcker *et al.*, 2017) proved valuable to detect similar priority areas in current S3 and identify TA for which the regions possess complementary strengths and needs. Based on the S3-synergy diamond targeting the EUSALP challenge of economic globalization (see Figure 5, upper left quadrant) a potential was identified for cross-regional cooperation between Bavaria, Franche-Comté, Slovenia and Upper Austria in the priority areas of manufacturing and new materials, and more particularly in new technological fields that may arise in combination of the respective priority areas. The cross-regional effort drew on complementarities in regional strengths (lightweight materials / Bavaria, lightweight technology / Upper Austria, circular-economy (materials circle, *e.g*. cascade use of materials/waste) / Upper Austria, second materials technology / Slovenia) and shared challenges and opportunities in lightweight materials, clean-technologies, bio-based composites and wood materials linked to the circular economy. Specifically, the entrepreneurial discovery process led to the identification of particular innovation capacities for the design, production and recycling of fibre composites for new lightweight materials as a TA to be developed cross-regionally based on complementary capacities and needs. In order to prepare the development of concrete actions for this TA the participating regions established in advance a brief documentation that was shared among the regions to establish an overview on the involved clusters and further stakeholders, current activities and initiatives, specific know-how, new developments, specific problems and challenges. The concerned cluster initiatives then met for an ADW to elaborate a joint action plan. The process consisted of 4 interactive rounds (round 1: identification of challenges and competences; round 2: matching challenges and solutions and prioritization; round 3: action development phase; round 4: drafting of action plan including next steps). At each step, participants were asked to document their contributions and ideas. The inputs were discussed after each round in a fruitful working atmosphere where ease of interaction was created. The cross-regional experience resulted in an action plan focusing on education efforts for mind-set change, training on company level and mapping of available technical solutions (Figure 6, left side).

In the Swiss canton of Fribourg, an EDW was conducted with regional cluster initiatives (Swiss Plastics Cluster, Cluster Food and Nutrition, Building Innovation Cluster), research institutions (such as the Plastics Innovation Competence Center of the School of Engineering and Architecture), enterprises and policymakers using the S3-synergy diamond addressing climate and energy challenges (see Figure 5, lower left quadrant). Strong existing capacities were identified in the fields of materials, food and nutrition and the construction sector. A systematic discussion of opportunities for structural transformation offered to these traditional strongholds by the trend towards a circular bio-economy led to the identification of a specific TA to prioritize in the regional development strategy[ii]: the TA should draw on and build up related innovation capacities necessary to develop bio-based inputs for the plastics industry. In the subsequent ADW, the key actors met to work on concrete actions to further develop the TA in the canton of Fribourg. An action plan was drafted to mount collaborative R&D projects, networking activities and development of critical skills between the clusters, research institutions and regional and extra-regional enterprises, *e.g.* to use waste-streams from the local food industry for protein-based barrier film packaging. Figure 6 (right side) summarizes this process. Note, in line with the definition of TA proposed in Chapter II (p.7), that the TA in question neither corresponds to the food sector, nor the plastics industry as such, but to the collection of innovation capacities from groups of companies, suppliers and research partners associated with these existing sectors and the concrete actions they need to undertake to specialize in the development of bio-based inputs for the plastics industry.

Both examples show instances of aspired cross-regional cooperation for the development of TA. In the case of Upper Austria, Bavaria, Franche-Comté and Slovenia, actions were specifically elaborated to make use of the complementarities among the four regions with respect to existing resources and needs. In the case of Fribourg, capacities from extra-regional actors were found crucial for the development of collaborative R&D projects. Both experiences also emphasized the difficulty to actually implement actions for the development of TA on a cross-regional basis. Neither between the regions from different European countries, nor between different regions of Switzerland did the participants of the workshop estimate the existing funding schemes to be sufficient to support the developed cross-regional actions. This finding is in line with Meier zu Köcker and Dermastia (2017) asserting that *"aligning S3 and related policy instruments among neighboring regions is still a challenge"* (p.24) and Meier zu Köcker *et al.* (2017) lamenting the absence of *"dedicated support schemes"* synchronized across regions for the development of cross-regional TA (p. 27).

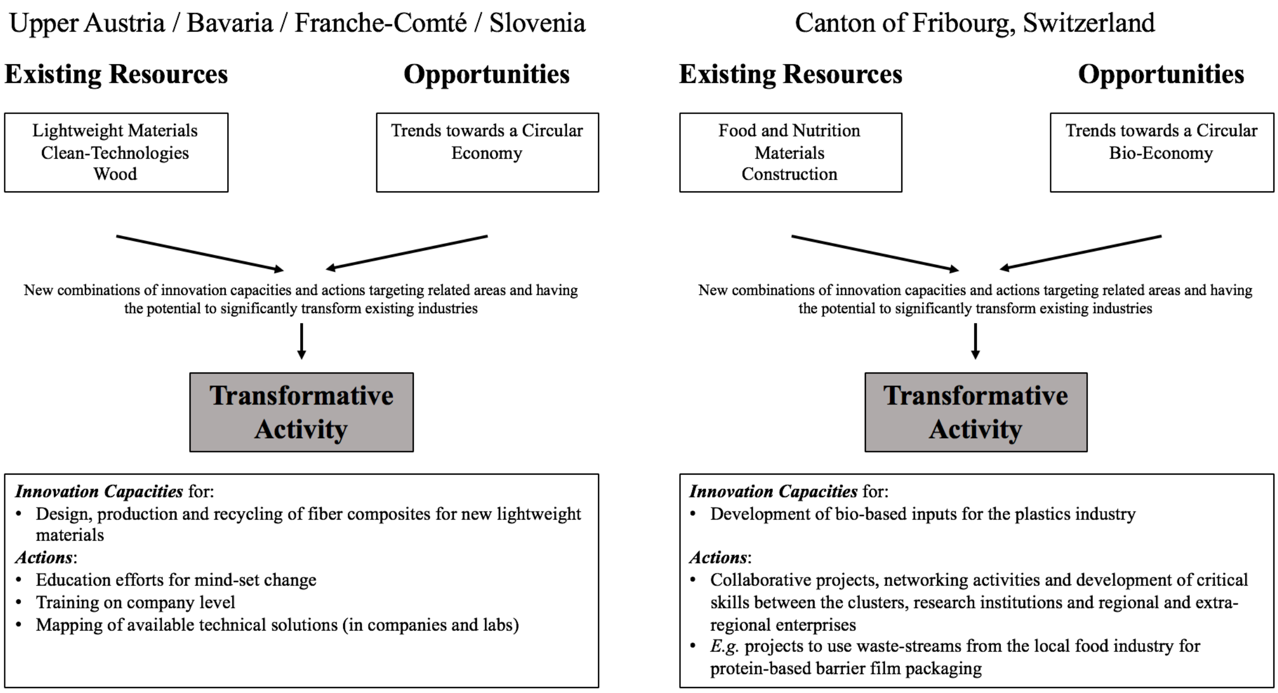



**Fig. 6:** Regional and Cross-Regional Entrepreneurial Discovery and Action Development
**Source:** Authors' elaboration.

# CONCLUSION AND RECOMMENDATIONS

This paper draws on the recent experience from the Interreg Alpine Space project S3-4AlpClusters[i], which put the interplay between S3 and clusters at the core of its conceptual and practical study of S3-implementation across the Alpine Space. While overlaps between the two concepts are evident and cluster initiatives are acknowledged in the relevant literature as tools in the context of S3 (see Ketels, 2013a), there is, to our knowledge, no comprehensive study on how clusters are currently involved in the practical development and implementation of S3. Moreover, practical implementation of S3 with cluster initiatives is found to be far from trivial and involves specific challenges (see Meyer zu Köcker and Dermastia, 2017). We therefore propose a novel focus on the interplay between S3 and clusters and introduce an innovation model as a practical effort to better integrate cluster initiatives in the S3 process. The model is a systematic process for the regional and cross-regional identification and development of transformative activities (TA). We define TA as a collection of innovation capacities and actions of a group of actors, derived from an innovative combination of existing structures, targeting related areas and having the potential to significantly transform existing industries (see Foray *et al.*, 2018). Cluster initiatives are recognized as key actors in the entrepreneurial discovery and action development process of the innovation model.

The ongoing debates on updating the S3 efforts in the European innovation policy framework and related regional innovation strategies focus mainly on practical implementation challenges. Potentially critical elements are identified at various levels ranging from a lack of understanding of the entire S3 concept to missing compatibility between S3 and policy tools for implementation, missing political commitment to focus on a limited field with high transformative potential, or a lack of critical mass in terms of innovation actors and public investments. New methodologies and tools are developed for future-oriented regional analysis and implementation of smart industrial transformation processes (see *e.g.* Nögel *et al.*, 2018). In a similar vein, the innovation model outlined in this paper is currently implemented with cluster initiatives across the Alpine Space within the S3-4AlpClusters project[ii]. Based on these first experiences, we conclude the paper with three recommendations we suggest to consider in current and future policy discussions on S3-implementation:

- The locus of S3-implementation should shift from existing priority areas to new transformative activities (TA)
- Cluster initiatives should be used as levers for regional economic development and take over active roles in a systematic process to identify and develop TA
- Cross-regional cooperation in the identification and development of TA should be further supported by cross-regional synchronized funding schemes

Tools and methodologies for S3-Implementation, such as S3-synergy diamonds, entrepreneurial discovery workshops (EDW) or action development workshops (ADW) are valuable individual contributions for future policy designs. Nevertheless, it is only by putting them into the context of a systematic innovation model, with a strong focus on transformative activities, and by making them the levers for cross-regional cooperation and a systematic involvement of cluster initiatives in regional economic development, that they become fully relevant for smart transformation processes leading to innovation within businesses, new value chains and jobs in innovative new areas.

# REFERENCES

**Antonioni, S., Mion, L., Keller, M. and Bersier, J.** (2018). Good Practice Report. Interreg Alpine Space – S3- 4AlpClusters. https://bit.ly/2PIGRtW.

**Bersier, J. and Keller, M.** (2018). Transformative Activities for Regional Development, Project Overview S3-4AlpClusters, School of Engineering and Architecture of Fribourg, University of Applied Sciences and Arts Western Switzerland. Interreg Alpine Space – S3-4AlpClusters. https://bit.ly/2Kz5MrT.

**Boschma, R.** (2017). Relatedness as driver of regional diversification: A research agenda. Regional Studies, 51(3), 351–364.

**Breschi, S., Lissoni, F. and Malerba, F.** (2003). Knowledge-relatedness in firm technological diversification. Research Policy, 32(1), 69–87.

**Christensen, T. A., Lämmer-Gamp, T. and Meier zu Köcker, G.** (2012). Perfect Cluster Policy and Cluster Program. Copenhagen/Berlin: Danish Ministry of Science, Technology and Innovation Competence Networks.

**Coffano, M. and Foray, D.** (2014). The Centrality of Entrepreneurial Discovery in Building and Implementing a Smart Specialisation Strategy. Scienze Regionali, 13(1), 33–50.

**Cortright, J.** (2006). Making Sense of Clusters: Regional Competitiveness and Economic Development, The Brookings Institution Metropolitan Policy Program.

**Delgado, M., Porter, M. and Stern, S.** (2014). Clusters, Convergence, and Economic Performance, Research Policy, 43 (10), 1785–1799.

**Delgado, M., Ketels, C., Porter, M. E. and Stern, S.** (2012). The Determinants of National Competitiveness, NBER Working Paper No. 18249. Cambridge, MA: NBER.

**Delgado, M., Porter, M. and Stern, S.** (2010). Clusters and Entrepreneurship, Journal of Economic Geography, 10 (4), 495–518.

**Enos, J.** (1995). In Pursuit of Science and Technology in Sub-Saharan Africa. UNU/INTECH Studies in New Technology and Development. Routledge: London, UK; New York, NY.

**Etat de Fribourg.** (2016). Programme de mise en œuvre de la Nouvelle Politique Régionale (NPR), 2016-2019. Fribourg.

**EU** (2013). Regulation (Eu) No 1303/2013 Of The European Parliament And Of The Council of 17 December 2013. *Official Journal of the European Union. L.347, 320-469.*

**European Commission** (2017). Strengthening Innovation in Europe's Regions: Strategies for resilient, inclusive and sustainable growth. COM(2017)376.

**Feldman, M. and Audretsch, D.** (1999). Innovation in Cities: Science-based Diversity, Specialization, and Localized Competition, European Economic Review, 43, 409–29.

**Feser, E., Renski, H. and Goldstein, H.** (2008). Clusters and Economic Development Outcomes, Economic Development Quarterly, 22 (4), 324–44.

**Foray, D., Keller, M., Bersier, J. and Meier zu Köcker, G.** (2018). Transformative Activities for Smart Specialisation: Considerations on a Workshop Methodology. Working paper EPFL / HES-SO HEIA-FR / ClusterAgentur. https://bit.ly/2yuTEAX.

**Foray, D.** (2017). Smart Specialisation, Edmund Phelps and the Palazzo Lombardia,' in J. Severijns (ed.), Solving Contradictions by Connectivity. Province of Limburg: Maastricht, The Netherlands.

**Foray, D.,** (2015). Smart specialisation: opportunities and challenges for regional innovation policies, Routledge.

**Foray, D. and Goenaga, X.** (2013). The Goals of Smart Specialisation. S3 Policy Brief Series. No. 01/2013. European Commission Joint Research Center.

**Foray, D., Goddard, J., Goenaga, X., Landabaso, M., McCann, P., Morgan, K., Nauwelaers, C. and Ortega-Argiléset, R.** (2012). Guide on Research and Innovation Strategies for Smart Specialisation. European Commission, Regional Policy.

**Foray, D., David, P.A. and Hall, B.** (2009). Smart specialisation: the concept, Knowledge for Growth: Prospects for Science, Technology and Innovation, Report, EUR 24047, Brussels, European Commission.

**Frenken, K., Van Oort, F. G., and Verburg, T.** (2007). Related variety, unrelated variety and regional economic growth. Regional Studies, 41(5), 685–697.

**Glaeser, E. and Kerr, W.** (2009). Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain? Journal of Economics and Management Strategy, 18 (3), 623–63.

**Hausmann, R. and Rodrik, D.** (2006). Doomed to Choose: Industrial Policy as Predicament, First Blue Sky Seminar, Center for International Development, Harvard University.

**Ketels, C. and Keller, M.** (2015). 25 Years of "The Competitive Advantage of Nations", Competitiveness Review, Vol. 25 Issue: 5.

**Ketels, C.** (2013a). Recent research on competitiveness and clusters: what are the implications for regional policy? Cambridge Journal of Regions, Economy and Society, 2013, 6, 269–284.

**Ketels, C.** (2013b). Cluster Policy: A Guide to the State of the Debate. In: Meusburger, Glückler and el Meskioui (eds.) Knowledge and Economy. Heidelberg: Springer.

**Ketels, C.** (2011). Clusters and Competitiveness: Porter's Contribution. In Huggins and Izushi (eds.) Competition, Competitive Advantage and Clusters: The Ideas of Michael Porter, 173–191. Oxford: Oxford University Press.

**Lindqvist, G., Ketels, C. and Sölvell, Ö.** (2013). The Cluster Initiative Greenbook 2.0. Ivory Tower Publishers, Stockholm.

**Lindqvist, G., Ketels, C., Sölvell, Ö.** (2003). The Cluster Initiative Greenbook. Stockholm: Ivory Tower.

**Meier zu Köcker, G., Keller, M., Dermastia, M. and Bersier, J.** (2018). Cluster Action Plan: Transformative Activities for Regional Development. Interreg Alpine Space – S3-4AlpClusters. https://bit.ly/2EFqwMP.

**Meier zu Köcker, G. and Dermastia, M.** (2017). StressTesting Regional Approaches Conducive to Implement S3 through Clusters. Interreg Alpine Space – S3-4AlpClusters. https://bit.ly/2ApFGl5.

**Meier zu Köcker, G., Dermastia, M. and Keller, M.** (2017). Strategic Alpine Space Areas for Cross-regional Cooperation. Interreg Alpine Space – S3-4AlpClusters. https://bit.ly/2OIBgi2.

**Neffke, F., Henning, M., and Boschma, R.** (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Economic Geography, 87(3), 237–265.

**Nögel, L., Sedlmayr, B., Wittpahl, V. and Meier zu Köcker, G.** (2018). Unterstanding regional transformation processes. New tools for regional development (German language), 2018, Institute for Innovation and Technology, Berlin.

**Porter, M. E.** (2011). On competition. Boston, MA: Harvard Business School.

**Porter, M. E.** (2008). Clusters and competition: New agendas for companies, governments, and institutions, in: Porter (ed.) On competition, 213–304. Boston, MA: Harvard Business School.

**Porter, M.E.** (2003). The Economic Performance of Regions, Regional Studies, 37, 549–78.

**Porter, M.E.** (1990). The Competitive Advantage of Nations, Free Press, New York.

**Rosenthal, S. and Strange, W.** (2004). Evidence on the Nature and Sources of Agglomeration Economies, in Henderson, J. and Thisse, J. (eds.) Handbook of Regional and Urban Economics, vol. 4., Amsterdam: Elsevier North-Holland.

## APPENDIX A

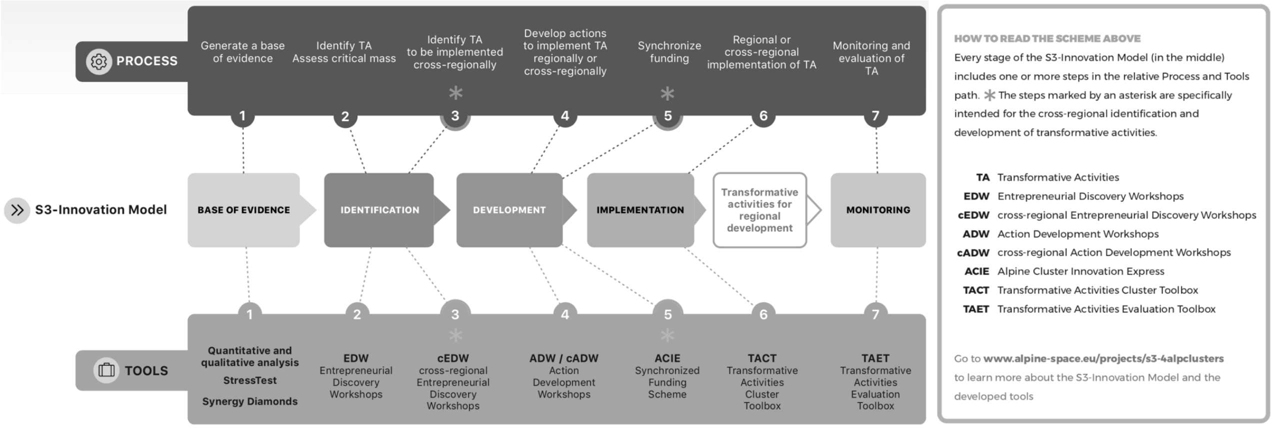



**Fig. A1:** The “S3-Innovation Model” of the S3-4AlpClusters Project
**Source:** ©S3-4AlpClusters

# AUTHORS

**MICHAEL KELLER, JACQUES BERSIER**
*School of Engineering and Architecture of Fribourg,*
*HES-SO - University of Applied Sciences and*
*Arts Western Switzerland*
E: michael.keller@hefr.ch

**IRIS REINGRUBER**
*Business Upper Austria – OÖ Wirtschaftsagentur GmbH*

**MATEJA DERMASTIA**
*Anteja ECG, Ljubljana*

**GERD MEIER ZU KÖCKER**
*ClusterAgentur Baden-Württemberg*

---

i **S3-4AlpClusters**
The project is co-financed by the European Regional Development Fund through the Interreg Alpine Space programme. It brings together 15 partners from 11 Alpine Space Regions (Piedmont, Lombardy, the Autonomous Province of Trento, Venetia, Slovenia, Upper Austria, Salzburg, Bavaria, Baden-Württemberg, Bourgogne-Franche-Comté, and the canton of Fribourg), as well as their clusters and 10 observers. Partners include private and public actors from business organizations, SMEs, regional and national authorities, sectoral agencies and academic and research institutes.

S3-4AlpClusters is led by Innosquare Clusters, the cluster platform of the School of Engineering and Architecture of Fribourg, member of the University of Applied Sciences of Western Switzerland.

All project reports cited in this paper are available on the project website:
http://www.alpine-space.eu/projects/s3-4alpclusters

Additional information is also available on the project’s YouTube channel:
www.youtube.com/channel/UCXf4dSJMZiTRCSSmaEGmMNg

ii The process laid out in this paper is currently implemented both regionally and cross-regionally under the label “S3-Innovation Model” in the 11 regions participating in the S3-4AlpClusters project (see Endnote i) above). For each step of the process, dedicated tools are tested and fine-tuned into a comprehensive toolkit for cluster initiatives. Appendix A, Figure A1 represents the “S3-Innovation-Model”, as it is currently tested in the project. The final toolkit will be published in March 2019 and presented at an international conference on March 14 in Venice.

iii The Swiss canton of Fribourg, as the only project partner outside the European Union, does not have a formal Smart Specialization Strategy (S3). Nevertheless, certain aspects of the cantonal strategy for competitiveness do reflect priorities similar to an S3. The latest specific formulation of this ongoing quest to define a cantonal competitiveness policy can be found in the cantonal implementation program for the 2016-2019 phase of the Nouvelle Politique Regionale (NPR; French for New Regional Policy), a nationwide policy framework for regional development (Etat de Fribourg, 2016).

iv More information on the EU Strategy for the alpine region (EUSALP): https://www.alpine-region.eu.

# RESEARCH IMPACT ASSESSMENT: FROM EX POST TO REAL-TIME ASSESSMENT

PIERRE-BENOIT JOLY, MIREILLE MATT AND DOUGLAS K. R. ROBINSON

DOI: 10.22163/fteval.2019.326

## 1. INTRODUCTION

This paper presents an ongoing research and development project to build research management tools based on real-time impact analysis (the toolset is labelled ASIRPA[rt]). The ambition is to use the lessons learned from ex post research impact assessment (RIA), building from the ASIRPA project which was launched in 2011 (Joly et al. 2015, Matt et al. 2017). The ASIRPA approach is currently implemented on a routine base at the French public research organisation INRA (Institut National de la Recherche Agronomique). Therefore, the project draws on lessons learned from ex post RIA and the experience of researchers and actors involved in research programming.

The aim of ASIRPA[rt] is to design an approach and tools to help conduct research projects or programmes with the aim to amplify impacts. The challenge of the current project is to develop management tools based on a better understanding of the mechanisms that generate research impact. These tools will be coproduced with potential users (Robinson and Rip 2013).

Given the uncertainty and complexity that characterise the transformation processes linked to research activities, we do not intend to design ballistic steering tools but to produce tools to foster learning processes, coordination and reflexivity of the actors involved. Our approach takes inspiration in different streams of literature.

First, based on Kuhlman (2003), we consider that such tools should foster competences of the actors involved in research activities and research programming, as well as strengthen collective learning and coordination. Second, our general representation of transformation processes linked to research activities is inspired by innovation studies, and more precisely actor-network theory (Callon 1986) and the innovation journey (Van de Ven et al. 1999). According to these theoretical frames, the processes involved cannot be steered and planned because they generate new knowledge, new socio-technical associations and their effect depends on the progressive alignments of many heterogeneous elements. To paraphrase Van de Ven (2016), one cannot control such complex and uncertain processes, but one can still learn to manoeuvre it.

With such tools and the interactions that it can generate, we also aim at favouring exchanges between users and thus contributing to the creation of communities of practice that will themselves contribute to improving the approach.

## 2. STATE OF THE ART, CONCEPTS AND APPROACHES

Since the beginning of the 2010's, RIA benefits from a renewed interest (Joly and Matt 2017). Although the field is still moving quickly, we know a lot on *ex post* RIA and relatively little on *in itinere* or *ex ante* assessment. In our project, we use the expression 'real time' to signify that what matters is the design of tools for continoues assessment of the transformative capacity of research and learning how to 'manoeuvre' for enhancing impact. Such tools take inspiration of *ex post* analysis and aim to enhance skills of actors involved for *ex ante* or *in itinere* conduct of research activities. Our idea is that these tools have to be designed to serve at different scales, from the project level to the programme level or any relevant cluster of projects. The landscape of RIA proposed by LERU in the context of the preparation of FP9 is relevant to our approach (Figure 1).

| Impact assessment and enhancement | Ex-ante | During | Ex-post | | |
|---|---|---|---|---|---|
| | | | short-term (end of project +1 year) | medium-term (+1-7 years) | Long-term (+7-15 years) |
| **Call and project level** | Pathways to impact statement | Follow-up/impact funding | | | |
| | | | Quantitative and qualitative assessment of outputs, outcomes and pathways to impact | | |
| **At the level of FP components (clusters, programmes, societal challenges, missions, …)** | Description of broad expected impact by the EC in WPs, or targets to meet (for missions) | Clustering | | | |
| | | | Assessment of interaction and communication patterns between research and societal context at the level of clusters or programmes | | |
| | | | Synthesising projects | | |

**Fig. 1:** A general overview of research impact assessment landscape
**Source:** LERU (2018)

If we consider the different approaches and tools available (Table 1), 'theory of change' and the various tools designed for its implementation are widely used. However, most of these tools poorly deal with complexity and uncertainty of research process. Moreover, design thinking may help to stimulate creativity and tools for co-design are well adapted to involve potential users in the innovation process. However, such tools do not take into account explicitly the growing information. Hence, process analyses such as ANT and innovation journey approaches are our best candidates. However, this does not exclude taking advantage of other approaches/tools where relevant.

| | **The impact path** | **The complexity** | **The uncertainty** |
|---|---|---|---|
| **Processual**<br>Impact pathway + +<br>Complexity + + +<br>Uncertainty + + + | Analysis dividing the process into sequences<br>Joint study of the configurations of the different variables (actors, etc), sequences, drivers (mechanisms), and bifurcations (turning points) | In social phenomena, causal relations are not constant, the results are unpredictable.<br>Relational, technological and temporal complexity | Time is a structuring dimension of the action that generates lock-in.<br>Different types of uncertainty.<br>Control of Innovation: learn to navigate in a set of possible future |
| Examples of tools that are affiliated with the approach | FORTH innovation expedition<br>Constructive Technology Assessment and tools to animate participatory workshops | | |
| **Theory of change**<br>Impact Pathway +<br>Complexity ++<br>Uncertainty + | Allows an analysis of the configuration of the actors and the context<br>If the nature of the intervention is predefined: expand the logical steps towards the objective of the programme according to a fairly linear approach (logframe). | Is suitable for complicated: takes into account the context, past experience, logical assumptions, between actions and between production of the results and achieving the objectives<br>Requires a consensus of actors on the general theory to be tested. | Formative dimension of the evaluation: learning based on experimentation, coordination of expertise, creation of relationships and common foundations to achieve a consensus<br>Takes into account a finite number of possible futures. |
| Examples of tools that are affiliated with the approach | LogFrames ; Logical frameworks with inventory of the conditions necessary for the impacts<br>Participatory Impact Pathway Approach (PIPA) ; Vianéo Approach | | |
| **Innovative design**<br>Impact pathway<br>Complexity +<br>Uncertainty + + + | Design process: define concepts C and knowledge K and characterise the operators necessary for the passage from C-space to K-space (expansion logic)<br>Methods that do not explicitly consider the phases of the impact pathway (e.g. generalization) | Concepts are tiered to reduce complexity.<br>New concepts must stimulate the production of new knowledge: the expansion of knowledge and concepts<br>The identities of objects and their ecosystems are not fixed. | The unknown is an engine for defining new design parameters.<br>The techniques AND functions of the objects to be conceived are unknown; and objects and their ecosystems have a variable identity. |
| Example of tools affiliated with the approach | Creativity Workshops: TRIZ an Creative Problem Solving on problem solving ; STUR method on analysing strategic partnerships ; CCAID method on the research-action-participation.... : KCP workshops : tools from the IDEAS network | | |

**Table 1:** A first overview of approaches and tools available for real time assessment

# 3. SKETCHING OUT THE ASIRPA (RT) APPROACH

## 3.1. WHAT DO WE DRAW ON? THE MAIN LESSONS FROM EX POST RIA (ASIRPA PROJECT)

The lessons learned from *ex post* RIA play a crucial role in building the real-time approach. The "impact pathway" is the core concept of the ASIRPA approach. We adapted this traditional framework (Douthwaite 2003) and shifted from a quite linear input/output analysis to an approach where:

- the process is divided into phases that are qualitatively different but that do not necessary follow a linear sequence;
- taking our inspiration in ANT, the dynamics are related to translations that allow to create new links between different elements (both human and non-human) and to transform and extend socio-technical associations;
- we do not primarily consider 'inputs' but what we call "productive configurations", a concept that aims at taking into account both the organisational complexity of the research activities considered and their embedding in a wider context;
- we focus on two key elements: (i) the role of intermediaries that play a key role in the dynamics of key translation processes; and (ii) the generalisation or scaling up/out, a phase that is often quite problematic.

The main lessons identified were the following:

- The complexity of the genesis of impacts, generally produced by a set of activities rarely brought together in a single project;
- The importance and diversity of configurations of actors and material resources that produce impacts;
- The identification of critical points along the impact pathway, with a special role of intermediaries and the process of generalization;
- The transformations of the network of actors during the process (an adoption network is generally different from a design network)

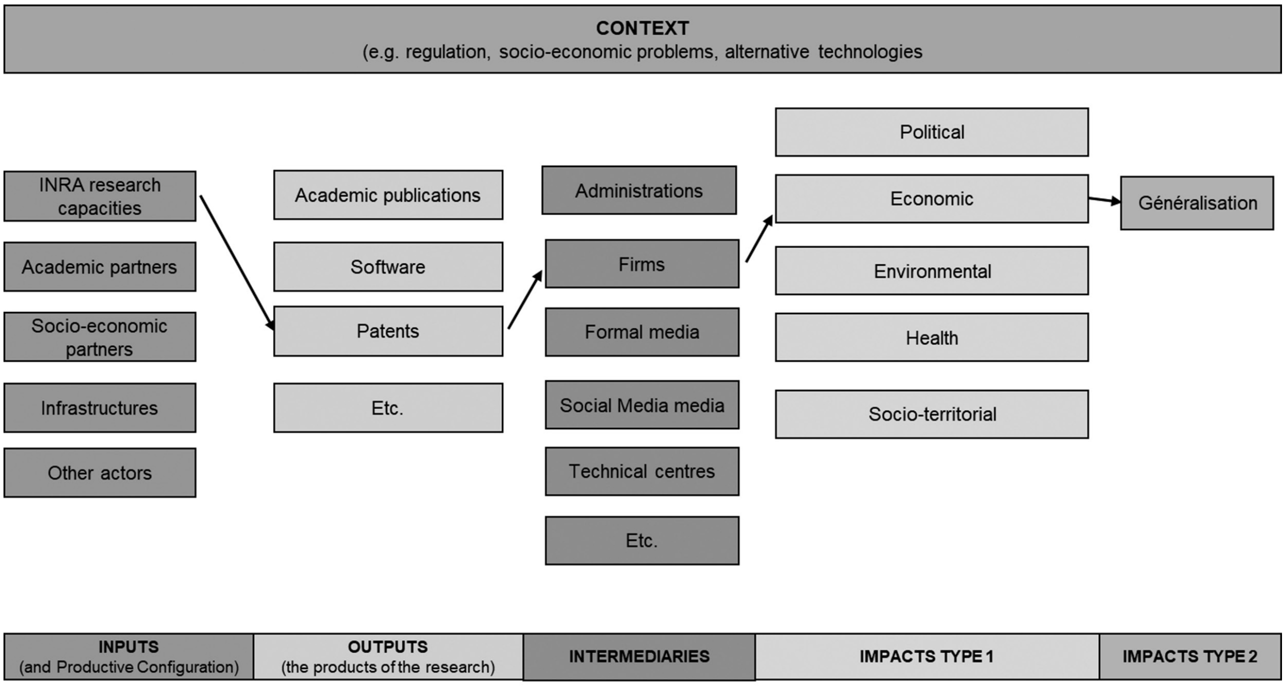



**Fig 2:** The impact pathway recast in the ASIRPA approach

- The role of the external context which can have facilitating or blocking effects, and open or close, sometimes suddenly, windows of opportunity
- The existence of 4 types of impact paths with different properties and effects;
- The long temporalities of the impact (20 years on average between the initiation of research and the first impacts, with important variations).

## 3.2. THE CONCEPTS

Research and innovation projects are characterized by high levels of complexity and uncertainty. Acknowledging these essential characteristics, the real-time evaluation approach aims to strengthen the capacities of the actors and the dynamics of collective action, thus drawing on two traditions:

- Developmental evaluation (Patton 2016) which informs and guides innovation and development actions that take place in dynamic and complex environments;
- Strategic intelligence (Kuhlman 2003) which aims to strengthen coordination and collective learning.

Moreover, the approach is attentive to the creativity of the actors and to serendipity. While it is necessary for the actors involved to be able to construct a theory of change, i.e. to form an image of the targeted transformations, to represent the impact paths, to identify the critical factors, etc., it is essential that the steering tools allow great flexibility and adaptability. The representation of the process at work is based on the analyses of the innovation journey. Basically, innovation is seen as a sequential, uncertain, complex and singular process. Nevertheless, knowledge of this process helps to identify facilitating and blocking factors. This knowledge must make it possible to design tools to manage the tension between a top-down direction and bottom-up explorations (Mazzucato 2018). The identification of the targeted transformations and the construction of an *ex ante* impact path must feed into explorations that may have many sources of surprise and that can lead to revising the targets (what Robinson 2009 has described as a reflexive strategy articulation support system). These tools should enable collective learning to be monitored; lessons (and data) from experiments should be collected and analysed.

## 3.3. THE PROPOSED APPROACH

The real-time evaluation process is based on an iterative model whose main lines can be outlined as follows.

## 1. TARGET IDENTIFICATION

What are the anticipated transformations that justify commitment to research? What are the different issues involved in these transformations? What is the magnitude of these transformations in the 5 dimensions of impact?Who are the potential users? How will they be interested in/affected by the transformations?

It should be noted that, given the uncertainties inherent in research, this target often constitutes what may be called a "rational myth": an objective in which we must believe but which we know from the outset is likely to change to a greater or lesser extent.

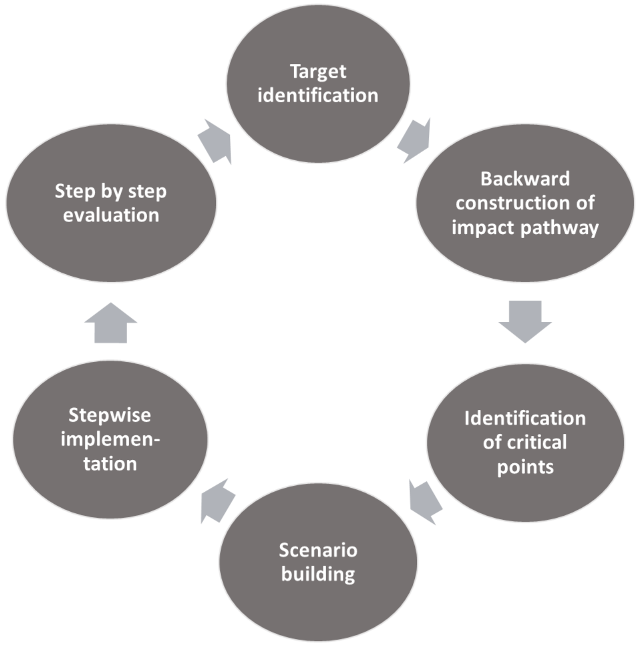



**Fig 3:** A first representation of the proposed approach

## 2. CONSTRUCTION OF THE IMPACT PATHWAY BY BACKWARD INDUCTION

Who are the key actors and mechanisms involved in the mainstreaming process? What are the blocking and facilitating factors?

Which intermediaries (organisations, technical objects, devices) will enable implementation by the first users?

Who are the knowledge-producing actors? what are the adjacent projects (ongoing, completed, in gestation), by whom are they carried out? what are the complementarities and competition between these projects?

## 3. IDENTIFICATION OF CRITICAL POINTS

What are the critical points associated with the different stages of the impact path? On whom do they depend?

What are the influences of these critical points on the envisaged process?

## 4. SCENARIO CONSTRUCTION (SCRIPTS - STEPS, BIFURCATION)

Taking into account the main elements above, construct the main scenarios of the project (or group of projects), with the main stages, critical points, bifurcations.

From this, deduce the main meeting points and the anticipated follow-up elements.

## 5. STEP BY STEP DECISION AND IMPLEMENTATION

## 6. EVALUATION AT EACH STEP AND NEW LOOP

This iterative model is constructed and used by the project manager or program manager, often supported by project engineering specialists. It is usually the result of a collective design, with the teams involved and, as necessary, with external partners and potential users. These interactions are based on tools accessible on an online platform.

**Essential point:** this is a sequential approach. The aim is not to resolve all the questions from the outset but to conduct a process in which the main stages, qualitatively different, are analysed and scattered by internal or external information gains from the project.

The principle is to identify the elements necessary to improve research contribution in the present sequence, bearing in mind the uncertainty about the future. This distinguishes this approach from traditional applications of theories of change. The sequential approach takes seriously the uncertainty, the gain of information during the process (on the state of the art, on the environment, because of relational learning, etc.) and the need to privilege flexibility and adaptability.

At each stage, we seek to optimize the approach by taking into account uncertainty and flexibility. For example, regarding the application of genomic selection methods to the estimation of the genetic value of bulls in milk cattle, proof of concept will be sought before making irreversible development investments; this requires developing a productive configuration to combine quantitative genetics, sample collection and high throughput sequencing skills. The proof of concept being done, we can enrol actors from the sectors who will be involved in the co-development of standardized techniques. For each project, it is therefore necessary to set a transformation objective and to determine the sequence of the main stages qualitatively different. This results in the determination of control points, which leads to the construction of project monitoring dashboards that are very different from the performance indicators generally used in change theory applications.

# 4. IMPLEMENTATION

ASIRPA[rt] uses a co-design strategy. We have formed a group of c. 15 experts who represent potential users, with strong experience in the coordination of big research projects (e.g. European H2020 projects) or coordination of research programmes of clusters of activities. This working group will be involved in the following process:

1. **Workshop 1**: User representatives are invited to share their experience (skills, possible tools...) in real time and express their needs;
2. Creation of the prototypes of tools;
3. **Workshop 2**: the prototype is proposed to the participants, the tools are selected, adapted or, collectively designed;
4. The methodology and its tools are tested by participants on pilot cases supervised by the ASIRPA team;
5. **Workshop 3**: Feed backs on first use, collective learning

After this first pilot phase, tools will be further developed and their use will be generalised.

# 5. CONCLUSION

Such an intervention research project is a major opportunity to both develop new knowledge on process approaches of research and innovation activities and contribute to key transformations along the ambition to address major challenges.

In the current context, where research impact is a major stake, this project runs the risk to strengthen managerial practices that cherish short-term efficiency, probably at the price of long-term inventiveness. A key challenge of the project is to find ways to articulate directionality and creativity, to favour a good balance between exploration and exploitation. Our choice is to interact strongly with actors who are directly confronted with research coordination issues, in order to strengthen their skills and competences with the hope that we will succeed in transforming, or at least managing, the contradiction between directionality and serenpidity.

# 6. REFERENCES


**Bamberger, M., Tarsilla, M. and Hesse-Biber, S.** (2016). Why so many 'rigorous' evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute, Eval. Program Plann., 55, 155–162

**Bjögvinsson, E., Ehn, P. and Hillgren, P.-A.** (2012). Design things and design thinking: Contemporary participatory design challenges. Design Issues, 28(3), 101–116.

**Callon, M.** (1986). The sociology of an actor-network. Mapping the Dynamics of Science and Technology, Macmillan. M. Callon, J. Law, and A. Rip: London.

**Chen, H.T.** (1990). Theory-Driven Evaluations. Newbury Park CA: Sage Publications, Inc

**Chen, H.T.** (2012). Theory-driven evaluation: Conceptual framework, application and advancement. In R.Strobl, O. Lobermeier and W. Heitmeyer (Eds.), Evaluation von Programmen und Projekten für einedemokratische Kultur (pp. 17-40): Springer.

**Cohendet P. and Llerena, P.** (1989). Flexilibité, information et décision, Economica, Paris 1989.

**Costello, G. J.** (2015). Review of innovation process frameworks: Working Paper, Galway: Galway-Mayo Institute of Technology Repository.

**Courtney H., Kirkland J. and Viguerie P.** (1997). Strategy under uncertainty, Harvard Business Review, 75(6): 66-82.

**Douthwaite, B., Alvarez, S., Cook, S., Davies, R., George, P., Howell, J., Mackay, R. and Rubiano, J.** (2007). Participatory impact pathways analysis: a practical application of program theory in research-for-development. The Canadian Journal of Program Evaluation 22, 127.

**Douthwaite, B., Kuby, T., van de Fliert, E. and Schulz, S.** (2003). Impact pathway evaluation: an approach for achieving and attributing impact in complex systems. Agricultural Systems 78, 243–265. doi:10.1016/S0308-521X(03)00128-8

**Ely, A. and Oxley, N.** (2014). STEPS Centre research: our approach to impact.

**Garud, R, Tuertscher, P. and Van de Ven, A** (2013). Perspectives on Innovation Processes. The Academy of Management Annals Vol. 7, No. 1, 773–817, http://dx.doi.org/10.1080/19416520.2013.791066

**Hatchuel, A.** (2000). Intervention research and the production of knowledge. Cow up a Tree. Knowing and Learning for Change in Agriculture. Case studies from Industrialised Countries 55-68.

**Joly, P. B., Matt, M., Gaunand, A., Colinet, L., Larédo, P. and Lemarié, S.** (2015). ASIRPA : a comprehensive theory-based approach to assess societal impacts of a research organization, Research Evaluation, 24 (4), 440-453, doi:10.1093/reseval/rvv015.

**Joly, P. B. and Matt, M.** (2017). Toward a new generation of research impact assessment approaches, Journal of Technology Transfer, DOI 10.1007/s10961-017-9601-0

**Jeuffroy M-H and Prost L.** (2015). Centre de ressources sur la conception dans les domaines de l'agriculture, de l'environnement et de l'alimentation: intérêt et conditions de faisabilité. Rapport du groupe de travail.

**Klerkx, L. and Aarts, N.** (2013). The Interaction of Multiple Champions in Orchestrating Innovation Netwokrs: Conflicts and Complementarities, Technovation, 33, 193-210.

**Kuhlmann, S.** (2003). Evaluation as a source of 'strategic intelligence', in: Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe. Shapira, P., Kuhlmann, S, pp. 352–379.

**Lorino, P.** (1989). L'Economiste et le Manageur, Eléments de Micro-économie pour une nouvelle Gestion, Editions La Découverte, Paris.

**Le Masson, P., Weil, B., et Hatchuel, A.** (2006), Les processus d'innovation: conception innovante et croissance des entreprises, Stratégie et management (Hermes science.), Paris: Lavoisier.

**Matt, M., Gaunand, A., Joly, P.-B., Colinet, L.** (2017). Opening the black box of impact – Ideal-type impact pathways in a public agricultural research organization. Research Policy 46, 207–218. doi:10.1016/j.respol.2016.09.016.

**Mazzucato, M.** (2018). Mission-Oriented Research & Innovation in the European Union: A problem-solving approach to fuel innovation-led growth European Commission. ISBN 978-92-79-79918-1 doi:10.2777/36546

**Meynard J.M. et al.** (2016). Designing coupled innovations for the sustainability transition of agrifood systems ? Agricultural Systems, http://dx.doi.org/10.1016/j.agsy.2016.08.002

**Patton M. Q.** (2016). Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and use, The Guilford Press, New York, London.

**Patton, M. Q.** (2011). Developmental evaluation: Applying complexity concepts to enhance innovation and use. Guilford Press.

**Pawson R** (2013) The Science of Evaluation: A Realist Manifesto. London: SAGE.

**Retolaza, I.** (2011). 'Theory of Change: A thinking and action approach to navigate in the complexity of social change processes'. Hivos/UNDP/Democratic Dialogue.

**Rip, A.** (2012). The context of innovation journeys. Creativity and innovation management 21, 158–170.

**Rip, A., Schot, J.W.** (2002). Identifying loci for influencing the dynamics of technological development.

**Robinson, D.** (2008). Complexity scenarios for emerging techno-science: adressing strategy-context fit by prospecting level dynamics of governance. Third International seminar on Future-Oriented Technology Analysis, Seville, 16-17 October 2008.

**Robinson, D. K. R.** (2009). Co-evolutionary scenarios: An application to prospecting futures of the responsible development of nanotechnology. Technological Forecasting and Social Change, 76(9), 1222-1239.

**Robinson, D. K. R., and Rip, A.** (2013). Indications of Socio-Economic Impacts of Nanotechnologies: The Approach of Impact Pathways. In: Konrad, K., Van Lente, H, Coenen, C., Dijkstra, A. and Milburn, C. (Eds.): Shaping Emerging Technologies: Governance, Innovation, Discourse. IOS Press, pp 153 - 166. ISBN 9781614993001.

**Rogers, P.** (2008). Using programme theory to evaluate complicated and complex aspects of interventions. Evaluation 14, no. 1, 2008: 29-48.

**Van de Ven, A,** (2016). The innovation journey: you can't control it, but you can learn to maneuver it, Innovation, DOI: 10.1080/14479338.2016.1256780

**Van de Ven, A. H., Polley, D. E., Garud, R., and Venkataraman, S.** (1999). The innovation journey, Oxford University Press. ed. New-York.

**Van de Ven, A. H., Poole, M. S.,** (1990). Methods for studying innovation development in the Minnesota Innovation Research Program. Organization science 1, 313–335.

**Van de Ven, A. H., Poole, M. S.** (1995). Methods for studying innovation development in the Minnesota Innovation Research Program. Longitudinal Fields Research Methods: Studying Process of Organizational Change, GP Huber and AH Van de Ven (Eds) 155–185.

**Van Wulfen, G.** (2013): The innovation expedition. A visual toolkit to start innovation. Bis Publishers, Amsterdam.

**Weiss, C.** (1995). Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families in 'New Approaches to Evaluating Community Initiatives'. Aspen Institute.

**Wigboldus S, with Brouwers J** (2016). Using a Theory of Scaling to guide decision making. Towards a structured approach to support responsible scaling of innovations in the context of agrifood systems. Wageningen University and Research, Wageningen.

# AUTHORS

**PIERRE-BENOIT JOLY**
*LISIS and IFRIS, Université Paris Est*
5 Bd Descartes, 77454 Marne-La-Vallée, France
E: joly@inra-ifris .org

**MIREILLE MATT, DOUGLAS K. R. ROBINSON**
*LISIS and IFRIS, Université Paris Est*
5 Bd Descartes, 77454 Marne-La-Vallée, France
E: mireille.matt@inra.f
E: douglas.robinson@esiee.fr

# NEW EVALUATION FRAMEWORK IN FINNISH INNOVATION POLICY

JARI HYVÄRINEN

DOI: 10.22163/fteval.2019.327

## 1. INTRODUCTION

The aim of this paper is to present ideas for the framework that should be used to evaluate the work carried out by a new Finnish innovation funding and export-promoting organization, Business Finland. The evaluation framework described in this paper includes both impact analysis at the agency level and its implications for decision making at the policy level. It is a challenge to modify traditional impact analysis of R&D and innovation funding into innovation policy actions that may improve the internationalization of the Finnish innovation environment. New terms in this context are export promotion, trial platforms and world-class ecosystems, and traditional terms are radical innovations, productivity and renewing.[1]

The structure of the paper is as follows. Section 2 describes the evolution of Finnish innovation policy from a technology-oriented policy to a broad view of innovations and, finally, to an innovation and internationalizing policy mix. Sections 3 and 4 explain the purpose and methodology of the paper. In section 5, we present the potential effects of Tekes impact assessments on the Business Finland model, showing new outcomes of the evaluation framework and evidence of the additionality of public R&D and innovation funding and export promotion. Section 6 concludes the paper.

## 2. BACKGROUND

**Finnish Innovation Policy in 1990-2010**

Innovation policy in Finland focuses on improvements in human capital and R&D that accelerate renewal and productivity in the economy. One target, established in the 1980s, was to build a national innovation system. In general, a system is run by public organizations that influence the development and diffusion of technology and innovations. During the 1980s and 90s, industrial R&D spending grew faster in Finland than in other OECD countries. In the 1990s, the policy targeted changes in technology and was called a technology policy. One target of the policy was material goods product innovations and technological process innovations. At the end of the 2000s, it turned into a broad-based innovation policy (TEM, 2009).

During the recession in the early 1990s, there was an acute need to find new tools, as Finnish industry became uncompetitive in Western markets and unemployment grew rapidly. In the mid-1990s, with recovery already on its way, the chosen policy concept was to adopt intensive technological growth, which became a guideline in the Finnish science and technology policy. Another concept was to combine the national innovation system with a knowledge-based society, which was also called "The Finnish Model" (Lemola, 2003).

At the end of the 2000s, a new innovation policy concept was launched as a broad-based innovation policy. This concept revisited the definition of technological innovations in particular and started to focus on non-technological innovations. The diffusion of technologies and service innovations in particular to society and the economy was considered a main driver of policy actions. The concept also concentrated on the capacity to absorb and utilize innovations produced outside Finland because only one percent of innovations are created in Finland, and small open economies should integrate their innovation actors in research and industry more deeply into global innovation networks (TEM, 2009).

In large-company-led networks, interventions were carried out using the cluster-based approach. The focus was on improving research-led competitiveness in rapidly integrating global markets. The main policy tool was the Strategic Centers for Science, Technology and Innovation (SHOKs) concept launched in 2007. SHOKs were cluster-based public-private partnership organizations. The main idea was to accelerate innovation processes and renew industrial clusters led by large companies from traditional industries. One idea of the SHOKs was to apply new methods of cooperation especially among applied research and industrial companies but also to improve international co-operation and support to develop absorption capacity in Finland (TEM, 2013). To support

---

1 Definitions used in the paper:

**Platform**

A platform is a model in which organizations diagnose problems, identify opportunities and find ways to achieve their goals together. A platform creates value by facilitating exchanges between two or more interdependent groups, usually multiple buyers and sellers. Successful platforms have a tendency to disrupt existing markets and institutions in significant ways.

**Ecosystem**

An ecosystem is a solution entity supported by interacting actors (public sector, companies, research organizations and individuals), which is self-organized around a focal idea, actor or platform – mainly digital – creating value for its clients and participants in the entity.

**Radical innovation**

A radical or disruptive innovation is an innovation that has a significant impact on a market and on the economic activity of firms in that market. This concept focuses on the impact of innovations as opposed to their novelty. The innovation could, for example, change the structure of the market, create new markets or render existing products obsolete (OECD, Innovation Policy Platform).

knowledge diffusion, the University Inventions Act came into effect in January 2007. The new legislation along with the introduction of the new University Law (2010) allowed universities to act more freely to acquire external funding and organize their activities. It also transferred IPR rights to universities, as before the act all inventions belonged to the inventors. The reform increased universities' incentives to co-operate with companies and motivated them to take action with regard to the commercialization of research.

### CHALLENGES OF INNOVATION POLICY IN 2010S

After the 2008 financial crisis, innovation policy faced new challenges as the Finnish economy was stuck in sluggish growth for 10 years. During the recession, neither fiscal nor monetary policy were able to solve the rigidities of the Finnish export sector. Moreover, the Finnish government made drastic cuts to public research and innovation funding during the period from 2011-2017. The financial cuts decreased co-operation between applied research and companies in particular. Moreover, the government budget cuts in 2015 included the termination of the public SHOK funding, and the SHOKs program was closed in 2016. Another change in the Finnish innovation system was to merge two public organizations, Tekes and Finpro. Tekes – the Finnish Funding Agency for Innovation – had been the most important publicly funded expert organization for financing research, development and innovation in Finland. The goal of Tekes was to boost wide-ranging innovation activities in research communities, industry and service sectors. Business Finland (BF) was created on January 1, 2018, with the aim of combining R&D and innovation funding with internationalization services and invest-in activities.

Since the financial cuts, Finnish innovation policy has focused more vigorously on the concept of innovation environment, which encourages companies to enhance innovations, renewal and international growth. Therefore, Finland should revive value added and enhance economic diversification in the future by improving the internationalization of SMEs. As the OECD (2017) puts it, "Finland needs to tap new sources of growth based on new and sustainable export strengths, as well as by revitalizing traditional industries, fostering their capability to compete globally through new economic competences and value added. This transformation will require Finland to engage more in 'radical innovation' and become more effective in utilizing its valuable knowledge capabilities and transforming them into globally competitive innovation." In the applied research and innovation sector the policy places particular emphasis on the fields of i) wellbeing and healthcare, ii) bioeconomy and clean technologies, and iii) digitalization as new sources of growth.

## 3. PURPOSE OF THE PAPER

A goal of this paper is to present a new impact analysis framework for the new Finnish innovation and internationalization-promoting organization called Business Finland. A challenge is to build tools to evaluate innovation policy actions to improve export and other global actions as well as productivity in the Finnish innovation sector.

Therefore, the aim is to combine three aspects: first, how to modify the Tekes impact model such that it measures both innovation and internationalization-promoting activities; second, to discuss impact goals by pointing out the changes in the model by questioning what should be taken into account in carrying out evaluations; third, the evaluation framework should reveal areas where new innovation policy tools make a difference, i.e., considering the increase of inputs and outputs defined as productivity and acceleration of company growth and internationalization. In the current innovation policy set-up, this is supposed to strengthen the economic performance of the business sector and provide the largest benefits to the economy and society in the long term.

## 4. METHODOLOGY

The aim of R&D&I funding is to generate sustainable economic, social and environmental development and improve the net wellbeing of society. To implement these impacts and outcomes, Tekes has a long tradition of impact assessment. The Tekes impact model includes three main theoretical factors. The first is market failures, i.e., when the private sector (especially startups and SMEs) does not receive sufficient funding to solve puzzles that the market economy cannot solve and moreover does not invest enough in climate change, health care, etc., to achieve societal goals. The second is additionality (inputs, behavioral, outputs, impacts), as expressed for example by Georghiou et al. (2002), namely that if the public sector intervenes it should have an additive impact on the private sector and society as a whole. The third is spillover theory, which highlights that there is a lack of ideas in the market and that the public sector can support R&D and innovations by carrying out co-operative projects that increase the creation of new knowledge and ideas in the economy. The genesis of spillovers indicates that the public sector should also correct system failures, as actors need sufficiently large networks to contribute to the formation of spillovers. When considering export-promoting services we need to add two assumptions, which should be taken into account when evaluating these services in the future. The first assumption is bounded rationality, namely that companies accelerate the costs of gathering and processing information and have no resources to generate it at the company level. Another assumption that is linked to bounded rationality is information failure: SMEs have biased information with regard to their export possibilities in the global market. By considering the costs and benefits of these outcomes, it is beneficial for the economy to produce such information by using public services. The goal of these services is to broaden the growth mentality and understanding of new global challenges in SMEs.

## 4. ANALYSIS OF EVALUATION FRAMEWORK

The next step is to describe a paradigm change in the evaluation framework. We use the main objectives of Tekes and Business Finland as an example to explain the changes in the framework. These changes can be interpreted as resulting first from the sluggish growth in the export sector and second from the government financial cuts that induced the modifications in the innovation system. If we look at the key areas in the strategies implemented by Tekes in 2005 and 2011, the focus was on cluster-based innovation policy. Industry dynamics, continuous renewal and co-operation, internationalization and impacts on the economy and society were explained through clusters. Clusters were seen as constantly renewing sets of actors looking for new partnerships and value creation at global and multidisciplinary levels. Based on the cluster policy, there were three main objectives for funding.

- The first objective was *productivity and renewal,* whose focus was to examine the impacts of Tekes activities on the productivity of Finnish companies and on the renewal of the business sector. The main findings on the productivity of SMEs were linked to time lags and spillovers. The direct results of innovation activity in companies can be found after a time lag. The results manifest themselves as impacts that promote productivity and renewal and as impacts that spread outside the company (Valtakari et al., 2010).
- The second objective was *wellbeing and environment,* where the aim was to measure Tekes's success in promoting its targets related to societal wellbeing, the environment, and climate change. It was reported that Tekes was able to promote innovations, which had a positive societal impact with regard to this objective. Nevertheless, it was stated that Tekes had little influence on the broader implementation of the outcomes. It is largely beyond the reach of the activities of Tekes to exert a direct impact on societal wellbeing, the environment, and the prevention of climate change (Hjelt et al., 2012).
- The third objective was *capabilities,* and it assessed the role of Tekes in generating innovation capabilities in the Finnish economy. Halme et al. (2015) found that Tekes succeeded well in improving different types of capabilities. On average, the highest impact was on networking, whereas the impact on internationalization activities was weak. However, the differences between impacts on various capabilities should be studied carefully and compared to general targets such as the development of renewing industries.

## STRATEGIC OBJECTIVES OF BUSINESS FINLAND

As of 2018, Business Finland has two strategic objectives: 1) Global Growth for Companies and 2) World Class Ecosystems and Competitive Business Environment (Invest-In). Business Finland's strategy is twofold: it enables companies to grow internationally and create world-class business ecosystems and a competitive business environment for Finland. Therefore, its first goal is to create new growth by helping businesses go global and by supporting and funding innovations. Top experts and the latest research data enable companies to seize market opportunities and turn them into success stories. When considering the second strategic goal, ecosystems and business environment, Business Finland's role is to support the creation and renewal of business ecosystems. Moreover, its focus is to strive to have the best competences and talent available. Finally, its goal is to drive co-operation between public and private players and facilitate joint industry actions for selected potential world-class ecosystems.

When considering objectives, one can remark that the innovation process has a long time span. Outputs and business results only manifest themselves a few years after the project has ended. Development of an idea into a product or service and its commercialization may take as long as over ten years, depending on the technology sector. However, the time span can also be short. For example in the ICT sector and especially in the mobile game industry, the innovation process can take only several months, and after this time span the opportunity to enter the market is over.

## INPUTS AND ACTIVITIES

In Finnish innovation policy, there are only two funding mechanisms: grants and loans. The third widely used mechanism, tax incentives, is missing in Finland. In recent years, increasing numbers of OECD countries have introduced tax incentives as a primary innovation policy tool. Many international studies show that tax incentives have achieved varying results. The general finding is that the increase in funds invested in companies' innovation activities has been at least equal to the tax incentive. In other words, tax incentives miss the link that would create incentives for companies to increase their own R&D funding more than the amount of the tax incentive. For example, Romer (2000) remarked that designed grants are better tools than tax incentives because government agencies need to identify interventions that are better than what the market would implement; then the targeted grant programs could be socially valuable. The only argument for using tax incentives is that they are easy to use.

The main question is to ensure that public R&D funding adds to the R&D inputs by the companies and does not even partially crowd out these inputs (Georghiou, 2002). Mostly the assumption of crowding out has been refuted by research results. Ali-Yrkkö (2008) and Einiö (2014) found that public R&D funding increases companies' own R&D investments. Moreover, Pajarinen et al. (2016) reported that public R&D&I funding to startups does not crowd out private venture capital funding. In addition, several international studies show that public R&D funding increases corporate investment in R&D instead of crowding it out. Mostly, input additionality can be explained by a market failure in SMEs.

At the input level, new services have been added to the Business Finland (BF) impact model. The aim of the company growth services is to increase company contacts abroad. When considering ecosystems and invest-in, BF services should recognize potential new ecosystems and attract new players to Finland. Goals have been set to attract both national and international companies to invest in Finland with renewed capabilities to act in value networks. Evaluation of BF services needs new tools because, first, intervention is a continuing process and is more unobserved than funding decisions, and second, there is a need to collect exact data on how services have direct or indirect influence on company behavior.

## DIRECT RESULTS AND IMPACTS ON SOCIETY

The ultimate goal for direct results in public R&D and innovation funding is to improve productivity in the private sector. Solid growth in productivity enhances companies' capability to compete in the market and accumulates wealth by increasing the country's ability to fund its welfare services. The rise in productivity is based on intangible investments and innovations. A new product, service or method that produces economic or social benefits defines success in innovation.

When public R&D and innovation funding has a positive impact on the number and quality of R&D projects in companies, the outputs of companies and their business performance ultimately also improve. Several outputs signal improved productivity. In the Tekes impact model, the outputs were measured as growth of new companies and business areas as well as utilization and spillovers of new knowledge. Moreover, outputs can take the form of publications, patents, licenses or new services and processes. The business performance of companies (measured as sales or turnover) is a result of these new products, processes or services, which may improve productivity.

Tekes impact assessments target the productivity of renewing industries and producing spillovers more than export-promotion services. Several research results and impact studies show that Tekes's activities have direct impacts on company innovation efforts and growth. In particular, they have improved the efficiency of innovation activity and have had a positive impact on the creation of innovations and the increase of intangible assets and growth of sales (Valtakari et al., 2010). Business Finland impact assessments focus on accelerating scale economies, trial platforms and ecosystems, and export growth of SMEs in the global market. Radical innovations are forerunners, especially in the internationally oriented SME sector, suggesting that new goals of improved global competitiveness can be fulfilled.

Impact studies that estimate the effects of SME innovation funding and export-promotion interventions have recently been carried out. Halme et al.'s (2018) econometric analysis measured the success of internationally oriented Finnish SMEs in 2009-2015, which were customers of the Business Finland organization (ex-Tekes, ex-Finpro) compared with non-customer SMEs. Overall, there are almost 4,000 internationally oriented SMEs in Finland. The results show that Business Finland customers were younger, more export-intensive and more likely to have workers in innovation-related tasks, on average. Moreover, Finpro or Tekes interventions improved employment growth in Finland, and Finpro customers experienced improved sales growth. Finally, there were also indications of improved export growth.

The international literature indicates varying results regarding the positive relationship between export promotion and internationalization (Halme et al., 2018, pp. 20-23): *"First, there is existing evidence in the literature that more efficient firms become exporters (Clerides, Lach, and Tybout, 1998). Second, there is a common view that firms self-select into export-promotion services; the decision to utilize such services is likely correlated with the unobserved ability to export (Munch and Schaur, 2018)."* Moreover, there is strong evidence in the literature that innovation affects internationalization. For example, Altomonte et al. (2013) found a positive correlation between innovation and internationalization in European countries. They suggest that internationalization goes beyond exports, and internationalization over the longer term is likely to be driven by innovation more than export promotion.

At the societal level, the main indication is how public interventions succeed in stimulating spillovers. The first priority of the interventions is to impact the growth of competence and human capital within actors in the innovation environment. Accumulation of new ideas, knowledge and open innovations determine the success of companies as well as economic growth in the long term. Investment in R&D and innovation are needed to increase competence in the private sector, and the spillovers are needed for society as a whole and the national economy. Several studies indicate that public R&D and innovation funding in Finland has generated spillovers of up to 50-70 % for the whole economy (Takalo et al., 2013; Valovirta et al., 2014). As noted, spillovers to society may accumulate substantially greater benefits than just the impacts on the individual funding recipient. Maliranta et al. (2016) found that *"an increase in innovation subsidies is typically associated with an inflow of innovators from high-productivity firms. These findings suggest that innovation subsidies contribute to economic renewal and the diffusion of knowledge between firms."* In the Business Finland impact model, societal benefits are dependent on ecosystem-level capabilities and furthermore on Finnish companies having a focal role in global ecosystems. Therefore, societal benefits are dependent on straightforward goals related to SME export growth and the multifaceted role of ecosystems.

Business Finland's impact model based on direct results and impacts on society needs several improvements. One of the challenges for assessing the impact of R&D and innovation funding and export promotion is related to the impact on broad changes in ecosystems of digitalization, cleantech, bioeconomy, health and finally wellbeing at the macroeconomic level. From the evaluator point of view, evaluations need new tools to understand vertical and horizontal interconnections of ecosystems and their relevancy at the level of the whole economy and society. Without these improvements, there is a danger that the final impact results and recommendations will miss the link between ecosystem-level spillovers and the strategic decision-making at the agency and policy level. Therefore, we need more explicit ecosystem-level methods whereby evaluators could focus on Business Finland strategic objectives. Moreover, ecosystems are based on the platform economy, and these platforms need to be sufficiently connected to global demand at the early stage that they can become competitive in the global environment.

# 5. CONCLUSIONS

The aim of this paper was to demonstrate the revisited evaluation tools that are needed to respond to the demands of internationally oriented innovation policy. These demands are challenging because traditional R&D and innovation-based impact models underline market failures and dynamic aspects of spillovers. Therefore, clear innovation and internationalization logics seem to be ambiguous. One main challenge is to verify the internationalization logic in the innovation-based additionality model. A question is whether we can solve the problem by adding theoretical aspects of bounded rationality and asymmetric information to describe a justification for intervention. Another pathway is to build a link between innovation and internationalization, such that export-promotion services boost growth-seeking innovative companies' access to the global markets. Once this puzzle is solved it will be easier to plan new services for innovation-based platforms and ecosystems.

# 6. REFERENCES

**Ali-Yrkkö, J.** (2008). Essays on the Impacts of Technology Development and R&D Subsidies, A-236, Helsinki School of Economics.

**Altomonte, C., Aquilante, T., Békés, G. and Ottaviano, G** (2013). Internationalization and innovation of firms: evidence and policy. *Economic Policy*, 28(76), 663-700.

**Clerides, S., Lach, S. and Tybout, J.** (1998). Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. *Quarterly Journal of Economics*, 113(3), 903-947.

**Einiö, E.** (2014). R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding based on the ERDF Population-Density Rule, *Review of Economics & Statistics,* Vol. 96.

**Georghiou, L.** (2002), 'Impact and Additionality of Innovation Policy' in Boekholt, P. (editor) *Innovation Policy and Sustainable Development: Can public Innovation Incentives make a Difference*? Contributions to a Six Countries Programme Conference 28 February – 1 March 2002. IWT-Observatory, Brussels.

**Halme, K., Haila, K., Barge, B., Dalziel, M., Lemola, T., and Hautamäki, A.** (2015). Impact of Tekes on Capabilities, Tekes Review 318/2015.

**Halme, K., Kotiranta, A., Pajarinen, M., Piirainen, K., Rouvinen, P., Salminen, V. and Ylhäinen, I.** (2018). Efforts of of Finnvera, Finpro, and Tekes in Promoting Internationally oriented SMEs, Tekes Reports, Business Finland, Report 3/2018.

**Hjelt, M., Luoma P., Pesola A., Saario M., Kämäräinen V., Maksimainen A. and Vesa J.** (2011). Hyvinvoiva yhteiskunta ja ympäristö – Katsaus Tekesin vaikuttavuudesta, Tekesin katsaus 292/2012.

**Lemola, T.** (2003). Transformation of Finnish Science and Technology Policy. *Science Studies,* Vol. 16(2003) No. 1, 52– 67.

**Maliranta, M., Määttänen, N., and Pajarinen, M.** (2016). Firm Subsidies, Wages and Labor Mobility. ETLA Reports No 60.

**Munch, J. and Schaur, G.** (2018). The Effect of Export Promotion on Firm-Level Performance. *American Economic Journal: Economic Policy*, 10(1), 357-387.

**OECD** (2017). OECD Reviews of Innovation Policy: Finland 2017, OECD Publishing, Paris.

**Pajarinen, M., Rouvinen P. and Ylhäinen I.** (2016). Kasvun nälkä – Pääomasijoitetut yritykset muutosagentteina, Taloustieto Oy, (2016).

**Romer, P.** (2000). Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers? *Innovation Policy and the Economy 1*: 221–252.

**Takalo, T., Tanayama, T. and Toivanen, O.** (2013). Estimating the Benefits of Targeted R&D Subsidies, *Review of Economics and Statistics*, 95, 255–272.

**TEM** (2009). Evaluation of the Finnish National Innovation System, Taloustieto Oy.

**TEM** (2013). "Licence to SHOK?" External Evaluation of the Strategic Centres for Science, Technology and Innovation, Publications of the Ministry and Employment and the Economy, 1/2013.

**Valovirta, V., Lehenkari J., Lehtoranta O., Loikkanen T., Suominen A., Bodewes H., Mostert B., Zegel S. and van der Veen G.** (2014). The Impact of Tekes Activities on Wellbeing and Environment, Tekes Reports, 308/2014.

**Valtakari, M., Riipinen T., Rajahonka M., Ilmakunnas P. and Väänänen L.** (2010). Katsaus Tekesin toiminnan vaikutuksista tuottavuuteen ja elinkeinoelämän uusiutumiseen, Tekes.

# AUTHOR

**JARI HYVÄRINEN**
*PhD (Econ.), Strategic Support Services, Director General's Office, Business Finland,*
Porkkalankatu 1, 00180 Helsinki, Finland
E: jari.hyvarinen@businessfinland.fi

# NEW RUNNING EXPERIMENTS IN INNOVATION AND GROWTH POLICY: WHAT CAN WE LEARN FROM RECENT EXPERIENCE?[1]

TEO FIRPO AND JAMES PHIPPS

DOI: 10.22163/fteval.2019.328

## INTRODUCTION

Although experimentation is not new to policy in general (Oakley, 1998), it has only recently started to be used in the field of innovation and growth policy. The case for using robust experimentation techniques, such as randomised controlled trials (RCTs), is clear (Bravo-Biosca, 2016), and yet the field has lagged others, such as development or education (Dalziel, 2018).

Nevertheless, the number of impact evaluations using the RCT design in the field of innovation, entrepreneurship and business growth policy has grown. The IGL Database, which attempts to collect all RCTs in this field, currently counts 130 such experiments[1], of which over two thirds took place in the past decade. IGL has played a role in this growth, through the IGL grants programme[2], funding in the past five years over 30 randomised impact evaluations in this field, and also assisting a number of government agencies in their own journey to experimentation[3].

This paper is an attempt to synthetize the findings from this wave of experimentation, with a focus on trials relevant to innovation policy. It does so in the context of IGL's work, drawing lessons both from experiments our organisation has been involved with and the work it has done directly with government agencies around the world. It aims to provide a primer on the lessons that policymakers and researchers can draw on the use of experiments to evaluate innovation and growth policies.

It is structured in two parts. The first section reviews recent experiments in this field and provides an account of the evidence that they have generated. A second section investigates what these recent experiences can teach us about the practice of running experiments to test interventions in innovation and growth policy.

### A BRIEF NOTE ON SCOPE AND METHODOLOGY

Neither of the two sections that conform this paper aim to provide a comprehensive review of the evidence in this field.

The first section covers a number of RCTs, starting from those funded or indirectly supported by IGL; it often draws from other experiments carried out recently. It makes reference to both published and forthcoming research. It does not attempt to conduct an evidence review on any of the particular sub-policy domains, as this would be beyond the scope of this paper[4].

The second section is based on the direct experience of the authors working on experimentation with public organisations in the field of innovation and growth policy.

### LEARNING FROM RECENT EXPERIMENTS

Innovation and growth policy covers a wide spectrum of instruments and goals. Here, we have organised lessons from experiments along two broad policy aims. The first one focuses on expanding the number of people who participate in innovation activities to include those who would not usually participate. The second category reviews a number of interventions that support existing innovators (entrepreneurs, firms, or researchers) through a variety of schemes aimed at facilitating collaborations, improving skills, and ensuring the best ideas are backed. A final section focuses on an overarching question that might be of special interest to policymakers, namely, what are the best ways to fund innovation?

### MAKING INNOVATION MORE INCLUSIVE

An often overlooked aspect of innovation is who gets to be involved in it – and who is excluded. In their recent research, Chetty and colleagues marshalled a large dataset of inventors in the US, providing data on test scores and the amount of innovators they were surrounded by during childhood (Bell et al, 2017). They showed that coming from a family or area with many inventors is a strong predictor of becoming one, but that for children from disadvantaged backgrounds – even the brightest ones – the path to a career in innovation is much more difficult; they posit that a lack of exposure to innovation early on is a big part of this story. This finding has important economic consequences (we could have many more inventions if more people had been exposed to innova-

1 This essay has benefitted extensively from prior work undertaken by the authors as well as other colleagues at the Innovation Growth Lab, including in particular Lou-Davina Stouffs and Albert Bravo-Biosca. The authors would like to thank them or their contribution to this essay, as well as two anonymous referees for their comments.

tors as children) but also a wider societal impact, since this mechanism is likely a contributor to inequality.

There have been a number of experiments that address this challenge. An intervention funded by the IGL Grants Programme in Denmark showed that a simple online course on entrepreneurship for 9-grade pupils could positively affect their sense of self-efficacy and intention to pursue a career in entrepreneurship[5]. Now another IGL-funded RCT, and led by the World Bank group, aims to expose 19,000 students in a Latin American country to entrepreneurship and STEM through an online intervention[6].

The research by Chetty and colleagues also pointed to the value of role models in promoting an innovation mindset (Bell et al, 2017). This effect seems to have a strong gender component. A recent trial by an IGL-affiliated researcher focused on this aspect and showed that women role models can help improve the entrepreneurial attitudes and intentions of young women; the experiment explored the mechanisms through which this effect operates, and sheds light on how role models can be leveraged in an educational context (Bechtold and Rosendahl Huber, 2018). The power of female role models seems to persist even for actual entrepreneurs, as shown by an earlier trial in Chile, and to be a cost-effective approach to boost income when compared to more expensive consulting services (Lafortune and Tessada, 2015).

Another set of experiments has explored other ways to include people who do not usually partake in innovation activities. Two RCTs funded through the IGL Grants Programme stand out. The first one, carried out within a corporate environment in the Netherlands, used behavioural 'nudges' to increase the number of 'intrapreneurs' – employees providing innovative ideas to the company. The changes introduced were small – eg making the submission of ideas a default, or using examples of previous company intrapreneurs – but had significant effects; crucially, the increase in the quantity of ideas proposed did not come at the expense of their average quality (Weitzel and Rigterig, forthcoming). A similar experiment with engineering students at a US university found that using monetary incentives increased participation in an innovation contest, also without a decrease in quality (Graff Zevin and Lyons, 2018).

What these experiments show is that there are a number of interventions that can be used to make innovation a more inclusive endeavour, bringing in new people who would have otherwise not participated. But what can be done to support those who are already trying to innovate? The next section turns to a number of experiments that focus on this question.

## SUPPORTING INNOVATORS

Having more people involved in innovation does not guarantee that they will be successful. Often innovators need support to be effective, eg by helping them find the right collaborators, or giving them training and advice. Here we focus on a number of experiments providing academic researchers, firms and entrepreneurs with the right tools to innovate.

A key component in the production of knowledge and innovation is collaboration (Wuchty et al, 2007; Santamaría and Nieto, 2007), and yet there is little robust evidence on how it comes about. An IGL Grants Programme trial is currently exploring the role of physical proximity by randomly assigning groups of researchers within a building (taking advantage of a temporary move); the authors aim to find out if being closer to a fellow researcher increases the likelihood of collaborating (Catalini and Ganguli, forthcoming). This follows a previous experiment by Boudreau and colleagues showing that bringing together medical researchers (who worked at the same institution) for a 90-minute session could raise their likelihood to collaborate on a grant application by 75% (Boudreau et al, 2017). When it comes to collaboration among firms, an experiment focused on a similar intervention – bringing business owners together on a regular basis to share information – also led to positive results, with increases in sales and knowledge sharing (Cai and Szeidl, 2018).

But how about collaborations between researchers and firms? This is a policy goal that is central to many policymakers, especially in Europe (European Commission, 2007). A policy instrument that has been increasingly used is 'innovation vouchers' – credits given to SMEs to connect with researchers. The implicit assumption is that once a connection is made, and the firm is comfortable reaching out to researchers, there will be long-term positive effects. Yet the evidence from experiments suggests this might not be the case. An RCT on a Dutch voucher programme found that an initial strong positive impact faded within the space of a few years – with firms that had received the voucher not performing any better than those that did not (Cornet et al, 2006)[7]. Nesta, where IGL is based, carried out an RCT to test the effects of 'Creative Credits', a voucher scheme focused on connecting SMEs with creative industry providers; it also found an initial impact that faded in the longer term (Bakhshi et al, 2015). More recently, IGL has supported one of its governmental partners with an RCT on their innovation vouchers programme[8]. These experiments point to a key advantage of RCTs: by comparing firms that were similar across all characteristics, but varied only in the randomly assigned reception of a voucher, and tracking outcomes for several years, they were able to go beyond the initial positive impacts and provide results that can inform a cost-effectiveness decision on the voucher programmes.

Another approach to support innovators has focused on giving advice and training to entrepreneurs to improve their skills and their ventures. Once again, however, there is not a lot of robust evidence on exactly what type of advice or training is most effective; a number of experiments have been recently run to investigate this question. An RCT conducted by the World Bank in West Africa compared two models of entrepreneurship training: one focused on 'traditional' business skills (eg financial management, marketing); the other on fostering a proactive mindset and entrepreneurial behaviours (Campos et al, 2017). It found that the latter was much more effective. Now an IGL Grants Programme trial is comparing similar training programmes in Jamaica[9]. Two other IGL Grants Programme RCTs are exploring how to improve the ways accelerators help new ventures through training. The first one, in Italy, has shown that teaching entrepreneurs to see their startups in scientific terms, framing each new move as a science experiment, can have a large positive impact on their customer activation and fundraising (Camuffo et al, 2017). The other, in Chile, is currently investigating whether 'structured accountability' – asking founders to periodically present on their strategy and progress on goals – can help startups succeed[10]. Previous experiments have also shown that small interventions can make a difference; for instance, an RCT presented at the IGL2017 conference showed that the value of providing founders with feedback already collected when the applied to be part of the Startup Chile accelerator: startups by founders who received the feedback were later on more likely to have survived and raised significantly more money (Wagner, 2016). Now IGL is supporting one of its partners in replicating these results, to help them decide whether internal evaluations of applicants should be shared with them as feedback.

There have also been a number of experiments testing the effectiveness of consulting services on firms. Bloom and colleagues demonstrated the strong positive impact on productivity of management consulting for manufacturing firms with an experiment in India (Bloom et al, 2013); the intervention was intensive and costly, but they found that the gains in productivity offset the costs. They also followed up several years later and found that many of the effects persisted (Bloom et al, 2018). A more recent example, using a less intensive consulting intervention for SMEs in Mexico, found strong effects on employment (Bruhn et al, 2018).

### WHAT IS THE BEST WAY TO FUND INNOVATION?

The two sections above highlight the range and variety of interventions to support innovation. Governments also often choose to directly fund research and innovation through grants. However, there is surprisingly little research on what are the best ways to evaluate and select the best proposals in grant funding calls. The details might matter significantly. Some experimental evidence, for instance, has shown that evaluators tend to give lower scores to proposals in their own areas of expertise and to highly novel proposals (Boudreau et al, 2016). Now an IGL Grants Programme trial is further exploring this question in the context of a matching-grant scheme for Mexican SMEs[11].

At IGL we are currently exploring a number of questions around this topic, and carrying out research with governmental partners on their grant-making processes. This work is part of our ongoing collaboration with a number of innovation agencies across the world. These collaborations have taught us a number of useful lessons on how experimentation can be applied to the work of public organisations. We now turn to these lessons.

## LESSONS FROM EXPERIMENTING IN INNOVATION AGENCIES

The rise in experiments in the field of innovation, entrepreneurship and growth has been primarily led by academic researchers. However, several public organisations have begun to use experiments to evaluate their own programmes, as well as better develop new ones. A key example is the UK's Department for Business, Energy and Industrial Strategy (BEIS). BEIS, one of IGL's original governmental partners, went from never having used RCTs in its evaluations of business programme, to running one of the largest business support RCTs to date[12]. More recently, BEIS has launched a large experimentation fund to support projects aiming to spread technology and management practice diffusion among SMEs[13]. This follows the announcement by the European Commission of a fund to support innovation agencies with their own RCTs[14]. Despite these positive examples, it remains difficult for public organisations to embrace the idea of experimentation. In this section, we review a number of practical lessons we have gathered from working with innovation agencies on experimentation.

### MESSAGING TRIALS ARE A POWERFUL ENTRY POINT FOR EXPERIMENTATION

Oftentimes, innovation agencies can find the process of running an RCT quite challenging at first. In our experience, a useful starting point is to run messaging trials – behavioural experiments to find out what language is most suitable to achieve a certain goal, such as convincing firms to take up a programme. However, running this type of experiments requires involving several teams from the organisation together, and they are most powerful when used as part of a larger strategy, rather than in an *ad hoc* fashion.

### NEW TYPES AND SOURCES OF DATA ARE NEEDED TO GET BETTER RESULTS

This lesson is probably applicable to all types of evaluation, although it is particularly useful for RCTs. We have found that relying exclusively on surveys can bring a number of problems – such as low response rates or survey bias. This can be a problem, especially in the context of RCTs, since surveying control group firms that received nothing from the organisation can be difficult. Novel data sources – such as web-scraping or other 'Big Data' tools – can be coupled with more traditional datasets to achieve better results. Moreover, better dataset matching – especially with administrative datasets such as tax data – can be a powerful tool in running successful experiments.

### EXPERIMENTATION AS A POLICY APPROACH

RCTs are a tool to validate a hypothesis – ie find out whether a certain policy intervention works as intended. However, experimentation is not restricted to validating, but can be used to explore new and innovative solutions to policy challenges, with techniques such as design thinking and horizon scanning[15]. In our experience, innovation agencies achieve the best results when they think experimentally throughout the policy cycle.

### MORE THINKING IS NEEDED ON OUTCOME MEASURES

Although RCTs are a robust method of causal inference, the value of the results depends on the quality of the outcome measurements used. Experiments, unlike retrospective studies, frontload the evaluator's work, so that the majority of the planning, decision-making, and analysis happen before the intervention even started. This has its advantages, but it also means that once the trial has begun it is very difficult to change any of its parameters. This is why more care is needed when selecting the outcome measures to be used. In our experience, evaluators need to think not just about first-order, but also second- and third-order effects, to ensure the indicators used capture the policy's real effects. For instance, an intervention connecting SMEs to research institutions might aim to foster better collaborations; researchers should think hard about how exactly this improvement will materialise: is it more connections, or higher frequency, or larger projects? A simple measurement, such as number of collaborations, might miss a more profound change taking place because of the intervention. Because the survey can only be run once, asking the wrong question can compromise the whole project. Wherever possible, we recommend using a logic model to understand what effects one might expect.

# CONCLUSION

This paper has focused on lessons from recent experiments on innovation, entrepreneurship and growth policy. It has reviewed the approaches tested, and the findings (where available), of a number of RCTs in this field. The first set of experiments surveyed show that it is possible to use interventions to expand the reach of innovation activities beyond current levels. Another group of RCTs described provides evidence on a number of ways to support innovators, from facilitating collaborations to providing intensive management consulting services. An open question remains on what are the best ways to structure evaluation and selection processes in grant-funding programmes.

These examples highlight two key elements of experiments. The first one is that RCT, when well run, can provide a clean estimate of the effects of a programme; these estimates can be used to investigate the cost-effectiveness of a programme and compare it to its alternatives. The second is that sometimes even small interventions – such as changing the language used to communicate, or sharing feedback that an organisation was already collecting – can really make a difference; this kind of inexpensive but impactful opportunities should be sought after and implemented wherever possible.

An important caveat to all these findings is that, despite their strong internal validity, RCTs do not necessarily have external validity. In other words, even though an experiment can give policymakers confidence that a programme worked in a particular context, this does not mean it would work elsewhere, or with different participants. This is not a limitation exclusive to RCTs – other impact evaluation techniques usually run into similar concerns. But it does point to the fact that evidence from these trials should be understood in context. Wherever possible, replications of these studies should be carried out to build more evidence on the effectiveness of the interventions studied in other contexts.

Moreover, the paper has also presented a number of practical lessons on how to experiment within public organisations working on innovation and growth. These include starting from small experiments, using new data sources, and devoting more consideration to the choice of outcome measures. In our experience, experimenting is not something that comes naturally to a lot of public organisations. Nevertheless, we hope this paper shows that it is an approach that holds immense potential – as one tool among many.

# REFERENCES

**Bakhshi H., J.S. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley et al.** (2015). Assessing an experimental approach to industrial policy evaluation: applying RCT+ to the case of creative credits. Nesta Working Paper 15/10.

**Bechtold, L. A. and Rosendahl Huber, L.** (2018). Yes I Can! - A Field Experiment on Female Role Model Effects on Entrepreneurship, Academy of Management Proceedings, Vol. 1.

**Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen** (2017). Who Becomes an Inventor in America? The Importance of Exposure to Innovation, NBER Working Paper No. 24062.

**Bloom, N., B. Eifert, A. Mahajan, D. McKenzie and J. Roberts** (2013). Does management matter? Evidence from India, The Quarterly Journal of Economics, 128(1), p. 1–51.

**Bloom, N., A. Mahajan, D. McKenzie and J. Roberts** (2018). Do management interventions last? Evidence from India, NBER Working Paper 24249.

**Boudreau, K. J., T. Brady, I. Ganguli and P. Gaule** (2017). A field experiment on search costs and the formation of scientific collaborations, Review of Economics and Statistics, 99(4).

**Boudreau, K., E. Guinan, K.R. Lakhani, and C. Reidl** (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance and resource allocation in science, Management Science 62(10).

**Bravo-Biosca, A** (2016). Experimental innovation and growth policy: Why do we need it? Innovation Growth Lab Policy Briefs, 1. Retrieved October 15, 2018 from: https://innovationgrowthlab.org/policy-briefs

**Bruhn, M., D. Karlan and A. Schoar** (2018). The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico, Journal of Political Economy 126(2), p. 635-687.

**Cai, J. and A. Szeidl** (2018). Interfirm relationships and business performance, The Quarterly Journal of Economics, 133(1), p. 1229-1282.

**Campos, F., M. Frese, M., Goldstein, L, Iacovone, H.C. Johnson, D.J. McKenzie and M. Mensmann** (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa, Science 357(6357), p. 1287-1290.

**Camuffo, A., A. Cordova and A. Gambardella** (2017). A Scientific Approach to Entrepreneurial Decision-Making: Evidence from a Randomized Control Trial, CEPR Discussion Paper No. DP12421.

**Catalini, C. and I. Ganguli** (forthcoming). Rethinking innovation spaces. Project page available at https://innovationgrowthlab.org/projects/rethinking-innovation-spaces

**Cornet, M., B. Vroomen and M. van der Steeg** (2006). Do innovation vouchers help SMEs to cross the bridge towards science? CPB Discussion Paper 58.

**Dalziel, M.** (2018). Why are there (almost) no randomised controlled trial-based evaluations of business support programmes? Palgrave Communications, 4(12).

**Edler, J., Shapira, P., Cunningham, P. and Gok, A.** (2016). Conclusions: Evidence on the Effectiveness of Innovation Policy Intervention. in J Edler, P Cunningham, A Gök and P Shapira (eds), Handbook of Innovation Policy Impact. Eu-SPRI Forum on Science, Technology and Innovation Policy series, Edward Elgar Publishing, Cheltenham, pp. 543-564.

**European Commission** (2007). Improving research transfer between research institutions and industry across Europe: embracing open innovation – embracing the Lisbon Agenda, Communication to the Council, the European Parliament, the European Economic and Social Committee and the Committee of Regions.

**Graff Zevin, J. and E. Lyons** (2018). Can Innovators Be Created? Experimental Evidence from an Innovation Contest, NBER Working Paper No. 24339.

**Giones, F.** (forthcoming). Accelerating science-based growth through co-innovation. Project page available at https://innovationgrowthlab.org/projects/accelerating-science-based-growth-through-co-innovation

**Lafortune, J. and J. Tessada** (2015). Improving financial literacy and participation of female entrepreneurship in Chile, Final Report to Global Development Network CAF/GDN Project.

**Wagner, R.A.** (2016). How does feedback impact new ventures? Fundraising in a randomized field experiment. Available at https://ssrn.com/abstract=2766566 or http://dx.doi.org/10.2139/ssrn.2766566

**Weitzel, U. and C. Rigtering** (forthcoming). Corporate Nudging & Intrapreneurship. Project page available at https://innovationgrowthlab.org/projects/corporate-nudging-intrapreneurship

**Wuchty, S., B. F. Jones and U. Brian** (2007). The Increasing Dominance of Teams in Production of Knowledge, Science 316, p. 1036.

# AUTHORS

**TEO FIRPO, JAMES PHIPPS**
*Innovation Growth Lab (IGL)*
Nesta, 58 Victoria Embankment, London EC4Y 0DS (UK)
E: teo.firpo@nesta.org.uk
E: james.phipps@nesta.org.uk

Endnotes

1 Available at https://innovationgrowthlab.org/igl-database. The most recent version of the Database is forthcoming.

2 The IGL Grants programme is funded by the Ewing Marion Kauffman Foundation, the Argidius Foundation, and Nesta to support randomised controlled trials (RCTs) that build the evidence base on the most effective approaches to increase innovation, support high-growth entrepreneurship and accelerate business growth. https://www.innovationgrowthlab.org/funding-opportunities

3 A list of our current partners is available at https://innovationgrowthlab.org/partners. Moreover, we have worked with the Dutch Ministry of Economic Affairs, Business Finland, Swedish Growth Analysis group, and the Danish Business Authority.

4 The Nesta supported 'Compendium of Evidence on Innovation Policy' provides a more comprehensive review http://www.innovation-policy.org.uk/compendium/. Further evidence reviews in this field, drawing on a number of studies with different methodologies, have been produced by the What Works Centre for Local Economic Growth, available at http://www.whatworksgrowth.org/policy-reviews/.

5 Cf K Moberg (forthcoming). The effect of online-based entrepreneurship programmes. Project page available at https://innovationgrowthlab.org/projects/assessing-effect-online-courses-entrepreneurship

6 Project page forthcoming.

7 For the original report (without longer term effects), see (Cornet et al, 2006); more recent results are currently being prepared, however, one of the authors presented some preliminary results in an IGL Webinar.

8 Cf Rosendahl, L., M. Kleine and J. Heite (forthcoming). The effect of Innovation Vouchers on innovation activity of SMEs in the UK: A randomized controlled trial. Project page available at https://www.ip.mpg.de/en/projects/details/the-effects-of-innovation-vouchers-on-innovation-activity-and-performance-of-smes-in-the-uk-a-randomized-controlled-trial.html

9 Ubfal, D. and A. Maffioli (forthcoming). Unbundling the effects of entrepreneurship education. Preliminary results available at https://innovationgrowthlab.org/blog/impact-soft-skills-training-microentrepreneurs-jamaica

10 Cf M. Leatherbee (forthcoming), Does structured accountability drive entrepreneurial performance? Project description available at https://innovationgrowthlab.org/blog/does-structured-accountability-drive-entrepreneurial-performance

11 Iacovone, L., D. Atkin, E. Verhoogen (forthcoming). Promoting high impact entrepreneurship in Mexico: An impact evaluation. Project page available at https://innovationgrowthlab.org/projects/promoting-high-impact-entrepreneurship-mexico-impact-evaluation

12 One of us was deeply involved in BEIS' (then BIS) early steps in experimentation, for an account of these early steps, see the following blog: https://innovationgrowthlab.org/blog/taking-first-steps-business-policy-experimentation

13 Cf https://www.gov.uk/government/news/government-launches-new-fund-to-support-small-business-growth

14 Cf http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/innosup-06-2018.html. IGL recently delivered two webinars to help potential beneficiaries to prepare their trial designs and implementation plans for the second stage of the selection process.

15 For a review of these techniques, see the following blog https://www.nesta.org.uk/blog/exploring-the-unobvious-an-overview/.

# THE ROLE OF CITIZENS IN SETTING THE VISIONS FOR MISSION-ORIENTED RESEARCH AND INNOVATION[1]

JULIEN CHICOT AND ALBERTO DOMINI

DOI: 10.22163/fteval.2019.329

## ABSTRACT

Horizon Europe aims to orient EU research and innovation policy towards bold and ambitious missions and to engage, as part of this process, a wide range of stakeholders. In presuming that the approach to public participation in policy-making is linked to the characteristics of each mission-oriented R&I initiative, this paper aims to investigate the role of citizens in the definition of missions and thus in building the (input and output) legitimacy of the related initiatives. On the one hand, a large sample of case studies provides evidence of the practices of citizen involvement in vision-setting and demonstrates that they are still primarily aimed at ensuring citizens' buy-in rather than involving them genuinely in the definition of missions. On the other hand, findings from stakeholder interviews and an expert workshop shed light on the challenges in engaging citizens in decision-making: besides designing an efficient procedure, the role of citizens in respect to other stakeholders should be clearly identified. Even though low involvement of citizens in vision-setting did not seemingly affect the effectiveness of most of the mission-oriented initiatives investigated, further efforts for engaging them in decision-making should be made in the light of the increasing complexity of challenges and the perceived democracy gap in Europe.

## INTRODUCTION

Political upheavals in recent years are symptoms of a significant and widening divide between politicians and their electorate, between the rulers and the ruled. With the exponential increase in the use of social media, participation and representation are acquiring new forms and pose new challenges to the functioning of even the most consolidated democracies. In a context where large parts of the population have access to education and information, decision-making does not any longer appear as the exclusive prerogative of the so-called 'establishment'. The views according to which elected officials take policy decisions while citizens express themselves only during elections are highly contested. Public institutions and especially those of the European Union, should keep up with this new phenomenon and demonstrate their good intentions to adapt to the new circumstances and address this democratic gap.

This trend is not a new phenomenon. Since the 1990s, if not before, democratic deficits and social exclusion have been in the spotlight (the democratic deficit of the European Union has been a topic of EU affairs already for a long time) and bottom-up policy-making mechanisms engaging citizens, such as participatory budgeting, have been experimented. By opening the decision-making process to external stakeholders, policymakers expect to reduce conflicts and favour societal acceptance of their decisions. In a context of a growing demand for transparency and participatory policy-making, policy evaluation should consider legitimacy, alongside other criteria such as effectiveness and efficiency.

These 'citizens' refer to individuals belonging to a social community ruled by recognised bodies and institutions. This broad definition embraces a wide array of actors, who may sometimes act, as individual experts or market actors. Despite this potential confusion between citizens and other categories of stakeholders, their distinction is especially relevant in the analysis of bottom-up policy-making. Unlike other actors, citizens should be involved to reflect on problems and potential policy responses based on the societal needs and values of the community to which they belong. While not neglecting the contribution of people when acting as services users, consumers of goods or individual experts, this research considers bottom-up approaches in policy-making as those allowing the involvement of citizens (also called hereafter 'common' and 'ordinary people').

Current trends in research and innovation (R&I) policy at the EU and national levels have given a renewed impetus to citizen engagement in policy-making. In June 2018, the European Commission proposed that "missions" form part of the future Ninth EU Framework Programme for

1 The authors wish to thank their colleagues in the Joint Institute for Innovation Policy (JIIP) and especially Jacqueline Allan and Robbert Fisher for their support in this endeavour and their valuable comments.
This research draws on two studies conducted for the European Commission (Directorate General for Research & Innovation) in the context of services contracts (30-CE-0880718/00-38 and 30-CE-0883606/0035) for the preparation of Horizon Europe. Both studies were coordinated by JIIP and involved its four members (Joanneum Research, Tecnalia, TNO and VTT) as well as the Danish Technological Institute (DTI) and Valdani Vicari and Associati (VVA). The views expressed in this paper are nevertheless exclusively those of the authors and do not reflect the position neither of JIIP and its members nor of the European Commission.

Research and Innovation, Horizon Europe. Following this approach, the EU R&I policy will increasingly concentrate efforts on the development and, in some circumstances, diffusion of new solutions to identified problems, and thereby on the achievement of ambitious goals. Missions would typically present clearly defined targets to be achieved within a specific timeframe, so progress can be measured against predefined milestones. While public administrations remain the main policy-makers, private organisations, such as businesses and foundations, have been also very active in identifying missions critical to them and their communities, most often with the support of public administrations (JIIP et al., 2018a).

Mission-oriented R&I initiatives may be divided into two broad categories depending on the nature of their goals: (1) programmes focused on achieving a single well-defined objective, often of scientific or technological nature (e.g. accelerating the development of a solar-powered aircraft able to revolutionise air transport), and (2) far-reaching initiatives aimed at (or implying) the transformation of systems to address complex challenges (also known as societal and/or transformative missions, e.g. climate change or the ageing society).

The orientation of R&I policy towards missions inherently requires that a vision is defined beforehand and that the actions of all relevant stakeholders are coordinated accordingly (Weber and Rohracher, 2012). Missions relate to complex challenges that isolated (traditional) policy-makers may have difficulty in grasping. Furthermore, their achievement may have large impacts affecting many actors. Due to this orientation, they are also more likely to be highly visible to citizens and therefore more sensitive to societal acceptance. Therefore, the main objectives of engaging actors other than traditional decision-makers are to ensure that the selected missions address the most pressing needs and that legitimate initiatives will result. For these reasons, full top-down approaches in the definition of visions for orienting R&I policies and efforts are raising growing criticism, and research on the rationales for, and the modalities of, citizen involvement in the development of missions, and in co-creating a vision for future R&I policies, becomes crucial.

The current work aims to contribute to an understanding of the relevance of an open approach to policy-making in the specific area of mission-oriented R&I, by distinguishing different levels of citizen involvement in the current practices, and the challenges that their implementation entails. By focusing on the rationales for citizen involvement and on the modalities in which this has been displayed, our research also aims to investigate the level of legitimacy that the general public entrusts, according to its societal values and needs, in the process leading to the launch of these broad policy interventions (input legitimacy) and in the pursued outcomes (output legitimacy) (Boon and Edler, 2018).

After an outline of the literature on the role of citizens in policy-making (Section 1) and the description of the methodology employed to gather evidence (Section 2), this paper proceeds in identifying the most common practices of citizen involvement observed in the vision-setting of mission-oriented R&I initiatives (Section 3). To complete the analysis, the impressions of stakeholders from academia, public administrations and industry help characterise three main challenges that policy-makers and researchers encounter in examining citizen involvement (Section 4). Finally, Section 5 concludes and provides the EU policy-makers with some recommendations.

# 1. ROLE OF CITIZENS IN POLICY-MAKING AND MODALITIES OF THEIR ENGAGEMENT

## 1.1 THE ROLE OF CITIZENS IN POLICY-MAKING

Policy-making is the process by which the responsible authorities determine an appropriate course of action to solve a problem and address an opportunity for their target group. It takes place in an environment which influences it (Flanagan and Uyarra, 2016), with stakeholders that range from recipients and providers of the solution (for a problem) and may include other interested parties (e.g. philanthropists or lobbyists). For these reasons, policy-making processes vary widely depending on the national, regional, sectoral or technological systems in which they occur. For the sake of clarity, they can be divided into three stages (Edquist, 2011): (i) setting the vision, i.e. defining the problem to be solved; (ii) identifying the causes of the problem and translating them into recognisable objectives and into sets of smaller, achievable and measurable goals; and, (iii) selecting the policy instruments. The visions that set the direction for policy interventions relate to problems that are considered social constructs. Their definition is influenced by a number of contingencies (Laranja, Uyarra and Flanagan, 2008) and has a political dimension that should not be overlooked (Borrás and Edquist, 2013).

Policy-making is a process orchestrated by the responsible authorities but with the concurrent intervention of a wide range of actors seeking recognition for their respective needs, and their inclusion in the policy agenda. These 'policy entrepreneurs' (Kingdon, 1984) may consist of industry and other interest groups, as well as individuals and citizens. Traditionally, policy-makers have been the elites, i.e. the dominant groups within specific communities, assuming and maintaining positions of power in governmental institutions as well as social movements (often under organised and militant minorities) that enforce their own thematic agendas. Citizens have been maintained in passive roles delegating their voice to elected representatives (convinced that further public participation could disrupt the functioning of public administrations) or being customers of public services (Vigoda, 2002).

The profound economic, demographic and social changes that emerged in most, if not all, OECD countries in the post-war period have led to a growing demand for the opening up of policy-making processes to public participation (OECD 2001). Since then, mechanisms – increasingly supported by the digitalisation of public services and social media – have been granting citizens the opportunity to mobilise, organise and influence priority setting. Increasing efforts have been made to improve access to public information (including explanations of the choice of the employed instruments). Policy-makers may also ask citizens, through a consultative process, to reflect on their decisions and to provide feedback and additional insights. Such actions may be used to support the selection of missions and associated policy instruments. Finally, citizens may be engaged earlier in the process to shape social and policy dialogues and identify the most pressing challenges, the missions and related policy interventions.

Nevertheless, the involvement of citizens in policy-making does not mean that they replace the public authorities in their role of designing and implementing policies. Formal policy formulation remains in the hands of traditional policy-makers, who should not be confused with 'policy entrepreneurs'. The paucity of citizen involvement is especially prominent in R&I policy, where it is believed that researchers should benefit from full freedom in the direction of their research (Bush, 1945). The emergence of mission-oriented R&I increases these tensions within the policy-making process between 'policy entrepreneurs' and traditional policy-makers and, more specifically, between the need to involve of a wide range of actors, to define the most relevant visions, and the need for leadership (as opposed to a self-organising process), to guide system transition (Bugge et al., 2018).

### 1.2 THE RATIONALES FOR CITIZEN INVOLVEMENT IN VISION-SETTING

Directionality is a core and differentiating feature of mission-oriented R&I policy (Weber and Rohracher, 2012), which implies the definition of a vision that will guide policy interventions towards the solution of identified problems. Because these problems are complex and "wicked" when they are linked to socio-economic challenges (Nelson, 2011), policy-makers may require, for their better understanding, knowledge from external stakeholders such as citizens.

Efforts to involve citizens in policy-making have increased also in response to their growing demand for transparency, accountability and participation (OECD, 2001). They are expected to have a direct and positive influence on the legitimacy of policy decisions, i.e. on their level of societal acceptance and (implicit) popular support. According to Dahl's definition (1998), legitimacy has to do with a "*general confidence among the public that a government's power to make binding decision for the polity are justified and appropriate*" (cited by Wallner, 2008, p. 422). Given that there is no universal criterion to assess whether, and to which degree, a policy measure is legitimate, such a functional definition of legitimacy highlights its subjective nature: it primarily relies on perceptions about the beliefs of individuals and groups.

Legitimacy is not static but varies throughout the policy cycle as the perceptions of, and popular support for, policy measures may vary during their implementation. However, even though opportunities for citizen participation may arise at any stage of this cycle, this research contends that policy-makers should pay particular attention to the engagement of citizens in vision setting. Indeed, the initial level of legitimacy has significant influence on the subsequent phases of the policy cycle (Jagers et al., 2016). Furthermore, a legitimate vision may ensure directionality of mission-oriented R&I initiatives and therefore reduce the risk of policy failure (Wallner, 2008).

Legitimacy has two dimensions: input legitimacy and output legitimacy (Boon and Edler, 2018). Input legitimacy designates the societal acceptance of the process through which needs are transformed into policy problems and the instruments to solve them are defined. It is influenced by the level of openness to stakeholders other than the traditional policy-makers, the efficiency of the process, as well as its transparency. Output legitimacy refers to the situation in which policies are societally accepted and supported by citizens because their outcomes are seen to contribute to addressing perceived societal needs (Scharpf, 2006).

Output legitimacy has traditionally been considered to be the most significant component of the overall legitimacy of policy measures. In other words, citizens may give their support to policy decisions, whose targeted outputs are in line with their needs and expectations (output legitimacy), even though they do not perceive the process, through which these decisions have been developed, as fully fair (e.g. as they were excluded from it) (Boedeltje and Cornips, 2004). Nevertheless, in the current context of growing criticism of the representative democratic model, the perception of fairness in policy-making should not be overlooked and top-down decision-makers must design and follow policy-making processes that satisfy citizens' expectations. Otherwise, R&I policy – particularly if mission-oriented – may fail.

## 2. METHODOLOGY

To investigate the engagement of citizens in setting the direction for, and building the legitimacy of, of mission-oriented R&I, evidence was drawn from two studies on mission-oriented R&I policy to support the European Commission in the preparation of Horizon Europe (JIIP et al., 2018a; JIIP et al., 2018b). The materials include a series of case studies (identified based on a global policy mapping), interviews with R&I stakeholders, and a final workshop with experts and stakeholders.

The current work uses multiple case studies to compile compelling and robust evidence for supporting the analysis of mechanisms (i.e. their rationales, context and instruments) to engage citizens in setting visions for mission-oriented R&I policy. 53 out of 140 identified mission-oriented R&I initiatives in the European Union, its 28 Member States and some of its main competitors[1] were analysed[2]. This selection takes into account their geographical and thematic coverage, whether they are public or private initiatives, the type of challenge they target (accelerators or transformers), their level of intervention (i.e. international, national, regional or local), and their scale. Information on their overall context, policy instruments, governance, drivers and level of engagement of citizens and stakeholders was collected[3] to ease their comparison.

To understand, based on these case studies, the practices used to engage citizens in the direction-setting process for mission-oriented R&I initiatives, the relevant information was identified and coded. It was firstly determined whether and how citizens were engaged; the timing of their engagement; and the instruments employed. Information was collected, where available, on the rationales for citizen engagement and complemented by desk research. A broad definition of citizens, including civil society organisations, was adopted. Even though no causal relationship between any feature of mission-oriented R&I initiatives and the degree of citizen engagement in vision-setting can be identified based on this methodology, the coding adopted enables a better understanding of such practices.

The case studies were complemented with insights from researchers, policy-makers (including EU agencies, national governments, local authorities and national research and innovation agencies), representatives of industry and of civil society organisations. Their perceptions of the challenges of, and solutions for, a higher level of engagement of citizens in the vision-setting process for mission-oriented R&I initiatives were collected by means of a series of interviews[4], and an expert and stakeholder workshop[5]. Their findings allow to flesh out the existing tensions between the need to have a clear orientation and directionality and the willingness to involve a wide array of stakeholders including citizens for improving the legitimacy of the policy interventions under consideration.

# 3. PRACTICES OF CITIZEN ENGAGEMENT IN VISION-SETTING

## 3.1 ABSENCE OF MECHANISM TO INVOLVE CITIZENS IN VISION-SETTING

A substantial share of the mission-oriented R&I initiatives analysed do not present any evidence of citizen involvement in the decision on their visions. These initiatives present some characteristics that may explain a perceived reluctance in engaging citizens.

The first explanation of a lack of public participation mechanisms is the high technological component of the missions. In these cases, decision-makers do not adopt an open approach either because they consider the contribution of citizens as hampering the fulfilment of their ultimate objectives or because they deem it irrelevant. To set such visions, there is no perceived need to involve other actors than knowledgeable experts or stakeholders in the relevant domains, e.g. governmental institutions, industrial representatives, technicians or scientists. No actions are therefore needed to inform citizens or involve them in decision-making, as, due to their low or lack of skills, they will not be able to make any valuable contribution. An example of such an approach is the development of the E-Fan electric aircraft. This technology-driven initiative, initiated and implemented by a consortium of private organisations, was conceived in a pure top-down manner in consultation with a few partners from several national and regional institutions, but without input from the general public, neither any active communication campaign nor participative mechanisms. In the view of the initiators of this initiative, the expertise and the feedback provided by experts and stakeholders in the aerospace and air transport sectors were sufficient to orient research efforts towards the development of a zero-emission aircraft. No clear evidence of actions aiming at increasing legitimacy of this endeavour among the general public has been found.

The absence of citizen engagement in vision-setting may also be explained by the qualities of their initiators. Desk research brings evidence that private (profit-oriented) companies[6] are less keen on involving citizens than public bodies pursuing societal goals. Given that businesses do not have a mandate (or the presumption) to deliver solutions to meet societal demands expressed by citizens, they do not feel accountable vis-à-vis the general public. For this reason, they tend to determine their missions in full autonomy without comments from citizens, preferring to rely on stakeholders with a recognised expertise and/or potential users. This approach is reflected, for instance, in the development of the pan-European aircraft manufacturer, Airbus. The governments that contributed to its inception did not involve citizens in setting up the objectives of this mostly technology-driven and commercial endeavour. The relevance and durability of this commitment to building a European aerospace consortium able to compete with US counterparts was not subject to either public consultations or an active communication strategy aimed at increasing its legitimacy.

Some public authorities share with private companies a strong reluctance in involving citizens in the development of policy interventions. Even in consolidated and longstanding democracies, both the government and the general public might indeed accept that a distance is maintained between them. This situation can be observed, for instance, in democratic regimes whose culture focuses more on the social role of communities than that of individuals (e.g. Japan) and/or whose policy-making practices are still heavily influenced by the legacy of preceding authoritarian regimes. The absence of citizen involvement in the e-Estonia initiative can be interpreted along these lines. Although this initiative, launched few years after the dissolution of the Soviet Union, aims to transform public services through their digitalisation and to enable more open policy-making mechanisms, the legacy of the Communist regime led the institutions to lack the necessary habit to engage citizens in the definition of these objectives and the overall vision of R&I policies.

The third explanation for the lack of citizen engagement is the fact that the public bodies that initiate the concerned mission-oriented R&I initiatives do not operate on the basis of democratic rules and do not therefore feel the need to be responsive to public needs and demands. This is the case of authoritarian regimes, where decision-makers are not democratically accountable. Here, the mission-oriented R&I initiatives prompted and managed by the Chinese government (a single-party authoritarian regime) illustrate such a top-down and authoritarian approach to vision setting. The (technology-driven) Work Station under Deep Sea project, the Five-Year Plans for Solar Energy and New Energy Vehicles are initiatives ruled by the Central Government without any evidence of efforts to build or strengthen the legitimacy of their overall objectives. Citizens were considered at most as potential consumers to be encouraged, via dedicated dissemination activities, to purchase the innovative solutions.

In sum, the category of missions for which vision and objectives have been set without foreseeing any mechanism for citizen involvement is particularly varied and heterogeneous. Evidence has nevertheless been found that there are several cumulative reasons to limit the communication and consultation processes to experts only, while explicitly excluding citizens. These top-down approaches in vision-setting are justified by the fact that the decision-makers do not seek to legitimise the targeted problems (output legitimacy) and because of the way in which these problems were selected (input legitimacy). These missions are defined by decision-makers who are not, or do not feel, accountable to public needs, or consider only scientific and technological dimensions, whose relevance may be estimated without the participation of citizens. Input legitimacy is neglected because of cultural factors and the non-democratic nature of the decision-makers.

## 3.2. INFORMATION SHARING TO STIMULATE BUY-IN

Most case studies gathered evidence of actions undertaken to inform relevant stakeholders and the public at large on individual mission-oriented R&I initiatives. Communication consists of the activities conducted to diffuse information on the rationales, implementation modalities and impacts of these initiatives. It pursues two interlinked objectives: to raise the public awareness of problems and to demonstrate the relevance of the (policy) interventions thereby stimulating buy-in. In such circumstances, citizens are passively involved to ensure (ex post) the legitimacy of decisions – already taken – on the missions to be pursued. The most commonly employed communication channels are dedicated websites, events (including conferences and workshops), social media, and education programmes.

The diffusion of information on policy interventions is a practice observed in various initiatives. However, three groups of mission-oriented R&I initiatives can be identified among those that involve citizens only through information-sharing activities.

The first category of mission-oriented R&I initiatives involving citizens exclusively through communication activities includes initiatives, whose mission is mainly of a technological or scientific nature, but with a political dimension. The legitimacy-building process that is perceived required in these instances is interpreted as bidirectional. On the one side, visions are considered legitimate because they were decided by legitimate decision-makers. On the other side, their popular support enhances, in return, the legitimacy of their decision-makers. Two particularly illustrative case studies are the US Apollo project with its goal of landing a man on the moon and returning him safely to the earth, and the British-French Concorde project to develop supersonic air transport. Despite their objectives to accelerate the development of new technologies, both were given high visibility in media (e.g. the live broadcast of Neil Armstrong from the Moon) and policy discourse. Besides its strategic importance in the then geopolitical context, the Apollo project had a high propaganda value and was aimed at demonstrating to the US citizens the national scientific and technological leadership. Reflecting its political dimension, President Kennedy asked for exploring different options for amendment when criticisms raised. Similarly, in the Concorde project, despite early reservations that investments would have zero or low return, the British government maintained its commitment, mainly for political reasons, among which avoiding the further reduction of the popular support for the government.

The second category of mission-oriented R&I initiatives in which citizen engagement is limited to communication activities and information sharing includes those that consider the perspective of citizens in the vision-setting process, to the extent that the missions relate to problems for which a consensus among citizens is assumed to exist. This approach is observed in situations of a shared sense of emergency in the aftermath of catastrophic events. Citizens are perceived highly likely to consider legitimate any policy interventions designed explicitly to solve a problem that affects most of them and may threaten their safety and/or well-being. Communication activities are conducted in order to maintain or even strengthen this initial level of legitimacy (guaranteed by the sense of emergency) to forestall any later loss of popular support and societal acceptance. An example can be found in the initiative taken by the Italian authorities to protect the Venetian lagoon, which is regularly exposed to exceptional tides (the so-called *'acqua alta'*) and floods with frequency and intensity increasing in recent years[7]. The MOSE project (Italian: *MOdulo Sperimentale Elettromeccanico*) aimed at the development of technologies for the protection of Venice and other cities in the lagoon from floods and other exceptional tides without affecting the economic activities of the commercial harbours. This initiative presents no evidence of any sort of active engagement of citizens in the decision on its vision. It was assumed that the mission would be considered legitimate by the inhabitants of the lagoon because of their vulnerability to this type of natural disaster.

The third category of mission-oriented R&I initiatives involving citizens in the selection of missions only via information-sharing actions consists of initiatives conducted by private foundations. As decision-makers are not elected officials in these cases, it could be expected that low efforts would be made to strengthen the legitimacy of decisions on the missions to pursue. However, some initiatives appeared to be very active in communicating to the large public, as their level of legitimacy has a direct impact on the success of their implementation. Indeed, they may partly rely on private donations (from individuals) and crowdfunding as well as volunteer work, and therefore on their capacity to convince citizens that they can contribute to legitimate missions (linked to the common good) through their financial and in-kind support. Making private and close connections between individuals' concerns and the problem to be solved and, by doing so, building or strengthening (ex post) the legitimacy of the concerned mission-oriented R&I initiative are assumed to be among the most effective ways to steer citizens in that direction. For instance, the Ocean Cleanup initiative aims at preventing, extracting and intercepting the plastic pollution of oceans. Its wide visibility is ensured by large media coverage as well as its active presence on social media. Information on the causes of plastic pollution is disseminated via the website of the Foundation, which is highly dependent on individual contributions. This strategy contributes to raising public awareness of this problem and the (urgent) need to tackle it.

The legitimacy-building processes within these three categories of mission-oriented R&I initiatives show similar patterns. Citizens are involved in building the output legitimacy of these mission-oriented R&I initiatives, as information shared with them aim to demonstrate the accuracy and magnitude of the problems to be solved. However, citizen engagement is not seemingly perceived as a relevant criterion in input legitimacy: decisions are made by legitimate decision-makers acting for the purpose of the national pride and global leadership or by private foundations, whose legitimacy does not derive from any popular election. Some missions relate to urgent societal needs that require a swift reaction of decision-makers. In such circumstances, it would be hardly feasible to mobilise citizens and consult them quickly enough. The definition of the vision is therefore made in a top-down manner and is justified afterwards through communication activities.

## 3.3 PARTICIPATORY INVOLVEMENT

Very few of the missions analysed include evidence of some degrees of citizen participation in the selection of the missions to be pursued. The most common way is through public consultations, whereby decision-makers ask citizens about their views on broad challenges or problems prior to designing policy interventions. This process is controlled and coordinated by the relevant public authorities, which decide on the issues on which citizens provide their feedback and on the procedures for this purpose. Even though public consultation engages more than information sharing, it cannot be yet considered as a genuinely active participation in the policy-making process, especially when the general public is involved late in the decision-making process and is asked to reflect on proposals already developed by policy-makers.

In the mission-oriented R&I initiatives considered, public consultations are organised in different manners, which affect the degree of citizen involvement and the level of openness of the policy-making process to externals inputs. Some initiatives remain open to (spontaneous) feedback and contributions from a wide range of actors, including individual citizens as well as representatives of industry, civil organisations or academia. For instance, feedback collection mechanisms, via public surveys, consultation webpages and interactions with public administrations through social media accounts, contributed to setting the objectives of the Finnish Bioeconomy Strategy. Open consultation tends nevertheless to favour the most vocal individuals, who usually have the knowledge to effectively engage in the decision-making process and the ambition to put their problems onto the policy agenda. Public consultations can be also organised by means of interviews and working groups composed of specific actors (including citizens and civil organisations) who have been identified by the decision-makers. For instance, citizens and other stakeholders were involved in Societal Advisory Boards for providing advices on relevant vision and objectives for the Dutch water management strategy to be implemented via the Delta Programme. Whereas open consultation may result in very low or no contribution, interviews and working groups can guarantee some bottom-up contribution to policy-making. Furthermore, the competent public authorities may define criteria for the selection of their participants, such that representativeness is guaranteed and that the voice of citizens is not captured by groups pursuing their own agenda.

Consultation mechanisms were set in mission-oriented R&I initiatives which are diverse in terms of the nature of the missions (either very scientific and technologic or rather transformative and societal), the type of challenges tentatively addressed (food, agriculture, bio-economy, environment, transport, health, etc.) and the level of policy intervention (supranational, national, local). Nevertheless, three groups of mission-oriented R&I initiatives in which citizen engagement is perceived particularly relevant can be identified.

Firstly, practices of vision-setting engaging citizens appear to be particularly relevant in initiatives aimed at solving important societal challenges, as the general public is assumed to be the best placed to identify and characterise the most relevant problems in this respect. For instance, the Clean Air London initiative, which aims at reducing air pollution in the city and at improving thereby well-being and quality of life of its inhabitants, set up mechanisms to collect feedback and suggestions from Londoners and to enable them to interact with the local administrations. Nevertheless, these mechanisms include surveys which were launched late in the decision-making process and aimed more at ensuring popular support for the initiative than at empowering citizens and fully engaging them in setting its vision.

Public consultation is also employed when the vision and objectives are established in the first place by the competent policy-makers, and are subsequently submitted to a panel of actors, including citizens and other types of stakeholders, for their refinement. For instance, to set the objectives of the Luxembourg 3rd Industry Revolution strategy, the initiators set up thematic working groups composed of more than 300 stakeholders, including companies, local administrations and civil society organisations. Their mission was to translate the concept of 'Third Industrial Revolution' into the Luxembourg context: this mainly refers to a transition of economic systems towards peer-to-peer models relying on the generation, distribution and use of renewable energy. The working groups identified and analysed ongoing trends and defined accordingly feasible and consistent objectives. Luxembourg residents were thereby given the opportunity to take part in the design of this transformative initiative.

Finally, where citizens can autonomously organise to trigger a specific policy which responds to a pressuring societal demand, traditional policy-makers might employ participatory schemes to regain control over the policy-making process and handle it for their own benefits. The collected evidence suggests that this situation may occur only in democratic communities whose social actors possess a fair awareness of the functioning of the policy-making process and demonstrate the capacity to prompt the introduction of societal demands in the public policy agenda. Provided that citizen movements are not seen as endangering the current balance of power or hampering the stability of the community, the traditional policy-makers might consider them as opportunities to (further) legitimise their actions. Consequently, also in this particular case, the decision-makers might attribute to the citizens the capacity to participate in setting the vision for mission-oriented R&I initiatives. An example of this case is the German Energiewende initiative, which consists in a long-term strategy for the development of a low-carbon energy system based on renewable energy and energy efficiency. The development of this mission would have not been possible without the long-standing activism of grassroots green movements advocating for energy transition policies and the phasing out of nuclear power plants. At first, policymakers underestimated the sense of urgency felt by the German citizens in regard to a green transition over decades, until the Fukushima nuclear accident convinced them to steer this long-date public movement in their favour. Participatory schemes were then not simply strengthened and institutionalised, but even encouraged and multiplied. Such shift in policy produced one of the most emblematic examples of citizen involvement in setting the goals and in designing the policy instruments of an mission-oriented initiative.

In regard to the level of legitimacy, the engagement of citizens in setting the vision of mission-oriented R&I initiatives via dedicated public consultations and other participatory mechanisms ensures first and foremost their output legitimacy. Furthermore, it demonstrates that traditional policy-makers are increasingly – but still marginally – attentive to input legitimacy. Citizen engagement in the decision-making process is interpreted as being not aimed only at ensuring that the ultimate decision will be in line with societal expectations and will have popular support. It can also contribute to building a policy discourse justifying the pursued missions by referring explicitly to the public participation mechanisms employed for their definition. The analysis of the considered case studies suggests that public consultations organised at the earliest stages in the decision-making process are nevertheless those that are the closest to genuine participation of citizens. The general public may be asked to contribute directly to the refinement of missions that were broadly defined by policy-makers, or to express their support to (rather elaborated) proposals of missions and objectives. However, in most cases, citizens are not asked to define, through any type of participatory process, which missions would be the most relevant in their views. Citizens are rather expected to give feedback on top-down defined proposals. Finally, in few instances, policy-makers may decide to translate problems already identified by grass-roots movements into well-defined policy interventions. By putting these problems onto the policy agenda, traditional policy-makers similarly ensure that the mission-oriented R&I initiatives prompted by autonomously organised groups of individuals are considered legitimate.

| Type and means of citizen engagement | Main observed features of missions | Rationale for the selected degree of citizen involvement | Consideration for legitimacy-building | |
|---|---|---|---|---|
| | | | Input | Output |
| **NO ROLE**<br>No instrument | Missions<br>• with high technological component,<br>• whose initiators lack culture and practices of citizen engagement,<br>• implemented in non-democratic regimes. | • Lack of responsiveness of policy-makers to public needs and demands.<br>• No valuable contributions expected from citizens.<br>• High reliance on experts, users and industry stakeholders.<br>• Fear of disruptive effects of citizen engagement on policy-making. | NO | NO |
| **COMMUNICATION**<br>Exclusive reliance on communication tools:<br>Dedicated websites<br>Events (conferences, workshops)<br>Social media<br>Education programmes | Missions<br>• related to societal challenges but initiated by private actors,<br>• with a high technological component and a major political dimension,<br>• related to urgent societal needs. | • Influence of the degree of popular support on the successful implementation of missions.<br>• Need to maintain the initial level of legitimacy throughout the policy cycle.<br>• Intertwined legitimacy of policies and their initiators.<br>• Trade-off between swift policy reactions to urgent needs and openness of policy-making to public participation. | NO | YES |
| **PARTICIPATION**<br>Public consultations<br>Interviews<br>Working groups<br>Surveys<br>Social media<br>Public Meetings<br>Stakeholder Forums | Missions<br>• with anticipated important societal impacts,<br>• broadly predefined and in need of refinement,<br>• defined by citizens and whose translation into the policy agenda serves the actions of the policy-makers. | • Influence of the degree of popular support on the successful implementation of missions.<br>• Perception of citizens as knowledgeable and capable of participating in policy-making.<br>• Democratic and transparent functioning of public administrations.<br>• Need to develop policy discourse justifying policy interventions. | YES | YES |

**Table 1:** Elements of citizen involvement and consequence on legitimacy

# 4. PERCEIVED CHALLENGES IN THE ENGAGEMENT OF CITIZENS IN VISION-SETTING

Interviews and a workshop collected the views from R&I actors and help flesh out the reasons of the resistance of policy-makers against further public participation in policy-making, and the modalities in which participation might take place.

## 4.1 WHY SHOULD POLICY-MAKERS INVOLVE CITIZENS?

It is widely admitted that all individuals should be given the opportunity to reflect on the relevance of policy interventions implemented in their community, and particularly on the direction given to mission-oriented R&I initiatives. Such positive opinions towards citizen involvement are mainly underpinned by the assumptions that private organisations and the traditional policy-makers may overlook societal needs while pursuing ambitions often related instead to their own needs; and that citizens have a better understanding than these actors about the most significant challenges to their communities. Furthermore, public participation in policy-making can increase the stability and the legitimacy of policy decisions and the level of transparency of decision-making processes. It may also contribute to making citizens feel responsible for the formulation and design of policies, while curbing the eroding of their trust in democratic systems. Finally, the consulted stakeholders advocating for public participation in decision-making refer to past experiences (e.g. the Irish Citizens' Assembly set up in 2016 to consider the most important issues of the country) to show that, if individuals are trained to effectively take part in participatory mechanisms and informed on the characteristics of the societal challenges, they might demonstrate abilities to grasp complex issues and make relevant policy proposals.

On the contrary, other consulted stakeholders express some reserve in respect to the engagement of citizens in decision-making, particularly in its initial steps, contending that it should be avoided, if not forbidden, and that other types of stakeholders with a better understanding of the challenges to be addressed (including experts and users) have more valuable inputs to bring in the policy-making process. In addition, citizens do not have the mandate nor the capacity to fully understand, represent and elaborate on the societal demands of their community. Because mechanisms for public participation may shed light on discrepancies between citizens and other stakeholders in terms of degree of participation in policy-making, they can also affect the community cohesion. Furthermore, citizens might not be aware of the complexity of the policy cycle to make relevant contributions, especially in countries or regions where mainstream media have a strong influence on their level of understanding of ongoing policy discussions. Social media similarly raise concerns in this respect, as they allow users to favour some information sources while outweighing others for any apparent sensible reason.

In the context of growing populism across Europe, consulted stakeholders worry that mechanisms for public participation in vision-setting divert policy decisions from the missions that are the most likely to accelerate system transformations for solving societal challenges. Because those missions may have disruptive (and negative) effects in the short term, citizens may be reluctant to fully support them. Moreover, the participation of citizens in the decision on missions is feared to benefit only easily understandable or 'fashionable' sciences or industries, i.e. those benefitting from high popularity and media coverage.

Finally, some actors are concerned that citizen involvement would be infeasible on practical terms. Mechanisms which are conceived to involve on a regular basis multitude of individuals are costly and their funding too difficult to be viable in the long run. An extra layer of complexity is added by the extreme variety of opinions on a number of issues which makes the finding of a consensus even more difficult.

## 4.2 WHICH KIND OF INDIVIDUALS (OR GROUPS OF INDIVIDUALS) SHOULD BE INVOLVED?

Citizens are perceived in competition with other categories of stakeholders in the policy-making process, especially with experts, users and civil society organisations.

In comparison with citizens, experts encompass all individuals with relevant and recognised knowledge in relation to the challenges to be tackled as well as the expertise and experience necessary for suggesting suitable, relevant and feasible answers to the identified problems[8]. Furthermore, experts may be assumed to be more likely to understand the specificities of the policy-making process. The expertise of these individuals and organisations can serve in crucial steps, including the definitions of the scope of the missions and of quantifiable and attainable objectives. For instance, in the Indian Electric Mobility Plan, the government decided to leverage on the participation of transport and automobile stakeholders which have specific interests in the development of the electric vehicle industry.

Some mission-oriented R&I initiatives attempted to use the contributions from both experts and the general public in a complementary way. For instance, in several of the analysed case studies, the feedback of citizens was used to reflect on the directions identified, in a first time, by individual experts. The EU Human Brain Project adopted this kind of approach: at first, scientists and industry representatives were asked to propose a set of projects; only in a second moment, citizens were engaged to verify the socio-economic and ethical dimensions of the selected proposals and validate the goals of the proposed projects.

What is mostly argued is nevertheless the rationales for involving citizens instead of stakeholders or experts, and whether the decision-makers should listen to the general public without mediation by any intermediate body and the support of skilled professionals. In the views of governmental agencies and the scientific community, individuals may have some understanding of the societal challenges, but clearly lack the knowledge required to solve them. On the contrary, some not-for-profit organisations argue that citizens, in comparison with experts, have a holistic vision much more focused on the future conditions of the next generations than on the scientific and technological challenges.

Citizens are also often confused and substituted (in the policy-making process) with users. However, while the former are defined in respect to their belonging to specific (social) communities, the essential feature of the latter is being economic and (demand-side) market actors integrating goods and services into their economic activities (by consuming or employing them) in order to obtain some benefits, including the solving of specific problems. Users are considered to hold specific knowledge that relates to their practices and habits and that allows to determine the feasibility of the orientation of R&I initiatives (especially when they are mainly aimed at the achievement of ambitious scientific, technological or economic challenges) and whether new goods and services are ready to be used. Therefore, the main rationale for user involvement in policy-making is to ensure that the vision is reachable and that the solution to be developed, for that purpose, could be used and diffused at a sufficient pace. Without diminishing this argument, civil society organisations claim that it cannot justify that users are involved in the place of citizens to decide the visions for mission-oriented R&I initiatives.

The particular role of the general public is indeed to consider the relevance of missions against societal values that are deemed of importance in their community. For these reasons, communication and dissemination activities should be clearly delineated, with the latter aimed at accelerating the uptake of (new) goods and services and not at raising public awareness of the targeted problems. This is the reason why the large events such as those organised in the Chinese New Energy Vehicle initiative for demonstrating newly developed electric vehicles are seen as targeting potential purchasers instead of involving the public at large, and having no influence on the legitimacy of mission-oriented R&I initiatives.

Given that all individuals are inevitably part an established community and jointly constitute the civil society, all the organisations that are made of individual volunteers have in common the aspiration of representing "a wide range of interests and ties" (OECD, 2006). However, the participation of civil society organisations in policy-making occurs differently than citizen involvement. In the first place, civil society organisations tend to be organised in a complex structure, with allocated responsibilities to group of individuals who are the real interlocutors vis-à-vis decision-makers. Secondly, they have their own communication channels and independently implement their outreach strategies. In the third place, no civil society organisation can claim to represent all the positions of their members and volunteers. They express official positions, which may be the result of internal mediation and compromise.

Moreover, there is a clear tendency in each organisation to focus on a theme or a set of issues, and to unavoidably advocate for the specific interests of limited groups of citizens. For the same reason, their involvement may lead to an excessive politicisation and polarisation of the policy-making process. Despite these concerns, the evidence collected in case studies, such as in the French Agriculture-Innovation 2025 strategy, suggests that some policy-makers prompt the participation of civil society organisations, like family associations, when in need of enlarging their legitimacy.

## 4.3 WHICH MODALITIES OF CITIZEN INVOLVEMENT?

Once citizens are recognized having a specific role in the design of mission-oriented R&I, policy-making procedures must adapt to allow their involvement. One of the few points on which most consulted stakeholders agree is that a fully bottom-up process is neither feasible nor advisable. Traditional policy-makers should not be excluded from vision-setting and the design of mission-oriented R&I initiatives, as they are the most capable actors to ensure policy coordination. The governance of mission-oriented R&I initiatives should instead rely on a multi-actors model where participants tend to complement, but do not substitute, each other.

Very few consulted stakeholders support the idea to create a dedicated and permanent body in charge of ensuring that citizens are involved in the general policy-making process. It is instead contended that existing institutions would be better positioned to translate popular preferences for specific challenges into concrete mission-oriented R&I initiatives. However, they would need to change their functioning by integrating citizens to their decision-making process. A solution is the establishment of multi-stakeholders groups with the mission to formulate recommendations to bodies in charge of taking decisions, such as the National Cancer Advisory Board (NCAB), which was established in 1971 to ensure the implementation of the War on Cancer initiative, and which contributed to the research plan for the Cancer Moonshot strategy. Public participation in decisions on the visions for mission-oriented R&I should rely, where possible, on existing practices of citizen engagement in the functioning of public institutions, in order to avoid having disruptive effects and inducing too high costs to public administrations.

Special attention should be paid to the criteria for the selection of the citizens authorised to participate in the vision-setting for mission-oriented R&I initiatives. A general principle should be that the selection should be balanced and avoid, where possible, organisations which may capture the voice of the general public to push forward their own agenda. However, institutional arrangements and adaptation of consultation tools for citizen engagement in decision-making will induce costs, as many consulted stakeholders highlighted. This is even more expected in the case of mission-oriented R&I orchestrated at the EU level, as such practices do not exist yet and may face linguistic and cultural barriers.

# 5. CONCLUSIONS

This research collects evidence and demonstrates that practices of citizen engagement in setting the vision for mission-oriented R&I initiatives are barely developed. Even though Horizon Europe aims to give them a renewed impetus, their diffusion might be hindered by resistance observed among a large range of R&I stakeholders. Therefore, this research highlights the need to promote new practices within the policy-making process in order to promote, facilitate and ensure the engagement of citizens in the decision on the missions and thereby improve both their output and input legitimacy, the latter being still rarely considered.

Citizen engagement in policy-making may contribute to solving the disenchantment many citizens currently perceive with the EU institutions. Mechanisms for public participation in decision-making are not aimed at replacing representative democracy, but instead at complementing it. They rely on the observations that the citizens participating in the policy-making process might feel more “committed” and become able to make well-grounded priority-setting. If involved in the policy-making at an early stage, they can also improve their understanding of how public institutions work. Ultimately, the citizens’ renewed feeling of responsibility for, and commitment to, the general interest of their community may reduce the distrust against representative democracy that jeopardizes the stability of institutions.

All these arguments are particularly valid for the EU institutions, as they are seemingly the most affected by the growing scepticism about the course of the traditional policy-making process. Eurosceptic feelings have strengthened and gained ground at a high pace over the past ten years in the founding and in the newer Member States, while the UK voters have voted to leave the European Union under the influence of nationalistic propaganda. In such an alarming context, the EU policy-makers must curb the perceived widening of the gap between EU institutions and the European citizens and to renew the decision-making processes in order to engage the general public in shaping the future of its policy interventions.

Since mission-oriented R&I initiatives are essentially aimed at solving problems that will help tackle pressing societal, economic, scientific and technological challenges, they are easily understandable by the general public and are conveniently communicable and justifiable in the public eyes. If handled correctly, missions conceived with citizen involvement will therefore contribute to increasing the legitimacy of the EU in R&I policy and possibly in other policy domains. Additional efforts for involving the general public in the definition of missions may help further reduce the perceived distance between citizens and the EU institutions. However, this requires the promotion of new practices within the EU policy-making process without lengthening it and increasing its costs (and while abiding by the institutional framework laid down by the Treaties).

Public consultation mechanisms have already demonstrated being valid and important means to ensure output and input legitimacy. The EU policy-makers may rely on these existing practices, enlarge their scope and scale them up at the EU level while overseeing their implementation in Member States. Moreover, further steps towards a more participative decision-process can be designed and implemented on the strengths of past experiences (such as the past Interactive Policy Making mechanism) and in view of the practices currently implemented and for which the European Union is advocating (notably, in the area of e-Governance). By considering the controversy that a suggestion for changing the current decision-making procedures may engender among the Member States and EU bodies, such attempts can be delimited, in a first time, to the field of R&I policies and initiatives, and, more specifically, to the mission-orientated pillar in Horizon Europe. Another possibility to ensure bottom-up participation in decision-making, as noted by several interviewees, are the platforms established for the definition of the Smart Specialisation Strategy[9]. These could be revived and used to involve the general public at a regional level, by giving the opportunity to define missions capable of meeting the needs of local communities. In addition to these options, civil society organisations suggest the establishment of “Citizens Conventions”, whose design and functions may address the challenges that this research identified[10].

The main issues with these tentative mechanisms relate to the substantial lack of experience in handling them in several Member States or at a transnational level. Moreover, even if correctly implemented, they would need to collect the opinions of citizens who are exposed to a wide variety of challenges or to similar ones but at varying intensities (e.g. ocean pollution may be less a concern for Central European countries). Citizen involvement mechanisms would have to cope also with the challenges traditionally faced by all sorts of exercises implemented at the EU scale, i.e. the presence of different political cultures and multilingualism. However, while not dismissing their importance, these challenges could be regarded as the *raison d’être* of an intergovernmental organisation such as the European Union and should not be considered as a hindrance to more participatory and democratic policy-making processes. On the contrary, given that missions aim at solving challenges which often have transnational impacts, the coordination of mission-oriented R&I should be guaranteed at EU level. As a consequence, the institutions of the European Union will be the best placed to find and implement renewed methods and practices to engage citizens in setting the missions to orient R&I policy.

# 6. REFERENCES


**Boedeltje, M. and Cornips, J.** (2004). Input and output legitimacy in interactive governance. Presented at the NIG Annual work conference, Rotterdam.

**Boon, W. and Edler, J.** (2018). Demand, challenges, and innovation. Making sense of new trends in innovation policy. Science and Public Policy, 45(4), 435–447. https://doi.org/10.1093/scipol/scy014.

**Borrás, S. and Edquist, C.** (2013). The choice of innovation policy instruments. Technological Forecasting and Social Change, 80(8), 1513–1522. https://doi.org/10.1016/j.techfore.2013.03.002.

**Bugge, M. M., Coenen, L. and Branstad, A.** (2018). Governing socio-technical change: Orchestrating demand for assisted living in ageing societies. Science and Public Policy. https://doi.org/10.1093/scipol/scy010.

**Bush, V.** (1945). Science The Endless Frontier (A Report to the President). Washington D.C.: United States Government Printing Office.

**Dahl, R. A.** (1998). On Democracy. New Haven: Yale University Press.

**Edquist, C.** (2011). Design of innovation policy through diagnostic analysis: identification of systemic problems (or failures). Industrial and Corporate Change, 20(6), 1725–1753. https://doi.org/10.1093/icc/dtr060.

**Flanagan, K. and Uyarra, E.** (2016). Four dangers in innovation policy studies – and how to avoid them. Industry and Innovation, 23(2), 177–188. https://doi.org/10.1080/13662716.2016.1146126.

**Jagers, S., Matti, S. and Nordblom, K.** (2016). How Policy Legitimacy Affects Policy Support Throughout the Policy Cycle. QoG Working Paper Series, 2016(15).

**JIIP, Joanneum Research, Tecnalia, TNO, VTT and DTI.** (2018a). Mission-oriented research and innovation: Inventory and characterisation of initiatives. Luxembourg: Publications Office of the European Union.

**JIIP, Joanneum Research, Tecnalia, TNO, VTT, DTI and VVA.** (2018b). Mission-Oriented Research and Innovation: Assessing the impact of a mission-oriented research and innovation approach. Luxembourg: Publications Office of the European Union.

**Kingdon, J. W.** (1984). Agendas, alternatives, and public policies. Boston: Longman.

**Laranja, M., Uyarra, E. and Flanagan, K.** (2008). Policies for science, technology and innovation: Translating rationales into regional policies in a multi-level setting. Research Policy, 37(5), 823–835. https://doi.org/10.1016/j.respol.2008.03.006.

**Nelson, R. R.** (2011). The Moon and the Ghetto revisited. Science and Public Policy, 38(9), 681–690. https://doi.org/10.1093/scipol/38.9.681.

**OECD.** (2001). Citizens as partners: information, consultation and public participation in policy-making. Paris: OECD Publishing.

**OECD.** (2006). Applying Strategic Environmental Assessment: Good Practice Guidance for Development Co-operation (DAC Guidelines and Reference Series). Paris: OECD Publishing.

**Scharpf, F. W.** (2006). Problem Solving Effectiveness and Democratic Accountability in the EU. Reihe Politikwissenschaft, (107), 43.

**Vigoda, E.** (2002). From Responsiveness to Collaboration: Governance, Citizens, and the Next Generation of Public Administration. Public Administration Review, 62(5), 527–540. https://doi.org/10.1111/1540-6210.00235.

**Wallner, J.** (2008). Legitimacy and Public Policy: Seeing Beyond Effectiveness, Efficiency, and Performance. Policy Studies Journal, 36(3), 421–443. https://doi.org/10.1111/j.1541-0072.2008.00275.x.

**Weber, K. M., and Rohracher, H.** (2012). Legitimizing research, technology and innovation policies for transformative change: Combining insights from innovation systems and multi-level perspective in a comprehensive 'failures' framework. Research Policy, 41(6), 1037–1047. https://doi.org/10.1016/j.respol.2011.10.015.


# AUTHORS


**JULIEN CHICOT AND ALBERTO DOMINI**
*The Joint Institute for Innovation Policy (JIIP)*
Boulevard de la Plaine 9, 1050 Brussels, Belgium
E: julien.chicot@jiip.eu
E: alberto.domini@jiip.eu

ORCid number of Julien Chicot: https://orcid.org/0000-0002-8498-956X



## KEYWORDS:

Mission-oriented R&I; Legitimacy; Citizen involvement


Endnotes

1 Brazil, China, Japan, Norway, Russia, Singapore, South Korea, Switzerland and the United States.

2 Even though this sample was aimed at being balanced, it does not have the ambition to be representative. In consequence, the outcomes of the case studies cannot be subject to statistical generalisation.

3 These case studies can be consulted on the online JIIP Global Observatory of Mission-Oriented R&I: www.jiip.eu/mop

4 In total, 40 organisations were interviewed. They were asked generic questions in relation to the characteristics and potential impacts of mission-oriented R&I initiatives.

5 The preliminary findings of both studies were presented to 20 experts and relevant stakeholders during a workshop co-organised with the European Commission in February 2018. A session was dedicated to the engagement of citizens in mission-oriented R&I initiatives.

Endnotes

6 In contrast with companies, private foundations may pursue missions that are often related to common goods and societal challenges. They may therefore be willing to involve, to some degrees, citizens in the legitimacy-building of their mission-oriented R&I initiatives.

7 In November 1966, a tide of 194 centimetres above the sea level submerged Venice and its surroundings and dramatically raised concerns for the safety of the inhabitants. In its aftermath, the first Special Law for Venice of 16th April 1973 (Law 171/73) declared the problem of safeguarding the city and its lagoon to be of "priority national interest".

8 This definition of experts includes organisations that can provide technical and managerial knowledge necessary for the successful development and diffusion of the solutions to the targeted problems, i.e. actors from industry (e.g. suppliers) as well as public administrations with a recognised expertise in the fields where they operate. The particularity of these actors (unlike independent experts) is that they may have a less neutral position on the problems to be tackled and therefore on the orientations of R&I policies and initiatives in their sector.

9 Smart specialization is a place-based approach, meaning that it builds on the assets and resources available to regions and Member States and on their specific socio-economic challenges in order to identify unique opportunities for development and growth.

10 The 'Citizens Conventions', as defined by Global Health Advocates and Sciences Citoyennes, will be composed of citizens randomly selected. For ensuring their accurate understanding of the challenges that they need either to prioritise or to translate into concrete missions, a balanced set of stakeholders with various (and preferably diverging) interests will explain to them their views, while experts will provide basic information on the underpinning scientific and technical aspects. The selected individuals will complete their training by requesting the intervention of organisations, which they deemed of interest to listen to. They will subsequently debate and decide among themselves about the most relevant (and therefore legitimate) missions to be pursued.

# IMPACT PATHWAYS: TRACKING AND COMMUNICATING THE IMPACT OF THE EUROPEAN FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION[1]

NELLY BRUNO AND MARTINA KADUNC

DOI: 10.22163/fteval.2019.330


## ABSTRACT

Since 1984, the EU investments in the successive Framework Programmes contributed to key scientific advancements and discoveries for the benefits of society and the economy. These impacts have been documented in multiple evaluation exercises and dedicated studies but still such assessments face common methodological challenges and limitations. A major difficulty is to identify and capture the direct and indirect effects that can be attributed to these risky investments in complex and open research and innovation systems over a long timeframe. For the post-2020 Programme, Horizon Europe, the European Commission proposed a revamped indicator framework built around a set of Key Impact Pathways. The paper shows how this new approach was developed and how it is expected to improve the monitoring and evaluation of the Framework Programme based on the latest technological developments.


## 1. INTRODUCTION

In June 2018 the European Commission adopted a proposal for Horizon Europe, the ninth European Framework Programme for research and innovation (R&I), with a proposed budget of nearly EUR 100 billion over 2021-2027 (European Commission, 2018b). Building on more than thirty years of European Framework Programmes, Horizon Europe is expected to strengthen the scientific and technological bases of the Union and foster its competitiveness, deliver on the Union strategic priorities and contribute to tackling global challenges, including the Sustainable Development Goals.

In a context of government austerity measures coupled with growing economic and social pressures, demonstrating and communicating the diversity of impacts and the European added value of R&I investments is crucial for the purpose of accountability, advocacy and learning. However, capturing these impacts is not straightforward and requires to deal with complexity. The questions of attribution/contribution, time-lags, and uncertainty/risk are among the key challenges faced for the evaluation of R&I investments worldwide.

The ex-ante development of an appropriate indicator system, based on the programme-theory approach and a reinforced use of latest technological advances can alleviate some of the difficulties faced. Such system would allow for more informed evaluations without further administrative burden, thereby reconciling methodological challenges and policy needs. In particular this paper sheds light on the rationale and the principles behind the development of the proposed revamped indicator system for Horizon Europe to track the progress of the Programme towards its objectives at any moment in time, along a set of Key Impact Pathways.

## 2. THE CHALLENGE OF CAPTURING THE IMPACT OF R&I INVESTMENTS

### 2.1 OVERVIEW

The EU Budget Focused on Results initiative was started in 2015 to join efforts of EU institutions, governments and civil society towards better spending, increased accountability and transparency, and maximum added value for EU citizens. This focus was further reinforced through the Better Regulation Guidelines (European Commission, 2017a), which cover the whole European policy cycle – including ex-ante impact assessment, monitoring and evaluation. In this context, programme evaluations are instrumental to assess the actual performance of the programme compared to initial expectations, in addition to helping improve its management and functioning. Evaluations are expected to go beyond an assessment of *what* has happened, and consider *why* something has occurred and, if possible, *how much* has changed as a consequence (i.e. quantification of change). In particular, evaluations have to look for evidence of causality – i.e. did the intervention (help) bring about the expected changes, and were there other unintended or unexpected changes?

However, evaluations are commonly confronted with a set of methodological challenges which are particularly strong when assessing R&I

1 For the development of the proposed Key Impact Pathways for Horizon Europe authors would like to acknowledge the contribution of many colleagues from the European Commission and the independent expertise of Peter van den Besselaar, Jose Ramon Flecha Garcia and Alfred Radauer.

policies and programmes, notably because of the nature of knowledge generation and its diffusion processes. In particular:

- **The time lag issue:** even if funding of very close to market activities can produce results within the timeframe of the R&I support programme, most R&I activities will generate impacts only in the very long-term. Twenty to thirty years may be required to be able to grasp the full spectrum of impacts from R&I investments (ICT applications are usually closer to the market, while drug development can take 15 years or more, see for example JIIP, 2016). A key issue for evaluators is thus to decide when to realistically capture the impacts of the programme.
- **Uncertainty and risk:** per definition, many R&I projects will fail. Innovation is the work of humans, it can never be predicted (Irvine and Martin, 1989). Some low risk programmes may have many incremental and short term effects whereas high risk programmes may have fewer but potentially more radical effects in the longer term. Comparing the two in the medium term would always favour the low risk programme and therefore lead to a certain risk averseness of public action, whereas the 'market failure' justification assumes that government acts when risks are too high for the private sector (Guellec, 1999). A key issue for evaluators is thus to decide how best to capture the impacts of the programme while acknowledging the need for trial and error in the R&I process.
- **The attribution/contribution problem:** Scientific progress builds on knowledge that cumulates over decades and spreads widely and unexpectedly into multiple domains and applications, as Issaac Newton put it 'standing on the shoulders of a giant'. Because of its inexhaustible nature and of the fact that it does not deplete when used unlike most resources, the positive spillovers of knowledge are not limited (Foray, 2000). Beyond the project funding, also other projects and factors influence positively or negatively R&I activities of programme's beneficiaries and the diffusion and uptake of the R&I results. Organisations are indeed not innovating in isolation but in the context of a system (Freeman and Lundvall, 1988; Lundvall, 1992; Nelson, 1993; Barré et al., 1997). A key issue is thus to decide how much 'credit' the programme should have for changes that occur after it is launched.

In addition, R&I policies are generally regarded as complex to evaluate because of the need to deal with multiple objectives (including solving societal challenges), implementation modalities, targets, instruments and target groups; evolving framework conditions; trial/errors processes and feedback loops.

Because of these challenges and the complexity of R&I processes, there is no gold standard in the methodologies and indicators to be used for the evaluation of R&I programmes. Typically evaluations are based on the intervention logic of the initiative and rely on the triangulation of quantitative and qualitative information from multiple sources, including surveys, interviews, case studies, expert groups, descriptive statistics, econometric analysis. In this context indicators should ideally cover the various sequences of a policy intervention. In the case of an R&I programme this translates into indicators on *inputs and activities* (*programme management data* on financial and human resources and the implementation of activities) that are expected to lead to *outputs* (such as reports, trained researchers, or new infrastructures), *results* (benefits for direct beneficiaries from their participation) and *impacts* which are the wider effects, i.e. spillovers or externalities beyond the direct beneficiaries, in particular for scientific progress, the economy and society. Collecting and monitoring *programme management data* is relatively straightforward. The challenge lies in devising an appropriate indicator systems that allows capturing the *outputs*, *results* and *impacts* over time, while minimizing the problem of attribution/contribution, time-lag and uncertainty.

## 2.2 LEARNING FROM OTHERS – LESSONS FROM R&I PROGRAMMES AROUND THE WORLD

Many monitoring and evaluation frameworks have been developed worldwide to demonstrate how public R&I funding organisations and their activities impact the economy and society. Guthrie et al. (2013) studied 14 such frameworks applied by different funders across the world and observed that the purpose of the framework (i.e. advocacy, accountability, learning, resource allocation) dictates their methodological choices: there is no one-fit-all solution. As a result, frameworks developed for accountability and allocation purposes are not suited for learning and vice versa. The former requires high level of transparency and comparability for which quantitative approaches are best, the later tend to use qualitative methods which are comprehensive and flexible but do not allow comparisons. The majority of current R&I monitoring and evaluation frameworks still mainly aim at accountability and resource allocation (Graham et al., 2018).

Accountability and resource allocation decisions are often based on quantitative approaches but the existing R&I statistics offer little or no information about the *'output'* side of the R&I process. Historically there have been only two established areas of indicators to support such measurement: scientific publications and citations (i.e. bibliometric data) to measure dynamics of science and data on patent applications, awards and citations to measure dynamics of innovation (Smith, 2005). The traditional assumption for R&I investment is that society derives most benefits when research is excellent, i.e. conducted at the highest level. Hence traditionally the only interest when measuring R&I impact was the impact on scientific knowledge and the ability to produce inventions (Bornmann, 2013).

However, in the last twenty years two major developments influenced the way R&I investment is perceived and measured. Firstly, assessing the economic impact of R&I became central, due inter alia to the increased austerity of public funding. As a result, company data on jobs and turnover are now commonly used for economic modelling to estimate the impact on productivity and growth (Ravet J. et al, 2018). Yet, the immediate statistics on innovation *'outputs'* remain narrowly focused on patents applications and do not sufficiently integrate other types of intellectual property rights (IPR) such as trademarks or standards. There are still missing data links to trace innovation outputs and their way to the market. The monitoring system of Business Finland (formerly: TEKES) is a practical example of the current state of the art in terms of R&I indicators to support economic impact measurement (van den Besselaar, Flecha, Radauer, 2018).

Secondly, there is a growing expectation that R&I programmes need to address the needs of society in general. But defining and measuring the societal impact is challenging and there are neither established indicators nor data or methodologies available. Most agencies and models do not consider societal impact at all (van den Besselaar, Flecha, Radauer, 2018). Some considerations are given to the use of alternative metrics ('altmetrics') to measure R&I outreach on social media and policy documents (European Commission, 2018c) but those are criticized as they confuse dissemination of R&I outputs with societal impact (van den Besselaar, Flecha, Radauer, 2018). Alternative approaches have been developed, which focus on indicators *'leading'* to societal impact. The SIAMPI project, for instance, showed that the key factor for societal impact to happen is to ensure that the R&I community and citizens interact with each other (Spaapen et al., 2011). It seems that quantitative indicators alone could never measure societal impact even in the narrowest sense. The proposed way forward is based on qualitative assessments of experts, researchers or citizens (van den Besselaar, Flecha, Radauer, 2018; European Commission, 2018c).

With these changing expectations on what R&I investment needs to deliver, it is now clear that policy-makers and experts look for an intelligent mix of qualitative and quantitative approaches to capture impact (European Commission, 2018c). Yet the underlying data challenge remains the main obstacle to deliver on such expectations (van den Besselaar, Flecha, Radauer, 2018; EC, 2018c). For instance:

- Even if bibliometric and patent data have been used for decades to monitor R&I impact, the inclusion of funder acknowledgments in publication and patent data is not widespread making it difficult to identify links with public R&I investment.
- The databases often exclude information on control groups such as non-successful applicants making it difficult to apply proper counterfactual impact evaluation methods.
- Existing data and databases operate in different silos and are not connected. This situation is changing rapidly, for instance the Star-metrics and Umetrics developments in the United States and the SMS Platform (RISIS project) in the EU are examples where different datasets are linked.

Overall there is not – and will probably never be - a perfect indicator framework for R&I programmes that would provide the required level of coverage, accuracy, simplicity and automation to generate the needed information to trace the diversity of impacts from R&I investments.

## 2.3 LEARNING BY DOING - LESSONS FROM PAST FRAMEWORK PROGRAMMES

Since 1984, the EU investments in the successive Framework Programmes contributed to key scientific advancements and discoveries for the benefits of society and the economy. These impacts have partly been documented in evaluation exercises and dedicated studies. As reported in a dedicated study on the impact of the Framework Programme (EPEC, 2011), these evaluations usually focus on specific parts of the programme or on specific instruments – with their own methodologies - whereas expert panels are typically asked to perform a meta-evaluation of the whole programme based on these inputs (see for instance Davignon E. et al (1997); Stampfer M. (2008); Fresco L. et al (2015)).

Overall Framework Programmes' assessments faced common methodological challenges and limitations. In particular, Framework Programmes lacked clear intervention logics from the design stage along with the appropriate monitoring system. This created a wide data gap in the identification of the contribution of the Framework Programmes to the diversity of impact streams. Many assessments focused on the analysis of output data such as scientific publications and patent applications but often faced difficulties in capturing longer term and wider effects, in particular for society or the economy. This is partly due to the early timing of most evaluations but also to the limited or unreliable data available beyond ad-hoc surveys, interviews or case studies. As indicated in the EPEC analysis (2011), *"the traditional evaluation record {of the Framework Programme} typically tells little about the achievement of high-level (policy) objectives, some things about specific or strategic objectives and quite a lot about operational objectives"*.

### ON ASSESSING THE LONG TERM IMPACTS FROM THE FRAMEWORK PROGRAMMES

In 2011 a specific attempt was made to look at the long-term impact of the Framework Programme through a set of in-depth case studies tracing projects and their contributions back to FP4 (EPEC, 2011). The study pointed to the *"existence of a range of longer term impacts of the Framework Programme that need to be understood in greater depth, in parallel with standard evaluation, in order to explore more policy options and allow the development of policies that are effective over the longer term". The study argued that "this will require continued experimentation and increased diversity in methods: first, because existing methodologies are not always able to address the different impact mechanisms involved in the longer term; and, second, because of the longer time constants involved. The complexity of the Framework Programme means that a single set of methods or a single pan-Framework study will not produce a simple, overall 'answer'. Rather, there is a need to explore the individual impact mechanisms in turn. Only when this has been done can we create a synthetic understanding of the Programme as a whole."*

A 2016 study on the impact of the Framework Programme on Major Innovations concluded that "*due to the complexity of innovation processes, individual projects or even the Framework Programmes by themselves cannot be turned into a systematic pipeline for Major Innovations. Major Innovations are triggered by a multitude of factors, of which the Framework Programme is one part of a bigger puzzle*" (JIIP, 2016).

Whereas the main objective of Framework Programmes' evaluations was to ensure accountability to the Council, the European Parliament and EU citizens, the evaluations became increasingly used also for advocacy and learning purposes. This required the development of an indicator system allowing to track progress along key dimensions of the Programme. For Horizon 2020 (2014-2020), the monitoring system of the Framework Programme thus underwent noticeable improvements. For the first time in the Framework Programme history, a set of Key Performance Indicators (KPI) was introduced and this data was made publicly available in close-to-real time through an interactive online dashboard. In order to report on the progress made towards the objectives for the interim evaluation of Horizon 2020 (European Commission, 2017b) these indicators have been complemented by other ad-hoc quantitative and qualitative indicators compiled through interviews, surveys, studies or internal analysis by European Commission services. Within the interim evaluation framework, an attempt was also made to classify and report on the expected impacts of the programme according to a set of three non-exclusive categories, based on the programme reconstructed intervention logic: scientific impact, economic/innovation impact and societal impact. Finally, the evaluation also reported on the longer-term impact of the previous Framework Programme, notably based on counterfactual analysis of research outputs and econometric modelling on jobs and growth (PPMI, 2017).

However, the interim evaluation of Horizon 2020 faced limitations due to:

- **Data availability**: most Horizon 2020 indicators focus on input/outputs but not on results and impact. Indicators to track progress on the societal challenges are not challenge specific, i.e. they relate to classical outputs from R&I projects - publications, patents, prototypes - but not to their impacts on e.g. decreasing CO2 emissions, improving health of citizen, or their security, often on the longer term. There is also no systematic collection of information related to the research results, innovations attained, impacts achieved on the market (e.g. sales, market shares, further investment received, efficiency gains obtained, etc.);
- **Reliability of data**: data are for many parts of the programme based on self-reporting by project coordinators (e.g. publications and patent applications) which while representing an administrative burden on the beneficiaries is not fully reliable; data on cross-cutting issues like gender equality and social sciences and humanities is based on manual "flagging" by project officers and is thus also subject to variations in interpretation.
- **Aggregation**: KPI are developed for specific parts of the programme but not for the programme as a whole making aggregation difficult;
- **Lack of benchmark**s: Worldwide there is no programme similar to the Framework Programme in terms of size, thematic coverage and depth making benchmarking difficult and no baseline data was collected.
- **Attribution/contribution assessment**: the headline indicators identified are not attributable to the programme and cover the European Union as a whole, such as the share of researchers as part of the active population or the share of the Gross Domestic Product (GDP) invested in research and development (Horizon 2020 funding represents less than 10% of public expenditures in R&D in Europe (European Commission, 2017b)). There is no established indicator/methodology to measure the contribution of the Programme to jobs and growth. The overall impact of the programme is thus mainly estimated based on econometric modelling analysing its contribution to European GDP growth.

Overall, the interim evaluation of Horizon 2020 identified a need for a further improvement and sophistication of the monitoring and evaluation framework to track and assess the impact of the Framework Programme in the short, medium and long term according to its wider set of objectives.

### 2.4 MEETING THE POLICY NEEDS: RECONCILING MEASUREMENT CHALLENGES WITH POLICY NEEDS

Evaluations are needed to inform the policy cycle. But evaluations also need to be informed by an appropriate monitoring system. Faced with complexity, little efforts are devoted in practice to try and monitor the diversity of impacts R&I programmes can trigger. However, even if there is no methodological solution readily available to handle complexity, a better communication of the impacts from R&I investments is necessary to inform budgetary arbitrations and policy decisions in the context of rapidly evolving socio-economic agendas. Policy makers cannot wait 25 years to say a policy intervention worked or did not work, there is a need for an early warning system. This means approaches should be developed to ensure the progress made can be captured.

As stated by Pawson (2003) evaluators are always left with the same question – complexity is inescapable, what can be done in the face of it? Pawson suggests a pragmatic approach for evaluators to deal with complexity in practice:

- **Stare it in the face** – map out the potential conjectures and influences that appear to shape the programme. Evaluation has to make sense of the collision of programme theories, rather than ticking off an agreed shopping list of hypotheses.
- **Concentrate your fire** – the only way to get to grips with complexity is to prioritise, by concentrating evaluation resources on those components of the programme theory that seem vital to its effectiveness and provide light monitoring elsewhere. It is better to draw out and test thoroughly a limited number of really key programme theories rather than achieve an approximate sketch of it all.
- **Go back to the future** – incorporate not only formative and summative elements in the evaluation but also design it so that it can contribute to future meta-analysis and policy development. This means adding 'systematic reviews' of the findings of previous evaluations to the multi-method shopping-list. Whilst the total package may be different, many of the components will be similar.
- **Stand on others' shoulders** – where some theories have been tested in evaluations of similar schemes, rely on these rather than repeating the work and create institutional memory that generates a progressive series of evaluation questions.
- **Criss and cross** – compare with the way similar programmes work in different contexts, in order to learn what works for whom in what circumstances.
- **Remember your job** – useful evaluations initiate a process of thinking through the tortuous pathways along which a successful programme has to travel, providing 'enlightenment' as opposed to 'political arithmetic'.

# 3. THE EMERGENCE OF IMPACT PATHWAYS - A MOVE TOWARDS INDICATOR SYSTEMS BASED ON THE THEORY OF CHANGE

It is impossible to forecast the trajectory of R&I activities and to know if the good path is being exploited at all: R&I activities usually do not follow a linear process (Freeman, 1987) but are based on "design and redesign" (Foray, 2000) and happen within systems. An innovation system is constituted by actors and elements which interact in the production, diffusion and use of economically useful knowledge (Lundvall, 1992). The specific global, regional, sectoral and technological system (e.g. Edquist, 2005) in which beneficiaries operate have an important indirect influence on the relative performance of R&I programmes, notably because of the regulatory, legislative, financial or political context but also because of the degree of availability of infrastructures or human capital or the level of consumer demand. As Edquist (1997) argues, the notion of optimality is irrelevant in a system of innovation context. The systemic nature of R&I processes make it difficult to isolate the impacts of a specific programme, notably when it comes to quantification.

However, approaches such as the programme-theory approach in the realist school of thought (Suchman, 2007; Chen, 1990; Weiss, 1987; Donaldson, 2007) provide an interesting way around the 'black box' of causation by providing testable hypotheses about how causes lead to effects (Technopolis, 2018).

Within theory-based evaluations, attention is paid to theories of policy makers, programme managers or other stakeholders, that are logically linked together. The objectives of the intervention are used to construct a set of logical steps via which the intervention is expected to lead to outcomes and impacts. The actual results will depend both on policy effectiveness and on other factors affecting results, including the context. The central thesis of the programme-theory evaluation is that the impact of the programme is expected to occur based on a logic set of events and interactions between the participants to the programme, the results of the projects funded and the wider environment. As reported in Rogers (2008) literature uses a variety of names for this concept including programme logic (Funnell, 1997), theory of change (Weiss, 1995, 1998), intervention logic (Nagarajan and Vanheukelen, 1997) and impact pathway analysis (Douthwaite et al., 2003b). The programme-theory approach became an evaluation standard in the European Commission with the introduction of the Better Regulation Agenda and related guidelines (European Commission, 2017a). Whereas application of this approach for programme evaluations is becoming common practice in many areas, including for European Structural Funds interventions (European Commission, 2014), it is not commonly used for tracking progress over time during the implementation of R&I programmes (van den Besselaar, Flecha, Radauer, 2018).

The Impact pathways concept falls under this approach: it looks for a simple and likely interpretation on how the project/programme/policy expects to lead to impact. Sketching impact pathways typically include the identification of a set of steps or intermediate signposts, in the short, medium or longer term which indicate that the outputs are likely transforming into wider aggregated impacts. Impact pathways are so far mostly used at the level of individual proposals and projects (see Douthwaite et al., 2003 for an example in the agricultural sector). Rogers (2008) based on a literature review gives practical guidance on how to apply it to complex settings. Overall key messages would be:

i. Keep the logic of the intervention sufficiently broad to encompass various and individual pathways;
ii. Refrain from using logic models to generate performance measures based on a set of quantitative indicators, without more in-depth qualitative and participative assessments.

The United Kingdom provides a concrete case of using impact pathways in the R&I area, at the level of the research proposal. In this framework a clearly thought through and acceptable Pathways to Impact demonstrating how the proposed research can make a difference is an essential component of a research proposal and a condition of funding. These Pathways to Impact cover academic, economic and societal impact and are updated and adapted as the context changes and the research trajectory unfolds (Tulley et al., 2018).

A key challenge in impact evaluation to deliver on policy objectives comes then in the *quantification* of the observed effects that can reasonably be attributed to the programme to identify its added value. The challenge facing the evaluator is to avoid giving a causal interpretation to differences that are due to factors other than the intervention. This is the essence of counterfactual impact evaluation, which aims to identify what would have happened if the programme had not existed (European Commission, 2013). In the case of R&I interventions, counterfactual analysis have typically been performed on business R&D support (European Commission, 2018c), or on publication outputs of individual research fellows (i.e. Jonkers et al, 2018). A key feature of most assessments using counterfactual methods is their reliance on microeconomic approaches based on data at the level of the firm or the individual researchers. These approaches are in principle open to peer scrutiny (subject to data availability) and reproducibility by other researchers (Klette et al., 2000).

This requires building data and indicators systems early on that allow for a tracking of progress over time at least for the key areas of expected impacts based on systematic and harmonised data collection. For instance, it also requires to collect data for those specifically stimulated by the initiative and data for those that were not stimulated by the initiative (i.e. *control groups*) to correct for external factors. The growing availability of (micro)data stemming from the current digitalisation age and the enhanced capacity to automate its treatment and link datasets make data collection easier to perform, although the use of such data in evaluation is still in its infancy.

# 4. A REVAMPED SYSTEM FOR THE FRAMEWORK PROGRAMME

## 4.1 SETTING OUT THE AMBITION: PATHS PRINCIPLES

Based on the lessons from past, international experience, the changing policy context and the evolving objectives of R&I investments, five key principles were identified for the development of a purposeful indicator framework for Horizon Europe (the **PATHS** principles):

- **P**roximity - Know who the individual researchers and companies are in order to better capture the impact the programme is having on the ground (e.g. by collecting unique identifiers such as VAT numbers, researchers IDs and funder ID), including through the use of control groups;
- **A**ttribution - Capture a diversity of impacts that can be attributed to the intervention from the Framework Programme, beyond classical indicators such as publications and patents, to seize the difference it is making for society, for the economy and for scientific progress;
- **T**raceability – Minimize the reporting burden on beneficiaries by developing automatic data harvesting from external public and private databases ("Once-Only"); using additional primary data sources such as project officers, evaluators and reviewers; and streamlining the reporting template;
- **H**olism - Tell the story of the progress of the programme as a whole at any moment in time, given the common long term objectives and cross-linkages of the different actions, while managing expectations on what can reasonably be reported by when;
- **S**tability - Build on the current systems (e.g. by ensuring maximum continuation and comparability with the previous Framework Programmes, in particular Horizon 2020) and increase data quality (e.g. by piloting different data collection and analysis methods already in Horizon 2020 monitoring and evaluation).

The application of the PATHS principles resulted in a proposal for a revamped indicator framework for the Framework Programme built around a set of Key Impact Pathways (European Commission, 2018c).

## 4.2 THE FRAMEWORK: HORIZON EUROPE KEY IMPACT PATHWAYS

Horizon Europe Key Impact Pathways, built around the Horizon Europe objectives, intend to structure the annual monitoring of Horizon Europe towards achieving its objectives. The Key Impact Pathways focus on the most typical changes that are expected to occur on a short, medium and longer term as a result of the Programme activities - allowing for a more realistic assessment and communication of the progress made over time and moving beyond the mere monitoring of programme management and implementation data. They are an integral part of the proposal adopted by the Commission for Horizon Europe (European Commission, 2018b).

In line with the typical impacts identified for past Framework Programmes and the underlying most salient impact pathways (see Arnold, 2012), the Horizon Europe Treaty-based objectives translate into three complementary and non-exclusive impact categories:

- Scientific impact: related to the creation and diffusion of high-quality new knowledge, skills, technologies and solutions to global challenges;
- Societal impact: related to strengthening the impact of R&I in developing, supporting and implementing EU policies, and supporting the uptake of innovative solutions in industry and society to address global challenges;
- Economic impact: related to fostering all forms of innovation, including breakthrough innovation, and strengthening market deployment of innovative solutions.

For monitoring purposes and to account for the multidirectional nature of R&I investments, it is proposed to track progress towards impact along three Key Impact Pathways each for the three types of impact identified. (Figure 1).

The nine Key Impact Pathways combine the latest developments in understanding, measuring and assessing the impact of R&I programmes. Each Pathway consists of a storyline, a time-sensitive indicator and data needs (Figure 2):

1. A storyline illustrates the typical message that can be communicated on the progress of the programme on this Pathway over time.
2. A time-sensitive indicator distinguishes between the short (typically as of one year, when the first projects are completed), medium (typically as of three years, and in time for the interim evaluation of the Programme) and long term (typically as of five years, and in time for the ex-post evaluation) to monitor the progress over time in a realistic way. To ensure the measurement focusses on the programme achievements (Attribution principle) the indicator starts from the projects' outputs to then look at their diffusion into results and impacts.
3. Data needs identify the main information needed and possible methodologies to collect the data, while minimizing the reporting burden on beneficiaries.

By design, the revamped indicator system appears simple and linear from a macro-perspective. However, this does not mean that the Programme is regarded as following a linear path towards impact. Instead the indicator system depicts the key dimensions on which impact is desired, and where information is needed. This should enable clear and straightforward communication of the main changes Horizon Europe as a whole is bringing in the longer term *(the principle of Holism*). With the use of storytelling, the indicator framework should bring about a much wider understanding of why the EU invests in R&I and how such investments generate value for society, for the economy and for scientific progress. Among the diversity of decision makers, stakeholders, implementers and beneficiaries a common general understanding is indeed crucial.

Looking more specifically at the key dimensions covered to track progress towards each type of impact, the Key Impact Pathways towards *scientific impact* focuses on the monitoring of: 1) the creation and diffusion of high quality new knowledge through high-quality scientific

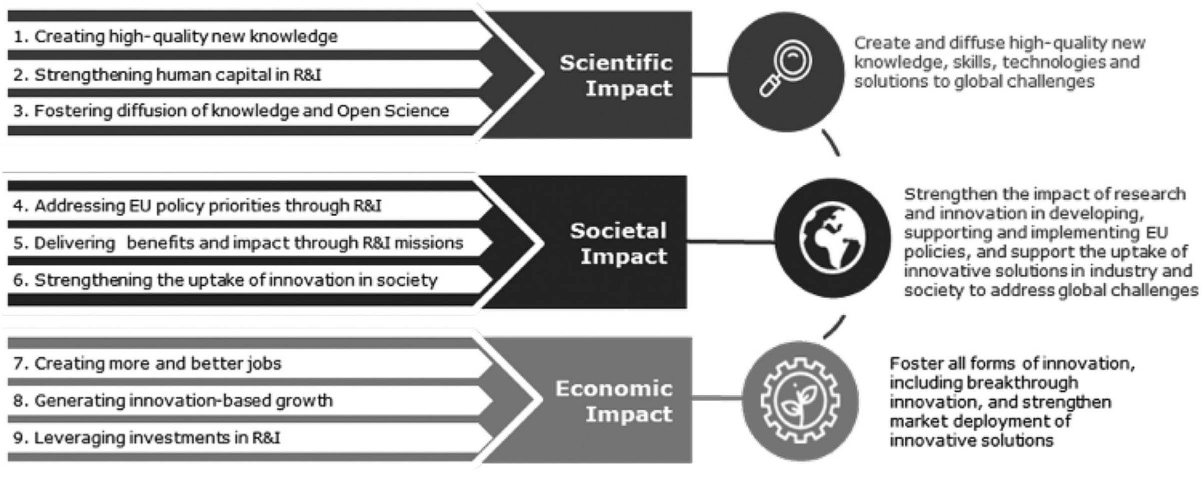



**Fig. 1:** Proposed Key Impact Pathways of Horizon Europe
**Source:** European Commission (2018), Impact Assessment accompanying the Commission proposal for Horizon Europe, the Framework Programme for Research and Innovation , SWD (2018) 307 final

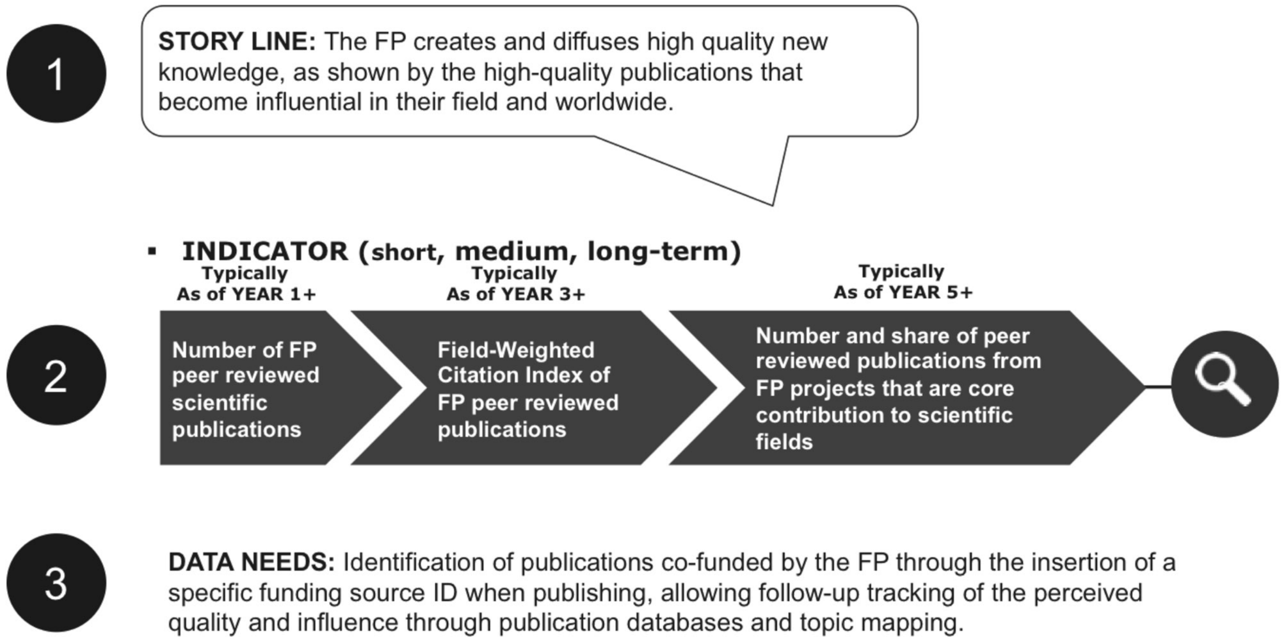



**Fig. 2:** Example of a Key Impact Pathway – Creating high-quality new knowledge
**Source:** Authors based on European Commission (2018), Impact Assessment accompanying the Commission proposal for Horizon Europe, the Framework Programme for Research and Innovation , SWD (2018) 307 final

publications that become influential in their field and worldwide; 2) the strengthening of human capital in R&I through evidence on improved skills, reputation and working conditions of researchers; and 3) the diffusion of knowledge and open science through evidence of open sharing and reuse of research outputs and later creation of new transdisciplinary/trans-sectoral collaborations. By collecting information on individual researchers involved in the Programme, including in the collaborative projects (including through an increased use of unique identifiers) the data system behind these Pathways is expected to harvest public data automatically from existing external public and private databases also after the projects' end (e.g. data on publications, citations, affiliations, patents) (*Traceability principle*) and allow for counterfactual analysis.

The Key Impact Pathways towards *societal impact* focusses on the monitoring of how and to what extent the Programme contributes to addressing EU policy priorities (including meeting the Sustainable Development Goals) by assessing portfolio of projects that generate outputs which aim to contribute to tackle global challenges or to achieve future R&I missions. A specific pathway also monitors the uptake of Framework Programme innovations in society by initially identifying whether end-users and citizens contribute to the co-creation of R&I content, to then look at the sustainability of these engagement mechanisms to then capture the level of uptake of the co-created scientific results and solutions. As there are currently no readily available methods to monitor societal impact on a scale as large as the Framework Programme, several methods will need to be tested before a full-scale implementation. One avenue is to use a portfolio analysis by dedicated reviewers (e.g. experts, evaluators) around specific topics, possibly based on the experience of the Innovation Radar methodology already implemented for the Framework Programme. New ICT tools and possibilities of semantic analysis will also need to be tested to inform the identification of relevant projects and outputs portfolios.

The Key Impact Pathways towards *economic impact* focusses on the monitoring of how and to what extent the Programme contributes to generating innovation-based growth, creating more and better jobs and leveraging investment in R&I. These are based on identifying and tracing individual outputs of projects (e.g. patents, trademarks and other IPRs) and public data (e.g. business registers, company databases) on participating as well as non-participating companies (e.g. turnover, employment). Such data will allow to build control groups for counterfactual analysis. Testing a possible introduction of funder identifiers for IPR applications and/or identifying ways to improve the identification of patents in patent databases will also be needed prior to a full scale implementation.

## 4.3 MEETING THE DATA NEEDS: ALIGNING DATA COLLECTION METHODS TO THE FRAMEWORK

The indicator system builds on the Horizon 2020 indicator system but indicators are streamlined and further specified to meet the objectives (*the principle of Stability*). Overall, a key vector for successful implementation of this revamped indicator framework is a much-increased reliance on microdata and unique persistent identifiers. This simple information will allow tracing e.g. the career paths of individual researchers involved, the growth of participating companies and the diffusion of knowledge through publications or patents in key areas of relevance for society. It will support the simplification agenda by minimising the reporting burden on beneficiaries. Furthermore the micro-level data collection methods will not only allow to report on Key Impact Pathway indicators but also to disaggregate indicators by type of actions, type of organisations, type of collaborations, sectors, disciplines, calls, countries and programme parts, when more granular information is necessary.

### THE ADDED VALUE OF COLLECTING UNIQUE IDENTIFIERS

The use of persistent unique identifiers, which can sort out different scientists/companies with the same names, and create a lifelong trace of their work, will allow to:

- Monitor the **number of researchers** supported through the programme and automatically access the publicly available information on their affiliation, mobility, career evolution, scientific production, IPR applications, etc. by linking the identifier to external databases.
- Monitor the **evolution of companies** supported through the programme and automatically access their scientific or innovation outputs, turnover, investment, etc. by linking the identifier to external databases.
- Build **control groups** to allow for counterfactual evaluation design (propensity score-matching, regression discontinuity design or difference-in-difference methods), e.g. tracing the differences between researchers and companies not benefitting from the programme and those benefitting from the programme

The indicator framework is overall expected to provide a solid basis for accountability in so that evaluations can focus on diving deeper into learning and identifying the necessary policy adjustments for the future. The indicators collected will be one of the many elements feeding into the interim and the ex-post evaluations of Horizon Europe together with other information sources and qualitative and quantitative indicators. Because of the time lags and the uncertainty of the R&I investments, the interim evaluation will typically provide first evidence on the relevance and coherence of the programme and the efficiency of the processes in place, to identify potential pitfalls or drivers early in the process. It will also include a longer-term assessment of past Programmes to shed light on longer-term impacts.

# 5. CONCLUSIONS ON THE EXPECTED RESULTS OF THE REVAMPED INDICATOR SYSTEM

The Key Impact Pathways underpinning Horizon Europe's monitoring system represent a novel, ambitious yet pragmatic approach for devising indicator frameworks for R&I programmes. It results from the identified need to start facing the complexity of R&I investments in monitoring and evaluation practices in order to deliver relevant and timely messages to policy makers. Based on a set of core principles (PATHS: proximity, attribution, traceability, holism and stability) this framework will ensure information is collected on a set of key dimensions on which impact is desired. Overall the Key Impact Pathways are expected to support a better capture and communication of the progress of Horizon Europe towards its objectives, including beyond its lifetime. The simplicity and storytelling nature of the Key Impact Pathways should bring a more immediate and continuous visibility of the European added value of R&I investments for science, the economy and society and allow to reach a wider audience beyond the R&I community.

To make best use of the potential of the Key Impact Pathways, data collection needs to match the ambition and pragmatism. The underlying richness and soundness of the analysis this will enable may well set a new trend for monitoring the impacts of R&I investments in the future. Policy makers will be able to better identify and recognise the multiple impacts of R&I investments, going beyond the mere identification of participation patterns, or the raw scientific and innovation production. A stronger focus on microdata collection and data linking will allow for an easier identification of concrete storylines at the level of individual researchers, projects or project portfolios, including on the drivers and barriers to impacts. This will be a key element in improving the quality of programme evaluations, and their usefulness for policy learning and policy design – thereby paving the pathway to impact.

# 6. REFERENCES

**Arnold E.** (2012). Understanding long-term impacts of R&D funding: The EU framework programme, Research Evaluation 21 pp. 332–343.

**Arnold, E., Clark, J. and Muscio, M.** (2005). What the Evaluation Record Tells us about Framework Programme Performance, Science and Public Policy, 32/5: 385–97.

**Barré R., Cresson E.** (1997). Science in tomorrow's Europe. Economica International, Paris.

**Bornmann, Lutz.** (2013). What is social impact of research and how can it be assessed? A literature survey. Journal of the American Society for Information Science and Technology, 64/2.

**Chen, H.** (1990). Theory-driven Evaluation. Beverley Hills CA: Sage.

**Coryn C. L.S., Noakes, L. A., Westine, C. D. and Schröter, D. C.** (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. American Journal of Evaluation, 32(2), 199–226.

**Davignon E. et al** (1997). 5-Year Assessment of the European Community RTD Framework Programmes, COM(97)151 final, European Commission.

**Donaldson, S. I.** (2007). Program theory-driven evaluation science, New York, NY: Lawrence.

**Douthwaite, B., T. Kuby, E. van de Fliert and Schulz, S.** (2003). Impact Pathway Evaluation: An Approach for Achieving and Attributing Impact in Complex Systems, Agricultural Systems 78: 243–65.

**Edquist C.** (1997). Systems of innovation: technologies, institutions and organisations, Pinter.

**Edquist C.** (2005). Systems of Innovation: Perspectives and Challenges in Fagerberg J., David C. Mowery D.C., Nelson R.R (eds.), The Oxford Handbook of Innovation , Oxford University Press, New York, US.

**EPEC** (2011). Understanding the Long Term Impact of the Framework Programme, Report to the European Commission.

**European Commission** (2005). Five-year assessment of the European Union Research Framework Programmes 1999-2003.

**European Commission** (2013). Evalsed Sourcebook – Methods and Techniques.

**European Commission** (2014). Programming period 2014-2020, Guidance Document on Monitoring and Evaluation - European Cohesion Fund, European Regional Development Fund.

**European Commission** (2017a). Better Regulation Guidelines, Commission Staff Working Document SWD (2017) 350, 2017.

**European Commission** (2017b). Interim Evaluation of Horizon 2020, SWD (2017) 220.

**European Commission** (2018a). Impact Assessment accompanying the Commission proposal for Horizon Europe, the Framework Programme for Research and Innovation, SWD (2018) 307.

**European Commission** (2018b). Proposal for a Regulation of the European Parliament and the Council establishing Horizon Europe - the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination.

**European Commission** (2018c). Evaluation of Business R&D Grant Schemes: behavioural change, mixed-method approaches and big data, Mutual Learning Exercise under the Horizon 2020 Policy Support Facility.

**European Court of Auditors** (2007). Evaluating the EU RTD FP – Could the Commission's approach be improved, Special Report No 9/2007, paragraph IV.

**Foray D.** (2000). L'économie de la connaissance, Repères, La Découverte.

**Freeman C.** (1987). Technology and economic performance: Lessons from Japan, Pinter.

**Freeman, C., Lundvall, B.-Å.** (eds.) (1988). Small Countries Facing the Technological Revolution, London: Pinter Publishers.

**Fresco L. et al** (2015). Commitment and coherence - Ex-post evaluation of 7th EU Framework Programme (2007-2013).

**Funnell, S.** (1997). Program Logic: An Adaptable Tool for Designing and Evaluating Programs, Evaluation News and Comment 6(1): 5–7.

**Graham K. E. R., Langlois-Klassen D., Sagal A. A. M., Chan L. and Chorzempa, H. L.** (2018). Assessing Health Research and Innovation Impact: Evolution of a Framework and Tools in Alberta, Canada. Frontiers in Research Metrics and Analytics (2)2018.

**Guellec D.** (1999). Economie de l'innovation », Collection Repères, La Découverte.

**Guthrie, S., Wamae, W., Diepeveen, S., Wooding, S., and Grant, J** (2013). Measuring Research: A Guide to Research Evaluation Frameworks and Tools. Santa Monica, CA: RAND Corporation.

**Irvine J. Martin B.,** (1989). Research Foresight: Priority-Setting in Science, Pinter Publishers, London.

**JIIP** (2016). The contribution of the Framework Programmes to Major Innovations, Report to the European Commission.

**Jonkers, K., Fako, P., Isella, L., Zacharewicz, T., Del Rio, JC, Sandström, U., Van den Besselaar, P.** (2018). A comparative analysis of the publication behaviour of MSCA fellows, European Commission – Joint Research Centre.

**Klette T.J., Moone J., Griliches Z.** (2000). Do Subsidies to Commercial R&D Reduce Market Failures? Research Policy, 2000, vol. 29, issue 4-5, 471-495.

**Lundvall B.A.** (1992). "National systems of innovation: Towards a theory of innovation and interactive learning, Pinter.

**Nagarajan, N. and Vanheukelen, M.** (1997). Evaluating EU Expenditure Programmes: A Guide. Brussels: Directorate-general for Budgets of the European Union.

**Nelson, R. R.** (1993). National Innovation Systems. New York:Oxford University Press.

P**awson, R., andTilley, N.** (1997). Realistic Evaluation. London: Sage.

**Pawson, R.** (2003). Nothing as practical as a good theory. Evaluation 9(4): 471 – 90.

**Pawson, R.** (2006). Simple principles for the evaluation of complex programmes,' in A Killoran and A Kelly (eds), Evidence based public health . Oxford: Oxford University Press.
**PPMI** (2017). Assessment of the Union Added Value and the economic impact of the EU Framework Programmes (FP7, Horizon 2020), Report to the European Commission.

**Ravet, J., Boitier, B., Grancagnolo, M., Le Moüel, P. Stirbat, L. and Zagamé, P.** (2018). The Shape of Things to Come: Ex-Ante Assessment of the Economic Impact of Horizon Europe. Paper presented at the Austrian Presidency of the Council of the European Union Conference on Impact of Research and Innovation Policy at the Crossroads of Policy Design, Implementation and Evaluation.

**Rietschel et al.** (2009). Evaluation of the Sixth Framework Programme for Research and Technological Development, Report to the European Commission.

**Rogers, P. J.** (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions, Evaluation, 14(1):29-48.

**Smith, K.H.** (2005). 'Measuring innovation', in Fagerberg J., David C. Mowery D.C., Nelson R.R (eds.), The Oxford Handbook of Innovation , Oxford University Press, New York, US.

**Spaapen, J., van Drooge, L.** [SIAMPI Consortium] (2011). Final Report on Social Impacts of Research. Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society.

**Stampfer M.** (2008). European Added Value of Community Research Activities - Expert analysis in support of the ex-post evaluation of FP6, WWTF, Vienna Science and Technology Fund.

**Suchman, E.** (1967). Evaluative research. New York, NY: Russell Sage Foundation.

**Technopolis Group** (2009). Framework Programme 6 – Meta-evaluation, Report to the European Commission.

**Technopolis Group** (2018). How should we evaluate complex programmes for innovation and socio-technical transitions. Technopolis group, June 2018.

**Tilley H., Ball L., Cassidy C.** (2018). Research Excellence Framework (REF) impact toolkit. Overseas Development Institute, March 2018.

**Van den Besselaar P., Flecha R., Radauer A.** (2018). Monitoring the Impact of EU Framework Programmes - Expert Report for the European Commission. Publications Office of the European Union: Luxembourg.

**Weiss, C.** (1987). Where Politics and Evaluation Research Meet, in D. Palumbo (ed.) The politics of program evaluation. Newbury Park, CA: Sage.
**Weiss, C. H.** (1998). Evaluation: Methods for Studying Programs and Policies. Englewood Cliffs, NJ: Prentice Hall.

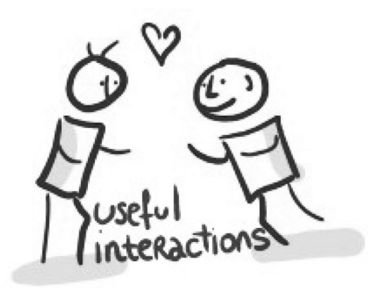
**Weiss, C.** (1995). Nothing as Practical as a Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families, in J. P. Connell, A. C. Kubisch, L. B. Schorr and C. H. Weiss (eds) New Approaches to Evaluating Community Initiatives, pp. 65–9. Washington, DC: Aspen Institute.

# AUTHORS

**NELLY BRUNO**
*European Commission - DG Research and Innovation*
1049 Brussels, Belgium
E: nelly.bruno@ec.europa.eu

**MARTINA KADUNC**
*European Commission - DG Research and Innovation*
1049 Brussels, Belgium
E: martina.kadunc@ec.europa.eu

## KEYWORDS:
Research policy, innovation policy, impact, evaluation, monitoring, Framework Programme, indicators, Europe, Horizon Europe, pathways





# IS A POLICY MIX MORE EFFECTIVE THAN INDIVIDUAL POLICIES FOR SME INNOVATION? AN EXPLORATORY ANALYSIS[1]

ANNALISA CALOFFI, MARZIA FREO, STEFANO GHINOI, FEDERICA ROSSI AND MARGHERITA RUSSO
DOI: 10.22163/fteval.2019.333

## INTRODUCTION

The provision of public funds to private firms for the purchase of services, particularly knowledge-intensive ones, has received so far little attention from the evaluation literature (Bakhshi et al., 2015; Bruhn et al., 2018 are notable exceptions). These interventions often target small and medium-sized enterprises (SMEs), providing them with a small amount of public funds that reduce their cost of purchasing services (Storey, 2003). Public funding can take the form of a direct subsidy or a voucher, which firms must use to purchase services from accredited service providers, or sometimes from any provider freely chosen by the beneficiary firm (OECD, 2000; Storey, 2003; IEG, 2013).

These interventions aim to help SMEs to access a variety of knowledge and competencies required for innovation, which are not available within the firm (Vossen, 1998; Storey, 2003). The implicit assumption is that SMEs primarily suffer from constraints on their financial resources, rather than on their capabilities. After receiving the subsidy, SMEs should be able to identify the services they need, as well as the suppliers that can best provide them. However, it is well known that SMEs, may not only lack the financial resources to invest in innovation, but also the capabilities to identify the competences and services they need, or the right suppliers that can provide them (Fontana et al., 2006; Ortega-Argilés et al., 2009). Subsidies for the purchase of knowledge-intensive services address the former problem, but not the latter.

As discussed by Shapira and Youtie (2016), to help SMEs increase their awareness of their needs and how to address them, they could be provided with complementary services, such as technology and innovation advisory services. We argue that such services could be usefully combined with innovation vouchers to increase the performance of SMEs. Technology and innovation advisory services are usually delivered by one or more experts, who carry out a thorough assessment of the firm's current knowledge and technology and an exploration of potential developments. This allows the people involved to undertake a highly customized process of mutual learning, which increases the firm's knowledge of its own innovation needs. Following the assessment, experts can direct the firm to other external service providers that will be able to deliver the specialized knowledge-intensive services it needs.

In several countries, technology and innovation advisory services are provided by, among others, publicly-funded innovation intermediaries, whose aim is to support innovation in SMEs by providing them with a variety of services. Precisely because the advisory services offered by intermediaries could improve SMEs' choice and use of knowledge-intensive services, we expect this combination of interventions to be more effective than the individual instruments.

This study presents an exploratory empirical analysis focused on two interconnected regional innovation policy interventions implemented in Tuscany (Italy). One was the provision of innovation vouchers that SMEs could use to buy knowledge-intensive services from accredited providers, while the other intervention was the creation of intermediaries that could help SMEs to access such services. Since firms could benefit either by only one of the two interventions, or by both, we use a dataset derived from administrative sources to assess whether the policy mix that includes both interventions was more effective than the voucher alone or even the technology and innovation advisory service alone. We adopt a propensity score matching approach applied to the case of multiple treatments, as proposed by Lechner (2002a, 2002b). In particular, we compare three different treatments: (i) the use of innovation vouchers for the purchase of knowledge-intensive services; (ii) the reliance on an intermediary's technology and innovation advisory service; (iii) the combination of the two treatments, i.e. the use of innovation vouchers for the purchase of knowledge-intensive services with guidance from the intermediary.

While policy mixes have been advocated as a response to complex problems (Flanagan et al., 2011; Cunningham et al., 2016), very little empirical evidence is available about the comparative effectiveness of policy mixes with respect to that of the single policies in the mix (Martin, 2016), and no other studies consider the particular combination of innovation vouchers and advisory services. This exploratory study captures an aspect that lies at the core of the policy mix literature, namely that the mix cannot be considered as the simple sum of the single instruments that are included in it (Magro and Wilson, 2013), but it can facilitate the emergence of synergies and complementarities among them.

1 This study was conducted in the framework of the research project 'Poli.in_Analisi e modellizzazione dei poli di innovazione in Toscana', co-funded by Tuscany Region and the Department of Economics 'Marco Biagi', University of Modena and Reggio Emilia. The authors wish to thank Albino Caporale, Emanuele Fabbri, Angelita Luciani, Serena Brogi, and the managers of the Innovation poles for their support in data collection. Opinions and conclusions expressed are attributable only to the authors.

## POLICY MIXES IN SME POLICY

The rationale for supporting the acquisition of knowledge-intensive services builds on the idea that, as innovation processes become more complex and the market environment becomes more turbulent, innovative firms need to mobilize a wide range of knowledge and skills, some of which are not available internally. SMEs, which have relatively scarce internal resources, may need support from external experts during one or more phases of the innovation process, from the realization of feasibility studies, to the marketing of innovative products or services (Vossen, 1998; Muller and Zenker, 2001; Storey, 2003; Toivonen, 2007; Shapira and Youtie, 2016). Innovation vouchers are gaining popularity because they are easier to administer than standard grants (Schade and Grigore, 2009) and help knowledge providers to better understand industry needs (Coletti and Landoni, 2018). Innovation vouchers have been found to promote firms' external relationships (BIGGAR Economics, 2010; Sala et al., 2016) particularly with public research institutions (Cornet et al., 2006), and their engagement in further innovation projects (Good and Tiefenthaler, 2011; Bakhshi et al. 2015), in particular for firms that had previously pursued innovative activities (Sala et al., 2016). As innovation vouchers lead firms to adopt a more structured approach to innovation, reducing the time-to-market (Sala et al., 2016), and to engage in more innovation projects and collaborations (Bakhshi et al. 2015), they can be expected to have a positive effect on firms' performance, both in terms of increased sales due to the introduction of innovative products, and in terms of greater efficiency thanks to improvements in internal processes.

Innovation intermediaries are organizations that support firm-level and collaborative innovation, often relying on public funding (Uotila et al., 2012; Knockaert et al., 2014; Caloffi et al., 2018; Russo et al., 2018). Intermediaries' activities frequently include the provision of expert advice on technology and innovation, particularly to SMEs in order to address their capabilities failures (Bessant and Rush, 2005; Knockaert et al., 2014). In fact, intermediaries, which by their nature are able to bridge different types of knowledge and competencies, are well placed to understand the features of the production and innovation processes that are implemented by the firm, the markets it operates in, and those it could enter. Drawing on their assessment of the firm's knowledge and technology, intermediaries identify the firm's strength and weaknesses, and advise it on the implementation of an appropriate innovation strategy (Shapira and Youtie, 2016).

To the best of our knowledge, the above policy instruments have always been investigated individually. However, they could be usefully implemented together, and, to test whether it makes sense to do so, we assess whether the performance of a firm participating in both policies improves more than if the same firm had participated in only one of the two policies. On average, we expect that the performance of firms that receive technology and innovation advice will be greater than that of firms that choose their external services without any particular help. Literature has shown that firms, especially the smallest ones, not only lack the knowledge and competencies that are needed to innovate, but also to understand what their needs are. If this is true, the intermediary can play an important role in guiding firms towards the best possible use of their vouchers, and therefore towards a greater improvement in performance than what would be achieved without such help.

## REGIONAL POLICIES IN SUPPORT OF SME INNOVATION: THE CASE OF TUSCANY

In Italy, regional policy interventions providing SMEs with incentives for the acquisition of knowledge-intensive services have been implemented since the devolution of enterprise policy to regions (Caloffi and Mariani, 2018). In Tuscany, a new policy was launched in 2008, whereby vouchers were issued to SMEs for the acquisition of one or more services drawn from a specific list (the "regional portfolio of knowledge-intensive services"). Forty-four different types of services were listed, including design or other technical expertise, quality testing and marketing of innovative products. Funding came from the European Union's European Regional Development Fund, and could be granted to firms operating in a wide spectrum of sectors. The voucher covered a percentage of the cost of the service, which varied from 60% to 80% depending on the type of service. The same firm could apply for more than one voucher both simultaneously and over time. The average voucher amount was relatively small (in line with international practice: OECD, 2008; Good and Tiefenthaler, 2011; Shapira and Youtie, 2016) and, in any event, the same firm could not get more than 200,000 Euros in three years.

In 2011 Tuscany's regional government launched twelve "innovation poles", which were specialized in specific technologies and/or sectors (Russo et al., 2015). SMEs that were members of an innovation pole received several visits from experts that worked for the innovation pole. These experts tried to understand the features of the firm's production and innovation processes, the markets it operated in, and those that it could target. Drawing on their assessment of the firm's knowledge and technology, intermediaries were able to identify the firm's strength and weaknesses and to identify appropriate innovation strategies that the firm could implement.

The two policies mentioned above could be combined. After having identified a feasible innovation strategy, experts provided SMEs with specific information on the innovation vouchers that they could obtain from the regional government. The experts could also help the firms choose the type of service that suited their needs, and support them in their funding applications.

## METHODOLOGY AND DATA

To discover whether the policy mix improves performance compared to the individual innovation policies, we recur to the multiple treatments setting, where the treated group is always formed by firms that are recipients of a specific innovation policy, and control groups are formed by firms treated with one of the two alternative policies in pairwise comparisons.

To identify the treated and control groups, we rely upon administrative data made available by the policymaker running the programmes. We consider two cohorts of treated groups by fixing the time to treatment respectively in 2011 and 2012, which corresponds to the first two years of activity of the innovation poles. In this period, the call for tender related to the vouchers for the acquisition of external services was also open. We consider only these early cohorts because we want to have a sufficiently long time frame to observe the ex-post results of these policies.

Time-varying data refer to three different time points. In particular, information on the firms' background characteristics refers to one year before the start of the policy, whereas information on the outcomes of interest refers both to the year in which the policy ended and one year after the end of the policy. As a whole: (a) 166 manufacturing SMEs only received vouchers for the acquisition of knowledge-intensive services; (b) 478 manufacturing SMEs only received technology and innovation advisory services thanks to their membership of an innovation pole; (c) 178 manufacturing SMEs participated in both policy interventions.

Given that the services we observe can be of various kinds and cover different phases of the innovation process, we consider a relatively wide range of outcome variables. In particular, we consider: labour productivity, measured as value added per employee; Total Factor Productivity (TFP)[i]; total value of sales (in log transformations); and number of employees. While the first two variables refer to measures of productivity or innovation capabilities, the latter can capture some evidence of firm growth. All the data used to build the outcome variables – except for the number of employees – come from the Aida Bureau van Djik database. Data on the number of employees comes from ASIA - Italian Institute of Statistics.

As the number of observations is relatively low (less than one thousand), we use two matching strategies. The first strategy implies that we retain the whole data without imposing the common support condition. Then, a second strategy is developed, according to which we bootstrap 200 samples of 450 firms (150 firms for each one of the three outcomes) and run the multiple propensity matching over bootstrapped samples, by imposing the common support condition. In both cases treated and control firms have been matched by adopting the Mahalanobis distance computed over the two propensities scores, and the set of outcome variables considered, i.e. firm age, log-transformation of sales, the number of employees, per capita value added and TFP at the pre-treatment year. We further impose the exact match by 2-digits NACE classification. Here, we will present results attained by the procedure run over bootstrapped samples by imposing the common support condition.

The variables we use in the matching protocol are presented in the following Table 1, which summarizes their averages in the three groups of treated firms. In particular, the couples treated-control are identified by looking at the lagged values of the outcome variables mentioned above (labour productivity, TFP, total value of sales, number of employees). Besides, we also consider firm age and sector (Nace sectors at 2 digits level), which we take from the database ASIA – Italian Institute of Statistics. All these variables are measured one year before policy participation.

**Table 1.** Averages of control (and outcome) variables by treatment in the pre-treatment period

| | Voucher<br>Mean | Advisory services<br>Mean | Policy Mix<br>Mean |
|---|---|---|---|
| Firm age | 25.6 | 26.6 | 27.4 |
| ln(sales) | 15.35 | 15.16 | 15.32 |
| Employees | 32.2 | 53.4 | 35.1 |
| Per-capita value added | 59.0 | 54.6 | 53.9 |
| TFP | 0.407 | 0.281 | 0.380 |
| N. of firms | 166 | 478 | 178 |
| Relative Frequency | 0.202 | 0.581 | 0.216 |

The table shows that few significant differences across groups emerge in term of pre-entry characteristics. In particular, firms that were treated with the voucher and those that are treated with the policy mix were very similar before policy participation, while firms that only received innovation and technology advisory services were larger and relatively less innovative than the firms in the other two groups.

# RESULTS

Table 2 displays the *sign and significance* of the average treatment effect on the treated (ATTs) estimated through the bootstrap procedure by imposing the common support condition. The table shows the signs of the ATTs, of the innovation policies on their respective participants during the post-entry period. Cells in dark grey indicate significance at the 5% level; cells in light grey indicate significance at the 10% level. Following Lechner (2002b, p.69), a positive ATT indicates *"that the effect of the program shown in the row compared with the program appeared in the column is an on-average higher rate of [performance] for [firms which] participate in the program given in the row"*. Compared with the matching procedure run over the whole sample without common support (whose results are not presented here)[ii], the bootstrapped matching procedure presents consistent results, but it is more conservative in finding significant impacts. Moreover, this procedure is a priori more consistent with theoretical aspects. This allows us to be confident of robustness of estimations based on the bootstrapped multiple matching.

**Table 2.** Average effects on Treated for participants, in rows, versus participants, in columns, measured as difference in outcomes

| ln(Revenues) | Time period | Voucher | Advisory service | Mix |
|---|---|---|---|---|
| Voucher | +1 | | + | - |
| | +2 | | + | - |
| Advisory service | +1 | - | | - |
| | +2 | + | | - |
| Mix | +1 | + | + | |
| | +2 | + | + | |

| Employees | Time period | Voucher | Advisory service | Mix |
|---|---|---|---|---|
| Voucher | +1 | | + | + |
| | +2 | | + | + |
| Advisory service | +1 | + | | + |
| | +2 | + | | - |
| Mix | +1 | + | + | |
| | +2 | + | + | |

| **Per capita value added (thousands euro)** | **Time period** | **Voucher** | **Advisory service** | **Mix** | **TFP** | **Time period** | **Voucher** | **Advisory service** | **Mix** |
|---|---|---|---|---|---|---|---|---|---|
| | +1 | | + | + | | +1 | | + | + |
| Voucher | +2 | | + | + | Voucher | +2 | | + | - |
| | +1 | + | | - | | +1 | - | | - |
| Advisory service | +2 | + | | - | Advisory service | +2 | - | | - |
| | +1 | + | + | | | +1 | + | + | |
| Mix | +2 | + | + | | Mix | +2 | + | + | |

Concerning the effects on revenues, the policy mix has been found to induce a significant increase, of about 30 to 38% higher than the ones induced by the single voucher or advisory service treatments. Differential significant effects are detected also when vouchers are compared to technology and innovation advisory services. The policy mix outperforms voucher and advisory service treatments also in term of the number of workers (the sum of employees and self-employed workers). In this case positive differentials of 7 to 9 workers are estimated. Also in this case SMEs treated with vouchers outperform those that only received advisory services.

Results in term of per capita value added, which is a measure of labour productivity, are less conspicuous but still positive and significant. Both firms treated with the mix and the voucher programmes outperform firms treated with the advisory services programme only. Further, the policy mix implies a higher labour productivity than the voucher programme only. Considering the TFP outcome variable, the mean effect of the policy mix compared to the advisory service treatment induces up to 16 percentage points of additional TFP for participants in the mix with respect to recipients of advisory services only, and up to 9 percentage points of additional TFP for participants in mix with respect to voucher recipients only. Vouchers are significantly superior to advisory services in term of TFP by about 15 percentage points.

# CONCLUSION

Our analysis finds some support for the claim that the mix of the two interventions works better than each one taken individually. The policy mix outperforms the technology and innovation advisory service alone, and the voucher alone, on all four outcomes. The technology and innovation advisory service seems to engender specific knowledge within the SME that triggers a number of internal mechanisms, which, in turn, generate a higher level of firm performance. Our results in terms of performance appear to be consistent with the mechanisms discussed earlier, which had already been partly described by Shapira and Youtie (2016).

In terms of comparisons between single instruments, vouchers outperform technology and innovation advisory services on all four outcomes. It must also be noted that the most innovative firms (those that have participated in the policy mix, and their matched samples) particularly benefit from the policy mix, compared with vouchers alone or the technology and innovation advisory service alone. Also in the comparison between the two individual policies, the more innovative firms (those that have participated in the vouchers and their matched sample) benefit from vouchers more than from technology and innovation advisory services. Instead, the less innovative firms (those that used the technology and innovation advisory services only, and their matched samples) do not have any additional benefits from using vouchers or the policy mix.

The mix of innovation vouchers supported by the provision of technology and innovation advisory services, appears to be a promising innovation policy in regard to the increase of revenues and employment, but also of labour and total factor productivity. This however only holds for firms that were more innovative to begin with.

This is a preliminary study building upon a combination of policy programme administrative data and outcome variables derived from widely used company and statistical databases. A more fine-grained investigation where administrative information is complemented with variables derived from a survey of programme participants in all treatment groups, is currently under way.

# REFERENCES

**Bakhshi, H., Edwards, J. S., Roper, S., Scully, J., Shaw, D., Morley, L. and Rathbone, N.** (2015). Assessing an experimental approach to industrial policy evaluation: Applying RCT+ to the case of Creative Credits. Research Policy, 44(8), 1462-1472.

**Bessant, J. and Rush, H.** (1995). Building bridges for innovation: The role of consultants in technology transfer. Research Policy, 24(1), 97–114.

**BIGGAR Economics** (2010). Evaluation of the first phase of the Scottish funding council innovation voucher scheme, Scottish Funding Council.

**Bruhn, M., Karlan, D. and Schoar, A.** (2018). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. Journal of Political Economy, 126(2), 635-687.

**Caloffi, A., and Mariani, M.** (2011). Shaping regional policy responses: the design of innovation poles. Policy Studies 32(4), 413-428.

**Caloffi, A., Mariani, M., Rossi, F. and Russo, M.** (2018). A comparative evaluation of regional subsidies for collaborative and individual R&D in small and medium-sized enterprises. Research Policy, 47(8), 1437-1447.

**Coletti, M. and Landoni, P.** (2018). Collaborations for innovation: a meta-study of relevant typologies, governance and policies. Economics of Innovation and New Technology, 27(5-6), 493-509.

**Cornet, M.B., Roomen, V. and Van der Steeg, M.** (2006). Do innovation vouchers help SMEs to cross the bridge towards science? CBP Discussion Paper n. 58: CBP.

**Cunningham, P., Edler, J., Flanagan, K. and Larédo, P.** (2016). The Innovation Policy Mix. In:

**Flanagan, K., Uyarra, E.,and Laranja, M.** (2011). Reconceptualising the 'policy mix' for innovation. Research Policy 40(5), 702-713.

**Fontana, R., Geuna, A. and Matt, M.** (2006). Factors affecting university–industry R&D projects: The importance of searching, screening and signalling. Research Policy, 35(2), 309-323.

**Good, B. and Tiefenthaler, B.** (2011). Innovation voucher - small is beautiful. Plattform Forschungs- und Technologieevaluierung, Vienna, December.

**IEG** (2013). Evaluation of the World Bank Group's targeted support for small and medium enterprises (English). Independent Evaluation Group (IEG) approach paper. Washington DC: World Bank.

**Knockaert, M., Spithoven, A. and Clarysse, B.** (2014). The impact of technology intermediaries on firm cognitive capacity additionality. Technological Forecasting and Social Change, 81, 376–387.

**Lechner, M.** (2002a). Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. Journal of the Royal Statistical Society. Series A: Statistics in Society 165(1), 59-82.

**Lechner, M.** (2002b). Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies. Review of Economics and Statistics 84(2), 205-220.

**Ortega-Argilés, R., Vivarelli, M. and Voigt, P.** (2009). R&D in SMEs: a paradox? Small Business Economics, 33(1): 3-11.

**OECD** (2000).The Bologna Charter on SME Policies. Proceedings of the Bologna Conference.

**OECD** (2008). A review of local economic and employment development policy approaches in OECD countries. OECD, Paris.

**Magro, E. and Wilson, J. R.** (2013). Complex innovation policy systems: Towards an evaluation mix. Research Policy 42(9), 1647-1656.

**Martin, B. R.** (2016). R&D policy instruments – a critical review of what we do and don't know. Industry and Innovation 23(2), 157-176.

**Muller, E. and Zenker, A.** (2001). Business services as actors of knowledge transformation: The role of KIBS in regional and national innovation systems. Research Policy 30(9), 1501-1516.

**Russo, M., Caloffi, A., Rossi, F., Fiordelmondo, V., and Ghinoi, S.** (2015). Evaluating the performance of innovation intermediaries: insights from the experience of Tuscany's innovation poles. fteval Journal for Research and Technology Policy Evaluation 41, 1-6.

**Russo, M., Caloffi, A., Rossi, F. and Righi, R.** (2018). Innovation intermediaries and performance-based incentives: a case study of regional innovation poles. Science and Public Policy.

**Sala, A., Landoni, P. and Verganti, R.** (2016). Small and Medium Enterprises collaborations with knowledge intensive services: An explorative analysis of the impact of innovation vouchers. R & D Management 46, 291-302.

**Schade, S. and Grigore, C.** (2009). Availability and Focus on Innovation Voucher Schemes in European Regions, Survey Report prepared by DG ENTR-Unit D2 "Support for innovation", November.

**Shapira, P., and Youtie, J.** (2016). Impact of technology and innovation advisory services. In Edler, J., Cunningham, P., and Gök, A. (Eds.). Handbook of innovation policy impact. Edward Elgar Publishing.

**Storey, D. J.** (2003). Entrepreneurship, small and medium sized enterprises and public policies. In Acs, Z. J., &Audretsch, D. B. (Eds) Handbook of entrepreneurship research. Springer, Boston, MA, 473-511.

**Toivonen, M.** (2007). Innovation policy in services: The development of knowledge–intensive business services (KIBS) in Finland. Innovation 9(3-4), 249-261.

**Uotila, T., Harmaakorpi, V. and Hermans, R.** (2012). Finnish mosaic of regional innovation system – assessment of thematic regional innovation platforms based on related variety. European Planning Studies, 20(10), 1583–1602.

**Vossen, R.W.** (1998). Research note—Relative strengths and weaknesses of small firms in innovation. International Small Business Journal 16(3), 88–94.

Endnotes

i Firm-specific TFPs are estimated at industry level using the semi-parametric Levinsohn and Petrin (2003) approach and, subsequently, they are scaled with respect to industry mean TFPs and log transformed. Log-transformed TFPs (hereafter, TFPs) provide relative measures on how firm-specific productivities deviate from the industry means.

ii Estimates are available from the authors upon request.

# AUTHORS

**ANNALISA CALOFFI**
*Dipartimento di Scienze per l'Economia e l'Impresa, Università di Firenze*
Via delle Pandette, 23, 50127 Firenze (Italy)
E: annalisa.caloffi@unifi.it

**MARZIA FREO**
*Dipartimento di Scienze Statistiche "Paolo Fortunati"*
Via Belle Arti 41, Bologna (Italy)
E: marzia.freo@unibo.it

**STEFANO GHINOI**
*Department of Economics and Management, University of Helsinki*
Latokartanonkaari 5, P.O. Box 27 00014, Helsinki, Finland
E: stefano.ghinoi@helsinki.fi

**FEDERICA ROSSI**
*Department of Management, Birkbeck, University of London*
Malet Street, London WC1E 7HX (UK)
E: f.rossi @bbk.ac.uk

**MARGHERITA RUSSO**
*Dipartimento di Economia, Università di Modena e Reggio Emilia*
Via Berengario 51, Modena (Italy)
E: margherita.russo @unimore.it

# THE ITALIAN INDUSTRY/ENTERPRISE 4.0 PLAN: EX-ANTE IDENTIFICATION OF POTENTIAL BENEFICIARIES AND EX-POST ASSESSMENT OF THE USE OF INCENTIVES

GIULIO PERANI, STEFANO COSTA AND STEFANO DE SANTIS

DOI: 10.22163/fteval.2019.334

## 1. INTRODUCTION

The Italian Institute of Statistics (ISTAT) in addition to its role as member of the EU statistical system – i.e. official provider of economic and social statistical data and indicators – is also the Italian largest public research institution in charge of undertaking economic policy evaluation. The launching in 2016 by the Italian Ministry of Economic Development (MISE) of an innovative program of industrial policy, with a strong focus on the support of the digitalisation of Italian firms, has given ISTAT a unique chance to test original analytical and policy evaluation methodologies.

By following the German model of an Industry 4.0 platform (Rüßmann et al. 2015; ZEW 2015), MISE has developed a policy to support the digital transformation of the Italian business sector (National Plan "Industry 4.0", eventually become "Enterprise 4.0"), so acting more as a process enabler than a leading actor. The key measure of such policy is, in fact, an increase of the depreciation allowance for investment in machinery. According to this incentive scheme, the depreciation allowance, i.e. the amount a business can reduce its profit by when taxes are calculated, will be a percentage of the 140% (rather than 100%) of the purchase cost of industrial equipment, which will increase to 250% if investing in digitally connected equipment.

The nature of such measure has forced policy makers to develop new methods of ex-ante and ex-post policy assessment, as the influence of public incentives on firms' investments in a given fiscal year can be appraised only after the process of financial reporting is finalised and the tax statements filed by the concerned firms (i.e. at least six months after the end of the year), so that the feedbacks are available to policy-makers more than two years after the launching of the incentive policy.

The issue discussed in this study is to what extent a new set of integrated microdata developed at ISTAT, combining statistical and administrative (mostly fiscal) sources, could help (a) to identify, ex-ante, the potential beneficiaries of the tax incentives and (b) to assess, ex-post, the degree of success of such policy measures.

As the exercise has been limited to a single incentive, implemented over a two-year period, only cross sectional data are available which do not allow for a proper modelisation of the relationship between potential and actual beneficiaries. Nevertheless, a comparison between these two sub-populations of firms will highlight, beyond the overlapping rate, similarities and differences which could be used both for a fine tuning of the policy measures under evaluation and for improving an ex-ante identification of potential beneficiaries in future rounds of policies supporting firms' digitalisation.

The rest of the paper is organised as follows: section 2 describes the main characteristics of the Italian National Plan Industry 4.0. Section 3 gives evidence on the use of the I4.0 incentives by Italian firms according to the ISTAT survey on the usage of ICT in the business sector and some recent ad hoc surveys carried out by ISTAT and other institutions. Section 4 analyses the firms' propensity to the digital transformation by adopting an innovative 5-group classification of firms that takes into account both their degree of digitization and their endowment of productive factors. Section 5 shows how the propensity to use the I4.0 incentives is spread across the new classification. Section 6 estimates what factors (beside digitalisation) affect the use of fiscal incentives. Section 7 draws some conclusions.

## 2. THE ITALIAN PLAN INDUSTRY 4.0

Early in 2016 a new industrial policy was designed which led MISE to launch, in February 2017, the Industry 4.0 National Plan (I4.0). The new strategy was aimed at integrating some "vertical" measures (mainly focusing on the support to specific sectors or technological areas) with a range of "horizontal" measures (accessible to all firms) with the specific objective to boost the investment in new technologies, as well as in research and development, and to increase the competitiveness of Italian firms (MEF *et al.* 2017).

In this respect, three criteria have been adopted in designing the policy:

- to implement non-discriminatory measures, i.e. leaving to firms the choice of whether investing or not in new technologies;
- to use almost exclusively indirect incentives, mainly fiscal ones, in order to reduce the administrative burden associated to applications for direct funding;
- to leave firms the choice across a range of different support measures, taking advantage of one or more of the incentives made available by MISE.

The main goal was that of encouraging Italian firms – mostly those in the manufacturing sector – to replace outdated production equipment with new machinery which could be possibly integrated with advanced digital technologies such as robotics and automation, cloud computing, big data, sensors, 3D printers, etc. The main incentive introduced by MISE was an increased depreciation allowance of the cost of acquisition of machinery embodying 'Industry 4.0' technologies (Nascia and Pianta 2018). In the 2016 Budget Law a distinction was introduced between:

- the increased depreciation allowance for investment in new machinery (as a fixed percentage[1] of the 140%, rather than the standard 100%, of the cost of the purchased equipment: "super-depreciation") and
- a specific support to investment in Industry 4.0 technologies, i.e. digitally connected devices and related software and services (increased depreciation allowance as a fixed percentage of the 250%, rather than 100%, of the investment spending: "hyper-depreciation").

Such an approach was virtually unprecedented in Italy and, as such, uncertain about its chances to be successful. Additionally, it was soon realised that this specific MISE policy was extremely difficult to be properly monitored and evaluated.

Since all Italian firms were eligible for most of the incentives included in the Plan, the identification – ex-ante - of a specific 'target group' was only a matter of speculation. On the other hand, the use of indirect incentives made it impossible to know whether the acquisition of new technology by a firm had been undertaken with the intention to use the fiscal incentives or not, thus making any on-going monitoring of this measure almost impossible to undertake.

In order to overcome the issue, MISE, also in co-operation with ISTAT and other research institutions, has been actively investigating over the last two years the behaviour of Italian firms about their use of the incentives made available by the I4.0 Plan.

Statistical surveys have been the main tool chosen by MISE in order to get the information needed for designing and monitoring its I4.0 policies. They have included:

- the ISTAT business confidence survey 2017;
- the MET (a Rome-based private research centre of economic policy) survey 2017-2018 (MISE-MET 2018);
- the annual ISTAT survey on the usage of Information and Communication Technologies (ICT) by Italian enterprises, 2017.

# 3. EVIDENCES FROM STATISTICAL SURVEYS ON THE USE OF INCENTIVES

## THE ISTAT BUSINESS CONFIDENCE SURVEY 2017

A preliminary evidence on the attitude of Italian firms towards the I4.0 incentives, as well as their use almost two years after the launching, was collected through the ISTAT business confidence survey carried out in November 2017 (ISTAT 2018). The results of such a qualitative survey on a representative sample of around 4,000 manufacturing firms allowed for shedding light on two key issues:

a. the role of incentives in encouraging the firms to invest in new technologies during the period 2015-2017;
b. firms' intentions to further invest in I4.0 technologies in 2018.

As expected, the survey pointed out that the super-depreciation incentive had either a "high" or "moderate" role in influencing the previous years' investment of the majority of Italian manufacturing firms (62.1%) as an average between 57.3% of small enterprises (less than 50 persons employed) and 66,9% of large ones (over 250 persons employed). Hyper-depreciation has been quite influential too: it had a "high" or "moderate" role to convince 53.0% of firms to invest in digital technologies, ranging from 34.2% for small firms to 57.6% for large ones.

When asked about investments planned for 2018, almost 46% of the surveyed firms reported the intention to invest in new software, 31.9% in communication technologies ("machine-to-machine" or internet of things), 27% in data processing (cloud, mobile, big data etc.) and in IT security. Firms' size emerged, of course, as a key factor influencing investments even though the needs to keep firms up-to-date with the technological progress and to increase the employees' skills (also through the recruiting of new personnel) are additional investments' drivers both in small and in large firms.

## THE MET SURVEY 2017-2018

Another sample survey, covering a population of 23,700 Italian firms including micro-enterprises (less than 10 persons employed) and service firms, was conducted by MET a few months after the ISTAT confidence survey by asking similar questions on the use of I4.0 tax incentives (MISE-MET 2018; *Cassa Depositi e Prestiti* and MET 2018). The comparison with the ISTAT confidence survey is hardly possible as both the reference population and the scope of the survey (the MET survey covers many different topics) were different. According to the MET survey only 15.2% of Italian firms asked for super- or hyper-depreciation incentives. This figure is remarkably low and appears to be strongly influenced by the inclusion of micro-enterprises (whose average is 12.1%) in the sample. Larger firms behave differently when accessing the mentioned tax incentives which, according to the survey, are indeed used by 32.8% of small firms (10-49 persons employed) and by 47.5% of medium and large firms (50 persons employed or more). Overall, the MET figures are lower than those produced by ISTAT but it is confirmed that at least 50% of the largest firms should have profited of available incentives.

## THE 2017 ISTAT BUSINESS SURVEY ON THE ICT USAGE

In addition to the potential use of occasional surveys to collect data for policy monitoring purposes (as in the surveys described above), ISTAT identified the survey on ICT usage in businesses (ICT survey) as the key source to assess the level of digitalisation of Italian firms and, in relation to it, their propensity to use public incentives to increase their technological assets.

1 Such percentage is set by the Italian Ministry of Economy for each single economic activity and category of investment goods. For instance, according to the 1988 standard, still in force in 2018, for a firm in the automotive sector the depreciation allowance, to be calculated for fiscal purposes, is 20% of the purchasing cost of computers and office automation equipment, 25% of the cost of auto-vehicles, 30% of the cost of testing machines, etc.

The Italian ISTAT ICT survey is part of the Eurostat's *Community survey on ICT usage and e-commerce in enterprises*, conducted on an annual basis since 2002, which collects data on the use of information and communication technology and the access to Internet, as well as on e-government, e-business and e-commerce activities, by Italian firms. The scope of the survey includes firms with 10 or more persons employed belonging to a broad range of economic activities. The reference population of the 2017 ISTAT ICT survey included about 185,000 firms whose behaviour has been estimated on the answers given by a realised sample of around 20,000 firms.

The results of the ICT survey are fully integrated in a broader ISTAT database of business data, also including data from other statistical and administrative sources, thus giving a chance for developing new and integrated indicators on the relationships between ICT usage and other features of the firms' activities.

# 4. THE DEGREE OF DIGITALISATION OF ITALIAN FIRMS

In order to support the monitoring of the I4.0 policies, ISTAT has developed an analytical approach, based on the profiling of firms according to their propensity to invest in digital technologies, to gauge if they could be interested or not in accessing the key I4.0 incentives.

A number of economic indicators are indeed available for the whole population of Italian firms by matching basic economic indicators from the Structural Business Statistics (SBS) annual survey and the administrative data, including tax statements, cost statements, balance sheets and firms' reports regularly collected by ISTAT. The realised sample of the Italian ICT survey has been matched with other data sources at micro-level in order to undertake a profiling exercise based on three steps and aimed at classifying the Italian enterprises, with no less than 10 persons employed, according to their propensity to invest in new technologies and to use public incentives for such investments[2].

The **first step** has been that of classifying the firms by adopting the indicators identified by Eurostat as essential to describe their level of digitalisation[3]. The Eurostat's *Digital Intensity Index*, calculated at firm level, has been adopted as the basis for an evaluation of how intense the digital investment by Italian firms is. This classification is, indeed, only a partial one, as it does not include any measure about the size of the technological investment by firms as well as about their ability to fully exploit the potential, in terms of productivity growth, made available by the investment in new technologies.

Then, the **second step** was that of integrating in the classification exercise an additional set of three digital indicators (all of them derived from the 2017 ISTAT ICT survey) describing the orientation of firms to implement advanced digital technologies (with reference to the period 2014-2016): (1) Investments on Cloud Computing, Web applications or Big Data Analytics; (2) Investments on e-commerce, social media; (3) Investments on Internet of things, addictive printing, robotics, augmented/virtual reality.

The **third step** introduced in the analysis two structural indicators about the firm's endowment of productive factors: (1) the availability of fixed capital (measured in terms of the monetary value of tangible and intangible fixed asset per person employed) and (2) the availability of human capital (based on the education and job tenure of the firm's workforce)[4].

It is assumed that the propensity to digital transformation might be influenced by the actual availability of fixed and human capital. The endowment (both quantitative and qualitative) of the factors of production (capital and labour) and their distribution among the firms' business functions directly affects productivity but also, indirectly, the digitization strategies. The availability of these indicators allowed for a firms' classification by level of capitalisation and by level of staff qualification.

**Table 1.** Breakdown of the population of Italian firms with at least 10 employees by degree of digitalisation (percentage). Year 2017.

| Degree of digitalisation | Indicators on fixed capital and workforce | |
|---|---|---|
| | **Medium-low capitalisation / medium-low staff qualification** | **Medium-high capitalisation / medium-high staff qualification** |
| **Low** | 1. Analogue (64.6% of firms) | |
| **Medium** | 2. Potentially digital-oriented (20.7%) | 3. Partially digitalised (2.3%) |
| **High** | 4. Digital-oriented (9.4%) | 5. Fully digitalised (3.0%) |

The combination of the indicators calculated as a result of the three described steps allowed for a classification of Italian firms according to their propensity to digitalisation (or digital transformation, see Table 1).

As shown in Table 1, five groups of firms have been identified. To the first group, which includes 64.6% of the population, belong firms with a very low level of digitalisation. The peculiarity of such firms is that in the ISTAT ICT survey they stated that ICT investments are not relevant for their current business activity. They can be seen as being still **analogue** ones. This group includes more than 90% of small-sized firms (10-49 persons employed), with a high share of firms belonging to traditional industries (metal products, food products, textiles and clothing, leather, wood), construction, horeca and some business services. The second group, that of **potentially digital-oriented** firms (20.7% of the population), is apparently interested in extending its digital activities but

2 Measuring the level of digitalisation of a firm is a difficult task and the proposed example is mostly a contribution to the literature on this topic (see Bley et al. 2016).

3 The Digital Intensity Index (DII) is a micro-based index developed by Eurostat to contribute to the EU Commission's monitoring of digital progress that measures the availability at firm level of 12 different digital technologies: (1) Percentage of employees connecting to Internet by PC; (2) Percentage of employees connecting to Internet by mobile devices; (3) Hiring of ICT consultants; (4) Internet connection's average speed; (5) Corporate web-site available; (6) E-commerce available on the corporate web-site; (7) Use of social media; (8) Intensive use of cloud computing; (9) Electronic invoicing; (10) Web advertising; (11) Percentage of online sales on total sales higher than 1%; (12) Percentage of online B2C sales higher than 10% of total online sales. The value for the index therefore ranges from 0 (the firm uses none of previous technologies) to 12 (it uses all of them). The DII is a component of the data scoreboard used in the Europe's Digital progress Report (https://ec.europa.eu/digital-single-market/en/news/europes-digital-progress-report-2017 ).

4 A detailed definition of this methodology can be found in ISTAT (2018).

this process may be hindered by low levels of capitalisation and human capital's qualification. Both SMEs and large firms are included in this group, mainly dealing with trade and manufacturing. The third group is that of the **partially digitalised** firms (2.3% of the population) which are units not yet able to complete their process of digitalisation despite their large availability of fixed and human capital. It is indeed quite a small group of firms, mainly belonging to trade and other service industries. The fourth group includes the **digital-oriented** firms (9.4% of the population) which are the largest share of firms with a high level of digitisation, but low levels of capitalisation and quality of the workforce. This group outperforms the other groups in terms of profitability. Finally, only 3% of the Italian firms with 10 persons employed or more are **fully digitalised**. Not surprisingly, they are the best performers in terms of productivity because of an effective combination of capital and labour.

### THE PROPENSITY TO USE THE INCENTIVES

It could be assumed that the higher the level of digitalisation, the higher the propensity of enterprises to use the incentives made available by the I4.0 Plan. In this perspective, the digitalisation can be seen as a process that builds upon itself by making available assets and competencies to allow for a constant grow. On the other hand, it is also a matter of fact that those firms that are not yet fully digitalised have a higher pressure, and potential, to catch up by using any available support.

These assumptions can be preliminarily tested by taking into consideration the answers given by the firms to the ICT 2017 survey's question about the three most important factors that could make digitalisation a driver of competitiveness and growth. Around 46% of firms with at least 10 persons employed consider **public incentives** one of the most important factors supporting the digital transformation but the impact of incentives is differently rated according to the level of digitalisation of recipients. By using the classification described above, Figure 1 provides for a comparison between the firms' attitudes towards public incentives to digitalisation by level of digitalisation. The partially digitalised firms, i.e. those which have in their medium-high level of capitalisation and/or qualification of employees a key driver to foster the digitalisation processes, are the most interested to receive public incentives (more than 75% of them). Also the fully digitalised firms are eager to further invest in technology with the support of public incentives (67.3%). The digital-oriented and the potentially digital-oriented firms are considering incentives less relevant even though the potential beneficiaries within these groups are numerous (respectively, 61% and 60%). Apparently, the availability of incentives cannot change the attitude of "analogue" firms towards the investment in digital assets and processes: only 36.9% of them consider public incentives and funding effective.

By interpreting the results of the 2017 ICT survey as an evidence to be used for an **ex-ante assessment** of the potential use of fiscal incentives made available by MISE for investments in new machinery and technologies (super-depreciation) or in digital technologies (hyper-depreciation) some hypotheses could be done[5].

1. With a specific reference to the acquisition of digital technologies in order to increase the level of digitalisation of a firm, public incentives seem to have higher impact on firms **partially or fully digitalised** – thus, the current degree of digitalisation could have a role in fostering investment in digital technologies.
2. As far as the asset availability is concerned, a **medium-high level of capitalisation** could encourage firms to exploit the support made available by the government to undergo a digital transformation.

# 5. USE OF FISCAL INCENTIVES: THE EVIDENCE

The results from statistical sources used to estimate the use of I4.0 incentives for the years 2015-2016 have been made available between October 2017 and March 2018. Late in spring 2018, the fiscal data on the actual use of the I4.0 supporting measures were released, for statistical and analytical purposes, by the Italian Tax Agency.

Fiscal data partially confirm statistical estimates[6].

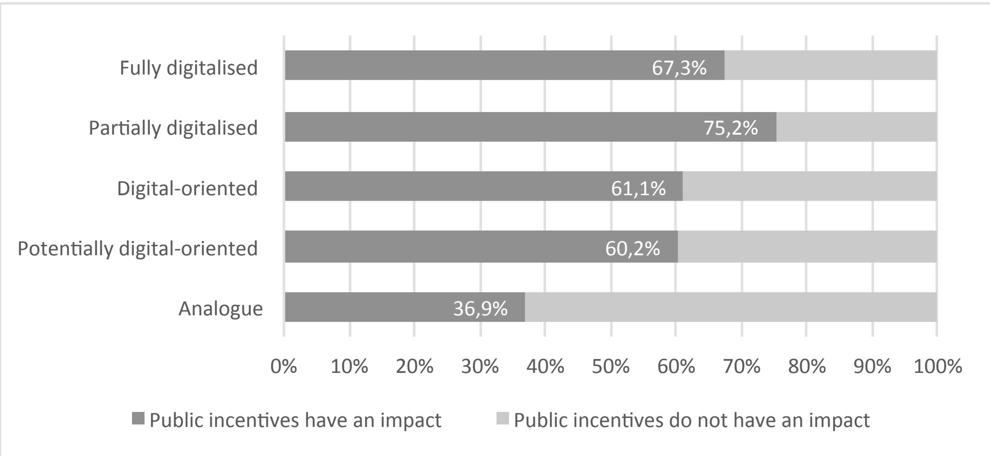



**Figure 1**. Firms' attitude towards public incentives to digitalisation. Italian ICT survey 2017.

5 Similar results can be found in the analysis by Centro Studi Confindustria (CSC 2018) based on the same set of data.

6 The analysis described in this paper is based on the tax statements of the firms surveyed by the ISTAT 2017 ICT survey in order to allow for comparing statistical and administrative (fiscal) data. In this respect, also sampled fiscal data, appropriately weighted, are representative of those of the population of Italian firms with at least 10 employees.

As shown in Table 2, the partially digitalised firms have indeed a leading role in the use of fiscal incentives for the acquisition of new technology (61.9% for super-depreciation and 5.9% for hyper-depreciation) even though at a lower level than estimated by the 2017 ICT survey (75% as a combination of both measures).

Not so far, in terms of percentage of beneficiaries, are the digital-oriented and the potentially digital-oriented firms (54.2% and 4.4% for the earlier and 58.8% and 3.6% for the latter, both above the average) by highlighting the role of public incentives to help firms to overcome financial and organisational barriers to technological innovation.

About the extreme cases, the fully digitalised firms seem interested in improving their technological capacity (54.1% used super-depreciation) but much less oriented (or needed) to get more digital equipment and software (only 2.6% used hyper-depreciation). Finally, the analogue firms confirm to be relatively reluctant to invest in new technology (45.7% used super-depreciation, 2.6% hyper-depreciation). An interesting point is that analogue firms are the only group that was under-estimated in the 2017 ICT survey (Figure 1) about its intention to use fiscal incentives.

**Table 2.** Percentage of Italian firms using I4.0 incentives for hyper- or super-depreciation. October 2015-December 2016. By digital intensity.

| | Percentage of firms investing in digital technologies (hyper-depreciation) | Percentage of firms investing in new machinery and technologies (super-depreciation) |
|---|---|---|
| Analogue | 2.6 | 45.7 |
| Potentially digital-oriented | 3.6 | 58.8 |
| Digital-oriented | 4.4 | 54.2 |
| Partially digitalised | 5.9 | 61.9 |
| Fully digitalised | 2.6 | 54.1 |
| All | 3.1 | 49.9 |

The information given in Table 2, focusing on percentages of firms using fiscal incentives, should be qualified by considering the actual size of the investments funded through hyper- and super-depreciation.

In Table 3, the average I4.0 annual depreciation per employee is displayed by comparing the five groups of firms by digital intensity (only applicant firms for the concerned incentive have been taken into consideration). Overall, it could be noticed that an average annual super-depreciation (roughly 140% of standard depreciation) of about 400 euros per employee corresponds to a total investment – for a firm with 100 employees, in a five year time-span – of around 600,000 euros: a substantial amount of money but not really sufficient to support a full restructuring of either a goods or services production line. In this perspective it has to be pointed out that partially digitalised firms, which have the highest percentage of incentives' use, also have, by large, the lowest average per employee expenditure in new technology (263 euros of hyper-depreciation and even 194 euros of super-depreciation).

The fully digitalised firms (at least those using fiscal incentives) have profited more than other groups by these measures with a yearly average of 824 euros of hyper-depreciation and 974 euros of super-depreciation per employee[7].

**Table 3.** Average yearly hyper- or super-depreciation per employee. October 2015-December 2016. By digital intensity.

| | Average yearly hyper-depreciation per employee | Average yearly super-depreciation per employee |
|---|---|---|
| Analogue | 349.2 | 281.3 |
| Potentially digital-oriented | 266.3 | 296.8 |
| Digital-oriented | 352.1 | 632.5 |
| Partially digitalised | 263.2 | 194.5 |
| Fully digitalised | 824.0 | 973.8 |
| All | 348.3 | 408.9 |

Additional information is needed to provide an assessment of the impact of the investments funded through the I4.0 Plan on the fixed capital of the beneficiaries. Preliminary evidence, shown in Table 4 (with reference to the same applicant firms as for Table 3), suggests that this impact could be fairly relevant, although not always sufficient to substantially increase the current capital per employee ratio. When the depreciation reported in the tax statements is compared to the current working capital per employee ratio, a net increase of the latter can be calculated ranging from 0.9% to 6.8% assuming, as an average investment, the acquisition of PCs or similar devices (depreciation coefficient of 20%, over 5 years, with hyper-depreciation). On the other hand, if the super-depreciation scheme applies, the increase of the working capital per employee ratio – for the same standard purchase – ranges from 3.8% to 11.7%.

**Table 4.** Average percentage increase of working capital per employee for investments with hyper- or super-depreciation.
October 2015-December 2016. By digital intensity.

| | Average % increase of working capital per employee (hyper-depreciation) | Average % increase of working capital per employee (super-depreciation) |
|---|---|---|
| Analogue | 3.5 | 4.6 |
| Potentially digital-oriented | 5.0 | 11.7 |
| Digital-oriented | 5.4 | 7.1 |
| Partially digitalised | 0.9 | 3.8 |
| Fully digitalised | 6.8 | 6.4 |
| All | 2.6 | 5.7 |

Some preliminary findings can be drawn from the evidence shown in the tables above. First, the groups of firms by digital intensity are ranked, in terms of actual incentives' use, consistently with the ex-ante estimations although forecasted and actual percentages of beneficiaries do not match exactly. From this perspective, both the ex-ante assessment and the on-going monitoring by using statistical sources have been quite successful.

Second, what was not possible to gauge from the available statistical sources is any forecast about how much investment firms were eager to make. A general comment, based on the available evidence, is that the impact of the I4.0 investments on the endowment of working capital and, even more, on the level of digitalisation, of Italian firms has been, on average, substantial but not disruptive.

7 Thus, on average, a firm in this group was allowed to claim for a tax deduction in 2016 of 1,798 euros rather than the standard deduction of 1,025 euros.

# 6. WHICH FACTORS AFFECT THE USE OF FISCAL INCENTIVES?

## THE ROLE OF DIGITALISATION

Evidence collected so far about the use of I4.0 fiscal incentives reflects, as expected, the high heterogeneity of the Italian business sector. By splitting the population of firms with at least ten persons employed, in five groups by digital intensity such heterogeneity has been partially reduced, as only some very key features (mostly technologically related ones) of firms' activity have been considered in profiling them. This does not exclude that other factors could have affected the firms' strategy as far as the investment in new technologies and the use of fiscal incentives to increase it are concerned.

A multiple regression analysis has been performed in order to compare the propensity of the five digital groups to use fiscal incentives (dependent variable) by excluding any spurious effect due to additional firms' characteristics (independent variables): productivity (value added per employee), capital intensity (working capital per employee), financial leverage (debt to capital ratio), vertical integration (value added/turnover), size (number of persons employed), job tenure of employees (years, average), education of employees (years of study, average), economic activity (2-digit NACE sectors), firm's age (years) and exporter status.

In Figure 2, the propensity to use the super-depreciation in a given measure is estimated for four groups of firms by digital intensity taking the group of analogue firms as a benchmark and any additional feature on a *ceteris paribus* basis[8].

The results of the regression confirm that the level of digitalisation and, even more, quantity and quality of fixed capital and human quality, affect the choices of firms about whether investing in new technologies – then using the available incentives – irrespective of size, economic activity or other features. The groups of firms shown in Figure 2, all outperform the benchmark group of analogue firms but, more relevant, the groups with medium-high capitalisation and medium-high workforce qualification have a remarkable advantage in implementing an innovation strategy with public support.

A similar analysis[9] is shown in Figure 3 with reference to the use of hyper-depreciation, i.e. the use of fiscal incentives to support investments in digital technologies. The overall propensity to use incentives is lower than for the super-depreciation but the pattern of the relationship between the firms' groups is almost the same. In addition to a minor role of analogue firms, those firms with higher capital intensity and workforce qualification display a higher propensity to use the fiscal incentives to digitalisation. It means that, *ceteris paribus*, a fully digitalised firm will use hyper-depreciation by a factor six times higher than a digital-oriented firm. In this case, the level of digitalisation plays also a role by giving a small advantage to digital-oriented firms on potentially digital-oriented firms and to fully digitalised firms on partially digitalised ones.

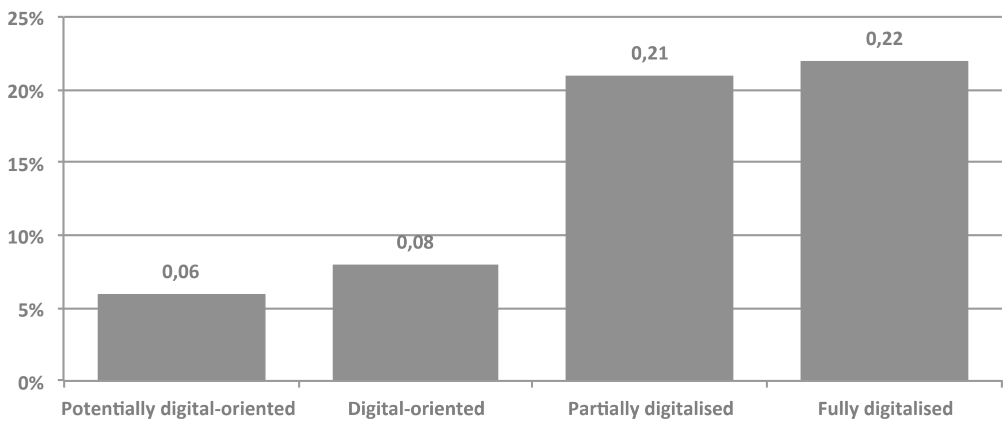



**Figure 2.** Propensity to use the super-depreciation incentive.
October 2015-December 2016.

8 Detailed information on the regression can be provided by the authors upon request.
9 Ibidem as footnote 3.

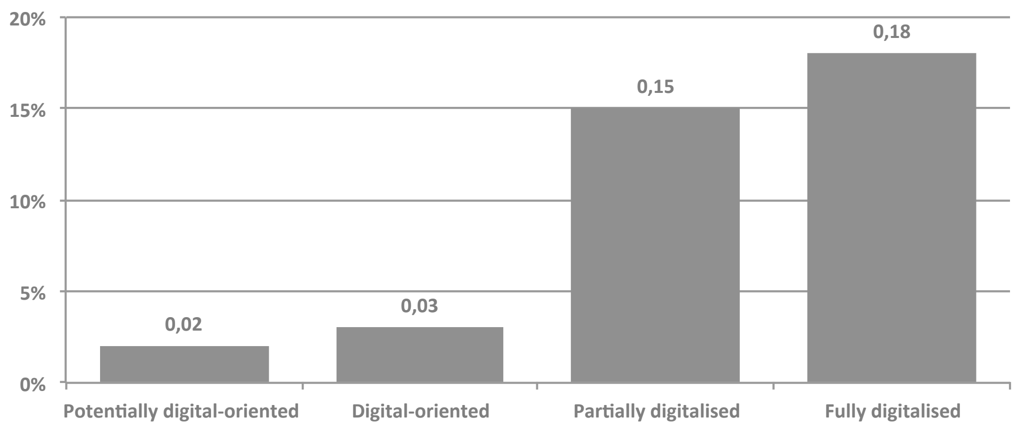


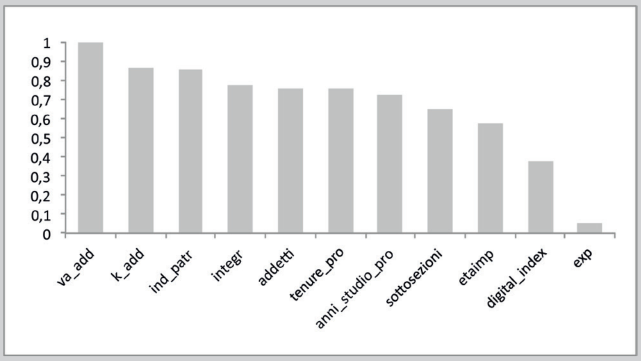

**Figure 3.** Propensity to use the hyper-depreciation incentive.
October 2015-December 2016.

These findings confirm the relevance of the classification by digital intensity proposed in the paper and provide for a new standard in the development of indicators on the digitalisation of the business sector: that of combining data on the use of digital technologies with information on the ability of the firms to effectively use such technologies, i.e. having developed both an appropriate infrastructure and the needed competences. Without these conditions to be fulfilled, even substantial incentives given by the public sector could be ineffective to support the firms' digital transformation.

## BEYOND DIGITALISATION

A further analysis, based on a random forest regression[10], allows for preliminarily exploring the role of non-digital factors to support the digitalisation of Italian firms through the use of fiscal incentives.

In Figure 4 and in Figure 5, a number of factors are ranked according to their role on framing a context where fiscal incentives can be effectively used by firms. The analysis considers the effect of each factor separately, thus avoiding any combined effect which could have influenced the data presented in previous paragraphs.




**Legend**

| Variable code | Variable name |
|---|---|
| va_add | Productivity (value added /p.e.) |
| k_add | Working capital per employee |
| ind_patr | Debt to capital ratio |
| integr | Vertical integration |
| addetti | Persons employed |
| tenure_pro | Tenure of employees (years, average) |
| anni_studio_pro | Years of study of employees (av) |
| sottosezioni | Economic activity |
| etaimp | Firm's age (years) |
| digital_index | Digital intensity (5 groups) |
| exp | Exporter status |

**Figure 4.** Factors moderating the use of fiscal incentives for investment in new machinery (super-depreciation). Year 2016.

10 Ibidem as footnote 3 .

Moreover, those data and analyses have emphasized the role of digital technologies as enablers of the adoption of more digital procedures and associated devices: a dimension not relevant in this new perspective. Finally, random forest is a machine learning algorithm that is not based on a pre-defined model about the role of each factor (variable) or the relationships among them but explores the moderating effect of each factor by selecting it randomly. This approach is very effective in a context where complex interactions among factors can be assumed and any information about their respective role is lacking.

In terms of results, both Figure 4 and Figure 5 show that the key factor influencing the propensity to invest in new technologies, thus to use fiscal incentives to do it, is the labour productivity. The more a firm is productive, the more it has an incentive to further increase efficiency and competitiveness.

Three additional factors strongly influencing the use of fiscal incentives are of structural nature: capital per employee, debt-to-capital ratio and vertical integration ratio. They are, respectively, the second, third and fourth most important factors to affect the use of super-depreciation and the third, fourth and fifth as far as the hyper-depreciation is concerned.

The most striking difference between the two incentives is about the role of the firms' size. Size is the second most important factor for the hyper-depreciation and the fifth for the super-depreciation. It seems that size is a significant condition to undergo a process of digitalisation with relevant investments in new technologies[11].

The quality of the workforce – both in terms of level of education and seniority at work – is also important for accessing both the incentives, as well as the economic activity.

Least relevant are three factors (for both incentives): the age of the firm, the level of digitalisation and the export propensity.

# 7. CONCLUSIONS

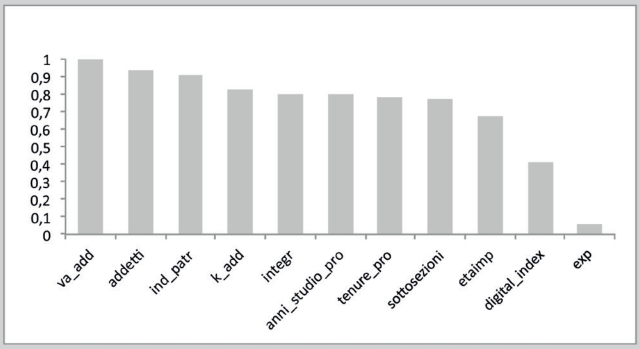
With reference to the question about the ability of statistical systems to provide for a useful knowledge base for designing effective fiscal policies in order to support innovation and digitalisation in the business sector, the answer from this paper is substantially positive. A detailed profiling of firms could allow for the ex-ante identification of groups of potential beneficiaries although additional work has to be done in order to develop suitable methods to improve estimations on the number of potential beneficiaries and on the amount of incentives potentially requested.

Another key issue is that of combining structural, financial and technological variables to identify the key factors enabling a firm to invest in technological innovation or, more specifically, in the digital transformation. Of course, the availability of digital competences is an essential asset but an innovation strategy that includes the acquisition of advanced technologies can be afforded only by an efficient firm with a high level of productivity, high quality workforce and which would be financially sound.

Three main findings of this study can be pointed out:

- The level of digitalisation does not affect the access to incentives, as a consequence, it does not affect the level of investment in new technology.
- Monitoring the use of incentives with surveys is a good starting point but survey results are clearly biased by an optimistic attitude of respondents.
- Technical, financial and human capabilities are the key factors boosting investment in new technologies.

To the extent some preliminary policy lessons could be drawn from the findings above, a few points have to be highlighted.




**Legend**

| Variable code | Variable name |
|---|---|
| va_add | Productivity (value added /p.e.) |
| k_add | Working capital per employee |
| ind_patr | Debt to capital ratio |
| integr | Vertical integration |
| addetti | Persons employed |
| tenure_pro | Tenure of employees (years, average) |
| anni_studio_pro | Years of study of employees (av) |
| sottosezioni | Economic activity |
| etaimp | Firm's age (years) |
| digital_index | Digital intensity (5 groups) |
| exp | Exporter status |

**Figure 5.** Factors moderating the use of fiscal incentives for investment in digital technologies (hyper-depreciation). Year 2016.

11 The relevance of firms' size in increasing the propensity to adopt new technologies is often emphasized by the literature (see Arvanitis and Hollestein 2001).

Digitalisation targets have to realistic and suitable for groups of firms with a very heterogeneous digital and productive structure.

The risk of opening up the access to fiscal (automatic) incentives to every firm is that such incentives could be used as a complementary source of funding for large firms already substantially investing in their digital transformation but, at the same time, as an occasional chance to support the acquisition of selected equipment by small firms or firms without a digitalisation strategy. This reduces substantially the potential impact of the I4.0 measures and call for an urgent intervention.

This raises the issue whether public support had to be mainly given to firms (mostly SMEs) only "potentially" digitalised by adopting the implementation of a two-steps approach: first, supporting the development of capabilities, then funding the digitalisation process.

# REFERENCES

**Akaike H.** (1974). A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19(6), 716-723.

**Arvanitis S and Hollenstein H.** (2001). The Determinants Of The Adoption Of Advanced Manufacturing Technology, Economics of Innovation and New Technology, 10:5, 377-414.

**Bley, K., Leyh, C. and Schäffer, T.** (2016). Digitization of German Enterprises in the Production Sector-Do they know how "digitized" they are? Paper presented at the Twenty-second Americas Conference on Information Systems, San Diego, 2016. https://www.researchgate.net/profile/Christian_Leyh/publication/305661673_Digitization_of_German_Enterprises_in_the_Production_Sector_-_Do_they_know_how_digitized_they_are/links/5798c14108aed51475e87572.pdf

**Cassa Depositi e Prestiti – MET** (2018). Il sistema produttivo italiano. Tra modernizzazione e Industria 4.0, Quaderni CDP, n.3.

**Centro Studi Confindustria** (2018). Imprese e politica insieme per l'industria italiana 4.0, Note CSC, n. 18/03.

**Götz M. and Jankowska B.** (2017). Clusters and Industry 4.0 – do they fit together?, European Planning Studies, DOI: 10.1080/09654313.2017.1327037.

**ISTAT** (2018), Rapporto sulla competitività dei settori produttivi, ISTAT, Rome www.istat.it/it/archivio/212438.

**MEF (Ministero dell'economia), MISE, MIUR, ML** (2017). Piano nazionale IMPRESA 4.0 Risultati 2017-linee guida 2018, September 2017, http://www.sviluppoeconomico.gov.it/images/stories/documenti/impresa_40_risultati_2017_azioni%202018_rev_eng.pdf.

**MISE – MET** (2018). La diffusione delle imprese 4.0 e le politiche: evidenze 2017, Rome, Italian Ministry for Economic Development.

**Neubig, T. et al.** (2016). Fiscal incentives for R&D and innovation in a diverse world, OECD Taxation Working Papers, No.27, OECD Publishing, Paris.

**Nascia L and M. Pianta** (2018). Research and innovation policy in Italy, Economia Politica, October 2018.

**Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P. and M. Harnisch,** (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group, 9.

**Sauter, R., Bode, M. and Kittelberger, D.** (2015). How Industry 4.0 is changing how we manage value creation. Horváth & Partners White Paper, https://www.horvath-partners.com/en/publications/featured-articles-interviews/detail/how-industry-40-is-changing-how-we-manage-value-creation

**Tao, F., Qi, Q., Liu, A. and Kusiak, A.** (2018). Data-driven smart manufacturing. Journal of Manufacturing Systems, 48/C, 157-169.

**ZEW** (2015), Industrie 4.0: Digitale (R)Evolution der Wirtschaft, ZEW IKT-Report Oktober 2015.

# AUTHORS

**GIULIO PERANI**
E: Giulio.Perani@istat.it

**STEFANO COSTA**
E: Stefano.Costa@istat.it

**STEFANO DE SANTIS**
E: Stefano.Desantis@istat.it

*Data analysis and economic, social and environmental research, ISTAT*
Via Cesare Balbo 16, Rome, 00184 (Italy)

# IMPACT EVALUATION OF R&D SUPPORT FOR SMES AND START-UPS AND ITS FEEDBACK ON PROJECT MANAGEMENT

MASAKI UEYAMA, MOTOSHI KUNUGI, TOSHIYUKI ISSHIKI, SHUMPEI MIYAJIMA AND SHIN UESAKA

DOI: 10.22163/fteval.2019.335

## INTRODUCTION

### ABOUT NEDO

Following the two oil crises of the 1970s, the need for energy diversification increased. Against this backdrop, the New Energy Development Organization was established as a governmental organization in 1980 to promote the development and introduction of new energy technologies.

Research and development of industrial technology was added in 1988, and today New Energy and Industrial Technology Development Organization (NEDO) plays an important part in Japan's economic and industrial policies as one of the largest public research and development management organizations with an annual budget for FY2018 of 159.6 billion yen (1.23 billion euro). It has two basic missions: addressing energy and global environmental problems, and enhancing industrial technology.

Drawing on its considerable management know-how, NEDO carries out projects to explore future technology seeds as well as mid- to long-term projects that form the basis of industrial development. It also supports research related to practical application.

### EVALUATION SYSTEM IN NEDO

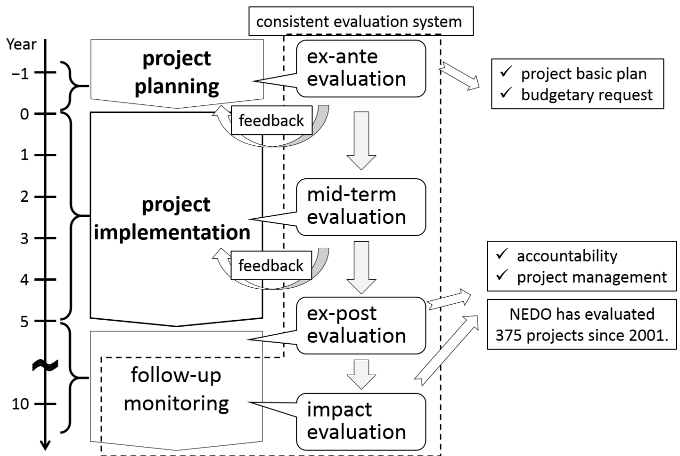
NEDO has established and been applying its own evaluation system for nearly two decades. Figure 1 shows the overall scheme of present NEDO evaluation and monitoring, at various stage of a 5-year project. Starting from the project planning stage, we have a set of four evaluation chances for each project.

1. Ex-ante evaluation, that is performed when it is still at the planning stage, to see how the project is worth being carried out. The results of an ex-ante evaluation are fed back for refining the project plan and requesting the final budget scheme.
2. Mid-term evaluation, that is performed typically once for a project, and the results are directly reflected to the management of the project for the rest of the period.
3. Just after the project is finished, an ex-post evaluation is performed. The results of ex-post evaluation are often used as a reference to planning of a related new project. For each mid-term and ex-post evaluation, an external subcommittee of typically 7 members, is organized.
4. In addition to these 3-step evaluation series, NEDO conducts follow-up monitoring in every other year for 6 years. Follow-up monitoring is done by NEDO evaluation department, supervised by an external specific subcommittee, using questionnaire and interview method. This survey is necessary for the impact evaluation, which assesses the post-project development by the participant companies, and the resulting impact of the project to the society.

Follow-up monitoring directly determines the present status of the project ranked as a 5-level TRL-like stage; 1) still under elementary research, 2) technology development, 3) already practically applied, 4) successfully commercialized, or 5) unfortunately terminated (abandoned). By applying this to all NEDO Projects, the success rate (expected probability of success) of NEDO projects will be estimated. The detail of follow-up monitoring is also described in the next section.

The overall results of ex-post and impact evaluation of all NEDO projects are then used for the accountability for tax payers, and for improving the project management system in general.

**Figure 1**. Overview of NEDO project evaluation



## EXTENDED FOLLOW-UP SURVEY

In 2009, we started an additional "extended follow-up survey" for selected NEDO projects that successfully created new innovative outputs - products or processes - utilizing core technologies that are developed by the project. We name these outputs "NEDO-inside products" (Yamashita et al. 2013), and a total number of 115 are registered at present. The extended follow-up survey is continued even after the end of the first 6-year monitoring period, and the data are used for estimating key indicators such as sales, return on investment and societal benefits by each product. Combining the results of follow-up monitoring and extended follow-up survey enables us to assess and disseminate the economic and societal impacts of NEDO R&D impacts, and then to reflect the knowledge in the improvement of project policy and management through success / failure factor analysis.

## NEDO'S SUPPORT FOR SMES INCLUDING START-UPS

In recent years, NEDO is focusing not only in promoting large-scale national projects based on national roadmaps, but also in supporting R&D of "small- and medium-sized enterprises ("SMEs" hereafter) including start-ups. The definition of SMEs and start-ups is shown in the next section.

This relatively new strategy of NEDO is set due to the fact that, in general, (1) faster development is expected by SMEs than by large companies, (2) innovation is liable to occur in so-called niche areas, where large companies dare not intend to do, and (3) societal impacts such as indirect economic effects may appear more directly on SMEs than on large companies. Some recent research (Farja, Y., Gimmon, E., Greenberg, Z. (2017), Foreman-Peck, (2013), Radas, S., Anic, I-D., Tafro, A., Wagner, V. (2015)) done in various countries shows, in general, that funding via subsidies is more effective and efficient for supporting innovations by SMEs than other measures such as tax incentives. From this viewpoint, a series of NEDO funding scheme have been reorganized to seamlessly support SMEs according to their present R&D phase (feasibility study, fundamental, development etc.).

Because the average size of R&D activities of SMEs is relatively small compared to that of large companies, the impact of their R&D onto the whole society is unlikely to show up clearly. On the contrary, as for the impact on the SMEs themselves, it is expected that the R&D results will have a greater impact on the growth and survival of the company itself, than in the case of large companies.

In this study, we used the data from follow-up monitoring and extended follow-up survey and analyzed three aspects as follows.

1. Commercialization rate
2. Success / Failure factors
3. Effects on the participating SMEs

# COMMERCIALIZATION RATE

SMEs (including start-ups) are defined, under the Small and Medium-sized Enterprise Basic Act of Japan (1963), as private sectors that fulfil either condition of the following.

**Table 1.** Definition of SMEs.

| Capital Stock | Number of employees |
|---|---|
| Not more than 300million yen* | Not more than 300 |

*approximately 2.3 million euro

There is no universally quantitative definition of "start-up companies", and in this study we conveniently set a start-up as an SME which is less than ten years old.

First, we checked the commercialization rate ("success rate") of SMEs using our follow-up monitoring data, and saw if it is significantly different from that of total commercialization rate of all NEDO projects including large companies.

Follow-up monitoring is done for all organizations that participated in NEDO projects (ca. 800 / year), at 1/2/4/6 years after the termination of each project. Web-based questionnaires set for the monitoring consists of four parts:

I. Present status of the post-project activities- using status of R&D subject ranked as a 5-level stage (TRL-like) defined above.
II. Possible factors of success or failure (Why success / failure?)
III. How was the project management provided by NEDO?
IV. Objective of participating in the project (process improvement, new business etc.)

The answers to these questionnaires are used not only for estimating the overall success rate of a certain group of projects (projects with SMEs in this case), but also to ensure accountability of the funding policy, to improve NEDO's project management and to assess social impact of the projects.

We analyzed 837 NEDO projects in which SMEs participated by applying the above mentioned viewpoints, using the data of follow-up monitoring and extended follow-up surveys. If the present status of the post-project activities falls into either 3) practical application or 4) commercialized of the 5-level stage, it is counted as a "success".

**Table 2.** Success rate of SMEs.

| Category | The number of projects (The number of companies) | The number of successful projects (The number of successful companies) | Success rate |
|---|---|---|---|
| SMEs excl. start-ups | 445 (351) | 150 (132) | 33.7% |
| Start-ups | 392 (293) | 129 (114) | 32.9% |
| Total SMEs | 837 (644) | 279 (246) | 33.3% |

Our overall results in Table 2 showed, SMEs achieve a practical application rate of around 33 %, which is remarkably higher than the average value of 25 % for all NEDO projects including those done by large companies. This result is consistent with other research for SMEs in other region of the world (Office of Extramural Research, National Institutes of Health, (2009); SQW Ltd. (2015)).

# SUCCESS/FAILURE FACTOR ANALYSIS

As an extended follow-up survey for this study, we conducted a series of individual interviews for 30 chosen SMEs that reached "success" stage with excellent results. Our preceding research (Kunugi et al. 2016) revealed some key factors leading to discontinuing / resuming projects, and further accumulation of data was utilized to analyse SMEs in this study.

The interview in this study consists of four parts:

I. Status of R&D results, practical application and commercialization
II. Specific activities taken by the company to achieve the results
III. Whether the company had enough resources to proceed those activities effectively
IV. Actions taken to complement resource deficit / to make good use of present resource

The entire set of interview answers are analyzed by extracting common tendencies and differences between companies. We found four tendencies summarised below.

a. Thorough ex-ante knowledge on the business environment and the strengths of the company: target customers, market/ technology region
b. Securing the resources: from both inside and outside of the company, including effective sharing of the resources
c. Adjustment by judging the change of the environment: continuous survey and search for the output market
d. Continuous effort for resources: resources are continuously needed after reaching practical application stage, for manufacturing and sales activities

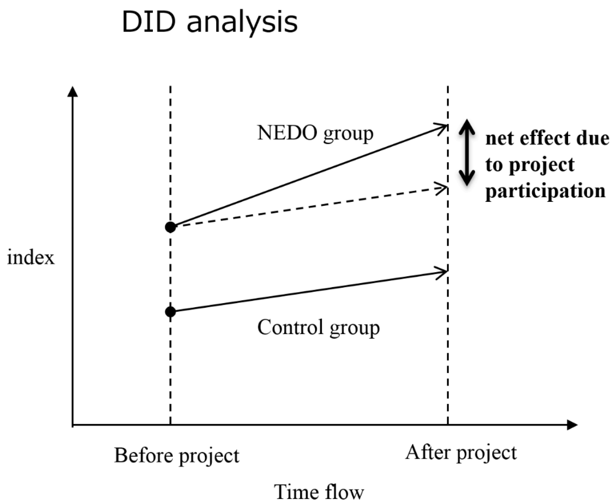
# EFFECTS OF PROJECTS ON THE PARTICIPATING SMES BY DID ANALYSIS

In order to obtain a reliable estimate of the effect of NEDO projects on participating SMEs, it is not enough to analyze only data for the companies who did participate the project, as this would not eliminate the external effects such as macroscopic economic trends on the results.

A Difference-in-differences (DID) analysis was also conducted in this study accordingly. Recent reports of DID analysis applied to the evaluation (Foreman-Peck, (2013), Ministry of Economic Development, New Zealand, (2011)) shows its reliability compared to traditional methods such as case studies, which tend to overestimate the additionality measurement.

In the DID analysis, a group of companies that participated in NEDO projects ("the NEDO group") and a group of companies that did not participate ("the control group") but similar to the NEDO group in terms of other attributes (e.g. sales, number of employees, type of business and region) were selected and examined. Details of the method we adopted in this study are described in Inoue, H. and Yamaguchi, E. (2017).

**Figure 2:** Schematic diagram of the DID analysis.




The population of our DID analysis comprises two groups: "NEDO group" that started a NEDO project between years 2007 and 2010, and the control group. Each company in the control group set is chosen for a corresponding "NEDO group" company, by comparing the location, type of industry, sales amount and number of employees.

Out of 442 "NEDO group" original companies, we found control companies and used the pair for the analysis.

The change in "performance indicators" such as sales amount and number of employees within six years - from the year each project begins and six years after that - was estimated for each company, and was then compared for the NEDO group and corresponding control group.

Average sales of each group for both year 0 and year 6 and their growth rate is shown in Table 3 and 4, respectively.

**Table 3.** Average sales amount of each group (million yen).

| NEDO group year 0 | NEDO group year 6 | Control group year 0 | Control group year 6 |
|---|---|---|---|
| 2,688 | 2,956 | 2,502 | 2,573 |

**Table 4.** Average sales increase rate of each group.

| | NEDO group | Control group |
|---|---|---|
| Average sales increase between year 0 and 6 | 48% | 12% |

From these results, the increase of sales looks larger for NEDO group than control group. We then tested these results statistically.

The Shapiro-Wilk test of the distribution of sales and the growth rate showed are both not on normal distributions. The Mann-Whitney's U test, which is a non-parametric method used for group comparison, showed significant differences at the significance rate of 5%.

Similar analysis on the number of employees did not show a significant difference at the significance rate of 5%.

# CONCLUSIONS

In recent years, NEDO is focusing not only on promoting large-scale national projects based on national roadmaps, but also on R&D support for SMEs including start-ups. In this study, the impact evaluation of NEDO's R&D support for SMEs and start-ups is investigated using NEDO's "follow-up monitoring" and "extended follow-up survey" data for all project participants.

The average "success" rate for SMEs was around 33% for both start-ups and SMEs excluding start-ups, which was significantly higher than the rate of large companies participating in NEDO projects, which was around 25 %.

Further series of individual interviews for 30 chosen SMEs that reached "success" stage with excellent results, sales increase for example, were conducted to identify common tendencies as keys to success. As a result, business environment around the expected products together with resource securing were found to be particularly important, both in the planning stage of the R&D, and the continuous period after the application stage.

A DID analysis was also conducted to clarify the effect of NEDO projects on participating SMEs, eliminating the external noise such as macroscopic economic trends. Careful choosing of DID controlled group and statistical tests revealed a difference on the average sales growth rate.

# REFERENCES

**Farja, Y., Gimmon, E. and Greenberg, Z.** (2017). The developing in the developed: Rural SME growth in Israel, The International journal of entrepreneurship and innovation, 18, 36-46.

**Foreman-Peck, J.** (2013). Effectiveness and efficiency of SME innovation policy, J., Small Business Economics, vol.41 no.1, 55-70.

**Inoue, H. and Yamaguchi, E.** (2017). Evaluation of the Small Business Innovation Research Program in Japan, SAGE Open, 2017:1-9. DOI: 10.1177/2158244017690791.

**Kunugi, M., Ueyama, M., Isshiki, T. and Tokuoka, A.** (2016). Follow-up Survey on Discontinued R&D Activities, session 1774, Evaluation 2016, American Evaluation Association, Atlanta, Georgia USA.

**Ministry of Economic Development, New Zealand,** (2011). Evaluation of the Impacts of Cross-Vote Government Assistance on Form Performance.

**Office of Extramural Research, National Institutes of Health**, (2009). National Survey to Evaluate the NIH SBIR Program -Final Report.

**Radas, S., Anic, I-D., Tafro, A. and Wagner, V.** (2015). The effects of public support schemes on small and medium enterprises, Technovation, 38, 15-30.

**SQW Ltd.** (2015). Evaluation of Smart; Impact and Process Evaluation.

**Yamashita, M., Yurugi, Y., Shishido, S., Yoshida, T. and Takeshita, M.** (2013). Impact evaluation of Japanese public investment to overcome market failure review of the top 50 NEDO Inside Products, Research Evaluation 1–21.

# AUTHORS

**MASAKI UEYAMA**
E: ueyamamsk@nedo.go.jp

**MOTOSHI KUNUGI**
E: kunugimts@nedo.go.jp

**TOSHIYUKI ISSHIKI**
E: isshikitsy@nedo.go.jp

**SHUMPEI MIYAJIMA**
E: miyajimaspi01@nedo.go.jp

**SHIN UESAKA**
E: uesakasin@nedo.go.jp

*Evaluation Department, New Energy and Industrial Technology Development Organization,*
1310 Omiya-cho, Saiwai-ku, Kawasaki City, Kanagawa 212-8554
(Japan)

# EVALUATION OF RESEARCH INSTITUTES OF THE NATIONAL ACADEMY OF SCIENCES OF UKRAINE: PROBLEMS WITH IMPLEMENTATION OF THE 'BEST PRACTICES'

IGOR YEGOROV

DOI: 10.22163/fteval.2019.336

## INTRODUCTION

Ukraine is going through difficult period of reforms, which comprise all spheres, including Science and Technology (S&T). It is important to stress that the economic situation has changed substantially in recent years. Some high-tech segments of the economy have disappeared along with design bureaux and research institutes, which worked for them. The branch sector has virtually collapsed without state financial support and the lack of orders from industry. In the past, attempts to conduct really profound reforms of the R&D sector were not systematic, as the country suffers from permanent political instability and changes of the governments (European Commission, 2016). The best part of Ukrainian science has been preserved within the National Academy of Sciences of Ukraine (NASU) and five other state-sponsored academies. The institutes of these academies of sciences have direct state financial support. These academies received more than three quarters of all state financing for R&D in recent years, while the National Academy of Sciences receives more than 50% of the state money on R&D alone (Naukova…, 2017). However, principles of management and criteria for evaluation of research establishments remained mainly unreformed; indicators of research efficiency went down in recent decades. That is why the state is very interested in a proper evaluation of research institutes and aims for changes within the national research system, based on new approaches, which could open the way for reforms in the R&D sphere.

The paper deals with the results of the evaluation of research institutes of the National Academy of Sciences of Ukraine in 2016-2017. The first results of evaluation are discussed and the ways for solving existing problems are proposed.

## BACKGROUND OF THE NEW EVALUATION PROCEDURE FOR NASU

National Academy of Sciences of Ukraine is the leading research organization of the country. It includes 153 research organizations, which form 3 sections and 14 departments according to the distribution of institutes on scientific disciplines. The Academy has 15.6 thousand Researchers; its total budget was 2.8 billion Hryvna Ukr. (2017).

NASU has a relatively high reputation in the country and abroad. The majority of Ukrainian journals from the Web of Science database are published by NASU. However, the Academy has also preserved some features from the Soviet bureaucratic organization, which provokes criticism in society and from foreign experts. Most critics refer to the obsolete managerial system and insufficient transparency in decision-making processes, including distribution of research funds. In fact, NASU is the last remaining part of the Ukrainian research system, which preserved some scientific potential, while Ukrainian science has shrunk substantially in the period of independence (National Academy, 2018). Overall, the number of researchers dropped by more than five times between 1990-2017, while GERD declined from almost 3% to 0.45% in the same period.

The idea of evaluation was to assess the real potential of research institutes, to pick up the best research organizations, to help to better understand problems of these organizations, and to develop corresponding recommendations for changes within NASU. However, the evaluation could have impact not on the Academy itself. If successful, a similar approach to evaluation could be extended to other research institutes of the state sector. At the same time, Ukrainian experience could be useful for some other countries, especially from Eastern Europe and Eurasia, which are trying to reform their research systems.

In 2015, the decision was taken to change the procedure of evaluation of the institutes of the National Academy of Sciences in the context of a general reform of the Ukrainian scientific system. The new evaluation procedure had a variety of aspects that were considered necessary to take into account. It was the intention to base the new evaluation scheme on international experience using both national and international indicators. Further it should have transparent and democratic procedures to exclude conflicts of interest, to give the evaluated research organization the possibility to appeal the evaluation results, to be more flexible by not using only one indicator for ranking. Further, the involvement of external evaluators was considered as a key precondition of success. In the course of time, Ukraine decided to utilize the German experience of the Leibniz Association due to a similar organization of the Leibniz Association and the National Academy in many respects. The German Leibniz Association and the National Academy of Sciences of Ukraine have also some similarities in their main directions of activities that open the way for implementation of the positive experience of organization of evaluation in Leibniz Association research organizations in Ukraine. The NASU and the Leibniz Association have research institutes in different scientific disciplines and institutes of multi-disciplinary profiles. Both rely predominantly on public funding as the main source of their activities. The NASU is larger than the Leibniz Association in terms of research personnel, and the number of institutes (Leibniz Association, 2016), while Leibniz Association has a larger budget (approximately 4-5 times larger in purchasing power parities) (World Bank, 2017). The institutes of NASU have a number of difficulties inherited from the Soviet times and greatly aggravated in the last 25 years, especially in the financial sphere. This means that they urgently need structural changes to provide a more rational distribution of scarce money to improve performance and to justify potential increase of state support. The Leibniz Association has substantial experience in transformation and integration of research organizations, because a number of its institutes stem from the research institutes of the GDR and the "Blaue Liste" institutes of West Germany. For these reasons, Leibniz Association serves as an international reference for establishing a new evaluation scheme for NASU in Ukraine.

Before reforming the evaluation procedure in Ukraine, the solely responsible actor for evaluation was the Presidium of the NASU. All research institutions which received public funding were subject to evaluation. The evaluations took place every five years, and were relevant for the institutions to be included in the state register of scientific institutions. The evaluation included a survey of scientific organizations and the supporting technical institutions, the evaluation at the department level collecting additional information, and the checking of the surveyed forms. At the level of the presidium, multidisciplinary expertise was taken into account and resulted in the ranking of the research institutions. The survey included information on aspects such as employee structure, main scientific outputs, applications of results in practice, financing, the extent of scientific and technical services, recognition of results on the national and international level, the number of foreign grants and embeddedness in the scientific community.

The indicators are surveyed on a quantitative level and were weighted using weighting factors, resulting in one final number as a result and a corresponding rank. The state certification of research institutions was the result of the evaluation procedure. The importance resulted out of the fact that this certification war necessary for the inclusion in the state register of scientific institutions.

## POLICY DESIGN, APPROACH, METHODOLOGY OF NEW EVALUATION PROCEDURE

The new NASU approach has some key principles (Metodika, 2017). One is that international experience as well as both national and international indicators are used. Secondly, the evaluation procedures are conducted more transparent, and potential conflicts of interest are systematically being avoided. Thirdly, the research organization has the possibility to question the procedure and results of evaluation. Further, the procedure is being made more flexible by not depending on a single indicator for ranking as it was the case before in the evaluation procedure of NASU. Finally, external, and in best case including foreign, evaluators are involved now (Evaluation Standards, 2015).

There are three stages of evaluation procedure. At the first stage, the expert group (first-level review board, which consists of 5-6 experts) evaluates the scientific activities of the institution. The members of the group inspect the institution's activities, analyze the inquiry form filled by the institution beforehand, verify whether the materials submitted by the institution are unbiased, and prepare their conclusion according to the selected criteria. At the second stage, the Permanent Expert Committee on a Relevant Field of Science (second-level review board) prepares a presentation on the institution activities in accordance with the report of the first-level group and after consultations with the institution. The second-level review board conveys the conclusion of the first-level group to the institution. The institution can make a statement concerning this conclusion. At the third stage, the Permanent Evaluation Committee of the National Academy of Sciences of Ukraine (third-level review board) considers the presentation of the second-level board, the conclusion of the first-level group, and the statement of the institution. The third stage of the evaluation should result in the report of the third-stage review board that should evaluate the scientific activities of the institution and contain recommendation on its further financing. The report of the third-stage review board should be based on the results of the first-level and second-level evaluation stages. The institution has the following opportunities to take part in the evaluation procedure: prior to the selection of experts of the first-level review board by the second-level review board; the institution can propose a list of main research fields to be covered by the evaluation procedure; the institution can propose experts in these research fields according to the criteria that avoid a potential conflict of interest; following the selection of experts of the first-level review board by the second-level review board, the institution can comment on whether the experts cover the research fields named by the institution; the institution can comment on whether it sees a potential conflict of interest among the experts selected. In case the second-level review board and the institution fail to reach an agreement after the discussion of the comments, the final decision should be made by the first-order review board. The institution obtains a mandatory copy of the first-level review board conclusion from the second-level review board and it is obliged to prepare its statement concerning the conclusion of the first-level review board.

Criteria for evaluation of the quality of work and the potential of an institution by the first-level review board are: development of the institution in previous years and its research strategy for the next years;

scientific results; scientific events and public outreach; appropriateness of facilities/financial provision. Special attention is paid to collaboration and networking (several positions are usually considered).

The second and the third level review boards take into account such criteria as importance of the institute for the development of the country, its role in the national economy, potential at the international level, perspectives and dynamics of research in corresponding scientific discipline and some other issues. Strategic significance of the institution is determined by answering the following questions as a result of the evaluation: *is the institution of strategic significance: for the further development of a specific scientific discipline and its environment? As a hub for specialists or regional clusters? For the further development of fields of technology, information and other services, consulting, socio-political tasks? for the profiling of programs of the NASU?*

Key quantitative indicators of evaluation are the following: number of publications (depending on the publication culture of the subject area, in particular in peer-reviewed journals, at peer-reviewed conferences, in monographs etc.); number of documents on commercial property rights and patents, the number of consulting contracts and expert reviews; the amount of third party funds raised for research, consulting, services, etc.; the income from commercial activity such as leasing. Other quantitative indicators could be also included into the evaluation procedure. Quality assurance of evaluation is provided by the a) internal quality management at the institution and b) by assessment of the institution by the relevant Department of NASU.

As a result of the evaluation, the institute could be assigned to one of four groups (in fact, 3 groups, as the last one deals with 'supportive' organizations) according to the level of evaluation results.

NASU created a special Evaluation Office to facilitate the process of evaluation in 2017, which was responsible for the organization of the evaluation and consultative services for expert groups and review boards.

The new evaluation procedure of NASU is strongly oriented towards avoiding conflicts of interest. Here NASU has tried to apply the same criteria as the Leibniz Association (no joint projects, no membership in the scientific boards and joint publications during the last five years etc.). However, due to limited monetary and competence resources, not every small conflict of interest can be fully avoided to maintain a high level of competence among the experts. To solve this, a wider scope of potential experts and more monetary resources for inviting experts would be needed. For example, it is simply not possible to invite best fitting experts from abroad (Western Europe, USA), because of budget constraints.

# RESULTS OF THE EVALUATION OF NASU RESEARCH INSTITUTES

In 2016 first 13 institutes (one from each Department of NASU) were reviewed. Twenty seven other institutes were added to this list in 2017. There are plans to conduct evaluation of 47 institutes in 2018 and the rest of the Academy in 2019. Thus, it is expected that more than half of the NASU institutes will be evaluated until the end of 2018.

It is too early to make final conclusions, but the results of the evaluation of 40 institutes in 2016-2017 and unfinished evaluation of approximately the same number of institutes in 2018, open the way for some important remarks.

Evaluation in 2016-2018 was useful for both the NASU and the institutes. Some objective information about the situation within the institutes was received, and corresponding recommendations on how to change it were made. This is definitely a positive moment.

However, a number of problems of evaluation have been revealed. More than half of the institutes received the highest mark for their scientific activities. In some cases, review boards had to correct the marks, made by the expert groups.

Evaluation itself revealed a number of barriers for the development of Ukrainian science that need to be overcome.

Like many scientific organizations in transition economies, NASU faces a problem of aging personnel caused by the ongoing emigration of young scientists mainly because of the limited attractiveness of the Ukrainian science system. This is not a particularity of the NASU, but a general difficulty of the science system. Low wages and unclear career tracks attract young and excellent scientist to other areas within the country or abroad. A strong challenge for the NASU is how to attract young scientists into the Ukrainian system of research. The solution requires a broader approach that includes wage policy and academic career tracks. It is an important field for coordination between different fields of policy making.

An example for a fundamentally problematic indicator for the Ukrainian situation is the generation of publication data from databases such as Web of Science or Google Scholar due to differences in the writing of names (transcription), which deteriorates the proper assignment of publications. Further, in different disciplines the assessment of publications and output has to be adjusted. For example, in some areas the revision, commenting and reprinting of classic writings is regular part of the scientific work and output, however hardly to be accounted for if e.g. mainly publications in journals are considered. In other areas the policy consultancy may be part of the regular work and output. These differences in the specific way of working have to be taken into account for a proper and expedient consideration of criteria and indicators. Just using certain publication types would be problematic to take differences between excellence and relevance into account, such as in the case of consultancy activities. Further, to not confuse quantity with quality the review of best publications should be taken into account, alongside with full publication lists.

The second problem is the implicit or explicit hierarchy of criteria and indicators. In several evaluation systems many indicators are imposed but only some really "count" in institutional evaluations. These are normally articles in refereed journals and third-party funding. It is very important to clarify these questions beforehand.

The key issue of the importance of different aspects peaks in the weighting question of indicators. To properly take into account the relevant qualitative and quantitative aspects, as well as the institutional individuality of each institute with corresponding institute specific goals, experts are strongly needed. Further, an involvement of peers may be beneficial to further balance out the weighting process.

The issue of problematic metrics for research assessment is already debated. A closer evaluation of appropriate and inappropriate use of quantitative indicators is regarded, including the conceptualization of "responsible metrics". A framework of five dimensions is available to assess appropriate uses of quantitative indicators:

- robustness: to base metrics on the best possible data regarding accuracy and scope

- humility: qualitative expert assessment should be supported by quantitative evaluation, but not supplanted
- transparency: transparency and openness of analytical processes , to allow verification by those who are evaluated
- diversity: usage of a variety of indicators to account for the variety of research fields
- reflexivity: recognizing systemic and potential effects of indicators and accordingly updating them (Wilsdon et al., 2015).

It should not be forgotten that the evaluation procedure can shape the mission, developments and working styles of institutes also in a negative way when obeying to certain indicators becomes more important than doing proper discipline specific work. These considerations call for a cautious application of quantitative indicators as well as an increasing importance of qualitative factors. In the evaluation of institutes, some structural factors must not be lost out of sight, such as if a context of structural reforming is given as well as the structural context of institutes for regions. If internal development processes are taking place, it is of major importance to not rise a trade-off situation between learning and evaluation, but instead take learning successes and learning processes which are put into place into account. Thus, evaluations which are only based on a certain point of time should be avoided, and the long-term development of the institute should be kept in mind. Hence, it is critical to take new orientations and priorities that the institute is setting into account, and check whether these are in line with national priorities. Also, short and long-term priorities of the institute and the NASU need to be identified and properly accounted for. Further, to strengthen the development aspect in evaluation the institute could, potentially in cooperation with NASU, conduct a SWOT-Analysis as one possibility for a self-assessment procedure to identify needs for the further development which should be put into practice. Hereby it should be made possible to take the developmental success and changes of the institute more explicitly into account at the next evaluation. Internal assessments are generally a very fruitful preparation for external evaluation. In-depth SWOT analysis could be useful for more precise evaluation of the Ukrainian research institutes.

There were also problems, which were identified with the procedure of evaluation:

1. Formally, experts had no conflict of interest in evaluating the research institutes. They had to sign special forms and the office of Evaluation checked all candidates on co-authorship and participation in joint projects. Unfortunately, it is almost impossible to provide real independence of experts within the relatively closed Ukrainian research system, while the country had no resources to invite a number of foreign experts. Usually, expert groups included not more than one foreign expert. Some of them could not take part in the evaluation procedure at all. Ukraine needs assistance in provision of independent experts for evaluation and participation of foreign experts in evaluation procedures. There are several options for solving this problem. First, initiation of a technical assistance project from the side the EU. The second is to involve representatives of Ukrainian scientific diaspora more actively. The third is to try to ask the government to provide extra funds for the evaluation. The office of Evaluation along with the management of the Academy has made some steps in these directions, including attracting experts from Ukrainian scientific diaspora. However, the results are still not clear.
2. There is a need to improve the list of specific indicators to make them more relevant to the reality of scientific activities of research institutes in different disciplines, as some important activities are not considered by the evaluation. This work is under way with the help of expert groups from different scientific disciplines.
3. The time for the preparation of the report of the institutes and the expert conclusions have to be extended. At the moment, it is 2-3 times shorter than in the Leibniz Association. Such extent could help to improve the quality of evaluation-related documents.
4. The focus has to be shifted to the research units. This will help to provide internal reorganization of research institutes.
5. A formal procedure has to be proposed to 'appreciate' the best institutes and units. At the moment, it is still not clear, what kind of extra benefits institutes could receive 'automatically' in the case of high marks.
6. Despite strong recommendations to consider the possibility of mergers of relatively small research organizations, this did not take place in the last two years. However, there is a clear need to continue to optimize the network of scientific institutions and organizations. In particular, the consolidation of institutions and the merger of institutions with similar profiles are relevant, as this could help to reduce administrative costs and to improve the general positions of the institute by reorganizations of weak units. Analysis of the existing situation within research institutes shows that a quarter of them have less than 20 researchers, some units have 3-5 persons only, including supportive staff. A number of them do not have specialists with highest academician degrees. Such reorganization could help to preserve important research areas and human resources, taking into account such aspects as the relevance of research topics, specific results – scientific publications, patents, licenses, etc.

The work on improvement of evaluation is under way now and the Ministry of Education and Science of Ukraine has announced plans to utilize the experience of NASU for other research organizations including those, subordinated to Ministry of Education and Science, in 2019.

# REFERENCES

**European Commission** (2016). Background Report. Peer Review of the Ukrainian Research and Innovation, Horizon 2020 Policy Support Facility.-Luxembourg: Publications Office of the European Union.

**Platform for Research and Technology Policy Evaluation** (2015) Evaluation Standards in Research and Technology Policy. Retrieved August 8, 2018 from https://www.fteval.at/.

**Metodika otsnuvannia dialnosti naukovyh ustanov NASU** (2017) - Retrieved August 10, 2018 from: http://www.nas.gov.ua/text/pdfNews/metodyka_text.pdf (in Ukrainian)

**Leibniz Association: key figures** (2016) – Retrieved August 22, 2018 from: https://www.research-in-germany.org/en/research-landscape/research-organisations/leibniz-association.html

**National Academy of Sciences of Ukraine in 2017** (2018) Kyiv: NASU (in Ukrainian)

**Naukova y naukovo-tehnichna Diyalnist v Ukraini** (2017) - State Statistical Service of Ukraine: Statistical Yearbook (in Ukrainian)

**Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I. and Wouters, P.** (2015): The metric tide: Executive summary – Report of the Independent Review of the Role of Metrics in Research Assessment and Management. HEFCE.

**World Bank data** (2017) Retrieved July 15, 2018 from: https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD

# AUTHOR

**IGOR YEGOROV**
*Office of Evaluation and Institute of Economy and Forecasting, National Academy of Sciences of Ukraine*
26 Panasa Myrnogo st., Kiev, 01011, Ukraine
E: Igor_yegorov1@ukr.net

# THE SHAPE OF THINGS TO COME: EX-ANTE ASSESSMENT OF THE ECONOMIC IMPACT OF HORIZON EUROPE

JULIEN RAVET, BAPTISTE BOITIER, MARCO GRANCAGNOLO, PIERRE LE MOUËL, LIVIU STIRBAT AND PAUL ZAGAMÉ

DOI: 10.22163/fteval.2019.337

## ABSTRACT

This paper provides an ex-ante assessment of the expected economic impact of the post-2020 EU Research and Innovation Framework Programme, Horizon Europe. A key novelty in the approach is the use of three different macroeconomic models for the assessment of the continuation of the current Programme, Horizon 2020: NEMESIS, QUEST and RHOMOLO. In addition, NEMESIS is used to assess different batches of policy options related to the budget, management and design of Horizon Europe. The paper also highlights key aspects and assumptions that policy-makers and researchers need to consider for this type of analysis such as budget allocation, performance, leverage and financing modes.

## 1 INTRODUCTION

EU-level investment in Research and Innovation (R&I) focuses on excellence through EU-wide competition and cooperation. Successive EU Framework Programmes have supported training and mobility for scientists, creating transnational, cross-sectoral and multidisciplinary collaborations, leveraged additional public and private investment, built the scientific evidence necessary for EU policies, and had structuring effects on national R&I systems. The political narrative has put more and more accent on 'shaping the future' through R&I policy and funding, thereby lending even more importance to the ex-ante assessment of the funding Programme's impact.

Horizon Europe, the 2021-2027 Framework Programme for EU R&I, will succeed the current Programme, Horizon 2020 (active between 2014-2020). This new programme will build on lessons learnt from previous evaluations[1], feedback from experts[2] and from other stakeholders. It will be an evolution, not a revolution, focusing on a few design improvements to further increase openness and impact. With Horizon 2020 well on track to deliver excellence, these changes in the design aim at making the successor Programme achieve even more impact (through the European Innovation Council and mission-orientation) and more openness (through strengthened international cooperation, a reinforced Open Science policy, and a new policy approach to European Partnerships).

Assessing the impact of the Framework Programmes ex-ante is crucial for policy-makers in order to inform their strategic decisions. There is a general consensus (Hall, Mairesse and Mohen, 2009; European Commission, 2017a; Di Comite and Kancs, 2015) that R&I are decisive in fostering productivity growth. However, putting a precise figure on the expected benefits of a large R&I programme is a challenging task with a lot of uncertainties, notably due to the ex-ante approach. This is made even more difficult by the long-term horizon that a proper analysis of these impacts requires.

This paper aims at providing an assessment of the expected economic impact of the post-2020 Framework Programme. It also highlights key aspects and assumptions that policy-makers and researchers need to consider for similar analyses, especially when they need to collaborate with each other.

## 2 MODELLING THE IMPACT OF THE EU FRAMEWORK PROGRAMME

The first ever ex-ante impact assessment of any EU policy initiative in the field of research was the impact assessment of the 7th Framework Programme (FP7) (Muldur et al., 2006; Delanghe and Muldur, 2007). The quantification of its economic impact relied on historical data (e.g. publications and patents) and on simulations based on macroeconomic modelling. The NEMESIS model was used for this impact assessment, and subsequently for the impact assessment of Horizon 2020 (European Commission, 2013).

Since FP7, macroeconomic models, including NEMESIS, have evolved and lessons from previous impact assessments can help policy-makers in using these models for current and future assessments.

In this context, macroeconomic modelling is an essential tool to support policy-making by quantifying the impact of the Programmes and assessing policy options. Depending on when the assessment takes place in the EU policy cycle (Figure 1), this can be done in an ex-post/interim (monitoring and evaluation of a programme) or ex-ante design (impact assessment), with policy options examined in impact assessements only in order to feed the preparation phase of the Programmes.

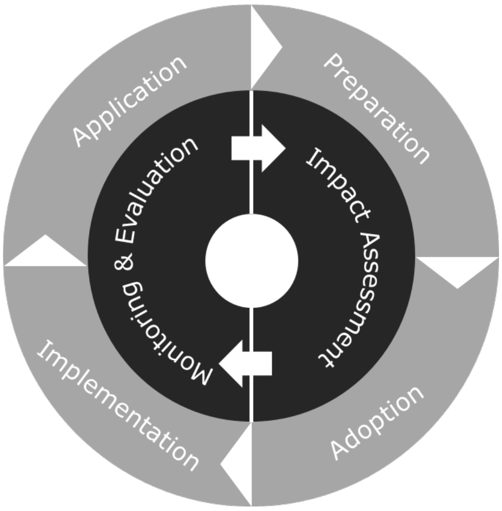



**Figure 1** EU policy cycle
**Source:** adapted from the EU Better Regulation guidelines (European Commission, 2015).

Expanding on the Horizon Europe Impact Assement (European Commission 2018), in this paper the NEMESIS, QUEST and RHOMOLO[3] models are used to assess the impact of the continuation of Horizon 2020 in order to triangulate the signs, patterns and sizes of the impact of continuing the current Framework Programme. To the authors' knowledge, this is the first time that results from different models are triangulated to assess the impact of EU R&I funding. By relying on these three models, the aim is to leverage on their respective strengths and compensate for their limitations. The strengths of these models rely on their specificities, and the differences between the models can help address specific needs of policy-makers. When using and interpreting results produced by these models, it is also essential to acknowledge their main limitations, as any model only allows for a partial representation of reality subject to the assumptions made.

NEMESIS is a macro-econometric model consisting of detailed sectoral models for every EU country. Measuring technical progress in NEMESIS is derived from the new growth theories where innovations result from the investment in R&D by private firms, and from R&D undertaken by the public sector. In the latest version of NEMESIS used for this paper, innovations still arise from private and public investments in R&D, but also from investments in two other complementary innovation inputs: ICT and Other Intangibles (including training and software). These enable improved accuracy in assessing R&I policies by considering the most up-to-date theoretical and empirical findings of economic literature (Le Mouël et al., 2016). Di Comite and Kancs (2015) consider that NEMESIS is the richest model in terms of innovation types and policy elasticities when compared to other standard macroeconomic models for R&D and innovation policies. This means that policy-makers can easily design options related to specific innovation types or innovation channels when using this model. However, NEMESIS is based on empirical observations of relationships among variables as well as on adaptive expectations instead of forward-looking ones, allowing for more degrees of freedom in behaviour than in other models. This may generate inconsistencies with recent developments in macroeconomic theory.

QUEST belongs to the class of micro-founded dynamic general equilibrium (DGE) models that are now widely used in economic policy institutions as the latest step in the development of macroeconomic modelling. The focus in these models is on the economy as a whole, as an integrated system of economic agents that base their decisions over a range of variables by continuously re-optimising while subject to budgetary, technological and institutional constraints. These models are forward-looking and intertemporal, i.e. current decisions account for expectations about the future. This analysis uses the semi-endogenous growth version of the Commission's QUEST model with an R&D production sector (QUEST3RD). The model economy is populated by households, firms producing final and intermediate goods, a research industry and a monetary and fiscal authority. The forward-looking dynamic approach of QUEST makes the model the most appropriate for assessing the impact of R&D and innovation policies over time. This is particularly important as effects of the initial investment are expected to last after the period covered by the Programme, which calls for a model that can precisely measure long-term impacts. On the other hand, QUEST III, being an aggregate macroeconomic model, groups all R&D activities in a unique R&D sector without capturing the complexity and diversity of the type of R&D investments (e.g. private and public R&D activities, product and process innovation, non-R&D, and disruptive innovations) or their extensive sectoral and geographical details.

RHOMOLO is a spatial DGE model that covers 267 regions at the NUTS2 level. Each region contains 10 economic sectors. A subset of these operates under monopolistic competition. The rest of the sectors operate under 'perfect' competition. Regional goods are produced by combining labour and capital with domestic and imported intermediates, creating vertical linkages between firms. By modelling regional economies and their spatial interactions, RHOMOLO is the most suitable model to address questions related to geographic concentration of innovative activities and spatial knowledge spillovers, which is also a crucial aspect for policy-makers. However, RHOMOLO trades off its detailed spatial dimensions with keeping the optimisation problems static and, hence, not capturing the inter-temporal consequences of innovation decisions. In addition, it does not distinguish between private and public innovation or between different types of endogenous innovation.

# 3 TAILORING THE MODELS TO THE SPECIFICITIES OF THE EU FRAMEWORK PROGRAMME

The three macroeconomic models do not initially reflect the reality of the EU Framework Programme. In order to adapt the models to the specificities of the programme, several parameters and assumptions need to be carefully considered.

The budget of the Programme is a first key element to specify. This entails the overall amount that will be spent, but also the temporal, national and sectoral allocation of the budget. Depending on the mechanisms of the model, additional dimensions can be added: the regional allocation (at NUTS2 level) for RHOMOLO, or the allocation between basic and applied research for NEMESIS. For the assessment of Horizon Europe, budget size and budget allocation[4] are assumed to be the same as in Horizon 2020 in the baseline scenario (i.e. the continuation of Horizon

2020 over 2021-2027), in constant prices and without the contribution of the UK (around 15% of the Horizon 2020 budget). This corresponds to about 85 billion euros in current prices over 2021-2027 based on the last year of Horizon 2020.

An essential aspect for all models is the mode of financing of the Framework Programme. Money spent for the Framework Programme can come from different sources, and in this regard, it is tempting but rather unrealistic to make it appear out of nowhere. In this paper, RHOMOLO and NEMESIS assume that the financing of the Programme can be reflected by lowered national expenditure. The mechanisms of QUEST can be used to assess two financing scenarios: (i) raising additional VAT revenues in the Member States and (ii) lowering national public investment.

A feature that is specific to the NEMESIS simulations is the use of different parameters for leverage[5] and economic performance[6] of EU R&I funding compared to national funding. The model assumes parameters that reflect a European Added Value of R&I funding: a better leverage of European funding when compared to national ones inducing more R&I expenditures for the same level of public funding (0.15 instead of 0.1 for applied research), and a higher research productivity (15%, also used in European Commission, 2013) of the European R&I Programme, explained by the higher competition at the European level than at the national one, and by the transnational collaborative aspects inducing more knowledge spillovers. This EU added value is supported by several studies (ECDG and Elsevier, 2017; Rosemberg et al.; 2016; Vullings et al.; 2014; Delanghe et al., 2011; PPMI, 2017). Values used for these parameters in the NEMESIS model are considered to be conservative with regards to the literature, including the following quantified results from PPMI (2017) based on data from the 7th Framework Programme (FP7) and Horizon 2020:

- Research organisations supported by FP7 are around 40% more likely to be granted patents or produce patent applications.
- Patents produced in FP7 are of higher quality and likely commercial value than similar patents produced elsewhere (70% more citations).
- Patents produced under the Framework Programmes are likely to be of higher technological value and more likely to be based on cutting edge scientific knowledge (11% more citations in FP7).

Horizon 2020 participants declare that the programme significantly improves their competitive position internationally (78 % expect a decrease in this area if they had not been funded) and access to new markets (71 % expect a decrease in this area if they had not been funded).

# 4 BASELINE SCENARIO: HOW MUCH IS THE CONTINUATION OF HORIZON 2020 WORTH?

NEMESIS, QUEST and RHOMOLO are used to assess the impact of the continuation of Horizon 2020 compared to a situation without a Framework Programme (i.e. discontinuation). This scenario assumes that EU funding for R&I will be carried over 2021-2027 with a similar budget as in Horizon 2020 (see Section 3). The three different models correspond to different approaches and present very different specifications and settings of parameter values. One should therefore not expect the three models to produce identical estimates of the economic impact of a given policy intervention. However, comparing the findings from the three models for the continuation of Horizon 2020 allows to triangulate results in order to assess the consistency of the impacts identified in each model. This triangulation is also essential for a better understanding of how the specific mechanisms of these models can affect the results they produce.

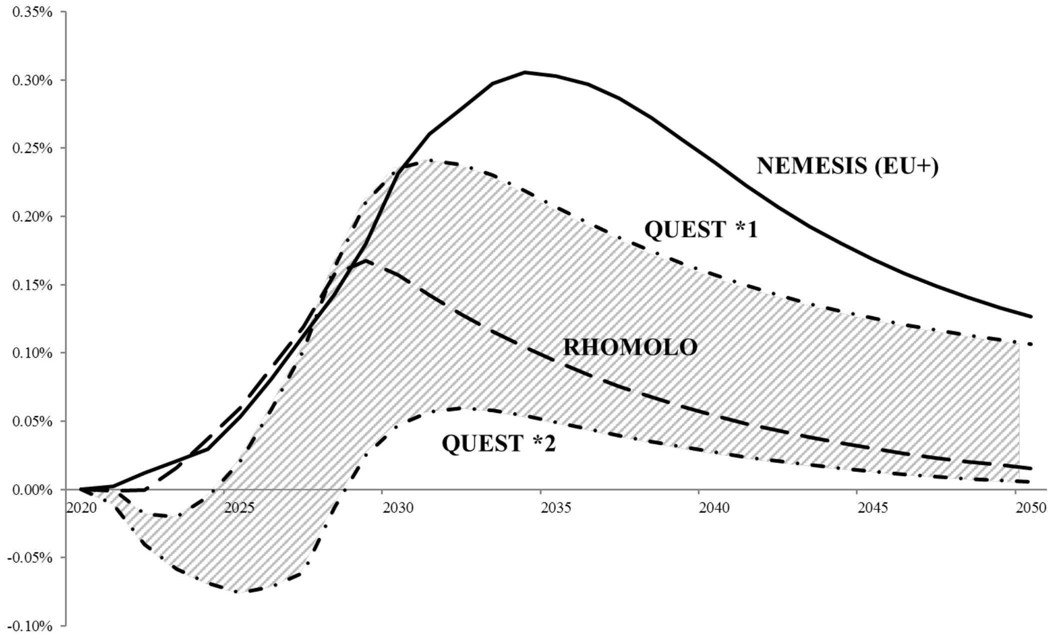



**Figure 2** GDP impact of Horizon 2020 continuation (deviation in % from a discontinuation)

**Source:** Seureco (NEMESIS) and European Commission (RHOMOLO and QUEST). Note: EU+ indicates that Nemesis uses higher performance and leverage for EU funding compared to national funding as a reflection of the EU added value of the Programme. QUEST *1 assumes that financing of the Programme relies on VAT increase. QUEST *2 assumes that financing relies on lowering public investment.

The models present consistent results in terms of sign and temporal pattern of the GDP gain from the Framework Programme (compared to the discontinuation of the Programme) over 2021-2050 (Figure 2). The three models show a strong increase in the GDP impact during or after the period covered by the Programme, with highest impacts expected between 2029 and 2034. The size of the GDP gain is the highest based on the NEMESIS results. This can be explained by the fact that the three models use different sets of innovation channels and elasticities. These results suggest that the continuation of the Framework Programme after 2020 is expected to bring an estimated average GDP increase of up to 0.19% over 25 years, which means that each euro invested can potentially generate a return of up to 11 euros of GDP gains over the same period.

The highest gains in the NEMESIS model can be partly explained by the fact that QUEST and RHOMOLO do not directly take into account the higher leverage and performance expected from EU funding of R&I compared to national funding, while this is acknowledged in the parameters of NEMESIS. As explained in Section 2, this assumption reflects the intrinsic EU added value related to the EU level investments due to factors that are not directly captured by these models, such as multidisciplinary transnational collaborations or critical mass.

Regarding the mode of financing, results from QUEST show that financing R&I investments from value added taxes produces higher economic benefits in the model in the medium and long run than with public investment cuts. This is because the financing mechanism in the model attributes potential productivity effects to public investments (e.g. roads, buildings) which are higher than for value added taxes.

The pattern in time is similar between the models. The NEMESIS model describes this pattern with the following three main phases. (i) An **investment phase** over 2021-2027 that is a 'demand phase' in which all the dynamics are induced by the change in the R&I expenditures, with or without moderated impacts of the innovations (as they take time to appear). This phase can be viewed as a Keynesian multiplier. (ii) The **innovation phase**: the arrival of innovations reduces the production cost of the new products or raises their quality, which induces an increase of demands for products. (iii) The **obsolescence phase**: new knowledge progressively declines due to knowledge obsolescence and, in the long-term, the macro-economic track goes back to the reference scenario.

The impact on jobs based on the NEMESIS model is also substantial (Figure 3). EU investments in R&I are expected to directly generate an estimated gain of up to 100,000 jobs in R&I activities in the "Investment phase" (2021-2027) and to foster through the economic activity generated by the Programme an indirect gain of about 200,000 jobs over 2027-2036, of which 40% are high-skilled jobs. However, during the investment period, while the Programme has a positive effect on jobs in R&I, the decrease in national public investment that is assumed by the model is mechanically accompanied by a comparable decrease in non R&I-related jobs. During this period, the increase in R&I investment raises the demand for employments in research activities. But the funds used to support R&I activities are taken from national public investments according to the assumption used in the model. This shift between both kinds of investments explains the decrease of high- and low-skilled employment while employment in research activities increases. Furthermore, the raise of the demand for employment in R&I activities increases the inflationary pressure on the high-skilled workers' wages (as employment in R&I activities are mainly provided by high-skilled workers). This reinforces the negative impact on high-qualified employment during the investment phase.

After the investment phase, total employment rises progressively to reach a maximum deviation of +228,000 employments in 2036 compared to a situation without Framework Programme. Between 2028 and 2036, around 60% of the cumulative EU employment gains relate to low-qualified workers, 30% to high-qualified workers and 10% to employments in research activities. After 2036, the declining economic gains resulting from the EU R&I Programme also reduce employment gains.

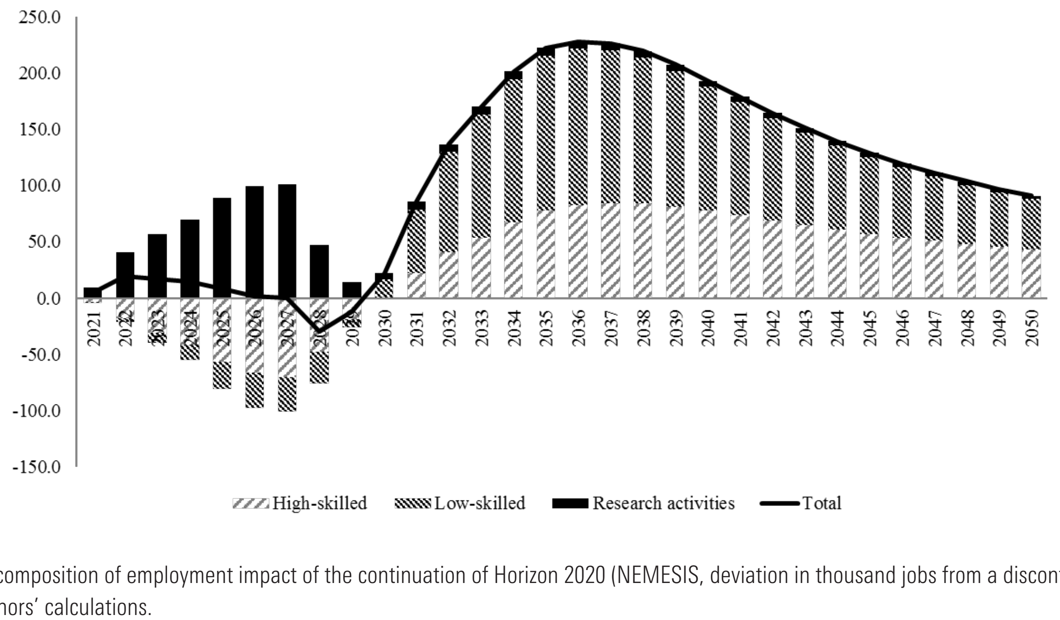



**Figure 3** Decomposition of employment impact of the continuation of Horizon 2020 (NEMESIS, deviation in thousand jobs from a discontinuation)
**Source:** Authors' calculations.

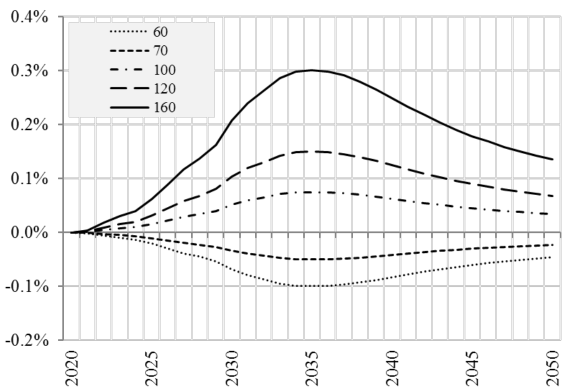
# 5 ASSESSING THE IMPACT OF POLICY OPTIONS

The NEMESIS model is used to assess different sets of policy options for Horizon Europe by changing specific parameters of the model. The impact of these options is assessed against the baseline, which is the continuation of Horizon 2020 (as described in section 4). Besides the parameters changed for each option, all assumptions are the same as in the baseline scenario.

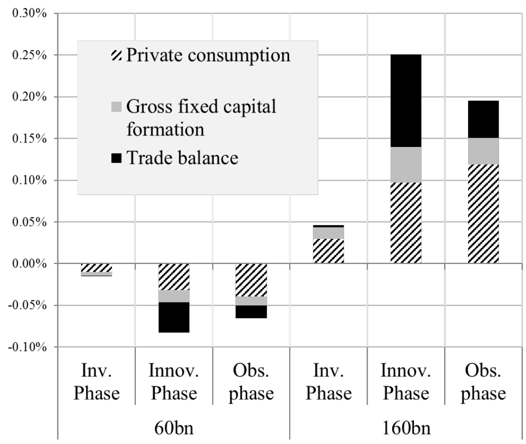
## 5.1 ASSESSING BUDGET OPTIONS

A first element that is critical for Horizon Europe is the budget allocated to the Programme. Different budget envelopes for Horizon Europe are assessed[7], ranging from EUR 60 billion to EUR 160 billion in current prices. Variations of the EU budget envelope around the baseline are compensated by equivalent variations of national public investments. In modelling options with lower budget compared to the baseline scenario, funds that are not used for Horizon Europe are "given back" to EU Member States (according to their contribution to the EU budget) and used in public investments (excluding research activities), i.e. gross fixed capital formation. Modelling options with higher budget assume that each Member State raises its contribution to the EU budget and finances this transfer to the EU by reducing its public investments accordingly. This does not correspond to a "centralisation" of R&I funds at EU level, as the budget variations of Horizon Europe are compensated by variations from national investments excluding R&I investments and not by variations of national public support to R&I (this aspect is addressed in section 5.2).

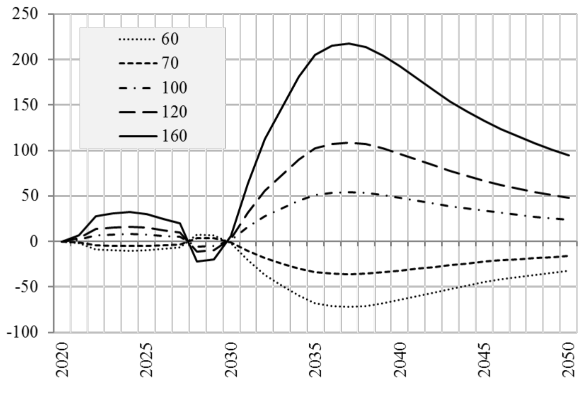
**(a) % GDP deviation from baseline**




**(b) Average contribution to GDP deviation from baseline**




**(c) Employment deviation from baseline (thousand)**




**(d) Average annual employment deviation from baseline (thousand)**

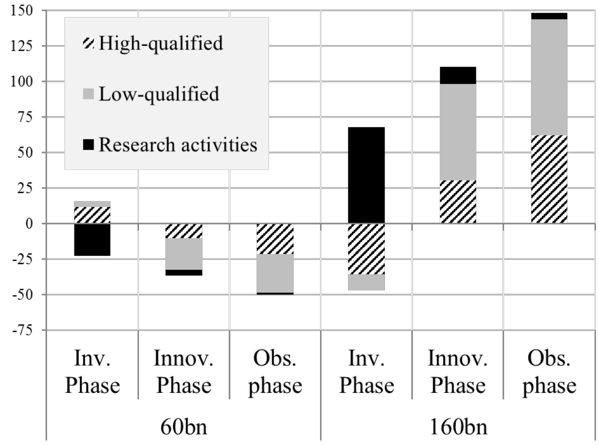



**Figure 4** Impact of budget options compared to the H2020 continuation scenario (baseline)
**Source:** Authors' calculations.

As shown in Figure 4a, the sign of the GDP variation follows the direction of the budget change, with options with largest and lowest budget envelope presenting respectively the largest and lowest impacts. The main driver of the impact size is the difference between the productivity of R&I investments and public investments (excluding R&I) as budget deviations are financed by an opposite deviation of the public gross fixed capital formation. During the investment phase (2021-2027), the EU GDP deviation under each budget option is relatively contained as the effects of the innovations resulting from R&I investments do not yet operate fully during this phase. After 2027, the GDP deviation with respect to the baseline scenario becomes increasingly important, and reaches a maximum around 2035. In the EUR 60 billion scenario, the EU GDP deviation reaches a minimum of -0.1% in 2035. In the scenarios with higher EU R&I budget, the deviations of the EU GDP can be significant, and reach a maximum of +0.3% in the EUR 160 billion scenario.

The decomposition of the EU GDP deviation (Figure 4b) in the lowest and highest budget options shows that the main contribution to the EU GDP deviation during the investment phase comes from the private consumption. After the investment phase, the trade balance becomes the most important contributor to the GDP deviation, with R&D intensive sectors being also the most open sectors to international markets. During the obsolescence phase, productivity gains are progressively spread to the overall economy, thus increasing real wages and reinforcing the contribution of private consumption to EU GDP deviation.

Regarding the impact on employment, (Figure 4c), the strongest deviation of total EU employment with respect to the baseline scenario is reached in 2037, with -72,000 thousand jobs in the EUR 60 billion scenario and +217,000 thousand jobs in the EUR 160 billion scenario. The employment deviation follows the same pattern as the GDP deviation after the end of the Programme, with an intensification of the deviation followed by a progressive decrease. However, in 2028 and 2029, due to the end of the EU support to R&I, combined with the effect of changes in real wages, especially for high-qualified workers, the impact on EU total employment is opposite compared to other periods. For example, in the EUR 120 billion scenario, the EU total employment is lower than in the baseline scenario, with -11,000 and -10,000 employments in 2028 and 2029 respectively.

In terms of types of jobs (Figure 4d), budget deviations directly impact R&I employment during the investment phase. Under the lower budget options, the reduction of EU support to R&I induces a decrease of R&I employment compared to the baseline scenario (with up to -40,000 in the EUR 60 billion option, and -20,000 in the EUR 70 billion option) but the increase in public investments (as a result of the reduction of EU budget to R&I) positively impacts high-qualified and low-qualified employment. The patterns are reverted for options with higher budget. After 2027, employment in R&I activities is close to the baseline level in all scenarios.

## 5.2 ASSESSING (DE)CENTRALISATION OPTIONS

While budget options in Section 5.1 considered that changes in the envelope of the Framework Programme can be reflected by corresponding changes in national investments, another approach is to shift R&I efforts between the different levels of intervention, i.e. national and EU level. This type of shift corresponds to a "centralisation" or "decentralisation" of the management of R&I funds. The impact of two options are assessed with respect to the central management of the Framework Programme: an option with more centralisation of EU funds for R&I at EU level and an option of more decentralisation at national level. The option with more centralisation is defined as a reinforcement of the Framework Programme after Horizon 2020 by centralising, at EU level, one third of the national competitive-based project funding (i.e. 8.75% of the national public R&D expenditures[8]). As a result, the total budget for Horizon Europe is EUR 160 billion (in current prices), which also corresponds to the highest budget option in Section 5.1. In the decentralisation scenario, the EU R&I programme is implemented at national level: EU funds for R&I over 2021-2027 are redistributed to Member States, who use them to support national R&I activities.

Figure 5 shows the impact of these options on EU GDP. Under the option with more centralisation, after 2027, innovation starts to diffuse widely into the economies and, as the amount of EU support to R&I activities is almost twice that invested in the continuation scenario, the positive impact in terms of innovation and, then, economic performance is higher. In this scenario, the EU GDP gain compared to the baseline scenario reaches a maximum of +0.21% in 2031 as the result of two main factors: (i) due to the initial allocation of funding at national and EU level used in the model, there is a shift towards more applied research (associated with more impact on absorption capacity and leverage); (ii) the economic performance is stronger when funds are used at EU level compared to national level. From 2028 to 2034, when more and more innovations enter the market, the EU GDP in the centralisation scenario progresses more rapidly than in the continuation scenario, with a maximum difference in 2031 of +0.21%. Under the decentralisation option, the observed deviation of the EU GDP is negative, but relatively limited.

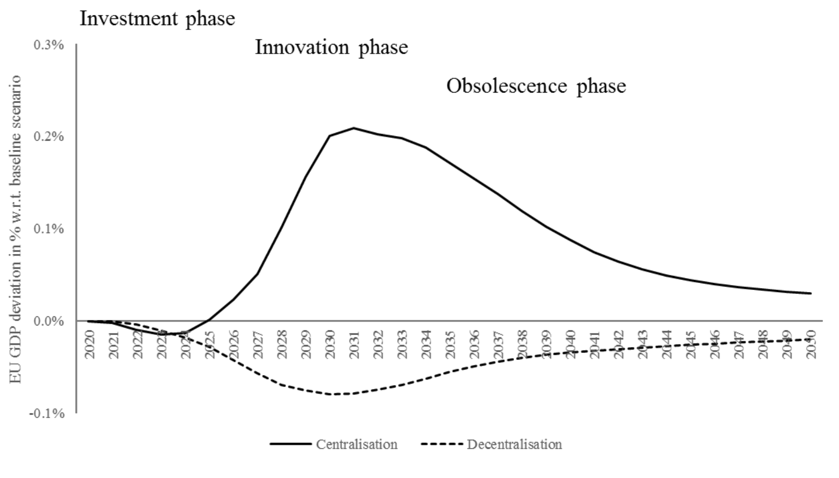



**Figure 5** EU GDP deviation in the de(centralisation) options (% w.r.t. the baseline scenario)
**Source:** NEMESIS model

Deviations in terms of employment are reported in Figure 6. With more centralisation, total EU employment is moderately lower than in the baseline scenario during the investment phase (2021-2027) with -59,000 jobs in 2026. This lower EU total employment is the combined result of three different mechanisms. (i) The higher investments in R&I activities (explained by higher crowding-in) increase the inflationary pressure on the high-qualified labour market. (ii) There is a reallocation of the funds between the two types of research (from basic to applied) that do not have similar labour contents. (iii) The reallocation of public R&I funding to beneficiaries through EU funding is not ex-ante neutral for all Member States: some of them lose funds whereas others win. After 2027, however, when innovation takes place, EU total employment in this centralisation scenario is higher than in the baseline scenario, with a maximum EU total employment gain of 175,000 units in 2034, of which 111,000 in low-qualified jobs, 60,000 in high-qualified employments and 4,000 in research activities. The decentralisation option shows a negative impact on the EU total employment in comparison with the baseline scenario: the maximum loss of EU total employment reaches 78,000 units in 2031 (49,000 low-qualified jobs, 27,000 high-qualified jobs and 2,000 jobs in R&I activities). In 2050, the difference is almost nil with -10,000 employments under the decentralisation option compared to the baseline scenario.

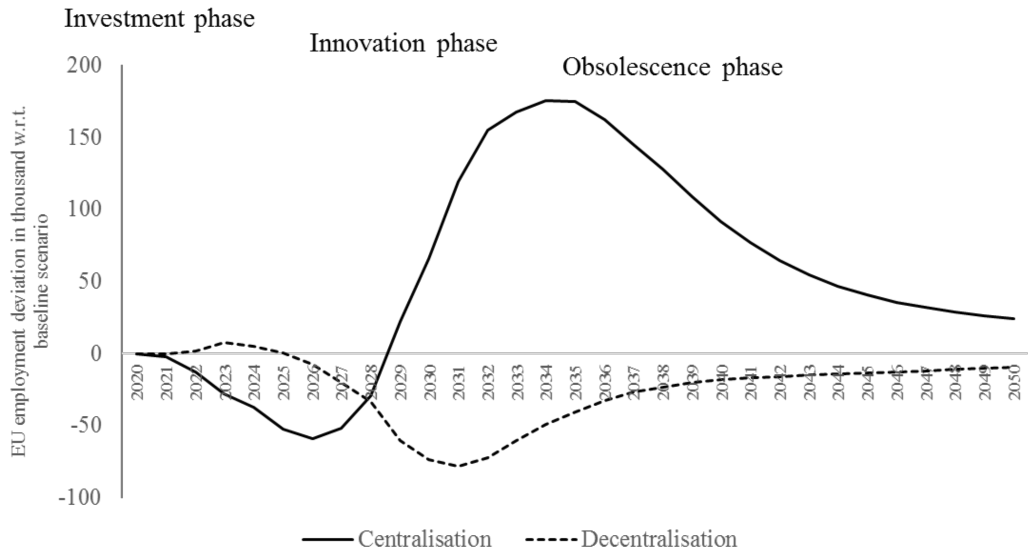



**Figure 6** EU employment deviation in the de(centralisation) options (% w.r.t. the baseline scenario)
**Source:** NEMESIS model

It is important to stress that these results strongly depend on the assumptions of higher leverage and economic performance of R&I at EU level compared to national level (see Section 3), which is the direct translation in the model of the EU added value of R&I funding. Sensitivity analysis carried out on the parameters of the model (see Boitier et al., 2018) shows that the higher direct leverage of EU financial support to applied research compared to national support has a moderate impact on EU GDP and total employment, while the higher economic performance of the R&I activities engaged at EU level compared to similar activities at national level is the key explanatory factor behind the impacts. As developed in section 3, evidence on this higher EU performance of R&I funding can be found in the literature. However, its precise quantification is not straightforward. This is why this paper uses conservative values with respect to existing evidence (see Boitier et al., 2018 for a survey of the related literature). However, it is worth stressing that the model cannot demonstrate as such that R&I at EU level performs intrinsically better than at national level (as, for example, due to multidisciplinary transnational collaborations or critical mass).

Another important aspect regarding the centralisation option is that it considers a total envelope of EUR 160 billion for the Framework Programme, which is also the highest budget option in Section 5.1. While budgets are the same under both options, the assumptions behind the budget increase compared to the baseline are different: in the centralisation scenario, funds for R&I are shifted from national to EU-level, while the budget increase considered in section 5.1 is compensated by a decrease in national investments. As a result, the centralisation scenario produces lower results compared to a scenario where national funds for R&I are not decreased. Hence, this result shows that an increase in EU budget for R&I is more beneficial if it does not crowd out national R&I support.

## 5.3 ASSESSING CHANGES IN THE DESIGN OF THE PROGRAMME

Changes in the design of Horizon Europe compared to Horizon 2020 (European Commission, 2018) aim at even more impact and openness. This will be achieved through several features such as the European Innovation Council, the mission-orientation, a strengthened international cooperation, a reinforced Open Science policy, and a new policy approach to European Partnerships. Assessing the overall impact of these changes is a very challenging exercise, as they correspond to several incremental improvements that are expected to affect different aspects of the Programme.

It is important to highlight the extent to which potential improvements in the design of the future Programme can enhance its impacts. This can be achieved in two steps. First, the impact of expected changes can be translated in changes in specific parameters of the model, which need to be identified. Second, the variation of these parameters needs to be quantified. This quantification is the most difficult task, as a lot of uncertainty encompasses the future response of key parameters such as performance or leverage to these changes. The impact of the changes also depends strongly on the effectiveness in their implementation in the future Programme. Hence, while the impacts of these changes is expected to be positive, their size is uncertain. Therefore different scenarios are considered, from low to high, by using ranges in the variation of the parameters. These ranges rely on plausible values found in the literature, with extreme values showing how impactful Horizon Europe can be in the most ambitious and optimistic conditions.

The following parameters were adjusted in order to reflect the impact of the changes that could be implemented in Horizon Europe. Adjustment of parameters that correspond to changes in the design of the Programme to increase impact and openness are the following:

- Higher economic performance: Horizon Europe will aim for higher economic impacts, including more market-orientation. This is translated in the model by modifying the performance of the EU R&I programme (from +0 in a 'low' scenario to +5 percentage points in a 'high' scenario compared to the baseline).

Lower knowledge obsolescence: Horizon Europe will focus on more breakthrough innovations and create more fundamental knowledge that could make innovations last longer in time. The NEMESIS model uses a depreciation rate of 15%[9], which is widely used in the empirical literature (see e.g. Corrado et al., 2016). In a 'low' case, we retain an obsolescence rate of 14%, increasing from 5 to 6 years the average life duration of the knowledge created. In the "high" case, this duration reaches 6.5 years.

- Stronger complementarities with other innovative assets: this should be reinforced by the the more cross-technological and cross-sectoral R&I supported, and more focus on breakthrough technologies and mission-orientation. Complementarities are reinforced by 5% a "low" scenario and 10% in a 'high' scenario.
- Higher direct leverage of private R&D: Horizon Europe should enable a better access to finance for breakthrough innovating start-ups. The main expected impact should therefore be an enhancement of the direct leverage of the EU support on private firms' R&I investment. In a 'low' scenario, leverage is the same as in the baseline for applied research (0.1), while it corresponds to the upper bound of the estimated range of the meta-analysis conducted by Dimos and Pugh (2016) in the 'high' scenario.
- Higher complementarities with national support to R&D: the programme is expected to increase complementarities between EU and national supports to R&I. This should reinforce national support to R&D, which is financed in the model by an equivalent reduction of other public investments (excluding R&I activities). This is translated indirectly in the model by an increased leverage of EU support on national support (adjusted here through increased leverage for basic research, set at 0.05 in a 'low' scenario and 0.1 in a 'high' scenario).
- Stronger knowledge diffusion: Horizon Europe should facilitate knowledge diffusion, encouraging multi-disciplinary collaborations, international cooperation and open science. Based on the literature, reasonable values, in the light of the progress achieved between FP7 and Horizon 2020 in terms of knowledge diffusion (see e.g. Vullings et al., 2014, or European Commission 2017a) should reflect increased knowledge spillovers compared to the baseline scenario: values used in this paper are +5% in the 'low' scenario and +10% in the 'high' scenario.

Results of all these changes in terms of GDP deviation according to the 'low' and 'high' scenario are presented in Figure 7. Compared to the continuation of Horizon 2020, changes in the programme's design can potentially generate an additional GDP gain up to 0.04% in a low scenario, and up to 0.1% in a high scenario. The impact of the changes is expected to be most significant after 2030. The total impact of the programme on EU GDP would be between EUR 800 billion and EUR 975 billion over 25 years[10].

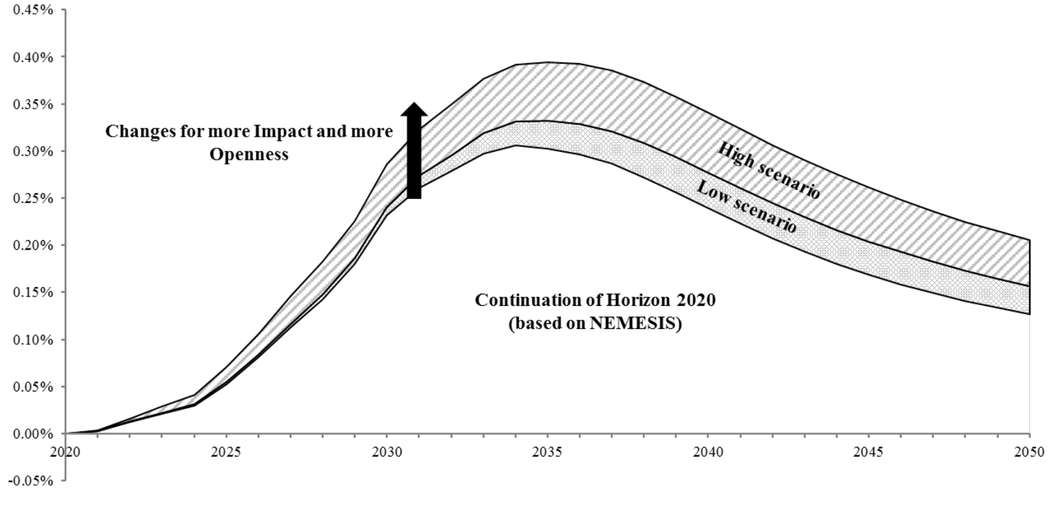



**Figure 7** Impact of the changes in design (GDP gain, compared to a situation without Framework Programme)
**Source:** Authors' calculations.

When combining the improved design of the programme with the budget proposed by the European Commission for Horizon Europe (EUR 100 billion in current prices), results from the NEMESIS model suggest that the estimated total impact of the Programme is even higher. In comparison to a case with no Framework Programme, the programme could generate up to ~ EUR 45 billion[11] per year over 25 years (~ EUR 1100 billion in total), i.e. up to 11 euros of GDP gains per euro invested over 25 years. Moreover, Horizon Europe could create up to 140,000 units of employment in R&I activities during its lifetime (2021-2027) and up to 340,000 units afterwards (figures based on maximum employment deviations estimated by the model). This corresponds to an average deviation of total employment of 170,000 units over the period compared to a situation without Framework Programme.

# 6 CONCLUSION

This paper shows how current models can be applied in assessing ex-ante the impact of a large and complex R&I Programme such as Horizon Europe. While the programme is expected to have various types of impacts along different impact pathways (European Commission, 2018; Bruno and Kadunc, 2018), including societal impacts, the paper focuses on economic impacts in terms of GDP and employment. A key novelty in the approach is the triangulation of results from three macroeconomic models (NEMESIS, QUEST and RHOMOLO) for the assessment of the baseline scenario, i.e. the continuation of Horizon 2020 over 2021-2027. This exercise shows that the models tend to agree on the pattern and sign of the impact of the Framework Programme. However, the size of the impact seems to depend on the specificities of the models (this includes their elasticities and intrinsic mechanisms) and on assumptions related to the EU added value of public investments in R&I and the way the programme is funded in the models. These assumptions also affect directly the analysis of options related to more centralisation or decentralisaiton at the national level of the management of EU R&I funding. Regarding EU added value, the higher performance of the Programme in the NEMESIS model is related to assumptions on higher performance and leverage compared to national funding. While there is strong evidence to support this EU added value, the quantification of its impact should be supported by further analyses. The empirical literature on this aspect is still poor and would benefit from additional contributions. Regarding funding, results from the QUEST model suggest that funding through VAT funding is more beneficial compared to lowering national investments. NEMESIS also shows that increasing the budget of Horizon Europe is much more impactful if it crowds out national public investments (except in R&I) instead of national public support for R&I.

Overall, past and current experience demonstrates the growing importance of macroeconomic modelling in the evaluation and impact assessment of EU R&I policy. Today, the need for state-of-the-art modelling approaches all along the policy cycle has never been so great. However, the complexity of the modelling exercise can make it challenging for policy-makers and modellers to collaborate with each other. It is not always simple to tailor a model to the specific needs of a precise R&I intervention. For example, while budget allocation and size can be easily translated into the mechanisms of a model, changes in the design, content or priorities of a programme require careful reflection as there is not always a straightforward adjustement of parameters in the models that corresponds to these changes. However, it is in general possible to find a way to proxy in these models the options considered by the policy-maker.

In this respect, modellers should help policy-makers understand the key features and assumptions of their models. More generally, policy-makers and modellers should collaborate closely with each other, hence allowing to better shape the things to come.

# REFERENCES

**Boitier, B., Le Mouël, P., and Zagamé, P.** (2018). Support for assessment of socio-economic and environmental impacts (SEEI) of European R&I programme. European Commission report.
Doi: 10.2777/038591.

**Bruno, N. and Kadunc, M.** (2018). Impact Pathways: Tracking and communicating the impact of the European Framework Programme for research and innovation. Paper presented at the Austrian Presidency of the Council of the European Union Conference on Impact of Research and Innovation Policy at the Crossroads of Policy Design, Implementation and Evaluation in Vienna, Austria.

**Corrado, C., Haskel, J., Jona-Lasinio, C. and Massimiliano, I.** (2016). Intangible investment in the EU and US before and since the Great Recession and its contribution to productivity growth, EIB Working papers, 2016/08.

**Delanghe, H. and Muldur, U.** (2007). Impact assessment of research programmes: The experience of the European Union's 7th Framework Programme. Science and Public Policy, 34(3), 169-183.

**Delanghe, H., Sloan, B. and Muldur, U.** (2011). European research policy and bibliometrics indicators, 1990-2005. Scientometrics, 87(2), 389-398.

**Di Comite F. and Kancs D.** (2015). Macro-Economic Models for R&D and Innovation Policies. IPTS Working Papers on Corporate R&D and Innovation, 03/2015.

**Dimos, C. and Pugh, G.** (2016). The effectiveness of R&D subsidies: A meta-regression analysis of the evaluation literature, Research Policy, 46(4), 797-815.

**ECDG and Elsevier** (2017). Overall output of select geographical group comparators and related FP7- and H2020-funded publication output. Final Report, https://frama.link/C7wPJhGp.

**European Commission** (2013). The Grand Challenge – The design and societal impact of Horizon 2020. Directorate-General for Research and Innovation. Doi: 10.2777/85874.

**European Commission** (2015). Better Regulation Guidelines. Commission Staff Working Document. SWD(2015)110 final.

**European Commission** (2017a). The economic rationale for public R&I funding and its impact, Policy Brief Series, Directorate-General for Research and Innovation, Doi: 10.2777/047015.

**European Commission** (2017b). Interim Evaluation of Horizon 2020. SWD(2017)220, book.

**European Commission** (2017c). LAB – FAB – APP Investing in the European future we want: report of the independent High Level Group on maximising the impact of EU research & innovation programmes. Directorate-General for Research and Innovation.

**European Commission** (2018). Commission Staff Working Document: Impact Assessment of Horizon Europe. SWD(2018)307 final.

**Hall, B.H., Mairesse, J. and Mohnen, P.** (2009). Measuring the Returns to R&D. NBER Working Paper No. 15622.

**Le Mouël, P., Le Hir, B., Fougeyrollas, A., Zagamé, P. and and Boitier, B.** (2016). Toward a macro-modelling of European Innovation Union: The contribution of NEMESIS model. Paper presented at the 9th Conference on Model-based Evidence on Innovation and DEvelopment (MEIDE) the 16-17 June 2016 in Moscow, Russia. https://frama.link/8Gmb7fnM.

**Muldur, U., Corvers, F., Delanghe, H., Dratwa, J., Heimberger, D., Sloan, B., Vanslembrouck, S.** (2006). A new deal for an effective European research policy: The design and impacts of the 7th Framework Programme. Dordrecht: Springer.

**PPMI** (2017). Assessment of the Union Added Value and the Economic Impact of the EU Framework Programmes (FP7, Horizon 2020). https://frama.link/o6oBPRZU.

**Rosemberg, C., Wain, M., Simmonds, P., Mahieu, B. and Farla, K.** (2016). Ex-post evaluation of Ireland's Participation in the 7th EU Framework Programme – Final Report. Technopolis Group, June. https://frama.link/9kXjUv_3

**Vullings, W., Arnold, E., Boekholt, P., Horvat, M., Mostert, B., Rijnders-Nagle, M.** (2014). European Added Value of EU Science, Technology and Innovation actions and EU-Member State Partnership in international cooperation – Main Report. Technolopis Group. doi:10.2777/1193.

# AUTHORS

**JULIEN RAVET**
*iCite, Solvay Brussels School of Economics and Management (SBS-EM, ULB), Belgium*
E: julien.ravet@ec.europa.eu

**BAPTISTE BOITIER**
*Société européenne d'économie (Seureco), France*

**MARCO GRANCAGNOLO**
*European Commission - DG Research and Innovation, Belgium*

**PIERRE LE MOUËL**
*Société européenne d'économie (Seureco), France*

**LIVIU STIRBAT**
*European Commission - DG Research and Innovation, Belgium*

**PAUL ZAGAMÉ**
*Société européenne d'économie (Seureco), France*

## KEYWORDS:

Research and Innovation Framework Programme, Horizon Europe, Ex-ante impact assessment, Macroeconomic modelling, NEMESIS, QUEST, RHOMOLO.

Endnotes

1 Notably the interim evaluation of Horizon 2020 (European Commission, 2017b).

2 A high-level group chaired by Pascal Lamy was set up by the European Commission in order to provide advice on how to maximise the impact of the EU's investment in research and innovation (European Commission, 2017c).

3 Courtesy of DG ECFIN and DG JRC of the European Commission for the results of, respectively, the QUEST and RHOMOLO models.

4 Horizon 2020 allocations were calculated based on data from CORDA.

5 The direct leverage effect is the difference between the total subsidy (to support R&I investments) received by an R&I entity and the total R&I expenditure engaged by this entity as the result of this support. For instance, if the direct leverage effect is positive (crowding-in), this means that the financial support received by the entity has a "multiplier" effect on the R&I investments of this entity. In this case, total R&I expenditures are higher than the financial support received.

6 Economic performance reflects the outcomes of investments in terms of sectoral value added or production (and GDP at national level). Economic performance of R&D investments in the NEMESIS model is based on the empirical literature (e.g. Hall et al., 2009).

7 With budget allocations being proportional to Horizon 2020 allocation.

8 According to GBAORD data (Eurostat), the share of project funding in total EU GBAORD is around 30% (Boitier et al., 2018). By converting this amount in percentage value of the government public expenditures in R&D, it amounts to about 26%. Therefore, the centralisation of a third of the national competitive-based project funding at EU level is, on average, equivalent to centralise, at EU level, 8.75% of the government public R&D expenditures in each member state, which represents around € 9 billion (constant 2014) per year.

9 With this 15% depreciation rate, more than half of the knowledge created today will become obsolete after 5 and half years.

10 In 2018 prices.

11 In 2018 prices.

# THE USE OF RESEARCH PORTFOLIOS IN SCIENCE POLICY

NICHOLAS S. VONORTAS AND ISMAEL RÀFOLS
DOI: 10.22163/fteval.2019.338

Acknowledgement: This research has been supported by the Science of Science and Innovation Policy program of the US National Science Foundation with the award SMA1734829 (2017) "Science Policy Research Report: The Use of Research Portfolios in Science Policy" to the George Washington University.

## ABSTRACT

Any research funding and/or research-performing organization in the public, private, and non-profit sectors needs to adopt a portfolio-wide perspective to R&D management to better align research project investments with the organization's overall strategic goals. Private sector firms have increasingly done so utilizing powerful new methodological tools and large amounts of data becoming available. In contrast, with relatively few exceptions, public R&D management still tends to base selection processes on the excellence of individual projects according to peers rather than considering the merits of the whole portfolio. There are good reasons for additional caution which, besides the usual inertia and the resistance by scientists trusting the peer review process, include multiple objectives of public programs, long-term accrual of results and associated uncertainties, and difficulty to monetize or value. This report argues it is high time for public R&D management to move forward. Portfolio analysis should not be applied similarly across the board. It will serve different purposes for different types of public R&D programs depending on risk/uncertainty, data availability, and target clarity (ability to define unambiguous program goals). Not all methodologies will be appropriate to all programs. Nonetheless, the toolkit, data depositories, and computing capability have expanded tremendously during the past couple of decades to render such experimentation possible and absolutely necessary.

## 1. INTRODUCTION

Public research and development (R&D) is widely believed to be important for improving knowledge, fostering economic growth and social well-being. Consequently, research and innovation policy can be seen as an investment and be designed, in part, in terms of expected socio-economic 'returns', their timing and degrees of risk taking (Borrás and Edquist, 2014). Analysts have, however, struggled to provide robust, widely acceptable methods to support decision-making for future investments. This has led to arguments of insufficient empirical or theoretical basis for making or justifying specific choices for investment (ITG, 2008, p. 1) and perceptions that public research is not adequately addressing societal needs such as global health (Sarewitz and Pielke, 2007). While there has been significant methodological progress during the past ten years or so – some of which is surveyed in this report – the issue is far from resolved.

A research portfolio is defined as the set of research activities supported by a funding and/or research-performing organization or a group of agencies/organizations. In large technology-intensive companies, portfolio-wide perspective to R&D management has long been applied as a means of better aligning research project investments with the firm's overall strategic goal of economic return maximization (Schilling, 2017). Still, with relatively few exceptions in the public sector (Ruegg, 2007), public R&D management still tends to base selection processes on the individual excellence of projects according to peers rather than considering the merits of the whole portfolio (Linton and Vonortas, 2015; Linquiti, 2015). Nascent attempts such as the Office of Portfolio Analysis at the National Institutes of Health are commentable but have yet to reach full acceptance, often due to resistance by the client community (scientists). In some contexts, research portfolios are described as 'profiles' (e.g. in German universities, Meier and Schimank, 2010).

There is no question that the appraisal of research portfolios is challenging across the board. It may be relatively more so in the public sector. Besides the usual inertia and the resistance by scientists trusting the peer review process, there are other serious reasons why this may be so: multiple objectives, project interdependency, difficulty to monetize or value. Public programs will frequently have multiple objectives requiring multiple (perhaps incompatible) performance measures for evaluation. Research projects and programs in public research portfolios can be interdependent. Their outputs are typically removed from the market, thus making monetary valuations arbitrary. Hence, accounting for public research investment in purely monetary terms is not advisable when looking at investments with uncertain evolution and payoff structure, as well as "fuzziness" in terms of the social desirability of the "impact" and associated values.

Improvements in data processing and visualization techniques (Börner et al., 2003; Van Eck and Waltman, 2014), coupled with conceptual developments in research and analytical methods better handling risk (Lo Nigro et al., 2016; Luehrman, 1998; Vonortas and Desai, 2007) in the last couple of decades, however, suggest that research portfolio approaches offer the possibility of improving the performance of R&D programs by identifying gaps and opportunities. They also help in making more transparent the multiple goals of most public R&D programs – thus facilitating the alignment of research with its various welfare, environmental, security and economic missions (Wallace and Rafols, 2015).

The rest of this paper runs as follows. Section 2 recounts the analytical literature concentrating on the modeling aspect of research portfolio analysis to quantify the returns to uncertain R&D. Section 3 recounts the literature using the notion of research portfolio as a heuristic for deliberation on research priorities and project selection in the face of incomplete quantifiable information, deep uncertainty, and lack of agreement on goals. Section 4 introduces recent advances in data availability, processing and visualization techniques which greatly facilitate portfolio management. We draw overall conclusions for policy decision makers in Section 5.

# 2. R&D PORTFOLIO MODELLING

## 2.1 ECONOMIC APPROACHES

Economic impact analysis is one part of an R&D program evaluation. Quantitative economic appraisals of public sector-funded R&D usually lean on capital budgeting methods extensively used in the private sector (Link and Scott, 2013). This analytical and theoretical framework has long been germane to the economics and business technology management literature. One critical aspect of this literature is the need to consider the counterfactual situation that would have existed should the R&D program in question had never materialized: the evaluation accounts for the incremental benefits between the two (additionality). A second critical aspect is the recognition of various types of spillovers, that is, circumstances where the (private) producer of knowledge cannot extract through the market system the full value the new knowledge adds to the economy. Such spillovers can be pecuniary, knowledge, and network spillovers, reflecting the different ways value escapes the original inventor. Spillovers do not necessarily imply inaction for the private sector. They do, however, imply market failure to some extent – underinvestment from society's perspective – and should be accounted for when calculating the social rate of return of the R&D program in question.

The classic approach to appraise economic returns to an investment is the net present value (NPV) (cash flow model) and the related internal rate of return (IRR). The model is expressed by the well-known function

$$\mathrm{NPV} = \sum_{t=0}^{T} \frac{F_t}{(1+r)^t},$$

where Ft is net cash flow at time t and T is the final time period. Link and Scott (2013) summarize a set of seventeen laboratory-based economic impact analyses of this type. While their analytical method arguably has portfolio characteristics – mainly by looking at effects throughout the supply chain rather than just to first tier beneficiaries – it also misses important others such as the explicit evaluation of interdependencies between R&D projects, of the greater strategic goals, and of effects beyond direct benefit/cost (public R&D typically has more than one objectives).

A good example of an explicit R&D portfolio analysis approach is National Research Council's development and application of an extended NPV methodology to estimate ex ante the net benefits of R&D projects of the US Department of Energy (NRC, 2005; 2007). This work was mandated by Congress which several years earlier had requested the NRC to produce a series of reports using quantitative indicators to appraise the effectiveness of applied energy R&D. The first report was a retrospective look of DOE's research on fossil energy and energy efficiency (NRC, 2001).

The most methodologically advanced of these reports (NRC, 2007) used a consistent methodology across six cases of applied energy research portfolios.[1] The study offered a significant advancement on prior practice by looking at all three perceived primary effects of DOE's programs: (1) to reduce technical risk; (2) to reduce market risk; and (3) to accelerate the introduction of the technology into the marketplace. The methodology uses expert panel reviews of the DOE R&D programs and estimates the expected economic, environmental, and energy security benefits in three different global economic scenarios. Decision trees are built to describe the technical and market uncertainties and the impact of DOE support in overcoming them. Finally, the acceleration effect was represented either by the change in the likelihood of a project to attain the program goals of completion by a critical date, or by the acceleration of their benefits vis a vis technology developing in the absence of the government program. The overall benefit of the DOE R&D program is given as the difference between the expected net benefits with DOE support and the expected net benefits without it (counterfactual). The expected benefits correspond to a probability-weighted average of the benefits in specific technical and market outcomes, within common scenarios and under common assumptions. Scenarios were built with the help of NEMS[2] forecasting the likely energy cost savings through 2030 from the deployment of the new technology generated by the program. The traditional discounted cash flow framework (NPV) was used for these calculations.

Linquiti (2015) has subsequently reevaluated one of those six cases – Chemical Industrial Technologies program – pointing out three shortcomings. The first relates to the use of point estimates, rather than a range (probability distribution), for the value of annual energy savings from each new technology. The second is the omission of interdependencies among R&D projects in the portfolio.[3] The third shortcoming is the use of the discount rates of 3% and 7% suggested by the Office of Management and Budget. The difference between the two is said to constitute a risk premium. As such, it is argued that the use of a 3% rate can be justified on the basis that public sector program administrators should not exhibit risk aversion. The use of a risk-adjusted discount rate (7%) is more difficult to justify.

The literature on project selection in the context of institutional R&D portfolio management is already extensive.[4] A good part of it focuses on the construction of portfolios of projects meeting certain merit criteria. However:

1 Integrated Gasification Combined Cycle Technology R&D program; Carbon Sequestration program; Natural Gas Exploration and Production R&D program; Distributed Energy Resources program; Light-Duty Vehicle Hybrid Technology R&D program; and Chemical Industrial Technologies program.

2 The National Energy Modeling System (NEMS) of the Energy Information Administration is a comprehensive computer-based system for modeling U.S. energy markets. It projects the production, consumption, imports, and prices of energy, subject to assumptions about macroeconomic and financial factors, world energy markets, resource availability and costs, behavioral and technological choice criteria, cost and performance characteristics of energy technologies, and demographics.

3 The NRC study notes the potential for such interaction.

4 See, for instance, recent accounts in Lo Nigro et al. (2016), Verbano and Nosella (2010), Baker et al. (2015), Vilkkumaa et al. (2015), Zschocke et al. (2014) and references therein.

*"Although rating the individual merit of a project is important, managers are increasingly seeking to maximize the overall value of their research portfolios by bringing the portfolios into alignment with strategic goals. This requires consideration of the relative merits of projects based on the overall capacity being generated by the sum of the combined individual projects. Taking a portfolio approach can also minimize unnecessary duplication of efforts and increase the synergy between inter-dependent projects. Measuring and anticipating these synergies is an exponentially difficult task since it requires a framework for gauging the relational importance of the inputs and outputs for a series of projects while at the same time considering the links between projects and their sub-activities in a portfolio. Projects are typically performed on varying time scales, have varying resource requirements, and have dissimilar goals. For example, some projects may not lead directly to monetary returns but may be invaluable for developing technical competencies and advancing the frontier of knowledge. The important concept to retain is that the combination of all of the individually good projects does not necessarily constitute an optimal portfolio (Chien, 2002)."* (Casault et al., 2013a, p. 89)

The idea of R&D portfolio analysis goes back to principles in finance and, in particular, the idea that assets should not be selected solely on the basis of their individual merits. Markowitz (1952) demonstrated that risks are not additive; neither are returns of financial assets. Evaluation of an asset's return should be in relation to other assets in the portfolio and overall market fluctuations.

$$E\left(R_p\right)=\sum_i w_i E\left(R_i\right)\ ,$$

where E stands for expectation, Rp is the return on the portfolio, and wi are weights on individual assets' returns, Ri. The risk associated with individual investments is managed through diversification: portfolio balancing combines assets that will be profitable as a group despite the uncertainties of individual assets and of the overall market. Financial portfolio managers diversify the investments in their portfolio to obtain a predetermined aggregate risk profile.

Much of the basic thinking of financial asset management applies to R&D project management. Both financial and "real" options give the option holder the right, but not the obligation, to take an action at a future date. Here too one deals with risky investments and uncertain markets. Real options are likely to be valuable when future outcomes are uncertain, there is flexibility to act in the future as the uncertainty is resolved, and the action can increase net benefits (Triantis, 2003). Here too one must consider relationships among projects, which can be both positive or negative (van Bekkum et al., 2009). Here too R&D project portfolio diversification enables achieving complex – and often conflicting – goals of an R&D strategy that cannot be attained by any single R&D project (Eilat et al., 2006).

The result has been the development of a quite extensive literature that has recognized the undervaluation by net cash flow techniques (NPV) of the managerial flexibility associated with real assets such as technical knowledge. Attempts for enhanced NPV applied in combination with decision trees has gone some way to account for this value as well as for addressing the deficiencies of the NRC work mentioned earlier. Still, there is a strong call for R&D investments to be analyzed as "real options" (Vonortas and Desai, 2007; Linquiti, 2015) – also including real compound options (Cassimon et al, 2011) – which more recently has been enriched further with an impressive (but still analytically difficult) literature on portfolios of R&D options.[5] An important reason for looking at portfolios of options is the realization that the optimal decision under uncertainty is not an average of the optimal decisions under certainty and it is not necessarily near the optimal decision under a core case (Baker et al., 2015). In short:

*"…[T]he certain absence of risk additivity in all investment portfolios, the frequent absence of return additivity in R&D portfolios, the value of purposively trading off risk and return, and the complex interaction of investments with conditional payoffs are all persuasive reasons to analyze and value not only individual R&D projects, but also the R&D portfolios they comprise."* (Linquiti, 2015, p.63-64).

Nonetheless, the application of financial portfolio theory to R&D project analysis is subject to difficulties (Casault et al, 2013a). For one, R&D projects and their outcomes (underlying assets) are very seldom traded in the market[6] and there is little information about the project's inherent value and expected future returns (on which the option valuation depends). Relatedly, R&D projects produce returns that are hard to monetize – the returns may arrive far into the future, they may relate to defense, security of natural resources, improvement of the natural environment, regulation, or reputation.[7] Monetary returns may not even be an important decision variable for R&D project selection. For a second, financial assets are typically assumed to behave in a Gaussian manner: expected returns have a defined mean and do not fluctuate much away from it (95.4% of all measurements will register within ±2Đ from the mean). Casault et al. (2013b) argue that this assumption is likely to be inappropriate for R&D projects where distinct milestones can greatly influence the expected value of the project. Long tail (large fluctuation) events define the system and cannot be ignored.

## MIXED-METHOD APPROACHES TO MODELLING

In order to account for multiple, difficult to monetize, and often conflicting program and project goals, a diverse set of alternative non-parametric methods to draw up real asset portfolios (including R&D) have been developed. They have been reviewed time and again in a burgeoning literature on mixed methods for constructing and analyzing R&D portfolios (Kurth et al., 2017; Gemici-Ozkan et al., 2010) and multi-criteria analyses (Kurth et al., 2017; Linton et al., 2002; Marafon et al., 2015). With multiple goals, the key question is to which extent the implicit prioritization of goals in research portfolios (science supply) fits with perceptions of socioeconomic demands or needs – as captured by experts (Sarewitz and Pielke, 2007). Recent reviews include Verbano and Nosella (2010), Casault et al. (2013a), and Linquiti (2015).

5 See, for example, Smit and Trigeorgis (2006), Brosch (2008), Magazzini et al. (2016), Montajabiha et al. (2017), van Bekkum et al. (2009).

6 Financial options are linked to traded financial securities whereas a R&D option is associated with non-tradeable (in the sense of fixed market prices) knowledge and information.

7 Nonetheless, there have been efforts to monetize such effects. See, for instance, the aforementioned studies of NRC (2005, 2007). Here is a need for further research.

The reader is referred to those sources for detail. Here we offer a summary view of some of the best known methods.

- **Peer review score**. Classic technique, it involves experts affixing a score on individual projects against a series of merit criteria. Projects are then rank ordered and the top projects selected. Despite serious deficiencies in systematic portfolio formulation, the process is useful in early stage activities ensuring the quality of projects that may form a portfolio.
- **Analytic Hierarchy Process (AHP)**. Technique to organize and analyze complex input from various sources. It helps structure a problem in terms of various quantifiable elements organized logically so that they can be measured against overall goals and alternative solutions. A hierarchy is structured starting with an overall project objective at the highest level that is decomposed into a series of uncorrelated criteria which can be further decomposed into a series of sub-criteria on as many levels as required by the problem. The lowest hierarchical level describes a series of alternative solutions for completing the criteria immediately above. Evaluation based on pairwise comparisons by experts which can be processed mathematically to determine overall project "efficiency". AHP is better viewed as an input to support decision making. It can be followed by a second optimization process for the overall portfolio.
- **Data Envelopment Analysis (DEA)**. Non-parametric methodology to estimate a frontier by estimating the relative efficiency of a number of producers. Efficiency is defined as the ratio of the sum of weighted outputs to the sum of weighted inputs. Advantages include avoidance of specifying mathematical functions and ability to compare quantitative and qualitative factors. The technique can also deal with a portfolio of projects with or without interactions.
- **Balance Scorecard (BSC)**. A model for analyzing strategy and performance information for all types of organizations (Kaplan and Norton, 1992). Widely adopted in the private sector to plan and align strategic initiatives, clarify and translate vision and strategy into action, and enhance strategic feedback and learning. The technique purports to provide a balance between (1) short- and long-term objectives; (2) financial and non-financial measures; (3) lagging and leading indicators; and (4) internal and external performance perspectives. Weaknesses include complexity of performance measurement, judgement biases, and the need to reach some synthetic metric that summarizes the whole set of multiple perspectives and indicators into success or failure. Multi-criteria decision-making frameworks are an appropriate approach to untangling these complexities in performance evaluation and decision-making.

Most of the techniques used by practitioners have been hybridized to help provide richer pictures of portfolios than any single technique. For instance, Eilat et al. (2006) combined BSC with DEA to establish a methodology to evaluate alternative portfolios of projects in order to choose the best combination. In another example, Kim et al. (2016) combined AHP and BSC to analyze the strategic fit of portfolio of national R&D programs with R&D policies.

# 3. RESEARCH PORTFOLIO AS A HEURISTIC FOR MANAGING RESEARCH PRIORITIES

There are growing concerns that research needs to become more responsive to societal needs and demands (Sarewitz and Pielke, 2007; Bozeman and Sarewitz, 2011). Posed in simple terms, the question is: "Are we doing the right type of science given current societal needs?" The answer to this question is often highly critical, as illustrated by widespread debate generated by Sarewitz' article in *The New Atlantis* in 2016: although research does contribute to wellbeing, it could be better aligned with societal needs or demands. Some empirical studies in health support the claims of misalignment (e.g. in prioritisation across diseases as shown by Evans et al., 2014 or Yegros and Rafols, 2018).

In order to improve alignment between research and societal needs, public R&D agencies have put in place a variety of initiatives for priority setting, such as grand challenges (Hicks, 2016) and participatory processes for setting research agendas (e.g. in health, the UK-based James Lind Alliance[8], or nationally in the Netherlands[9]). In this broader and more political discussions on priority setting, given high uncertainty and lack of value consensus (ambiguity), R&D portfolio analysis serves different purposes and requires different management strategies.

## 3.1 R&D PORTFOLIO ANALYSIS UNDER CONDITIONS OF HIGH UNCERTAINTY AND AMBIGUITY

The quantitative techniques and mixed-methods for portfolio modelling presented earlier are useful for applied research in conditions in which there is a reasonable understanding of the potential outcomes of projects and in which value or goal disagreements regarding priorities are relatively minor. In the context of research that is not applied downstream, making estimates of project success in the face of multiple and ambiguous goals becomes very difficult.

There are two types of limitations regarding knowledge, as illustrated in Figure 1, following Stirling and Scoones (2009). On the one hand, there is the uncertainty about possibilities of research success in achieving the expected goals. When the probabilities of success can be estimated, as in finance, one can use the concept of 'risk', meaning that there is some statistical information about expectations of success and portfolio modelling is possible. Under conditions of multiple, but well-defined goals (shifting towards the right to 'Ambiguity'), mixed methods such as Peer Review Score or Data Envelopment Analysis can be helpful. However, when probabilities cannot be estimated we should stay with the notion of 'uncertainty'. On the other hand, there is the ambiguity, or lack of knowledge, or lack of agreement regarding the goals of a project, particularly in the very common situations of public R&D in which there are multiple goals. In summary, under conditions of high ambiguity and/or high uncertainty, modelling becomes problematic.

8 http://www.jla.nihr.ac.uk/
9 See Knowledge Coalition (2016) The National Research Agenda. Knowledge Coalition. https://wetenschapsagenda.nl/?lang=en

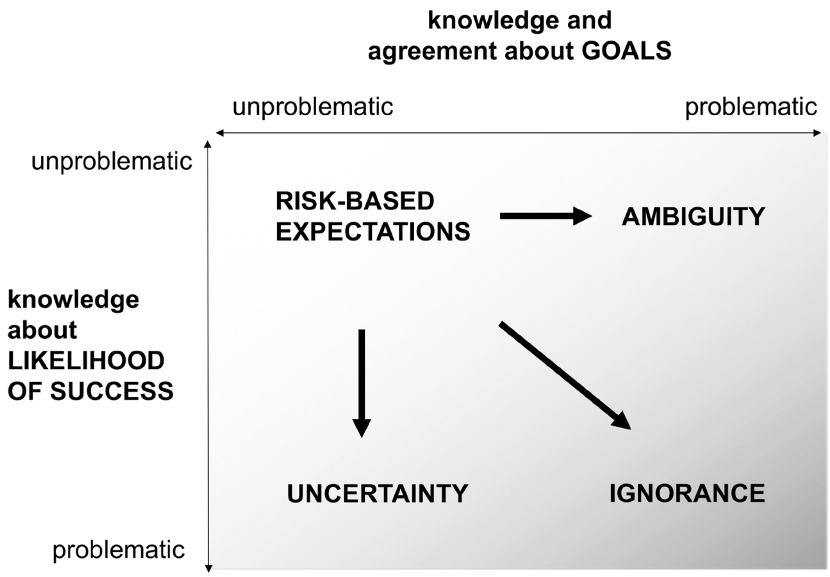



**Figure 1.** Types of knowledge limitations in relation to project management.
**Source:** Adapted from Stirling and Scoones (2009).

Under these conditions of ambiguity and contested nature of the goals (given multiple desirable outcomes) and high uncertainty, the analogy with the financial portfolios breaks down to a large extent (Wallace and Rafols, 2015). The techniques reviewed in section 2 of portfolio modelling can still play an important role at illuminating the value of diversity and seeking positive interactions or complementarity between projects in resource allocation. However, under uncertainty and ambiguity R&D portfolio analysis can be particularly helpful as a tool to coordinate collective reflexivity on the goals and the expected outcomes of research programs. For example, in agreement with calls for mapping the public values of research (Bozeman and Sarewitz, 2011) and responsible innovation (Stilgoe et al. 2013), R&D portfolios are explored by the UK BBSRC[10] as a means to foster "anticipation, inclusion, reflexivity and responsiveness" in research management through participatory processes (Smith et al., 2016).

The opening up of portfolio analysis to a broader set of participants – from scientific experts and policy maker to wider forms of expertise and lay people — is consistent with Pielke's (2007) view that under conditions in of high uncertainty and lack of value agreement, one cannot separate analysis and decision-making as two separate, consecutive processes. Since technical assumptions used in modelling analyses can depend experts' values and can be biased towards quantifiable evidence, portfolio analysis should ideally be examined by diverse stakeholders bringing in contrasting perspectives on uncertainties and ambiguities. In this way, it is possible to build-up evidence-based policy making while trying to include those sources of evidence that are less quantifiable, formalized or institutionalized (Saltelli and Giampetro, 2017).

## 3.2 COMPARING SCIENCE SUPPLY AND SOCIETAL NEEDS

There can be many heuristics or strategies for mixed-methods or qualitative analysis of research portfolios depending on the goals, organizations and contexts of the research programs. In general, it involves the comparison of the composition of a portfolio (science supply) with the distribution of desired or expected outcomes (societal needs).

Hage et al. (2007) provide a useful and pragmatic framework to qualitatively assess the composition of a portfolio. The key questions to be posed are: "Where to invest? What capabilities are needed and where? Which coordination mechanism should be used and where?". Building up capabilities for a certain portfolio focus involves thinking about the personal skills and technological instruments needed and providing training programs, whether new kinds of organizations or coordination activities are needed (e.g. new technology platforms). Emphasis in capabilities reminds us that societal impact is often not achieved directly through the research carried out, but through the capabilities created, particularly in terms of human resources (Bozeman and Rogers 2001).

In portfolios for issues around large scale societal problems or grand challenges, it will be particularly important to pay attention to coordination mechanisms between different arenas of research – whether more basic, applied, commercialization, etc. The ensemble of programs or policy actions within a given R&D portfolio can be thought as the 'policy mix' that will implement it (Flanagan, Uyarra and Laranja, 2011).

10 Biotechnology and Biological Science Research Council.

**Appraisal of science supply**

The first and paramount question to be addressed in portfolio management is 'Where to Invest'. The contents within an R&D portfolio, which will define the options or choices to be made, can be understood from different perspectives. Typically, they are defined in terms of disciplines, technologies, application or problems (Hage et al., 2007; p. 733). The choice of the specific perspectives used is very important as it will determine the type of priority setting, e.g. whether the choice is among disciplinary topics or among types of problems. Once a perspective is chosen with type of classifications (or 'ontologies') that describe the portfolio, the next step is to explore the distribution of research over categories, for example with a cognitive map or research landscape, as illustrated in Figure 2. This allows to begin asking questions such as *Where are there gaps? Where a small investment can make a noticeable impact?* (Hage et al., 2007, p. 734).

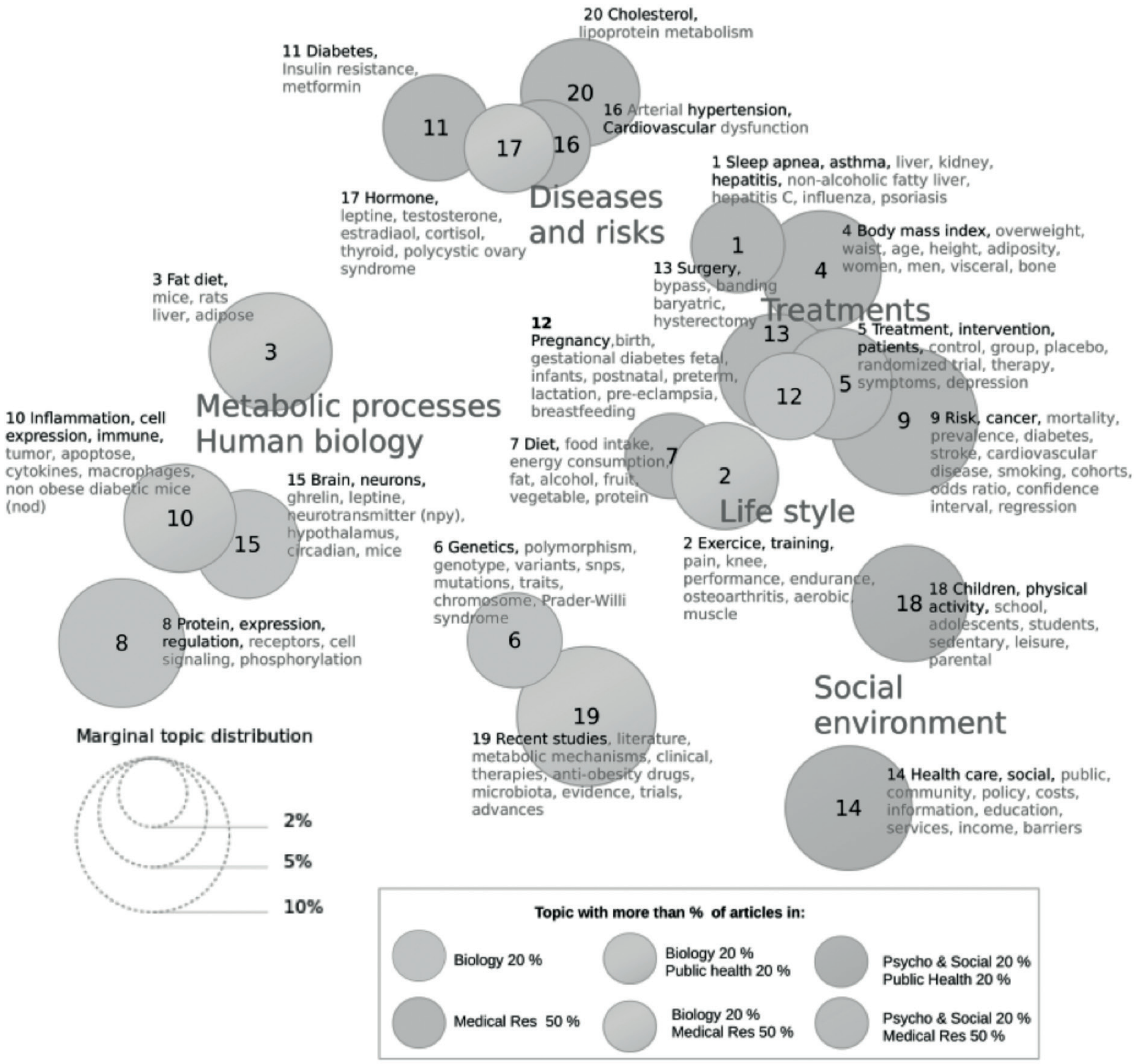



**Figure 2.** Relative distribution of publications related to obesity over various topics.
**Source:** Cassi et al. (2017).
**Note:** This figure illustrates the research landscape of obesity. The obesity portfolio of a given funding agency is defined by its distribution of topics over this landscape. The size of the circles is proportional to the number of publications in a given topic. Colours indicate main disciplines: basic biology (green, left), medical research (orange, top), public health and social sciences (purple, bottom right).

Improvements in data availability, data processing and science mapping have resulted in major advances in research portfolios visualization facilitating the task of portfolio mapping. These advances are detailed in Section 4 below. Although these new techniques are very helpful, they rely on decisions on classifications which often have important effects yet are poorly understood. It is thus important to keep a critical eye on classification schemes used.

**Appraisal of societal needs or demands**

The other key issue is to map societal needs or preferences about expected research outcomes. This is possibly the most challenging factor in portfolio management. Generally, there is no quantitative information about societal needs. Health is an important exception since one can use public estimates on burden in terms of years lost due to disease (e.g. DALYs Disability Adjusted Life Years) or in terms of labor or healthcare costs (Evans et al., 2014; Yegros and Rafols, 2018). Increasing availability of digital healthcare (big) data is quickly enhancing the possibility of making more fine-grained estimates of health needs. For example, the NIH shows the comparison between its research expenditure and disease burden in a dedicated webpage.[11] Also in the case of agriculture, one can make exploratory estimates of 'revealed demands' on the bases of data on crop exports, imports, cultivated area, food consumption or processing, and crop use in animal feed (Nature Plants, 2015; Ciarli and Rafols, 2017).

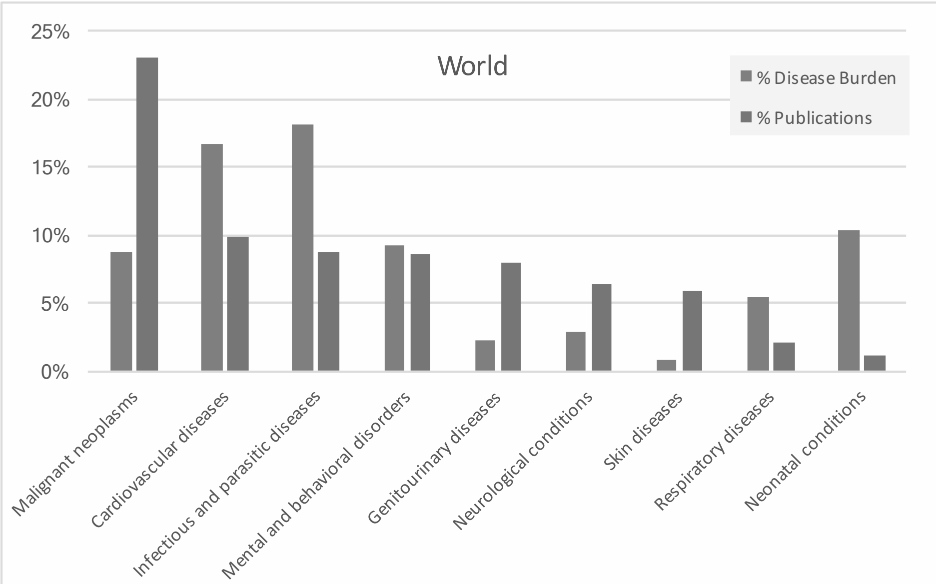



**Figure 3.** Comparison between relative disease burden and associated research output for the world.
**Source:** Yegros and Ràfols (2018).
**Note:** Percentage of disease burden (in blue, left) is based on WHO estimates in terms of Disability Adjusted Life Years (DALYs). Percentage of research outputs per disease (in red, right) are estimated from Web of Science publications using as disease classification MEDLine's Medical Subject Headings. Only selected categories are shown.

In the absence of data on societal needs or demands one alternative is to use scoring and multi-criteria methods such as those listed in Section 2.2. However, recent science policy initiatives point towards the importance of deliberative processes with a wide participation of stakeholders (e.g. the agenda developed by the EC Scientific Panel for Health[12]). The UK James Lind Alliance[13] is an example of a program of stakeholder engagement in priority setting of health needs for specific diseases.

In summary, there are now established and complementary methods – including institutional data, mixed approaches such as multi-criteria methods, and stakeholder deliberation – for making estimates of societal needs and preferences regarding research outcomes, even if results may always be interpreted as controversial.

## 3.3 IMPLEMENTING PROCESSES OF R&D PORTFOLIO APPRAISAL

R&D portfolio analysis can be used as a heuristic tool to appraise research priorities against societal needs or demands. It should be noted that the perspective on research portfolios focuses at program level within agencies, institutes or divisions.14 Various agencies are already using technical tools of portfolio analysis for reporting and information purposes, generally based on publication and funding data. However, R&D portfolio analysis requires not only various technical efforts, but institutional learning at implementation (Hellström et al., 2017).

On the basis of an experience in the UK BBSRC, Robert Smith and colleagues (2016) propose four management stages for implementing

11 https://report.nih.gov/info_disease_burden.aspx
12 https://ec.europa.eu/programmes/horizon2020/en/h2020-section/scientific-panel-health-sph
13 Established in 2004 and is supported by the UK National Institute of Health Research http://www.jla.nihr.ac.uk/
14 The distribution of resource across at higher levels shaped by political processes of budget allocation across agencies or divisions, is beyond the scope of this study.

portfolio deliberation participation in funding agencies. The first phase involves clarifying the aim and scope of stakeholder participation. The second phase mobilizes internal human resources in the agency in order to understand the scientific topics of the portfolios and the expected societal outcomes. The third phase involves a critical analysis of the knowledge base, while phase four identifies the stakeholders to participate. The deliberation process can follow methods thoroughly tested in engagement practices.

During the process of portfolio analysis aimed at funding, one should also be aware that public funding is only one of the determinants of *de facto* research priorities. Other factors having major influences include private funding, preferences (biases) implicit in research evaluation, and institutional goals, particularly in mission-oriented organizations funded via block grants (such as health research centers or agriculture institutes) (Wallace and Rafols, 2016).

# 4. DATA AVAILABILITY, PROCESSING AND VISUALIZATION OF PORTFOLIOS

Government policies of data transparency and accountability as well as technical advances in data availability, processing, classification and visualization are progressively facilitating the quantitative analysis of research portfolios. However, these developments are still in early phase and portfolio analysis has yet to overcome some technical hurdles. For example, a report by the Rathenau Institute notices that:

*"One of the most important initial results of this study was our observation that there is a major shortage of hard data on the allocation of research funding. That shortage makes it virtually impossible to develop informed policy, estimate policy effects and know whether the priorities set by a funding body will have an impact."* (Koier et al, 2016, p.11)

In spite of these difficulties, the technical support for portfolio analysis is quickly advancing. We present below developments in terms of data availability, processing and visualization following the steps in knowledge domain analysis (Borner, Chen and Boyack, 2003, p. 189).

**Data availability and infrastructure**

Knowledge infrastructure of project funding is now publicly available and keeps improving. US StarMetrics[15] (with Federal Reporter) or the UK Gateway to Research[16] contain details of publicly funded research, allowing large scale analysis of the performers, the contents and the contexts of research projects. Data providers such as the Web of Science now include acknowledgement of publications since 2009, though the data is based on self-reporting and has limitations (Costas and Van Leeuwen, 2012). Information services analysing these data are now being offered by academic analysts (e.g. at universities in Indiana, Leiden, or Montréal) and consultancies (e.g. ChalkLabs, SciTech Strategies and Uber Research). Funding agencies such as the NIH are creating internal information infrastructure and capabilities to manages portfolios (Srivastava et al., 2007; Haak et al., 2012).

**Data processing and classifications**

Data processing and classification is often the most opaque technical step in portfolio analysis. However, it deserves careful attention since the use of specific classification schemes and the subsequent categorization of projects has major implications. Large scale disciplinary classifications are based on journal classifications offered by data providers such as the Web of Science or Scopus, which show important differences (Rafols, Porter and Leydesdorff, 2010). In the last decade, more fine-grained and thematically accurate classifications based in article-level classifications have been developed (Waltman and Van Eck, 2012; Klavans and Boyack, 2017). However, these classifications rely on citation data and are thus problematic for grants. Co-word maps (Ciarli and Rafols, 2017) and new semantic algorithms, such as topic modelling (Blei, 2012), allow the construction of research landscapes and portfolios using only text (e.g. Cassi, 2017). The robustness of these semantic methods is yet open to debate (Leydesdorff and Nerghes, 2017).

**Visualizations**

Novel visualisation techniques greatly facilitate the portrayal of cognitive landscape and social networks in which the projects of portfolios are embedded. The literature is rife with examples of visualization techniques which offer a portfolio view of projects (see Börner's scimaps.org), as visualization tools such as VOSviewer or Gephi become easier to use. These maps are useful for mapping purposes – portfolio spread and an overall picture of the relationship to strategic research objectives of the institution – which, in turn, are more consistent with how decision makers conceptualize qualitative traits in their own judgement. Weaknesses include the potential lack of stability of visualization and that these techniques do not generally address portfolio-level issues such as project or thematic relationships and synergies, although it is feasible (e.g. Rafols et al. 2012).

# 5. CONCLUDING REMARKS

The use of research portfolios in science and innovation policy depends on the type of research and policy context. In cases where there is manageable degree of uncertainty and some value consensus, one can apply modelling techniques. In cases, with high uncertainty and lack of consensus on agreement on goals, portfolio analysis can feed into and enrich qualitative processes of priority setting.

A set of conclusions emerges from our discussion on R&D portfolio modeling:

1. It is feasible to estimate the risk and potential return of applied R&D projects. However, discounted cash flow methods (NPV) are increasingly recognized as inadequate in characterizing public applied R&D investments, much as they have been recognized in the private sector for some time now. Alternative methods such as 'real options' allow better appraisal of the value of R&D management flexibility in the presence of risk and of the differential effects on each R&D project depending on the level of risk and the size of the upside payoff.

15 https://www.starmetrics.nih.gov/
16 gtr.rcuk.ac.uk

2. Accepting that both risk and return are important compels the use of portfolio-level analysis. Risk is not additive. Frequently the returns of R&D projects are not additive. Absent portfolio analysis there is no way to discern the overall risk and overall return of a set of R&D investments.
3. Portfolio analysis requires the explicit evaluation of the relationship between individual R&D projects. Both technical connections and market connections are important. One can think, for example, of the opportunity to share results and the possibility of economies of scope across projects. Inter-project relationships affects the risk of the overall R&D portfolio. Furthermore, managers may also consider the possibility of non-obvious "deep" connections through independent variables (e.g., energy prices).
4. Public R&D programs typically have multiple objectives (ambiguity). In order to account for multiple, difficult to monetize, and often conflicting program and project goals, a diverse set of alternative mixed methods to draw up real asset portfolios (including R&D) have been developed. These methods can stand on their own bottom; they can be combined with economic quantitative approaches for addressing portfolio optimality.

Using research portfolios as a heuristic for managing research priorities leads to complementary insights:

5. In R&D programmes characterised by multiple objectives, mixed methods for exploring ambiguities in goals and uncertainties in outcomes can be used as part of wider participatory and/or deliberative processes for priority setting in funding agencies.
6. These broader priority setting practices require not only the development of technical expertise, but also institutional learning for managing processes of deliberation and integration of diverse knowledge sources.
7. In recent years there has been a rapid increase in data sources that facilitate portfolio analysis of science supply in terms of funding and publication data, classification schemes and visualization techniques. However, attention must be paid to adapt data management to portfolio goals and contexts – especially the analytical categories (classifications) used.
8. Expected outcomes of portfolios can be compared with estimates of societal needs or demands. These estimations are very challenging. Increasing data availability in sectors such as health, agriculture and sustainability allows to develop new estimates of the societal needs addressed by research programs. Mixed method approaches can assess the diversity of views of stakeholders.
9. Novel data and visualizations of portfolio outputs (science supply) together with new data and methods for assessing societal needs (demand) can improve priority setting processes, in particular facilitating the participation of stakeholders in deliberations.

Our policy recommendations in short:

i. Use of portfolio methods for R&D program appraisal is generally recommended.
ii. Modeling of research portfolios is recommended for cases of agreement on program goals where value estimates are possible (presence of risk, low uncertainty)
iii. Portfolios can also be a useful tool to assist in deliberative processes aimed at aligning science supply with social needs or demands for cases without agreement on program goals, and when uncertainty is rampant.

Finally, while not treated explicitly in this report due to space limitations, "big data" exploiting unconventional sources of information may hold a big promise in terms of estimating not easily monetized goals of public R&D programs, thus deserving research attention.

**Acknowledgements.** The authors thank the participants of the "Workshop on Government Decision-Making to Allocate Scientific Resources" organized by the Innovation Policy Forum of the US National Academies in Washington, DC, January 8-9, 2018, for useful comments and recommendations. We also would like to thank several interviewees for opinions on the use of portfolio analysis in funding agencies. Both authors acknowledge funding support by NSF's Science for Science and Innovation Policy program. All remaining errors and misconceptions are ours.

# REFERENCES

**Baker, E., Olaleye, O. and Reiss, L. A.** (2015). Decision Frameworks and the Investment in R&D, Energy Policy, 80: 275-285.

**Blei, D. M.** (2012). Probabilistic topic models. Communications of the ACM 55(4): 77-84.

**Borrás, S. and C. Edquist** (2014). Innovation Policy for Knowledge Production and R&D: the Investment Portfolio Approach, Papers in Innovation Studies, Paper no. 2014/21. CIRCLE, Lund University.

**Börner, K., Chaomei C. and Boyack, K. W.** (2003). Visualizing knowledge domains. Annual review of information science and technology 37.1 (2003): 179-255.

**Börner, K., Klavans, R., Patek, M., Zoss, A. M., Biberstine, J. R., Light, R. P., ... and Boyack, K.W.** (2012). Design and update of a classification system: The UCSD map of science. PloS one 7.7 (2012): e39464.

**Bozeman, B. and J. Rogers** (2001). Strategic management of government-sponsored R&D portfolios. Environment and Planning C: Government and Policy, 19:413-442.

**Bozeman, B., and D. Sarewitz** (2011). Public Value Mapping and Science Policy Evaluation. Minerva 49(1): 1–23. doi:10.1007/s11024-011-9161-7.

**Brosch, R.** (2008). Portfolios of Real Options, Springer.

**Cassi, L., Lahatte, A., Rafols, I., Sautier, P. and Turckheim, É. De.** (2017). Improving fitness :Mapping research priorities against societal needs on obesity. Journal of Informetrics, 11(4), 1095–1113. https://doi.org/10.1016/j.joi.2017.09.010

**Ciarli, T. and Rafols, I.** (2017). The Relation between Research Priorities and Societal Demands: The Case of Rice. Available at SSRN: https://ssrn.com/abstract=3093285

**Casault, S., Groen, A. J. and Linton, J. D.** (2013a). Selection of a Portfolio of R&D Projects, in Link, A. N. and Vonortas, N. S. (eds) Handbook on the Theory and Practice of Program Evaluation, Edward Elgar.

**Casault, S., Groen, A. J. and Linton, J. D.** (2013b). Examination of the Behavior of R&D Returns Using a Power Law, Science and Public Policy, 40: 219–228.

**Cassimon, D., De Backer, M., Engelen, P. J., Van Wouwe, M. and Yordanov, V.** (2011). Incorporating Technical Risk in Compound Real Option Models to Value a Pharmaceutical R&D Licensing Opportunity, Research Policy, 40:1200-1216.

**Costas, R., and van Leeuwen, T.N.** (2012). Approaching the "reward triangle": General analysis of the presence of funding acknowledgments and "peer interactive communication" in scientific publications. Journal of the American Society for Information Science and Technology, 63(8): 1647–1661. https://doi.org/10.1002/asi.22692

**Eilat, H., Golany, B. and Shtub, A.** (2006). Constructing and Evaluating Balanced Portfolios of R&D Projects with Interactions: a DEA Based Methodology, European Journal of Operational Research, 172(3): 1018–1039.

**Evans, J. A., Shim, J. M. and Ioannidis, J.P.** (2014). Attention to local health burden and the global disparity of health research, PLoS One, 9(4): e90147.

**Flanagana, K., Uyarra, E. and Laranja, M.** (2011). Reconceptualising the 'policy mix' for innovation. Research Policy 40:702-713. doi:10.1016/j.respol.2011.02.005

**Gemici-Ozkan, B., Wu, S. D., Linderoth, J. D. and Moore, J. E.** (2010). R&D Project Portfolio Analysis for the Semiconductor Industry, Operations Research, 58(6): 1548-1563.

**Haak, L. L., Ferriss, W., Wright, K. Pollard, M. E., Barden, K., Probus, M. A., Tartakovsky, M. and Hackett, C. J.** (2012). The electronic Scientific Portfolio Assistant: Integrating scientific knowledge databases to support program impact assessment. Science and Public Policy 39: 464–475. doi:10.1093/scipol/scs030.

**Hage, J., Jordan, G. and Mote, J.** (2007). A theory-based innovation systems framework for evaluating diverse portfolios of research, part two: macro indicators and policy interventions. Science and Public Policy, 34(10): 731–74.
DOI: 10.3152/030234207X265385

**Hellström, T., Jacob, M. and Sjöö, K.** (2017). From thematic to organizational prioritization: the challenges of implementing RDI priorities. Science and Public Policy 44(5):599–608.

**Hicks, D.** (2016). Grand Challenges in US science policy attempt policy innovation. International Journal of Foresight and Innovation Policy 11.1-3 (2016): 22-42.

**Interagency Task Group on Science Policy** (2008), The Science of Science Policy, Report to the Subcommittee on social, Behavioral and Economic Sciences, Committee on Science, US National Science and Technology Council, Office of Science and Technology Policy.

**Kim, J., Jo, H., Kim, Y. and Jeon, J.** (2016). Managing the Strategic Fit of National R&D Programme Portfolios: The Case of Creative Economy Policy in Korea, Innovation: Management, Policy & Practice, 18(2): 99-144.

**Klavans, R. and Boyack, K. W.** (2017). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? Journal of the Association for Information Science and Technology 68.4 (2017): 984-998.

**Koier, E., van der Meulen, B., Horlings, E. and Belder, R.** (2016). Spinning plates – Funding streams and prioritisation in Dutch university research. The Hague, Rathenau Instituut 2016. Available at https://www.rathenau.nl/en/news/spinning-plates-funding-streams-and-prioritisation-dutch-university-research.

**Kurth, M. H., Keisler, J. M., Bates, M. E., Bridges, T. S.,Summers, J. and Linkov, I.** (2017). A Portfolio Decision Analysis Approach to Support Energy Research and Development Resource Allocation, Energy Policy, 105: 128-135.

**Kurth, M. H., Larkin, S., Keisler, J. M. and Linkov, I.** (2017). Trends and Applications of Multi-Criteria Decision Analysis: Use in Government Agencies, Environment Systems & Decisions, 37(2): 134-143.

**Leydesdorff, L. and Nerghes, A.** (2017). CoÐword maps and topic modelling: A comparison using small and mediumÐsized corpora (N< 1,000). Journal of the Association for Information Science and Technology 68(4): 1024-1035.

**Link, A. N. and Scott, J. T.** (2013). The Theory and Practice of Public-Sector R&D Economic Impact Analysis, in Link, A. N. and Vonortas, N. S. (eds) Handbook on the Theory and Practice of Program Evaluation, Edward Elgar.

**Linquiti, P. D.** (2015). The Public Sector R&D Enterprise: A New Approach to Portfolio Evaluation, Palgrave-Macmillan.

**Linton, J. and Vonortas, N. S.** (2015). From Research Project to Research Portfolio: Meeting Scale and Complexity, Foresight and STI Governance, 9(2): 38-43.

**Linton, J., Walsh, S. T. and Morabito, J.** (2002). Analysis, Ranking and Selection of R&D Projects in a Portfolio, R&D Management, 32(2): 139-148.

**Lo Nigro, J., Morreale, A. and Abbate, L.** (2016). An Open Innovation Decision Support System to Select a Biopharmaceutical R&D Portfolio, Managerial and Decision Economics, 37: 392-406.

**Luehrman, T. A.** (1998). Strategy as a Portfolio of Real Options, Harvard Business Review, September-October.

**Magazzini, L., Pammolli, F. and Riccaboni, M.** (2016). Real Options and Incremental Research in Pharmaceutical R&D Project Portfolio Management, Creativity and Innovation Management, 25(2): 292-302.

**Marafon, A. D., Ensslin, L., Rde Oliveira Lacerda, . T. and Ensslin, S. R**. (2015).The Effectiveness of Multi-Criteria Decision Aid Methodology: A Case Study of R&D Management, European Journal of Innovation Management, 18(1): 86-109.

**Markovitz, H.** (1952) "Portfolio Selection", Journal of Finance, 7(1): 77-91.

**Markowitz, H. and Markowitz, H.M.** (1991). Portfolio Selection: Efficient Diversification of Investments, Cowles Foundation for Research in Economics at Yale University, Blackwell.

**Meier, F. and Schimank, U.** (2010). Mission Now Possible: Profile-Building and Leadership in German Universities. In R. Whitley, J. Gläser, & L. Engwall (Eds.), Reconfiguring Knowledge Production. Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation. (pp. 2011–2236). Oxford: Oxford University Press.

**Montajabiha, M., Khamseh, A. A. and Afshar-Nadjafi, B.** (2017) A Robust Algorithm for Project Portfolio Selection Problem Using Real Options Valuation, International Journal of Managing Projects in Business, 10(2): 386-403.

**National Research Council** (2007). Prospective Evaluation of Applied Energy Research and development at DOE (Phase Two), Committee on Prospective Benefits of DOE's Energy Efficiency and Fossil Energy R&D Programs, Board on Energy and Environmental Systems, Division on Engineering and Physical Sciences, Washington, DC: National Academies Press.

**National Research Council** (2005). Prospective Evaluation of Applied Energy Research and development at DOE (Phase One): A First Look Forward, Committee on Prospective Benefits of DOE's Energy Efficiency and Fossil Energy R&D Programs, Board on Energy and Environmental Systems, Division on Engineering and Physical Sciences, Washington, DC: National Academies Press.

**National Research Council** (2001). Energy Research at DOE: Was It Worth It?, Washington, D.C.: National Academy Press.

**Nature Plants Editorial** (2015). Neglected tropical crops. Nature Plants. doi: 10.1038/NPLANTS.2015.204

**Pielke, R. A.** (2007). The honest broker. Making sense of science in policy and politics. Cambridge: Cambridge University Press.

**Rafols, I., Porter, A.L. and Leydesdorff, L.**(2010). Science overlay maps: a new tool for research policy and library management, Journal of the American Society for Information Science and Technology, 61(9):1871–1887.

**Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P. and Stirling, A**. (2012). How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management, Research Policy, 41: 1262–1282. doi: 10.1016/j.respol.2012.03.015

**Ruegg, R.** (2006). Bridging from Project Case Study to Portfolio Analysis in a Public R&D Program: A Framework for Evaluation and Introduction to a Composite Performance Rating System, Economic Assessment Office, Advanced Technology Program, National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg MD (NIST GCR 06-891).

**Ruegg, R.** (2007). Quantitative Portfolio Evaluation of US Federal Research and Development Programs, Science and Public Policy, 34(10): 723-730.

**Salter, A. J., and Martin, B. R.** (2001). The economic benefits of publicly funded basic research: A critical review. Research Policy 30: 509–532. doi:10.1016/S0048-7333(00)00091

**Saltelli, A. and Giampietro, M.** (2016). What is wrong with evidence based policy, and how can it be improved? Futures. https://doi.org/10.1016/j.futures.2016.11.012

**Sarewitz, D. and Pielke Jr., R. A.** (2007). The neglected heart of science policy: Reconciling supply of and demand for science, Environmental Science & Policy 10: 5–16. doi:10.1016/j.envsci.2006.10.001.

**Schilling, M. A.** (2017). Strategic Management of Technological Innovation, New York, NY: McGraw-Hill Education (5th ed).

**Srivastava, C. V., Deshmukh Towery, N. and Zuckerman, B.** (2007). Challenges and opportunities for research portfolio analysis, management, and evaluation, Research Evaluation 16: 152–156. doi:10.3152/095820207X236385.

**Smit, H. T. J. and Trigeorgis, L.** (2006). Strategic Planning: Valuing and Managing Portfolios of Real Options, R&D Management, 36(4): 403-419.

**Smith, R.D.J., Hartley, S., Jewitt, T. and Middleton, P.** (2016). Integrating diverse knowledges into research funding organisations: consequences for operationalising responsible innovation. Paper presented in the 4S Conference, Barcelona 2016. Journal version to be available soon. See also report of the study. Available at https://kclpure.kcl.ac.uk/portal/files/48727275/KITo_v1_0.pdf

**Stilgoe, J., Owen. R. and Macnaghten, P.** (2013). Developing a framework for responsible innovation, Research Policy 42:1568– 1580. http://dx.doi.org/10.1016/j.respol.2013.05.008

**Triantis, A.** (2003). Real Options, in Logue D. and Seward J. (eds) Handbook of Modern Finance, New York, NY: Research Institute of America (p. D1-D32).

**Van Bekkum, S., Pennings, E. and Smit, H.** (2009). A Real Options Perspective on R&D Portfolio Diversification, Research Policy, 38: 1150-1158.

**Van Eck, N. J., and Waltman, L.** (2014). Visualizing bibliometric networks. In Measuring scholarly impact (pp. 285-320). Springer International Publishing.

**Verbano, C. and Nosella, A.** (2010). Addressing R&D Investment Decisions: A Cross-Analysis of R&D Project Selection methods, European Journal of Innovation Management, 13(3): 355-380.

**Vilkkumaa, E., Salo, A., Liesio, J. and Siddiqui, A.** (2015). Fostering Breakthrough Technologies – How Do Optimal Funding Decisions Depend on Evaluation Accuracy?, Technological Forecasting & Social Change, 96: 173-190.

**Vonortas, N. S. and Desai, C.** (2007). Real Options Framework to Assess Public Research Investments, Science and Public Policy, 34(10): 699-708.

**Wallace, M. L. and Rafols, I.** (2015).Research Portfolio Analysis in Science Policy: Moving from Financial Returns to Societal Benefits, Minerva, 53(2): 89-115.

**Wallace, M. L. and Rafols, I.** (2017). Shaping the agenda of a grand challenge: Institutional mediation of priorities in avian influenza research. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2699844

**Waltman, L. and Eck, N. E.** (2012). A new methodology for constructing a publicationÐlevel classification system of science, Journal of the Association for Information Science and Technology 63(12): 2378-2392.

**Yegros, A. and Rafols, I.** (2018). Re-assessing priorities in health research. How public R&D can improve its alignment with health needs, Newsletter of "La Caixa" Foundation. Available at: https://observatoriosociallacaixa.org/en/indicador/-/asset_publisher/ATai9MyKZiYq/content/responde-la-investigacion-a-las-necesidades-de-salud

**Zschocke, M. S., Mantin, B. and Jewkes, E. M.** (2014). The Effect of Competition on R&D Portfolio Investments, Production and Operations Management, 23(8): 1439-1449.

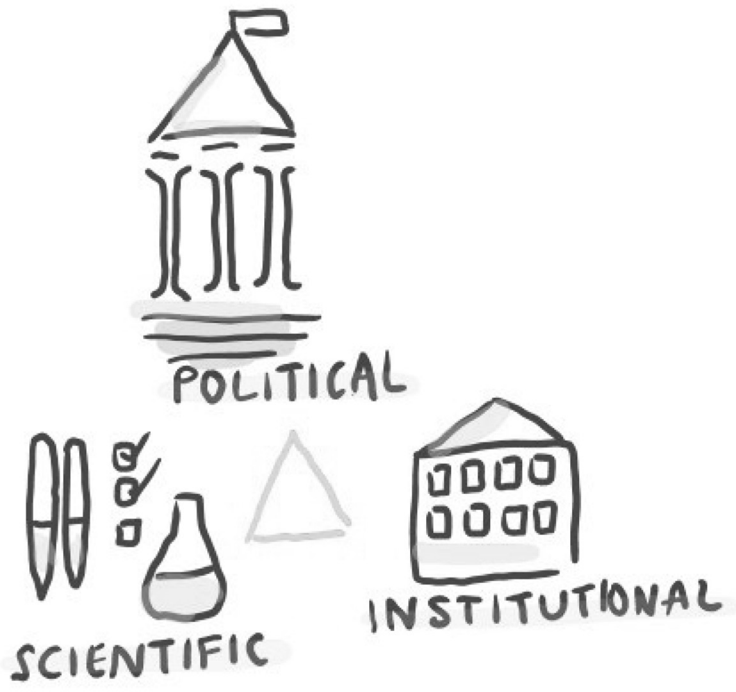
# AUTHORS

**NICHOLAS S. VONORTAS**

*Professor, Institute of International Science and Technology Policy & Department of Economics, The George Washington University, USA; São Paulo Excellence Chair, University of Campinas, Brazil; Leading Research Fellow, National Research University Higher School of Economics, Russian Federation.*

**ISMAEL RÀFOLS**

*Research Fellow, Ingenio (CSIC-UPV), Universitat Politècnica de València, Spain; Visiting Professor, Centre for Science and Technology Studies (CWTS), University of Leiden, The Netherlands; Associate Researcher, SPRU (Science Policy Research Unit), University of Sussex, UK.*



# IMPACT IN THE AGRO-FOOD AND BIO-ECONOMY DOMAIN

UNO SVEDIN AND BETTINA HEIMANN
DOI: 10.22163/fteval.2019.339

## ABSTRACT

The focus of this paper is to reflect on the issue of "impact" of R&I in the agro-food and bio-economy domain (AF-BE) – fields that have a long tradition for such concerns. However, the relevant approaches have changed over time due e.g. to technical developments, globalization and related changes and preferences in society at large. Accordingly the respective features of impacts - and connected indicators to assess them - have to find new forms as well. These considerations are clearly seen in the implementation of the SDGs within the chosen domain. The resulting needs for R&I policy and connected impacts in the AF-BE fields are discussed.

## 1. INTRODUCTION

The issue of "impact" in relation to national and EU-level research and innovation has been present for quite some time, but attention in policy circles has gradually increased year by year often targeting specified policy aims. It is a common procedure to assess investments in research - not least in fields of applied research which rely heavily on the innovation capacity with regard to new technologies. This, includes integrating technologies developed in other fields in an innovative way as well as structural innovations (e.g. institutional and organizational) - to increase efficiency and competitiveness in a given sector. These considerations have since long had an important role when formulating research policy at European and national levels in agro, food and bio-economic fields (also involving forestry and fisheries). This research and innovation domain is at the focus of this paper, both providing a brief outline of the tendencies within these realms, but also as a contribution to the more general debate about the use and role of such considerations in research and innovation policy, i.e. about "impact" at large.

## 2. BACKGROUND

After WW II – and onwards – promoting food production and securing availability were the overriding policy aims in Europe with a strong focus on:

a. the optimization of the production systems, based on crop production as well as animal production, including their relationship to varying conditions; This traditional field is characterized by pin pointedly identified research and innovation objects often related to optimization of various parts of a production systems, e.g. crop development,under certain soil conditions or new efficiencies related to animal production and connected animal health aspects. Not least quickly expanding micro biological and genetic methodological capacities have been strong drivers for innovation and associated impacts in these fields. Many of these types of results of research and innovation efforts have been brought into farm practice through specifically designed agricultural knowledge and innovation structures organized at national and regional levels. These tools and mechanisms are still valid, but in a modified way since the character of the challenges for the sector have expanded over the decades and thus targets have changed accordingly. New and partially different research and innovation backgrounds have come into the picture, and thus also partially new targets and methods have been needed such as those outlined below:
b. the maintenance of the resilience of these systems facing changing weather and climate conditions, water availability and pest infestations and diseases.
c. the consideration of the systemic inter sectoral connectivity of the agro-food and bio-economy fields, e.g. with energy, land use, food and health domains.
d. the relation between agricultural activities and their societal embedding, i.e. trade, socio-economic aspects, consumer behavior, urban-rural connections in planning etc.

There has been a growing interest in all these categories.

## 3. GRAND CHALLENGES – EMERGING INTEREST ARENAS

Due to the quickly expanding need to address "the grand challenges of our times" – a strong interest has grown to invest in research explorations of large webs of phenomena such as climate change (e.g. IPCC, 2018); tightening supplies of energy, water and food; public health and pandemics and herby widening their general understanding and injecting new and fresh perspectives into the sector related research. This has also become a very important part of the EU's research strategy during the last decade (not least after the Swedish EU presidency conference with the Lund declaration, 2009). One exemplification on how research in the traditional arena of food and agriculture is changing is demonstrated through its presence as a central part of the new bio-based economy concept and ongoing reflections about emerging possibilities of a circular economy. These arenas are examples of quickly expanding efforts in binding together all sorts of considerations about bio-based resources on land (including both agriculture and forestry) as well as of the aquatic environments. In the following some examples are indicated:

a. the new connections as e.g. the link between food and forestry domains are introducing food (and feed) possibilities from forestry products – or as new types of support functions to the food industry by increasing systemic recycling from one domain to another (Sandeep, 2018).
b. the innovations due to digitalization and various forms of computer guided "precision farming" and evolving artificial intelligence (AI) practices are providing an expanding set of new tools - but also new challenges
c. the above described considerations are also embedded in new challenges in regional development when re-connecting urban and rural relations in new forms (e.g. more urban oriented food production lines) (Svedin and Liljenstrom, 2018). As an overriding reflection about these and other new aspects with regard to this particular field of R&I it could be said that the appropriate style and balance has to be discussed between on the one side finding more specific strict targeting in relation to pin pointed research and innovation objects - and on the other hand widening the exploration about the contexts, drivers and transformative features of broad sets of phenomena.

# 4. DIFFERENT AIMS WITH REGARD TO "IMPACT IDENTIFICATION"

The balance issue in a policy sense has implications for the considerations about how to handle quite diverse forms of ambitions and targets – also in terms of potential impacts.

### 4.1. "IMPACT" IN TERMS OF PIN-POINTED EFFICIENCY DRIVEN RESEARCH TARGETS

The more partially targeted aims of a technical nature have as research and innovation objectives a different character than many of the systemic oriented 'grand challenges' targeted ones. Concerning pin-pointed research efforts e.g. the technology readiness level (TRL) scale has been employed to provide a structure to the innovation efforts and processes in the EU framework program Horizon 2020 (European Commission, 2014). Concerning the AF-BE domain – where several players contribute to the introduction of an innovation – the scale helps to clarify their respective positions in the range from concept to adoption. This method could also help to define the competencies, funding mechanisms, drivers and deliverables related to each position on the TRL scale. Thereby it could be used to highlight cooperation opportunities between different types of actors and the management of those opportunities. Thus the TRL scale can help researchers to define the 'end-users' of their future research results and to clarify which type of partners they should collaborate with to achieve the highest and most relevant impact with their respective research in order to solve e.g. a particular technological bottleneck or develop an innovative procedure for increased efficiency. It could thus also be used as a means to illustrate more clearly to a variety of research funders what the particular contribution could be in terms of intended impacts.

### 4.2 "IMPACT" IN TERMS OF EX-ANTE CONSIDERATIONS

The broad systemic oriented objects of inquiry are different in terms of aims, institutional embedding, as well as the style of methods and approaches applied. Consequently, there is a strong need for further development of indicators of relevant impacts in all the traditional, but in particular also with regard to the more recently emerged and emerging cross linked phenomena since all these areas of concern have different criteria of success and failure. Target setting for impact is a constant process where obtained knowledge in the form of achieved results and development of new and emerging technologies from earlier development cycles are the basis for the next step of target setting and strategizing. Therefore the question of impact can be looked at either from an 'ex–ante' or 'ex-post' position. In the case of an ex-ante approach "impact" means "potential impact" and thus depends on what priorities will be taken, which strategies are set in motion and which decisions are made that will influence the development of a particular outlined research program (at different institutional levels:, subnational, national, European and global). Hence it will influence the opportunities of technological development in a certain field in various ways depending on design, operational approaches and context. It also might influence the management of future research institutions and the systemic effects framing entire sub-branches of a certain policy complex. Thus a research/ innovation proposal needs to consider what could be addressed immediately and what is less urgent, e.g. with reference to various measures to handle alternative. In the operational ex ante phase the selection of key research and innovation structures to investigate the prioritized issues are of strong importance. They will define what the chances are that a certain framing of a systemic challenge can deliver adequate answers in a solution oriented manner and at the right time. They will also influence what kind of measures and structures are needed to interact with stakeholders and how to disseminate possible outcomes (which activity potentially might need its own financing).

### 4.3 EX-ANTE AND FORESIGHT EFFORTS

In the last decades 'foresights' have become an important tool to scan and define the overarching issues and concerns that need to be addressed and can be used to create a common prioritizing. It is usually the research policy community together with research institutions and with other important actors and stakeholders - often industry but also civil society representations - that are part of the process. Within the broader agro food and bio-economy field for example a series of foresights have been conducted. Some have been driven by EU related bodies as the Standing Committee of Agricultural Research (SCAR) (e.g., the EU-SCAR 3rd foresight report in 2011 and the 4th in 2015 and the EU/ JRC Science and Policy reports (e.g. "Global Food Security 2030 – Assessing trends with a view to future EU policies", 2015) . At the global level studies conducted by UN related bodies could be exemplified (e.g. FAO, "The future of food and agriculture – Trends and challenges", 2017) and the OECD (e.g. "Alternative futures for Global Food and Agriculture", 2016). These types of bodies undertake regularly such scanning as do national bodies (e.g. UK Government Office for Science: "The Future of Food and Farming", 2011) (and the Irish research body Teagasc study from 2016: "Teagasc Technology Foresight 2035"). In all these studies the central aim is to define:

- What are the core issues?
- What are the major drivers?
- Who are the key actors?

Foresights are used as an instrument to reflect on the most pressing challenges at the respective level of investigation and the type of impacts being of interest.

### 4.4 EX-POST ISSUES WITH REGARD TO AGGREGATES

In the case of an ex-post approach the point of departure is the assessment of a set of already created aggregates of research and innovation investments and institutionalizations that have been materialized within a certain past time frame, e.g. somewhat longer than the length of a "normal" research program at national and European level evaluating the various impacts on society achieved by the particular selected set of activities. It is important to note that it is not easy to identify immediately any profound impact within a short time frame after the formal end of the activities under scrutiny, in particular not any changes of a transformative kind that influence the ways things will be done differently or how structures deeply have been transformed. Therefore evaluation investigations have to be undertaken in a sequence of steps (e.g. after 3-5-10 years) that map and put in perspective what have been the outcomes and why or why not the initial aims were achieved. Strong reflection capacities are required and structures have to be available to make such reflections. Causality flows for research investments have to be investigated, i.e. comparing the reasoning at the input side why certain impacts at that time were expected (given the organizational and financial set up) with the outcome much later. This should also explore something about the societal dynamics, i.e. through which efforts aiming for some transformational steps later emerged as manifested changes (including non intentional ones). Numerous evaluation and assessment reports of research programs at national as well as European level (e.g. H2020 Interim Evaluation, 2018) are based on such ex-post approaches. However they are often conducted as mid- and end-term evaluations and therefore do not catch the longer-term impacts - neither those intended nor the undesired ones – thus being beyond the immediate research results (including processes and management). One example for a systematic approach to map impact against investment (financial as well as intellectual) at a longer term is the "asirpa" approach developed by INRA to assess the institution's research efforts against socio-economic impact gained (M. Matt et al., 2017).

### 5. THE IMPLEMENTATION OF THE SDGs – IMPACTS WITHIN A GLOBAL CONTEXT

Since the UN adopted the global Sustainable Development Goals (SDGs) for the period until 2030 and beyond up to 2050, these have become a core concept of European and national funding strategies within a global context. Science, technology and research within the field of the bio-economy and the agro, food, aqua, and forest sector(s) are key means to the overall implementation of the SDGs and thus provide ways to reach these goals. One particularity of the sector under discussion in our context is that it both provides challenges to some of the SDGs, but also at the same time is vital to the possibility to reach many of these other goals. As the SDGs are interlinked in many ways and are operating at different levels, their implementation calls for scientific and technological solutions that match such considerations. There are and will be a multitude of actors with different interests, perceptions and backgrounds involved in the process - also at different levels (IIASA, 2018). A systems approach to sustainable agriculture needs to be further developed in the service of finding overarching solutions in the SDG context. It should take into account the diversity of interactions among humans and the environment, so much at the needed core of the considerations for the future of the agricultural sector. Such reasoning is reflected e.g. in a paper by Patrick Caron et al. (2018) looking at food systems to ensure sustainable development since they link climate, agriculture and food.

The challenge how to measure impact in such a broader frame has to be given much and extended attention. A first step is to find relevant impact indicators for the different levels. But the reflection has to go beyond the multi layered analysis since complex systems are dynamic and technological developments and their societal framings - depending on context - might temporally have to be strongly in tune with the dynamic requests of the solutions, as e.g. the climate challenges so clearly demonstrate. Follow-up questions are

- How to adapt – and even construct - relevant indicators when new practices are starting to be established and new knowledge is emerging?
- How to ensure that policies focusing on global priorities such as the SDGs do consider that these needed actions may have unintended or unexpected consequences in an array of other sectors than those connected with agriculture?

## 6. CONCLUDING REMARKS

Thus systems considerations have to be kept in mind as well as the various effects on regional/local realities in a globalized world; also the multiple functions of and impacts from agriculture with regard to socio-economic and ecological resilience need to be highlighted. Research and innovation strategies have to take all these aspects into account. Thus there is a need for a much broader array of disciplines and transdisciplinary efforts to be engaged. In addition new types of project partners are called for in service of new research approaches. In this context a few principal questions need to be addressed:

- How to integrate an increased reflexivity capacity into the overall research system?
- How to mobilize a sufficiently broad set of relevant actors?
- How to understand the different roles of actors?
- How could we create relevant frameworks of exploration of these issues and provide mechanisms for societal experimentation?

At the same time also strategic funding has to reflect these needs. Policies have to be designed and implemented that permit the mobilization of the necessary innovative capacities. Also there is a need to enhance reflective processes around these systemic concerns in the research community at large. This also implies finding new platforms to address the design aspects of relevant assessment processes. With the formulation of the grand challenges at European level and the adoption of the 2030 UN Agenda for Sustainable Development at global level the systemic and nexus based policy targets have become more widely - but not totally - accepted, as has the understanding that research approaches and programs have to reflect these concerns. However, traditional research areas are still of continued importance, although new methodologies and approaches need to be developed also for their purposes.

The necessary indicators to measure transformational progress are still underdeveloped and need much more methodological thought, practice development and new institutional innovations and strategic support.

# ACKNOWLEDGEMENT

The two authors would like to recognize the strong influence and the discussions and policy formations by the European NGO EURAGRI (www.euragri.aau.dk). During the years of EURAGRI activities (especially during the last 10 years) considerable contributions around themes as those covered in this article have been voiced in annual conferences, supportive workshops, and through informal networking. We are deeply indebted to all our colleagues, because without their contributions and skills our understanding would have been much more limited. This involves also the contacts over years with central decision makers at many levels who have contributed understanding of the policy world and shown how the handling of important issues for our common future needs the combination of knowledge production and societal decision making with strong focus on the challenges involved. However, the responsibility for the present paper lies entirely with the two authors.

# REFERENCES

**Bergeret, P.; Svedin, U. and Valescini, E.** (ed.) (2018). Bio-economy challenges and implementation – the European research organisations' perspective. Éditions Quæ, Paris.

**Caron, P. et al.** (2018). Food systems for sustainable development: proposals for a profound four-part transformation. Agronomy for Sustainable Development 38: 41 https://doi.org/10.1007/s13593-018-0519-1

**European Commission** (2018). Horizon 2020 Interim Evaluation – Maximizing the Impact of EU Research and Innovation, https://ec.europa.eu/research/evaluations/index.cfm?pg=h2020evaluation

**European Commission** (2014). Horizon 2020 – work program 2014-2015 General Annexes https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

**European Commission, Standing Committee on Agricultural Research (SCAR)** (2015). The 4th SCAR Foresight Exercise (FEG4) (expert report), "Sustainable Agriculture, Forestry and Fisheries in the Bioeconomy - A Challenge for Europe", http://ec.europa.eu/research/scar/pdf/feg4-draft-15_may_2015.pdf

**European Commission, Standing Committee on Agricultural Research (SCAR)** (2011). The 3rd SCAR Foresight Exercise (FEG3) (expert report), "Sustainable food consumption and production in in a resource constrained world", https://ec.europa.eu/research/scar/pdf/scar_3rd-foresight_2011.pdf

**FAO** (2017). The future of food and agriculture - Trends and challenges, http://www.fao.org/global-perspectives-studies/resources/detail/en/c/458158/

**Huyghe, C.; Bergeret, P. and Svedin, U. (ed.)** (2016). Diffusion and transfer of knowledge. Éditions Quæ, Paris.

**IIASA** (2018). TWI2050 Report: Transformations to achieve the Sustainable Development Goals, http://www.iiasa.ac.at/web/home/research/twi/Report2018.html

**IPCC** (2018). Global Warming of 1.5 0C, Special report, IPCC, 2018.

**JRC Science and Policy reports** (2015). Global Food Security 2030 – Assessing trends with a view to future EU policies, https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/global-food-security-2030-assessing-trends-view-guiding-future-eu-policies

**Matt, M.; Gaunand, A.; Joly, P.-B. and Colinet L.** (2017). Opening the black box of impact - Ideal-type impact pathways in a public agricultural research organization. Research Policy, 46(1), 207-218. DOI: 10.1016/j.respol.2016.09.016.

**OECD Publishing** (2016). Alternative Futures for Global Food and Agriculture. https://read.oecd-ilibrary.org/agriculture-and-food/alternative-futures-for-global-food-and-agriculture_9789264247826-en#page1

**Sandeep, S.** (2018). Production of microbial protein from brown seaweed and spruce wood and its use as a novel feed ingredient in aquaculture. : NMBU 2018 (ISBN 978-82-575-1518-8).

**Svedin, U. and Liljenström, H.** (2018). A Multilevel Approach to Urban Regional Agglomerations: A Swedish Case of Transition Paths toward a "Fossil-Free Society" by 2050 in: Urban Agglomeration Mustafa Ergen, IntechOpen, DOI: 10.5772/intechopen.73104.
https://www.intechopen.com/books/urban-agglomeration/a-multilevel-approach-to-urban-regional-agglomerations-a-swedish-case-of-transition-paths-toward-a-f

**Swedish EU Presidency Lund Declaration** (2009). https://era.gv.at/object/document/130

**Teagasc** (2016). Teagasc Technology Foresight 2035, https://www.teagasc.ie/media/website/publications/2016/Teagasc-Technology-Foresight-Report-2035.pdf

**UK Government Office for Science** (2011). The Future of Food and Farming, Final report, London, https://www.gov.uk/government/publications/future-of-food-and-farming

# AUTHORS

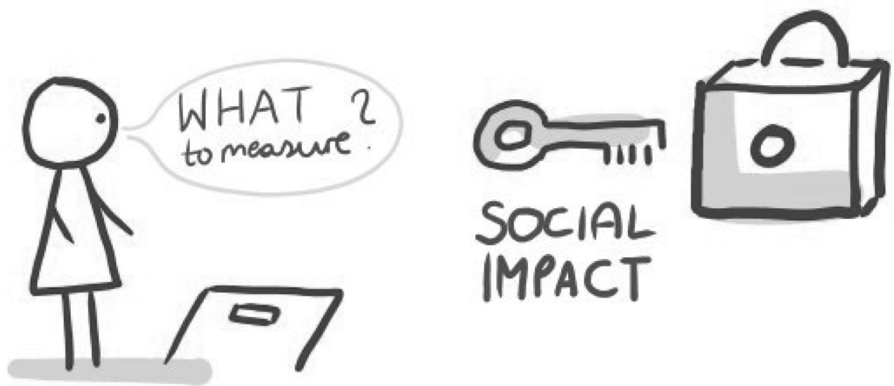
**UNO SVEDIN**
*Stockholm Resilience Centre, Stockholm University*
Kräftriket 2B, SE-10691 Stockholm, Sweden

*EURAGRI, Aalborg University*
Fredrik Bayers Vej 7H, 9220 Aalborg Øst
E: uno.svedin@gmail.com

**BETTINA HEIMANN**
*Department of Chemistry and Bioscience, Aalborg University*
Fredrik Bayers Vej 7H, 9220 Aalborg Øst, Denmark
E: bhe@bio.aau.dk

*EURAGRI, Aalborg University*
Fredrik Bayers Vej 7H, 9220 Aalborg Øst

**Keywords:** Impact of R&I - Agro-food sector - bio-economy - Circular economy - Grand challenges - Ex ante and ex post assessments - EU policy - Sustainable Development Goals



# RESEARCHERS' PERSPECTIVES ON IMPACT OF RESEARCH & INNOVATION: A STRUCTURAL TOPIC MODEL APPROACH TO COST ACTION PARTICIPANTS

ELWIN REIMINK
DOI: 10.22163/fteval.2019.340

## ABSTRACT

The concept of "societal impact" has in the recent decades started to play a defining role in the (political) debate on the position of R&I and science funding. In this paper, we add to this debate by exploring researchers' perspectives on the impact of their participation in research networks. We discover that these perspectives can differ between researchers in different roles and career stages, and that these perspectives do not always correspond with "linear" perspectives on societal impact. We conclude that impact assessment might benefit from a more comprehensive focus, with an equal focus on defining project elements.

## INTRODUCTION

The last decennia have seen somewhat of a shift in the relationship between science and society (see Mostert et al., 2010; Bornmann, 2013). There have been several different ways to describe this shift, each with its own particular focus: for example, the "Triple Helix" model focusses on shifting institutional arrangements in knowledge production (Leydesdorff and Etzkowitz, 1998; Etzkowitz and Leydesdorff, 2000). Meanwhile, "Mode 2 knowledge production" sees a paradigm shift amongst the principal actors of knowledge production (Gibbons, 2000; Hessels and van Lente, 2008), while "post-normal science" focusses on the somewhat broader question of a shifting relationship between science and society (Ravetz, 1999). While the respective perspectives of these approaches – as well as their envisaged consequences – differ, they all have in common that they question the traditional role of science as solely focussed on scientific production. Summarizing: the "old" idea of science as relatively isolated from society-at-large and as a linear producer of scientific output is replaced by a somewhat "messier" model of science having a deeper interaction with other parts of society.

This shift in perspective has not only been extensively discussed in the "science on science", but has also seen extensive follow-up in policy-making, notably through the idea of "societal impact" (Bornmann, 2013). Traditionally, "impact" in research was perceived by the community as focussed on science: hence the term "impact factor" and related metrics, like the h-index (Hicks et al., 2015). Yet, as some authors argue, this might leave us with research which is not necessarily the most useful to all societal stakeholders. As Nightingale and Scott (2007, pp. 543) put it: *"Long-term changes in knowledge production can produce mismatches between the research society requires and the research society produces"*. Moreover, the "traditional" perspective with a great emphasis on scientific metrics has also seen validity issues, for example concerning self-citation (e.g. Fowler and Aksnes, 2007).

The concept of societal impact comes, however, with its own particular issues. First, there is no particular encompassing definition of societal impact which goes beyond the definition that societal, economic or ecological goals are (ultimately) served by the proceedings of research. Moreover, when stakeholders from different backgrounds are specifically asked, they appear to have very different concepts in mind concerning societal impact (Van der Weijden et el., 2012). Second, and partially a consequence of the first issue, there is no structured way of measuring societal impact which goes beyond case studies, either in a comparative or in a stand-alone form (e.g. Bell et al., 2011). These two related limitations have consequences for the role of impact in the daily practice of research, especially where it concerns research evaluation. Societal impact (or just "impact") has come to play a major role in the evaluation of proposals of research projects (Holbrook and Frodeman, 2011). Yet, due to a lack of standards concerning what "impact" actually implies, there is a threat that evaluators will struggle to hold descriptions of (potential) impact in research proposal to a uniform yardstick. Similar problems of definition and measurement can complicate ex post evaluation of the success and impact of research projects.

Two broad types of potential solutions to this lack of both conceptualisation and measurement of "societal impact" have been proposed. A first strain of thinking emphasises the innate link between science and societal values. In other words: science is not funded by the public for the very sake of performing science, but rather because ultimately, science serves societal goals. Hence, these societal goals should be debated and ultimately pronounced, and "societal impact" should, as a consequence, be measured as the extent to which scientific programs contribute to these goals. This "public values" perspective proposes a strong conceptualisation of societal impact, based upon values, and deduces the measurement from this concept (Bozeman and Sarewitz, 2005; Sarewitz and Pielke, 2007).

A second strain of thinking starts, instead, from research itself. This perspective emphasises the diversity of the field of science and, as a result, the complexity of formulating concepts of impact which are both inclusive and concrete. Instead, this perspective argues that all forms of societal impact start with the proliferation and dissemination of researchers' knowledge to a wider audience. This mechanism is defined as "productive interactions" by Spaapen and van Drooge (2011) and as "research uptake" by Morton (2015). While research interactions are not a sufficient condition for impact, they are arguably a necessary condition: if research remains completely isolated from the broader (scientific or societal) community, it cannot be applied to the problems it might be supposed to tackle. We can formulate this as follows in a single sentence: the more research interactions take place, the more pathways to impact, previously blocked by a lack of cooperation and/or shared knowledge, are opened up. Moreover, the nature of these interactions can tell us something about the (potential) ultimate impacts caused by the interactions (De Jong et al., 2014).

In this paper, we look at impact from the second, "bottom-up", perspective. We ask ourselves how "interactions" in research and innovation and their concrete benefits are perceived by the researchers and research community themselves.

# COST ACTIONS

To study the phenomenon of "productive interactions" from first-hand observations, we turn to the particular case of COST Action networks. COST Actions are bottom-up science and technology networks, open to researchers and stakeholders with a duration of four years. They are active through a range of networking tools, such as workshops, conferences, training schools, short-term scientific missions (STSMs), and dissemination activities.  However, COST does not fund research itself. COST Actions are managed by a Management Committee, in which all countries who have accepted the Action Memorandum of Understanding are represented by researchers relevant to the Action topic. The Management Committee (MC) is itself led by a leadership group, encompassing the Action Chair, Action Vice-Chair, the leaders of the different Action Working Groups and the STSM Coordinator. Each Action also has a Grant Holder, who is charged with the management of the financial and administrative side of the grant

An average COST Action can easily encompass over 200 participants, in somewhat different roles. At the "core" of the network are the **researchers in the Action leadership**, who are (relatively) heavily involved in the management of the network and can often be supposed to already have relatively strong ties to other Action participants. The **other members of the Management Committee** can be expected to be more varied in their integration in the Action network,. Finally, the experience of **regular participants** might depend on the networking tools they participated in (Meetings, Short-Term Scientific Missions and/or Training Schools), as well as on their frequency of participation. This leads to our first expectation:

**Expectation I: Researchers in COST Action leadership positions might have a different view on the impact of COST Actions than other participants**

COST Action participants do not only vary in the role they play in their respective Action networks, but also in the places they occupy in the broader system of research and innovation. COST Action participants differ in age from the mid-twenties all the way up to the high sixties. And while age might not be a perfect proxy for the career stage of individual researchers, on average we expect younger researchers to differ in their expectations and effective application of networking tools when compared to more advanced age cohorts. Researchers in an earlier career stage might have a different view on how COST Actions impact both their own research and the research in their field at-large. This leads to our second expectation:

**Expectation II: Researchers who are earlier in their career might have a different perspective on impact than researchers who are more advanced in their career**

Finally, COST Actions mostly involve researchers with a background in higher education and/or academic research, but they also incorporate participants from other backgrounds, notably from government agencies, from non-profit organisations or from business. Given the particular nature of academia and academic careers, academic participants might see the benefit and impact of COST Actions in a different light than other participants. This leads to our third expectation:

**Expectation III: Researchers from non-academic backgrounds might have different perspectives on impact than researchers from an academic background**

The fact that in COST Actions, researchers with very different backgrounds participate in networks with supposedly similar objectives renders these networks ideal "petri dishes" to gauge perspectives on "bottom-up" views on impact. All target groups as identified above participate in the same networks, with the same objectives[1] and in a similar management and strategic context. Yet, different participants might come with different expectations to COST Actions, and they might also have different experiences when participating in the Action networks.

# METHOD

In order to study differences in (perceived) perspectives on societal impact, we apply a Structural Topic Model (STM) approach. The STM approach finds it origins in political science, where it is used to study both cognitive and emotional attitudes towards political actors and objects – and the difference between those two. In broad terms, it allows distinguishing the topics which different target groups mention when asked the same open question, but it also allows to differentiate how different target groups talk about the same topic.

The technical background for Structural Topic Models can be consulted in Roberts *et al.* (2014) and Lucas *et al.* (2015). For the purpose of this paper, we will make an attempt to explain the method in layman's terms.

In general, pieces of text that pertain to the same topic will look like each other. The vocabulary of any language is limited, which means that when discussing a certain topic, an interlocutor will have to rely on repetition of certain words, or even sentence constructions. They will either repeat their own words or the words used by somebody else discussing the same topic. Hence, if certain words pop up in an unexpectedly high frequency in two separate texts, it is probable that these two texts so-

1 Which can, of course, differ from network to network.

mehow discuss the same topic[2]. This allows the classification, identification and eventual clustering of topic-relevant texts without necessary having to read these.

Two particular strategies can be followed in such a classification exercise. On the one hand, you can start out with established categories – for example, "texts concerning cats" and "texts concerning dogs". Each evaluated text is subsequently screened on words commonly associated with cats, and words commonly associated with dogs. This strategy has the main advantage that it will produce an outcome which goes in the lines of what you are exactly looking for. However, it also requires that you know the categories you want to distinguish upfront. This general approach is commonly known as *supervised learning*.

The opposite of supervised learning is (naturally) *unsupervised learning*. In unsupervised learning, the groups in which different objects are categorised are not *a priori* defined. Instead, the algorithm defines the groups itself, based upon observed similarities between groups. The Structural Topic Model is of this second category; based upon words occurring in a more than average frequency in several objects of study at the same time, "topics" are created. The Structural Topic Model allows the attribution of topics towards individual texts according to probability – for example, the text "Bacon rejected by critics" might be associated with both the topics of "cinema" and "food" according to a certain probability. Hence, in STM, a given body of text is not "definitely" grouped into a single topic, but rather has a distribution of probabilities of belonging to different topics.

The Structural Topic Model, finally, allows the testing of the different identified topics with metadata. In other words: are certain topics more prevalent than others in texts with a certain characteristic or background? This is crucial to see to what extent our expectations hold true, since we can differentiate between responses from our different target groups (participants in leadership positions, younger researchers and non-academic participants).

To conclude: the Structural Topic Model allows the identification of underlying topics in (unstructured) text data. It has proven its value in the context of analysing the flux of topics discussed on the internet in the run-up to the 2008 US presidential elections (Roberts *et al.*, 2014). It is particularly a useful tool when a large amount of (unstructured) text data has to be analysed without strong a priori expectations. This makes it an interesting tool to test on researchers' attitudes towards impact of scientific activities; a topic hitherto only sparely studied.

# DATA

In February-March 2018, survey company GfK Belgium executed a "Customer Satisfaction Survey" amongst around 43,000 participants and stakeholders in the COST framework. These participants had participated in the COST framework at least somewhere over the years 2016 and 2017, either as an Action participant, an Action main proposer or an Action grantholder. Of these, 14,384 participants responded for a response rate of 33%.

In the Customer Satisfaction Survey, one particular open question was asked which touches upon (perceived) impacts of the COST Action networks. It was formulated as following:

**"What was your direct benefit related to your participation in a COST Action?"**

The question does not directly invoke impact, but this might not be a pressing issue; after all, the exact wording of "impact" might not be fully understood by all respondents, while "personal benefit" is relatively unequivocal in its meaning. A more crucial issue concerns the focus on "personal" in "personal benefit". Such an individual focus might induce the respondent to "automatically" think of impact on the strictly individual level, as opposed to broader, societal impact. Indeed, when analysing the results of this analysis, we should take this particular caveat into account.

In total 6168 respondents gave some kind of answer to this question (the question was optional – respondents could leave it blank). From these 6168, stop words (like "I", "can", "the", "are" etc.) were removed, and the remaining words were stemmed[3]. For the words remaining in the answers, we checked the number of answers in which the stemmed word appeared. Only words which appeared in at least 1% of the responses (i.e. in 62 responses) were retained. This has two advantages: it removes non-sensical answers (or answers not rendered in the English language) and it makes the eventual identification of the Topic Model easier, since many sparse observations are removed. 123 responses did not contain any stemmed word which reached the 1% threshold, and were therefore completely removed from the analysis, leaving 6045 responses for the identification of the Structural Topic Model.

The 6045 responses are as follows distributed over our variables of interest:

**Table 1**. Observed frequencies of independent variables.

| | | | |
|---|---|---|---|
| Younger researcher | 2013 | Leadership position | 636 |
| Other researcher | 4032 | Other position | 5409 |
| Non-academic participant | 886 | Female | 2870 |
| Academic participant | 5159 | Male | 3175 |

A particularly sensitive step in the identification of Structural Topic Models is the number of topics to choose. The "unsupervised" method does not have a naturally defined number of topics, since the categories are not a priori known. For this particular run, we have chosen to limit the number of topics to 5, which is a relatively modest number of topics (for example, Roberts et al. explored 20 topics). There are three reasons to do so.

First, unlike the Roberts paper, which concerned political campaigns, there is no temporal dimension to our analysis. In political campaigns, events of any kind can influence what people are talking about during the course of the campaign. We do not expect any such effect on perspectives of impact; at least not within the confines of our population of interest. Second, this paper is intended to give an indication of the extent to which perspectives on impact vary between target populations. For this end, a full description of all the possible topics discussed is not necessary, and a first step better involves less rather than more complexity.

1 This has not necessarily be the case, well understood. A text with a high frequency of the words "cat", "roof", "struggle", "worries" and "health" might be the story of a cat owner trying to get their pet out of an awkward situation, but it might also be a discussion of theatre night.

2 Stemming means that verbs, nouns, adjectives and adverbs with a similar origin (and supposedly similar meaning) are grouped together. For example, "collaboration", "collaborating" and "collaborative" are all grouped under the stem "collabor".

In subsequent steps, a more complex model can still be studied. Third, we briefly looked into involving more (either 10 or 15) topics, but quickly realised that this would create a situation in which some topics started to overlap, at least on a contextual level.

# RESULTS

Using the data and methods described above, we came to the following 5 topics. Each topic is described in Table 2 by the 15 words which are most exclusive to this topic (hence, they are relatively used the most in relation to this particular topic vis-à-vis other topics). Alternatively, we can find responses which are typical to the 5 distinguished topics. In Table 3, we display for each of the five topics two reactions which are "typical" of the topic at-large.

**Table 2**. 15 words most exclusive to structural topic, per identified topic.

| | |
|---|---|
| Topic 1 | „public" „joint" „project" „propos" „collabor" „build" „applic" „paper" „develop" „increas" „activ" „creat" „research" „intern" „lead" |
| Topic 2 | „scientist" „differ" „interest" „peopl" „get" „work" „countri" „know" „meet" „field" „european" „discuss" „colleagu" „similar" „expert" |
| Topic 3 | „network" „knowledg" „improv" „exchang" „share" „gain" „experi" „scientif" „idea" „connect" „inform" „skill" „profession" „expertis" „access" |
| Topic 4 | „action" „school" „cost" „train" „stsm" „particip" „confer" „phd" „student" „workshop" „benefit" „attend" „support" „abl" „organ" |
| Topic 5 | „learn" „contact" „futur" „partner" „establish" „start" „make" „met" „new" „techniqu" „possibl" „lot" „method" „problem" „open" |

As can be seen in Table 2 and Table 3, the topics identified have distinct characteristics. Topic 1 is strongly oriented towards outputs – papers and common projects, most prominently. Topic 2, to the contrary, is more oriented towards other researchers and meeting new people. Topic 3 is mostly oriented towards networking and sharing knowledge. Topic 4 is very concretely oriented towards the activities deployed within the framework of COST Actions and the benefits of being able to attend these activities. Topic 5, finally, is somewhat more complicated, and seems to be a mix of different perspectives. This can either be due to respondents truly seeing different benefits of participating in COST Actions, or it can alternatively be due to the nature of our data collection (web-based surveys). In any case, topic 5 has mostly a general orientation on benefits of participating in a COST Action.

**Table 3**. 2 examples of two typical on-topic responses, per identified topic.

| | | |
|---|---|---|
| Topic 1 (Output orientation) | New collaboration with 2 other researchers that led to joint papers being published and a new research project externally funded | An ongoing collaboration that has resulted in few publications and other collaborations that resulted with a EU grant proposal |
| Topic 2 (People orientation) | the interaction with experts from different countries and similar areas and the efforts to search global solutions at the European level | Getting to know groups working in the same field from other European countries |
| Topic 3 (Knowledge orientation) | Networking, collaboration development, expertise improvement, exchange of knowledge | network, sharing knowledge and data, improvement of the quality of research output |
| Topic 4 (Activity orientation) | It provides me chances to go to international training school, conference and another institute for short term visiting. Without the support from COST Action, at least half of them will be impossible. I appreciate it very much. Thanks. | I got invited into an ERA net project proposal as a result of participation in the COST action. I have three PhD students that benefited greatly from participating in workshops and training schools organized by the COST action. |
| Topic 5 (General orientation) | Learned to use new equipment and technology. Met new colleagues with whom there is a possibility for future collaborations. | I met new colleagues and started new collaborations, which are exciting new directions. |

In a second step, we test our expectations as formulated earlier on by seeing whether the tendency to talk about the five respective topics we identified is related to background characteristics of the respondents. We do so by regressing, for each topic, the respective chance that a body of text belongs to this particular topic onto the three independent variables of interest, which correspond to the three expectations (concerning younger researchers, non-academic researchers and researchers in leadership positions) as formulated in the Introduction. Additionally, as a control variable, we include gender. The four independent variables are all coded as binary variables: researcher younger than 40 years vs researcher of 40 years or older, researcher with a non-academic background vs researcher with an academic background, researcher in a leadership position vs researcher in another position, female researcher vs male researcher.

In the case of age, gender and professional background, the characteristics have been self-reported by the respondents through their e-COST (the COST online platform) profile. In the case of professional background, we have observed some misreporting (e.g. somebody from academia reporting that they are from a "governmental agency", which is non-academic). Hence, estimators might be slightly biased for this particular variable, although we do not have a strong a priori expectation concerning a potential direction of this bias; some underestimation of effects might occur, in any case.

We will consider the five different topics individually.

**Table 4**. OLS regression on Topic 1: Output orientation.

| Independent variable | Coefficient | Std Error | Significance |
|---|---|---|---|
| Intercept | 0.23354 | 0.00444 | *** |
| Younger researcher | -0.03378 | 0.00456 | *** |
| Non-academic | -0.01725 | 0.00596 | ** |
| Leadership position | 0.02894 | 0.00776 | *** |
| Female researcher | -0.00792 | 0.00498 | N/A |

Concerning output-orientation, we observe substantial differences between the groups of interest. Researchers in leadership positions tend to be more oriented towards outputs like common projects and proposals. For both younger researchers and non-academic participants, however, this tends to be less the case. Possibly this is due to younger researchers not yet being in the position of seniority which enables effective participation as co-leaders of major projects in R&I. For non-academic participants, common projects might sometimes be less attractive due to the academic focus of some research projects.

**Table 5**. OLS regression on Topic 2: People orientation.

| Independent variable | Coefficient | Std Error | Significance |
|---|---|---|---|
| Intercept | 0.24681 | .00408 | *** |
| Younger researcher | -0.02013 | .00551 | *** |
| Non-academic | 0.01009 | .00643 | N/A |
| Leadership position | -0.01951 | .00476 | ** |
| Female researcher | 0.00208 | .00844 | N/A |

Concerning orientation towards meeting other people, we see that both researchers in leadership positions and younger researchers are less likely to veer towards this particular topic in their response. For researchers in leadership positions, this might be explained by the fact that they possibly already have the right contacts, and do not have to use COST Action networks to create such links. For younger researchers, this phenomenon is slightly more puzzling; it would seem to make sense that younger researchers still have to build up their network. It might be that younger researchers are more looking for the actual activities deployed in COST Actions than for the people they meet during these activities.

**Table 6**. OLS regression on Topic 3: Knowledge orientation.

| Independent variable | Coefficient | Std Error | Significance |
|---|---|---|---|
| Intercept | 0.19901 | .00306 | *** |
| Younger researcher | -0.00162 | .00420 | N/A |
| Non-academic | 0.01940 | .00581 | *** |
| Leadership position | -0.00871 | .00371 | * |
| Female researcher | -0.00408 | .00646 | N/A |

Concerning orientation towards gaining and sharing (general) knowledge, we see that participants with a non-academic background are more likely to refer to this topic when describing their benefit of participating in a COST Action network. Researchers in leadership positions are, on the other hand, less likely to refer to such benefits. Again, we could say that for researchers in leadership positions, access to knowledge might be less "attractive", since they are already in the centre of this body of knowledge, figuratively speaking. For non-academic participants, we could hypothesise that access to knowledge might be particularly appealing since they are not necessarily in touch with academic knowledge on a daily basis. COST Actions, which can incorporate up to 300 researchers from different backgrounds and disciplines can, in this respect, form a "glossary of knowledge" for outside researchers.

**Table 7**. OLS regression on Topic 4: Activity orientation.

| Independent variable | Coefficient | Std Error | Significance |
|---|---|---|---|
| Intercept | 0.15380 | 0.00424 | *** |
| Younger researcher | 0.05690 | 0.00543 | *** |
| Non-academic | -0.02054 | 0.00699 | ** |
| Leadership position | 0.01543 | 0.00532 | ** |
| Female researcher | 0.00563 | 0.00863 | N/A |

Concerning orientation towards Actions' activities, we see that researchers in leadership positions and, in particular, younger researchers are more likely to indicate an orientation towards activities deployed by COST Actions. Non-academic participants are, on the other hand, less likely to mention this topic. The tendency of younger researchers to mention this topic is particularly pronounced. A potential reason for this remarkable result is that younger researchers are more oriented towards gaining skills, which are transmitted through COST Action activities like Short-Term Scientific Missions and Training Schools, although other hypotheses might equally be offered.

**Table 8**. OLS regression on Topic 5: General orientation.

| Independent variable | Coefficient | Std Error | Significance |
|---|---|---|---|
| Intercept | 0.16710 | 0.00365 | *** |
| Younger researcher | -0.00195 | 0.00463 | N/A |
| Non-academic | 0.00866 | 0.00645 | N/A |
| Leadership position | -0.00195 | 0.00436 | N/A |
| Female researcher | -0.00920 | 0.00754 | N/A |

Finally, the general orientation does not correlate with any of our variables of interest. This is not a particularly surprising finding; there is no particular reason why one target group would be less (or more) likely to mention particular benefits of participating in COST Actions.

# CONCLUSION

We started this paper with the assessment that the concept of "societal impact" has come to play a more dominant role in assessing the value of research and innovation, especially from the perspective of public funding. We equally observed that there is, as yet, no clear convergence on how to define or further conceptualise "societal impact". In order to make a (very modest) step towards a bit more clarification, we applied a "bottom-up" perspective on the question, by asking researchers themselves what they saw personally as an impact of their involvement in research projects.

One important conclusion is that it depends on who you ask: in the context of COST Action networks, we observed differences in perspectives between different participants in the projects. Perceived impact depends on the position in the COST Action, with researchers in the core having different perspectives than other researchers, but the perceived impact also varies with career stage and with the background of participants.

Notably, participants with a non-academic background turned out to be less oriented towards projects and collaborations, and more oriented towards general knowledge sharing. This is a particularly intriguing result given that one particular strain on "societal impact" stresses the need for "productive interactions" as the basis of achieving this impact (de Jong *et al*., 2014). The orientation of non-academic participants of COST Actions towards knowledge sharing seems to confirm this notion. Yet, impact evaluation still has a – somewhat understandable – orientation towards easy-to-measure, direct results of impact (Donovan, 2007). It might actually well be that the actual pathways towards societal impact – by opening channels between researchers and societal actors – are not fully grasped by this "traditional" orientation towards concrete outputs and results.

From a broader perspective, we can maybe see some seeds of a more encompassing way of assessing impact. The four specific topics we identified (beside the "general" topic) all point towards specific aspects of COST Actions: the individuals involved (the "people" orientation), the activities deployed (the "activity" orientation), the common projects spinning off from the Action (the "project" orientation) and the general topic of the Action (the "knowledge" orientation). In other words, the perceived benefits of participation in a COST Action do, in the eyes of the participants, originate from a mix of different aspects of the Action they participate in. Hence, although this is still a very tentative conclusion, it might be worthwhile to evaluate the impact of research projects – including Actions - on different accounts. Spin-off projects and results can surely play an important role in this evaluation, but so might whether the project was broad and inclusive concerning the partners involved, whether adequate activities were deployed in the scope of the project and whether the topic and the general set-up of the project stimulated an exchange of knowledge and practices. All of these aspects seem at least to play some role in making sure society ultimately benefits from research.

Of course, there are some limitations to this paper. We asked participants of one specific instrument in the landscape of R&I funding to formulate their thoughts on the benefits of participating in this benefit. Given that this instrument – COST Action – revolves around research networking rather than research itself, we should be careful with extrapolating the results to conclusions about the impact of R&I funding at-large. Moreover, the method deployed in this study was intentionally explorative, and follow-up research would still have to confirm – or amend – the initial patterns observed in our study. Nonetheless, the finding that (societal) impact of research might be approached from very different perspectives looks relatively robust, and surely needs more consideration in future studies. In this respect impact might be much like beauty: it is all in the eye of the beholder.

# REFERENCES

**Bell, S., Shaw, B. and Boaz, A.** (2011). Real-world approaches to assessing the impact of environmental research on policy. Research Evaluation, 20:227.

**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, 64:217.

**Bozeman, B. and Sarewitz, D.** (2005). Public values and public failure in US science policy. Science and Public Policy, 32:119.

**De Jong, S., Barker, K., Cox, D., Sveindottir, T. and van den Besselaar, P.** (2014). Understanding societal impact through productive interactions: ICT research as a case. Research Evaluation, 23:89.

**Donovan, C.** (2007). The qualitative future of research evaluation. Science and Public Policy, 2007:585.

**Etzkowitz, H. and Leydesdorff, L.** (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. Research Policy, 29:109.

**Fowler, J. and Aksnes, D.** (2007). Does self-citation pay? Scientometrics, 72:427.

**Gibbons, M.** (2000). Mode 2 society and the emergence of context-sensitive science. Science and Public Policy, 27:159.

**Hessels, L.K. and van Lente, H.** (2008). Re-thinking new knowledge production: A literature review and a research agenda. Research Policy, 37:740.

**Hicks, D., Wouters, P., de Rijcke, S. and Rafols, I.** (2015). The Leiden Manifesto for research metrics. Nature, 520:431.

**Holbrook, J.B. and Frodeman, R.** (2011). Peer review and the ex-ante assessment of societal impacts. Research Evaluation, 20:239.

**Leydesdorff, L. and Etzkowitz, H.** (1998). The Triple Helix as a model for innovation studies. Science and Public Policy, 25:195.3

**Lucas, C., Nielsen, A., Robets, M.E., Stewart, B.M., Storer, A. and Tingley, D.** (2015). Computer-Assisted Text Analysis for Comparative Politics. Political Analysis, 23:254.

**Morton, S.** (2015). Progressing research impact assessment: A 'contributions' approach. Research Evaluation, 24:405.

**Mostert, S., Ellenbroek, S., Meijer, I., van Ark, G. and Klasen, E.** (2010). Societal output and use of research performed by health research groups. Health Research Policy and Systems, 8: 30.

**Nightingale, P. and Scott, A.** (2007). Peer review and the relevance gap: Ten suggestions for policy-makers. Science and Public Policy, 34:543.

**Ravetz, J.R.** (1999). What is Post-Normal Science. Futures, 31:647.

**Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J. Kushner Gadarian, S, Albertson, B. and Rand, D.G.** (2014). Structural Topic Models for Open Ended Survey Responses. American Journal of Political Science, 58:1064.

**Sarewitz, D. and Pielke, A.J.** (2007). The neglected heart of science policy: reconciling supply of and demand for science. Environmental Science & Policy, 10:5.

**Spaapen, J. and van Drooge, L.** (2011). Introducing 'productive interactions' in social impact assessment. Research Evaluation, 30:211.

**Van der Weijden, I., Verbree, M. and van den Besselaar, P.** (2012). From bench to bedside: The societal orientation of research leaders: The case of biomedical and health research in the Netherlands. Science and Public Policy, 39:285.

# AUTHOR

**ELWIN REIMINK**
*Data & Impact Analysis Officer, COST Association, Avenue Louise 149, Brussels, Belgium*
E: elwin.reimink@cost.eu

# IMPLEMENTING EVIDENCE-BASED STRATEGIC DECISION MAKING IN HER INSTITUTIONS: THE CASE OF PARIS SCIENCES ET LETTRES UNIVERSITY[1]

EMMANUEL BASSET AND MATTHIAS HEUSER

DOI: 10.22163/fteval.2019.341


## ABSTRACT

The design and implementation of impact-oriented R&I policies depends on the capacity of contextualizing the expected impact of such policies. However, most of the available data sources for R&I policies are still fragmented, highly heterogeneous and not interoperable. Public policy and strategic decision making suffer from lack of integration of existing data, which are available from separate sources, follow different definitions, have disparate time scales. In such a scenario, decision makers and policy makers can still easily stumble on spurious correlations when indicators and relevant data are deprived of context and maintained in vertically separated containers or 'silos'. Silos of data remain separate, non-integrated, often not even interoperable.

In other terms, to inform R&I policies and design impact-oriented approaches, data that are currently dispersed and highly heterogeneous, need to be accessed in an integrated, unified and semantically consistent way. To achieve this, a combination of semantic-based technological solutions and open government approach must be increased. This, however, does not imply that the problem of quality and pertinence of specific indicators for the uses retained is solved by this technological choice.


## INTRODUCTION

This paper presents as a case study the project of establishing a comprehensive Research Information System (RIS) in University Paris Sciences & Letters[2] (PSL), a key institution in the Parisian Higher Education, Research and Innovation (HERI) landscape. The system has been developed for integrating distributed and heterogeneous data sources for (1) informing top-level strategic decision-making at university level in the context of a period of radical change in the French HERI system as well as (2) open up the university to other quadruple helix actors by providing a detailed research portfolio and (3) generally increase the availability of pertinent data to mid-level management and individual researchers, fostering a culture change towards data use. It relies on the Semantic Web technological framework to extract meaningful insights from extensive and heterogeneous data and aims at a powerful contribution to the Open Science movement.

Our research question is therefore: how to tap the potential of current developments to overcome the difficulties attached to the implementation of a RIS at university-level and how to maximise its chances of success?

## CONTEXT

Like many areas of public and private administration, HERI institutions are taken in the recent revolutions in the use of data for policy design and strategic decision-making. The "University 4.0", as we might mockingly call it, is suggested as the new model for rational and evidence-based development of research and higher education.

Measuring scientific production through several categories of quantitative indicators (e.g. number of publications, number of citations related to these publications, number of patents related to research, number of research grants obtained) has become common practice in all fields of academic life: students and staff may choose universities according to international rankings that rely heavily on quantitative indicators (ARWU, CWTS Leiden, THE, QS are the most famous examples); expert committees may look at values such as the h-index to hire or promote researchers; national agencies may allocate funds to universities or research programs proportionally to the number of publications accepted by journals with high impact factors[3]. Although the use of such indicators

---

1 The authors acknowledge the STI 2018 Leiden conference, from which this template was adapted.

2 www.psl.eu

3 The first forms of evidence-based policy-making in the context of HERI as it is relevant today is linked to the introduction of systematic performance assessment sometimes correlating directly performance to funding allocation, the UK Research Assessment Exercise (RAE), established in 1986, being the first national research funding system based on past research performance. The UK reform served as reference for some other countries who introduced elements of performance-based funding in the 2000s, generally run either directly by Ministries (Norway), dedicated agencies (Flanders in Belgium, Italy and Spain) or research councils (Australia). In 2010, at least fourteen performance-based research funding systems were found across Europe and other countries like Australia, New Zealand, Hong Kong and China. Further information in Hicks (2013).

is still highly debated, they are used increasingly by stakeholders, from students to faculty and administrative staff, so that the impact of such indicators has somewhat short-circuited debates about their pertinence.

In this situation, Higher Education Institutes (HEI) will need to find a middle ground between a purely pragmatic stance, prioritising making the most of the situation by profiting from the levers for visibility the indicators might provide, and a critical one, prioritising the improvement of the system. But as said above, HERI institutions are not alone in these developments. At regional, national, and international levels, sophistication in the use of evidence for defining research and innovation strategy and implementation has been growing, as have demands for public accountability. This means that the dialog between policy makers and actors of the regional/national/supranational ecosystems becomes more and more sophisticated as well. We see only the beginning of this in France, where national evaluation and funding organisms still rely on somewhat crude information systems and demands are often ad hoc and fairly disorganised. Still, with an eye on the developments in Great Britain and the Research Excellence Framework[4], it does seem reasonable for an HERI actor to prepare for future developments in the direction of organised use of data for policy making across different levels and actors.

Especially since the publication of the Shanghai ranking in 2003, the pressure on universities increased to adapt to new quantified standards of excellence which are different (and often in part contradictory) with established procedures of quality-insurance mostly based on peer-review. As Thoenig and Paradeise put it: *"Excellence rankings induce significant consequences for the very definition of academic quality. Universities are split between two quality regimes: a traditional one based on reputation and expressed through cardinal judgements delivered by insiders, and a new one based on quantitative ordinal scales invented by rankings of excellence as defined by outsiders"*[5]. One should not underestimate the significance of this shift: does generalisation of use of externally constructed data undermine universities' "strategic capacity" (Thoenig and Paradeise 2016)?

This point has been a key motivation for PSL to build a performant research intelligence system. Instead of continuously finding oneself confronted with externally defined indicator systems, PSL wished to build up internal competence and internal data to reclaim if not sovereignty, at least a strong voice on the modalities of evaluation of its work. More broadly, PSL saw the implementation of a RIS as an opportunity to foster acceptance of data within the academic community, and to find the right balance between quantitative evidence with existing qualitative procedures of quality-assessment.

A main point of diagnosis that constitutes a second key motivation and that will explain quite some decisions made concerning the design of the system at PSL is that institutions should take care of their own data. Data must (in most cases) be curated at its source and there must be a proper feedback loop with its users, that are, in our situation also its main producers who are a key element in the production of quality data. Else, data is likely not to reflect the reality of the research activity it is supposed to represent. Everyone needs good data for their own purposes.

However, if data systems do not connect between them, aggregation and collaboration become difficult. This goes for the inner organisation of universities as well as for their relations to the exterior. Therefore, PSL chose a scalable system based on a technology mix apt to prepare it for interoperability with other systems. This allows the integration of data internally, but also, potentially, towards the outside (other HERI actors, our region, the national ministry, the EU, or the public). Moreover, the Linked (Open) Data (LOD) approach allows us to contextualise internal data with external sources. This way, we can weight e.g. internal information on publications against external bibliometric data, clinical trial data, national or EU projects or against patents. LOD enables multiple interoperable perspectives.

# THE PROJECT *PARIS SCIENCES ET LETTRES*

Paris Sciences et Lettres University (PSL) is a research-intensive Parisian University system or sometimes coined "collegiate university" established in 2010. It is not a fully integrated university at this point but engaged in a densification process from which it shall emerge as "one" university in some years' time. This transformation has been engaged during the excellence initiative, which provided PSL with a 750M€ endowment on the condition that it becomes a single university. PSL comprises 9 full members, all small and highly selective grands écoles and a range of associates, many of which are strong institutions with their own history and independent reputation like the École Normale Supérieure de Paris, the Paris School of Mines or the Paris Observatory, the most important research centre on astronomy in France. As said, it is amid a profound institutional reconfiguration, which on the one hand explains the present initiative and on the other makes it difficult because of a complicated political and administrative situation.

# OBJECTIVES

Penfield and al. (2014) identify four primary purposes of measuring research impact:

1. Monitoring: a need of HEIs to monitor their performance and visibility in the local, national or regional environment;
2. Accountability: the growing importance of accountability to demonstrate to non-academic stakeholders (government, industry, wider public) the value of research and of the public investment in it;
3. Decision-making: the need to help decision-making, especially in case of resource allocation;
4. Understand: the new capacity to understand how research leads to impact thanks to data.

We propose to adapt this broad framework to specific needs of Higher Education Institutions: effective RIS can have positive impact on various levels of activity.

- On the level of external partners (quadruple-helix-actors) and integration in regional and national policy definition: the system can increase transparency of research activity, making the

4 www.ref.ac.uk
5 Thoenig and Paradeise, In Search of Academic Quality, Palgrave Macmillan, 2015, p. 4

university more accessible from the outside, and encouraging partnerships. It can also serve for instance as a tool supporting regional Smart Specialisation Strategy and the related Entrepreneurial Discovery Process.
- On the level of top-management: evidence-based definition of strategy can in principle support an effective piloting of HEI – for reasons we will detail below, the direct use of quantitative indicators for decision making has been however put aside for the short term.
- On the level of mid-level management, units and individual researchers: if the system provides a sufficiently fine granularity and the capacity to explore output from the content side (which is the perspective of the researchers), this can lead to an increased capacity to meaningfully gauge one's competences and ways to communicate them.
- On the level of accountability: a performant RIS enables a stronger evidence-based communication generating more value for internal and external transparency and communication by thorough publicly visible analysis and reporting.

Beyond these benefits, a performant RIS may allow the university to actively contribute to the Open Science movement by building up potentially transferable expertise and infrastructure.

# MANAGING BIASES OF THE INDICATORS

However, designing a pertinent Research Information System (RIS) requires a clear view of biases induced by existing quantitative measures of research production and quality.

Criticism against the measure of research impact via standardised quantitative indicators have accompanied their rise. As exemplary, we briefly relate the debate surrounding impact assessment via bibliometric indicators, notably citation counts.

Citation counts have become the reference in bibliometrics for impact assessment. Hicks and Merkels (2013, p. 4) recall the main areas of controversy: *"Disagreement centers on what citations measure (quality would be ideal, but impact is more realistic), distortions of citation counts (high rate achieved through negative citation, citation circles, self-citation etc.) or the value of contributions that do not appear in papers (database curation, creation of new materials or organisms, or increased human capital for example)"*. Other biases have been identified: cumulative effect in favour of advanced researchers (for the specific case of h-index - Gingras 2014, Penfield and al. 2014, Bornmann and Marx 2013); difficulties to measure long-term impact (Gingras 2014), thus discouraging risky upstream research whose impact may, if ever, be visible only after years (Wang et al. 2017); biases due to constitution of existing databases (overrepresentation of journals from United Kingdom and United States; underrepresentation of certain fields - SSH - or formats - monographs, conferences, etc.).

Therefore, we considered carefully the indicators that were to be included in the PSL Research Information System. We prioritized normalized indicators (FWCI by Scopus; proportion of publications in the Top 1%/Top 10%). We have chosen to complement bibliometric indicators with other types of metrics (e.g. ERC and Marie Curie grants and other types of European projects, which are a good proxy for measuring research quality at European research-intensive universities). We also added metrics which were not focused on impact (volume of internal and external collaborations). The system provides direct access to all corpuses of publications related to the quantitative indicators featured to meet the requirement of transparency and to enable in-depth analyses and criticism.

# MAIN USES IDENTIFIED

The main uses that are currently implemented are:
- Monitoring: as a young university which still has to be confirmed by French government (cf French policy of "Idex" – mergers of best HEIs in France – since 2011), PSL's ability to collect reliable and precise data on various subjects (academic and research staff by laboratories / fields; consolidated budgets for research, number of research projects, etc.) is key to enhance its own institutional credibility towards external stakeholders (International Idex jury, government). Moreover, the complexity of the Parisian landscape of Higher Education and Research (one lab may be affiliated simultaneously to PSL, Sorbonne Université and to the national research organism CNRS) makes it necessary to have precise data on the status of the several researchers within each lab.
- Accountability: added-value of the merger-process must be proven. Specific indicators were identified to illustrate synergies made possible by the merger (increase of co-publications between merged institutions, increase of national and regional research projects, etc.). More generally, strategic dialogue with French government relies partly on quantitative analyses; it is crucial for PSL to provide reliable data proving PSL's position as university of excellence in France. By now, we have focused on available indicators (increasing internal collaborations to demonstrate synergies, number of ERC grants per researchers, number of publications in the Top 1% / Top 10%, etc.).
- Communication: quantitative indicators based on transparent data sources help PSL to position itself as a major Higher Education Institution in France (and even in Europe). Quantitative indicators, rather than reputation, objectify (or: seem to objectify) the scientific potential of PSL (and, then, the interest of the merger). Objective results (through two international rankings: THE and QS, but also through consolidated bibliometric indicators) are likely to increase significantly researchers' and students' sense of belonging.

For the first stage of development and implementation, we focused on metrics and uses that were to increase PSL's strategic capacity by providing reliable and transparent data.

# DIFFICULTIES

The difficulties to overcome are numerous:
- Some of the metrics are either themselves of doubtful quality, e.g. the biases introduced by bibliometric citation counts, so that their use in policy-definition is problematic or are too complicated to be sensibly used beyond a small circle of experts.

- The grasp of central management on the activities of units and individual researchers is typically low, so that the formulation of strategy and the conscious development of the research portfolio is difficult. The interest which researchers have for such matters is equally low. However, when evidence-based strategy definition is used as a tool for collective strategic decision making, it can support strategic thinking throughout the organisation. This is especially efficient when the whole process is fully transparent and co-designed with all actors at different levels: e.g. unit directors, research programs, schools. Building tools not only for top management, but that researchers can use themselves strengthens the engagement of communities towards the creation of a comprehensive RIS.
- Quality granular data on science and innovation activities and results is typically lacking, so that data must first be painstakingly acquired, and results can often only be published much further down the line (3 years of data gathering without any output visible to the community are not rare – enough to make a project lose impetus).
- A partial and segmented view of the research and innovation process, the value chains and of the overlapping institutional and public policies, due to a lack of integrated or interoperable data.
- A difficulty to define pertinent priority areas because of the difficulty to classify R&D activities and results beyond journal taxonomies provided by bibliometric data providers (try to grasp "Cryptography" or "Breast Cancer" via those taxonomies, especially since the Scopus taxonomy is established on the level of the journals and not on the level of the individual publication).
- A gap between experts and stakeholders, between users and providers of data and analysis, due to different vocabularies, knowledge and experiences.
- Concretely in our case: PSL is a complex entity: it is a university system on the path of becoming a single university. This means, that heterogenous institutions with varying levels of data quality and availability are collaborating in this project accompanied by political unrest during a time of deep change in the French and especially Parisian HERI landscape. Related difficulties include:
  - lack of pressure from the relevant ministries;
  - lack of interest of the individual researchers whose engagement is yet essential at least for data curation and quality management purposes;
  - lack of infrastructures at PSL member institutions;
  - the complex and intricate Parisian HERI system with many research units being shared among more than one actor entailing the need for collaboration.
- In France, no strong evaluation procedures are in place that could put pressure on institutions to adopt a data policy as it happened in the UK and in Italy.

# METHODOLOGY

We are sceptical towards pharaonic data projects at the national level, of which we have seen a few and which for now have not led to the breakthroughs they promised.[6] We do not wish to imply that such projects are nonsensical per se. We simply wish to say, that institutions should not wait for such projects to move forward. Institutions must take up the initiative themselves whilst ensuring that their actions can be adapted a posteriori to other initiatives or overarching standardisation efforts.

We adopted a methodology based on the presumption that engaging the community and keeping it engaged are key, if we want the system to be adopted and useful. We therefore adopted two principles:

- Do not place more work on people for data gathering than you absolutely must, because increasing the workload of people without providing quick return endangers engagement.
- Provide tangible returns pertinent for the stakeholders as early as possible in the process to establish and maintain legitimacy of the project. Define most needed uses as a first step. Target uses that can be profitable also for faculty staff (not only for the top management).

Building on these principles, instead of starting to build a comprehensive data warehouse, we started with a single dashboard as a pilot based principally on open data and bibliometric data to show the community what the capacities of the RIS might be with minimal supplementary workload on the community.

From this initial exercise, we go on to the definition of specific fields of interest and defined by focus groups and discussed with a wider audience during workshops and integrate further data (open if available or internal if not) to provide further pertinent indicators or to increase data quality by integrating internal data. However, the idea is to develop the RIS step by step:

→ Definition of an indicator → integration of data → rendition of results to stakeholders → restart.

# TECHNOLOGY

From the technological point of view, the requirements were the following:

- The system should be a lightweight and minimally invasive data federation and integration tool, that can be plugged into existing sources. It should not require the adoption of a specific new data curation system by each of the original source curators.
- The system should link internal data with external and ideally open data sources, since this achieves 2 goals:
  - it is a highly effective means for quality assessment since it allows us to compare the data we received internally to external quality sources. E.g. we compare internal information on European projects to the EU's own CORDIS database.
  - we can reach a high level of synergy by using available information from different sources.

This way, we can ensure the quality of the data whilst enriching the external sources with detailed information only available from local providers. This approach allowed us for instance to track down the lab level participation in European projects (CORDIS only provides institutional level information) by combining the EU's base CORDIS[7] with information from the Open Data service ScanR[8] operated by the French Ministry of Higher Education.

6 The latest being Conditor, a large-scale initiative federating many actors to establish a French bibliometric database.
7 http://cordis.europa.eu/home_en.html
8 https://scanr.enseignementsup-recherche.gouv.fr

- It should rely on open standards, be open source and under active development by a reliable community, because this ensures the independence of the client from any service provider (including SIRIS[9]). It also should be implemented among a relevant range of actors to ensure its sustainability.

The system is based on Semantic Web technologies, using Ontology-Based Data Access and of Linked Data approaches. This is a state-of-the-art framework which enables it to federate heterogeneous sources under a common vocabulary (ontology) without reforming the data curation at the local level.

The system conforms to the standards of the World Wide Web Consortium (WC3)[10] and to the Europe 2020 Strategy of the European Union[11] that advocates and promotes the use of Linked Open Data and Semantic Web technologies.[12] It includes an endpoint in the standard SPARQL language, which will allow the user to participate in Linked and Open Data initiatives in the future in order to increase visibility. The ontology is entirely compatible with VIVO-ISF[13] originally developed at Cornell University and with its European counterpart, the Common European Research Information Format (CERIF)[14]. This is essential for the future development and for the independence of the system from specific providers and to maintain sovereignty over its data at all times. VIVO-based systems are already implemented by over 140 academic actors[15] across 25 countries from Cornell and the MIT to UCLA, the presumption of durability is therefore warranted. Institutions that have implemented Linked Open Data approaches include National Statistical Agencies like the French INSEE[16] or the Italian ISTAT[17], as well as publishers, most notably Nature (Springer)[18].

# RESULTS

The current dashboard provides numerous elements for strategic analysis. The perimeters of analysis can be freely defined on the level of the research units (of which PSL has roughly 180), so that any combination of units can be aggregated for analysis. Available indicators include

- Bibliometric indicators for around 300 categories weighted against France, Europe or the world including benchmark modules with other French and European universities → useful to gauge strengths and weaknesses in the overall profile.
- EU projects: participation and funding filterable by programme weighted against France or Europe → top-level performance on the international scene
- Internal networks between units: co-publications and internal projects → useful to track real collaborations (and their respective intensity) in all the fields.
- External networks: Co-publications and EU projects Ð evaluate partnerships

More importantly, however, we have had numerous meetings with internal and external stakeholders and were able to assess more in detail the requirements and expectations. The transition from a first quickly done, largely top-down phase to a slower pace led by participatory design principles sees the project now on presumably much firmer position.

The added-value of this project until now consists in:

- Providing University's top management with consolidated data (staff, number of publications and impact, number of projects, range of internal collaborations, etc.), thus enhancing its credibility as an organisation able to define a strategy (both towards internal - the schools composing the University - and external stakeholders)
- Providing reliable data that may support lobbying and communication strategy
- Favouring a cultural change towards data use, also within academic community
- Supporting the development of other data-based tools for alternative uses (showcase PSL's areas of scientific expertise, build scientific maps on general or specialized topics, etc.)
- Helping to better define the way bibliometrics could be implemented to support decision-making.

# REFERENCES

**Bornmann, L. and Marx, W.** (2013). How good is research really? EMBO Reports, 14, 226-230.

**Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Poggi, A., Rodríguez-Muro, M. and Rosati. R.** (2009). Ontologies and databases: The DL-Lite approach. In Tessaris, S., Franconi, E. (Eds.), Reasoning Web. Semantic Technologies for Information Systems – 5th International Summer School Tutorial Lectures RW Lecture Notes in Computer Science, vol. 5689 (pp. 255-356). Cham: Springer.

**Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, Rosati. R. and Ruzzi. M.** (2009). Using OWL in data integration. In: Semantic Web Information Management – A Model-Based Perspective (pp. 397–424). Cham: Springer.

**Deem, R., and Brehony, K. J.** (2005). Management as ideology: The case of 'new managerialism' in higher education. Oxford review of education, 31, 217-235.

**Dunleavy, P., Margetts, H., Bastow, S., and Tinkler, J.** (2006). New public management is dead—long live digital-era governance. Journal of public administration research and theory, 16, 467- 494.

9 www.sirisacademic.com
10 www.w3.org/standards/semanticweb
11 https://ec.europa.eu/info/strategy/european-semester/framework/europe-2020-strategy_en
12 See also the EU's own portal https://data.europa.eu/euodp/linked-data
13 www.vivoweb.org
14 www.eurocris.org/cerif/main-features-cerif
15 http://duraspace.org/registry/vivo
16 http://rdf.insee.fr/sparql
17 http://datiopen.istat.it
18 www.springernature.com/gp/researchers/scigraph

**Ferlie, E., Musselin, C. and Andresani, G.** (2008). The steering of higher education systems: A public management perspective. Higher education, 56, 325-348.

**Foray, D. and Goenaga, X.** (2013). The goals of smart specialisation. European Commission, Joint Research Centre, S3 Policy Brief Series No. 1.

**Gianelle, C. and Kleibrink, A.** (2015). Monitoring Mechanisms for Smart Specialisation Strategies. European Commission, Joint Research Centre, S3 Policy Brief Series No. 13.

**Gingras Y.** (2014). Les dérives de l'évaluation de la recherche: du bon usage de la bibliométrie. Paris: Raisons d'agir.

**Haustein, S. and Larivière, V.** (2015). The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects. In Welpe, I.M., Wollersheim, J., Ringelhan, S. and Osterloh, M. (Eds), Incentives and performance (pp. 121-139). Cham: Springer.

**Hicks, D. and Melkers, J.** (2013). Bibliometrics as a tool for research evaluation, in Link, A.N. and Vonortas, N.S. (Eds), Handbook on the Theory and Practice of Program Evaluation, (pp. 323-349). Cheltenham: Edward Elgar Publishing.

**Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I.** (2015). Bibliometrics: the Leiden Manifesto for research metrics.

**Mosca, A., Rondelli, B. and Rull. G.** (2017). The OBDA-Based "Observatory of Research and Innovation" of the Tuscany Region. Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology, Bozen-Bolzano.

**Mosca, A., Roda, F. and Rull, G.** UNiCS - The ontology for Research and Innovation policy making. Submitted.

**Penfield, T., Baker, M. J., Scoble, R. and Wykes, M. C.** (2014). Assessment, evaluations, and definitions of research impact: A review. Research Evaluation, 23, 21-32.

**Radosevic, S.** (2017). Assessing EU Smart Specialization Policy in a Comparative Perspective. In: Radosevic, S., Curaj, A., Gheorghiu, R., Andreescu, L., Wade, I. (Eds) Advances in the Theory and Practice of Smart Specialization. London: Academic Press.

**Thoenig, J. and Paradeise, C.** (2016). Strategic Capacity and Organisational Capabilities: A Challenge for Universities. Minerva. 54, 293-324.

**Thoenig, J. and Paradeise, C.** (2015). In Search of Academic Quality, Palgrave Macmillan.

**Wang, J., Veugelers, R. and Stephan, P.** (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. Research Policy, 46, 1416-1436.

# AUTHORS

**EMMANUEL BASSET**
*Director of Research & Graduate Education, PSL University Paris,*
60 rue Mazarine, Paris, 75006 (France)
E: emmanuel.basset@psl.eu

**MATTHIAS HEUSER**
*Strategy Consultant at SIRIS Academic SL*
Av. Francesc Cambó 17, Barcelona, 08003 (Spain)
E: matthias.heuser@sirisacademic.com

# THE POLITICAL CONTEXT OF RESEARCH INFRASTRUCTURES: CONSEQUENCES FOR IMPACT AND EVALUATION[1]

ISABELLE VAN ELZAKKER AND LEONIE VAN DROOGE

DOI: 10.22163/fteval.2019.342


## ABSTRACT

Research Infrastructures (RIs) face big expectations regarding their societal impact. As a consequence, there is a need for methods to monitor and assess impact. But expectations differ between funders, organisations and countries, they change overtime and they are not always clear. In addition, expectations often relate to other functions, roles or ideas of an infrastructure than that of an organisation that enables excellent research. It is clear that a standard set of impacts and indicators doesn't do justice to these differences. Yet, at the same time, there is a need for a harmonized approach to impact monitoring. In this paper we describe the development of such an approach for a consortium of RIs as part of the H2020 ACCELERATE project.


## INTRODUCTION

Large Research infrastructures (RIs) have become an objective of a variety of policies, both on the regional and national, as well as the European level. Governments and public organisations across all these levels make large public investments to construct and operate RIs. Moreover, different countries and organisations are increasingly cooperating in funding and managing RIs. To legitimize these efforts, both funders and RIs themselves expect RIs to have various beneficial impacts, ranging from scientific breakthroughs to regional innovation and the development of new technologies.

The evaluation, as well as the governance of societal impact, is a challenge. There have been reports and studies of impact, yet there is no common agreement on impacts to expect, or approaches to evaluate impact. However, improved governance and evaluation of societal impacts is expected to contribute to the long term sustainability of RIs.

The context of this paper is the ACCELERATE project, dedicated to the long term sustainability of RIs in the field of materials research. The RIs have articulated the need for a proactive governance of societal impact. The RIs indicated that the studies and methods available, do not respond to their needs. We are involved in the project in order to develop a societal impact approach for use by the RIs themselves. The paper addresses the question: How to understand societal impact of a Research Infrastructure?

In this paper we describe how the (European) RI policy landscape developed in the past two decades. It is in the context of these broader developments that the question of societal impact is brought to the fore. We describe core elements of methods used for societal impact assessment of RIs. We relate this to the practice and needs of the RIs involved in the ACCELERATE project. In the discussion, we reflect on the implications of the political context in which RIs operate, for the understanding of societal impact of RIs.

## BACKGROUND: DEVELOPMENT OF THE RI (POLICY) FIELD

Since the turn of the century, RIs have gained a significant position on the European (science) policy agenda. The memorandum 'Towards a European Research Area', published and approved in 2000, positioned RIs as a policy objective on the agenda (European Commission, 2000). RIs are, according to the Memorandum, important for scientific progress. They are tools for European cooperation and integration. The notion of a European strategy for RIs that was introduced in the memorandum, offered individual member states the prospect of reducing costs, by sharing the capital and operational investments accompanying the establishment of RIs (Papon, 2004).

Research Infrastructures, according to the European Commission, are facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields. They include major scientific equipment or sets of instruments; knowledge-based resources such as collections, archives or scientific data; e-infrastructures such as data and computing systems and communication

1 ACCELERATE is funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant agreement 731112.

networks; and any other infrastructure of a unique nature essential to achieving excellence in research and innovation (European Commission, 2018a: part 4, p.5).

To enable the development of a European strategy on RIs, the European Strategy Forum on Research Infrastructures (ESFRI) was established in 2002. ESFRI aims to support a "coherent and strategy led approach to policymaking on research infrastructures" (ESFRI, 2018). One of the tools is roadmapping; a strategic, long-term, policy-relevant planning exercise between member states (OECD, 2008), resulting into roadmaps for the construction and development of pan-European research infrastructures. In 2006 ESFRI published its first roadmap and ESFRI has updated the roadmap multiple times since then. Each new roadmap includes new projects and initiatives, as well as projects from earlier roadmaps. What is on the roadmap is an RI.

The efforts to establish a coherent European RI strategy influences the national processes in its member states. ESFRI expects member states to develop their own national roadmaps. These need to include national facilities as well as participation in international RIs. In the Netherlands, for example, the national roadmap is harmonized with the European roadmap (NWO, 2016). This means that national RIs, or consortia of RIs, have to link to an international initiative on the ESFRI roadmap, to be eligible for a place on the national roadmap. A substantive amount of Dutch public funding for RIs is tied to the national roadmap. This financial incentive thus encourages national RIs to connect with pan-European RI initiatives.

Another development in the creation of a unified European RI landscape, is the introduction of the European Research Infrastructure Consortium (ERIC) legal framework (EC, 2009). The ERIC framework provides consortia the possibility to act as a European legal entity. The consortium can consist of – and is funded and governed by - EU Member States, associated countries, third countries and intergovernmental organisations. The ERIC framework provides a blueprint for a structure and it allows for a faster process than creating an international organisation. An ERIC needs to represent added-value in the development of the European Research Area (ERA). It needs to contribute to significant improvement in the relevant scientific and technological fields, to the mobility of knowledge and/or researchers within the ERA and to the dissemination and optimisation of results (EC, 2018b).

In the past two decades, the number and variety of facilities that are identified as RIs and that are included on roadmaps have grown. Currently, a large variety of facilities is identified as an RI: from single sited physical buildings with equipment for scientific experiments and measurements to distributed testbeds for crops and from virtual and networked datasets for social sciences and humanities research to mobile facilities for marine research.

## "LONG TERM SUSTAINABILITY" AND SOCIETAL IMPACT

The long term sustainability of Research Infrastructures has received attention in recent years. The Conclusions of the Council of the European Union in 2016 underlined the importance and urged the European Commission to develop an action plan. In response, ESFRI established the Long-term Sustainability Group. The OECD as well as various H2020 projects address the issue. The lifecycle of a RI often covers multiple decades. To ensure the sustainability of a RI throughout its lifecycle, ESFRI identified different aspects and issues concerning sustainability (ESFRI 2017). They include the effective governance of RIs, the (lack of) coordination between the national and European level, and the availability of people with the right skills and experience.

One of the obstacles for ensuring the long term sustainability of RIs is the lack of a sound methodology for identifying and assessing the societal impact of RIs. Some RIs require substantive public investments. The expectations driving these investments have shifted in the past decades and RIs. Even RIs that do not require such investments are now expected to contribute to the needs of contemporary society (Hallonsten, 2017). As a consequence, there is political and social pressure to identify, monitor and evaluate the contribution that RIs make to society in general, or to regional and national economies, or through the science the RIs deliver, such as better healthcare, a cleaner environment or developments to communications and transport (ESFRI, 2017). However, clear articulation of expectations regarding societal impact, or regular monitoring, are not yet common practice, neither among funders, members and stakeholders, nor at RIs (ESFRI, 2017). Still, the need to develop a standard methodology for assessing the societal impact of RIs is widely shared (cf. European Commission (2017), OECD (2017), ESFRI (2017)).

## THE SOCIETAL IMPACT OF RESEARCH INFRASTRUCTURES

Despite the call for a methodology, there is no lack of studies dedicated to the societal impact of RIs. They cover a broad range of methods, from ex-post qualitative case studies, to ex-ante cost-benefit analyses (Giffoni et al, 2018b). Most focus on a specific RI (e.g. on ISIS (Simmonds, 2016), European Social Survey (Kolarz, 2017) and ICOS ERIC (Van Belle et al, 2018)), but there have been attempts to develop a more generic framework for the assessment of societal impacts as well. Examples include the work of Technopolis (Greniece et al, 2015), the FenRIAM guide (Roschow et al, 2014), the ongoing work of the OECD Global Science Forum (OECD, 2017; OECD, 2018) and the recently started H2020 project RI-PATHS.

There are similarities between the studies, such as the use of a model for impact – or of elements such as inputs, activities, outputs, outcomes and impacts. Another similarity is the articulation of the differences between RIs, including in studies dedicated to a specific RI. Yet the studies differ in how they analytically "pull apart" impact and RIs. The questions "impact of what?" and "impact on what?" are addressed in different ways.

## DATA

We have used desk research to study the evolving policy context of RIs. We focused on policy documents concerning RIs, the ERA and roadmaps, as well as policy documents addressing societal impact (assessment) of RIs. We studied literature on RI impact, including consultancy reports and case studies. Through our project, we had access to official as well as internal documents of the member RIs, including Statutes and Annual Reports, as well as monitoring documents.

We interviewed representatives of the RIs involved in ACCELERATE and organised joint workshops to identify questions, interests and needs regarding impact, to identify relevant stakeholders that are involved in the evaluation of societal impact, as well as to discuss expectations and practices.

Finally, we have been involved and invited in a number of meetings regarding societal impact of RIs. This helped us relate our project to ongoing developments in the field of societal impact of RIs.

# RESULTS - WHAT IMPACT AND IMPACT OF WHAT?

The studies of and reports on societal impact of RIs use different interpretations of impact and RIs. We describe three common trends. We then use these interpretations to describe the RIs of the ACCELERATE project.

## IMPACT OF WHAT

Some studies use observable characteristics of an RI as a starting point. For example the scientific domain or discipline the RI serves, its scope (single-sited, distributed, mobile or virtual) or phase (construction, operation, decommissioning) (ESFRI, 2017). RI PATHS proposes a more holistic approach, where the taxonomy is based on type of research (Giffoni et al, 2018a). Technopolis uses a typology of characteristics that makes the phase explicit (Technopolis, 2015).

## IMPACT ON WHAT

Some of these studies propose a typology of impacts as well. Technopolis (Greniece, 2015) distinguishes between impacts on the economy, on innovation, on human resource capacity and on society.

## IMPACT AS A CONTRIBUTION TO A GOAL

More recent impact studies (Kolarz et al. 2017), policy documents (ESFRI 2017, OECD, 2018) and impact approaches developed by RIs themselves (ESS, 2018), use a different approach. They relate impact to other features of the RIs. They state that a one-size-fits-all approach will not do, and that "there seems to be no "silver bullet" for capturing the impacts of RI" (Berger et al, 2018: 55), precisely given the heterogeneity of RIs and of impacts. They relate societal impact to the objectives or goals of a specific RI (ESS, 2018). Some include that impact also relates to expectations of stakeholders (OECD, 2018: 1), since RIs face multiple stakeholders, that have different strategic visions and expectations. Studies relate impact to goals, missions and expectations. Impact can be understood as a "contribution to".

## THE ACCELERATE RIS

The ACCELERATE consortium consists of five RIs: CERIC (Central European Research Infrastructure Consortium), ESS (the European Spallation Source ERIC), FRM II, HZG-GEMS (Helmholtz Gesellschaft) and ELI (Extreme Light Infrastructure). They are all dedicated to enabling materials research: the characterisation of matter, from subatomic to supramolecular scale. The research that the RIs enable is done with equipment called beamlines or instruments. These are connected to a powerful source, an accelerator, spallation source, or laser.

## IMPACT OF WHAT?

The members can be further described using some basic characteristics:

**Phase:** ELI is currently under construction. It will enable materials research in the future, however at present it is a building project. FRM II is in operation since 2015.

**Scope:** FRM II is a single sited RI. The research facility is located on one specific site, in Garching, Germany. The spallation source of ESS is built on a single site in Lund, Sweden. However the Data Management and Software Centre (DMSC) is located in Copenhagen, Denmark. In the other cases, the RI consists of multiple physical sites that together make up the RI. HZG-GEMS manages instrumentation at different sites, ELI consists of three research facilities and CERIC coordinates between a number of facilities.

**Governance:** Three of the RIs, ELI, ESS and CERIC, are ERICs. This means that they have a European legal status. They are governed by European countries that are a member of the ERIC, and that fund part of its construction/operation, either in-kind or in cash. In contrast, the two German RIs are part of existing research organisations: FRM II is governed by Technical University Munich and HZG-GEMS that operates instruments at distant facilities, is operated by the Institute for Materials Research, which is part of the Helmholtz Gesellschaft. The latter two are funded through national and regional scientific funds.

**Span of control:** With span of control we refer to the responsibilities and possibilities of the RI regarding the facilities it offers access to. CERIC does not own any physical instrument, beamline or source; it offers access to beamlines operated by representatives of the member states, at different partner facilities. On the other hand, FRM II manages both source, as well as part of the instruments. FRM II both facilitates research as well as does in house research and it is a source of medical isotopes.

## IMPACT ON WHAT

All ERICs (CERIC, ESS and ELI) need to represent added-value in the development of the European Research Area. The ERA focuses on five key priorities. The ERICs are expected to report on their contribution to these priorities. The priorities are negotiated by different political actors within the EU. They can change overtime in a response to new issues arising or others becoming less relevant.

In some RIs, the statutes provide some information on what they should impact on. CERIC- ERIC's objective shall be to *"stimulat[e]… beneficial impact on the scientific, industrial and economic development.* (CERIC 2014: 6) and CERIC *"shall proceed to the periodical evaluation of […] its impact on the European Research Area, on the Regions hosting its Partner Facilities and at international level."* (CERIC 2014: 12)

## IMPACT AS A CONTRIBUTION TO A GOAL

Some members have defined contributions to goals These (strategic) goals are used as, in other words to define, impact categories.

ESS for example, uses their strategic goals as impact categories: (1) World-Class RI Enabling Scientific Breakthroughs and Addressing Grand Societal Challenges (2) Supports and Develops Its User Community, Fosters a Scientific Culture of Excellence and Acts as an International Scientific Hub. (3) Is Built on Time and on Budget, Operates Safely, Efficiently and Economically, and Responds to the Needs of Stakeholders and (4) Develop Innovative Ways of Working (ESS, 2018).

# DISCUSSION AND CONCLUSIONS

We discuss the above in the context of the experiences and needs of the members of ACCELERATE. They require an approach, or an understanding of societal impact, that improves the possibility to pro-actively govern impact. These RIs, just as others, face a complex stakeholder community, consisting of their members, funders, users and beneficiaries; each with different expectations regarding societal impact, each with different requirements, and each with a different interpretation of impact, if any.

The members of ACCELERATE are all dedicated to enabling materials research by providing access to instruments and beamlines. Apart from these similarities the RIs differ considerably: from building projects to up-and-running organisations; from institutionally or nationally governed organizations to former EU project consortia turned ERICs, funded by member states; from an annual budget of 3 million euro to an 1.8 billion euro investment. It is evident that the impacts will differ given these characteristics, and that different impacts are expected given the different stakeholder communities. At the same time, supranational European initiatives call for coherent policies for RIs, including for the (assessment of) societal impact of RIs. There is an inherent tension here, given the different characteristics and contexts.

The ACCELERATE members are expected to have an impact on the economy, innovation and other societal sectors, as well as on the societal challenges. Yet these expectations are in most cases not concrete or specific. It is often unclear to the RIs what is expected in terms of the nature of the impact, the contribution by the RI or the evidence of impact. Regarding the impact on ERA priorities, for instance, the ERICs merely contribute to. This is in line with the more recent development, where impacts relate to objectives. As mentioned above, one of the members has developed a societal impact approach that uses its strategic objectives as impact categories.

However, the ACCELERATE members report that there is more to impact than is formally agreed and communicated in statutes, mission statements and strategic objectives. Different stakeholder groups have different perspectives of what an RI is. For a hosting member country, the seat of an ERIC is a prestigious project; for the ministry of economic affairs of a member country, the same RI is an opportunity for high-tech industry; for a ministry of science, the very same RI is the opportunity to collaborate with excellent scientists from abroad. For a regional government, the RI is a high-tech employer; the reactor of that RI is perceived by some of the local population as a potential danger; for doctors and patients in a different country, the RI is a provider of medical isotopes.

Every stakeholder seems to have a different perception of an RI. And each perception relates to a different impact or contribution. Pro-active governance of societal impact includes pro-active governance of the image or perception the stakeholder has of an RI. What an RI does, or what it monitors, depends not only on the impact expected, but also on the perception of a stakeholder, and the RI, what the RI is.

These perceptions are not just “out there”; what an RI is, and what impact to expect, can be discussed and negotiated between the RI and its stakeholders, and among different stakeholders. An RI is in that sense a boundary object (Star and Griesemer, 1989). It is adaptable to viewpoints of a wide variety of stakeholders including stakeholders that are not commonly involved in science and innovation, such as local communities and regional employers. What a Research Infrastructure is, is influenced by the perception and expectations of the stakeholder, the activities and strategy of the RI as well as the negotiations, or the lack thereof, between these actors.

An RI is as a boundary object from another viewpoint as well. Policies regarding RIs have opened up possibilities and opportunities to include, develop or identify facilities as Research Infrastructure or in other words: to put them on the map. What is commonly referred to as a Research Infrastructure has been negotiated, expanded and stretched resulting into a large variety of projects and activities that are nowadays identified as an RI.

Revisiting current initiatives and practices in assessing societal impact and relating that to the practice of a number of RIs, aids in understanding the challenge regarding societal impact. The diversity of RIs, the large variety of stakeholders of an RI and the different views they have about an RI, suggests indeed that a standard taxonomy, or a standard set of indicators, does not do justice. However, it does provide a direction. It starts from the observation that an RI operates in a complex context, and that it is perceived differently by different stakeholders. Both RIs as well as stakeholders search for points of reference. This searching provides the opportunity to develop a joint view on what the RI is, and what can be expected of it.

# REFERENCES

**ACCELERATE** (2018). Retrieved October 11th, 2018 from http://www.accelerate2020.eu/

**Berger, F. et al.** (2018). International Comparative Study: Appraisal and Evaluation Practices of Science Capital Spending on Research Infrastructures. Final Report. Brighton: Technopolis Group

**BrightnESS** (2018). private communication

**CERIC** (2014). STATUTES OF CERIC-ERIC. 24 June 2014 Decision no. 2014 2014/392/EU.

**ESFRI** (2018). Retrieved October 2018 from https://www.esfri.eu/

**European Commission** (2000). Communications “Towards a European research area” COM (2000) 6 of 18.1.2000.

**European Commission** (2009). Council Regulation (EC) on the Community legal framework for a European Research Infrastructure Consortium (ERIC) No 723/2009 of 25 June 2009.

**European Commission** (2010). Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions - Towards a European research area COM/2000/0006 final.

**European Commission** (2017). Sustainable European Research Infrastructures – A call for action. Commission Staff Working Document, SWD (2017) 323 final. Luxembourg: Publications Office of the European Union.

**European Commission** (2018a). Horizon 2020: Work Programme 2018-2020. Part 4. European research infrastructures (including e-Infrastructures). European Commission Decision C(2018)4708 of 24 July 2018.

**European Commission** (2018b). retrieved October 21st, 2018 from http://ec.europa.eu/research/infrastructures/index.cfm?pg=eric

**European Strategy Forum on Research Infrastructures (ESFRI)** (2017). Long-Term Sustainability of Research Infrastructures. ESFRI SCRIPTA Volume II. Milan: Dipartimento di Fisica - Università degli Studi di Milano.

**Giffoni, F. et al** (2018a). Deliverable 3.1 - Working note on RI typology. Retrieved September 24, 2018 from https://ri-paths.eu/

**Giffoni, F. et al** (2018b). Task 3.2 State of play – literature review. Retrieved September 24, 2018 from https://ri-paths.eu/

**Griniece, E. et al.** (2015). Evaluating and Monitoring the Socio-Economic Impact of Investment in Research Infrastructures. Tallinn: Technopolis Group.

**Hallonsten, O.** (2017). Big Science Transformed. Science, Politics and Organization in Europe and the United States. London: Palgrave Macmillan.

**Kolarz, P. et al.** (2017). Comparative impact study of the European Social Survey (ESS) ERIC. Final report. Brighton: Technopolis Group.

**Netherlands Organisation for Scientific Research (NWO)** (2016). National Roadmap Large-Scale Scientific Infrastructure. Den Haag: NWO.

**OECD Global Science Forum** (2008). Report on Roadmapping of Large Research Infrastructures. Paris: OECD.

**OECD** (2014). The Impacts of Large Research Infrastructures on Economic Innovation and on Society: Case Studies at CERN. Paris: OECD.

**OECD** (2017). "Strengthening the effectiveness and sustainability of international research infrastructures." OECD Science, Technology and Industry Policy Papers, No. 48.

**OECD Global Science Forum** (2018). Draft Assessment Framework Initial Consensus Indicators as of 13 March 2018. Paris: OECD.

**Papon, P.** (2004). European scientific cooperation and Research Infrastructures: Past tendencies and future prospects. Minerva. 42, pp. 61-76.

**Roschow, R. et al.** (2014). FenRIAM full guide. Proposal for a Foresight-enriched Research Infrastructure Impact assessment Methodology. Retrieved September 9, 2017 from http://www.fenriam.eu/overview.html

**Simmonds, P. et al** (2016). ISIS Lifetime Impact Study. Volume 1 – Full Report, Technopolis Group.

**Star, S.L. and Griesemer, J.** (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science. 19 (3), pp. 387-420.

**Van Belle, J. et al.** (2018). ICOS Impact Assessment Report. Amsterdam: Technopolis Group.

**Zubascu, F.** (2018). Are research infrastructures the answer to all our problems? Science Business Retrieved September 19, 2018 from: https://sciencebusiness.net/news/are-research-infrastructures-answer-all-our-problems

# AUTHORS

**ISABELLE VAN ELZAKKER**
*Rathenau Institute*
Anna van Saksenlaan 51, 2593 HW The Hague (the Netherlands)
E: i.vanelzakker@rathenau.nl

**LEONIE VAN DROOGE**
*Rathenau Institute*
Anna van Saksenlaan 51, 2593 HW The Hague (the Netherlands)
E: l.vandrooge@rathenau.nl

# EVALUATION FRAMEWORK FOR PROMOTING GENDER EQUALITY IN RESEARCH AND INNOVATION:

## HOW DOES GENDER EQUALITY INFLUENCE RESEARCH AND INNOVATION OUTCOMES AND WHAT IMPLICATIONS CAN BE DERIVED FOR SUITABLE EVALUATION APPROACHES?[1]

SUSANNE BÜHRER, SYBILLE REIDL, EVANTHIA K. SCHMIDT, RACHEL PALMEN, CLEMENS STRIEBING AND DORA GROO

DOI: 10.22163/fteval.2019.332

## BACKGROUND AND PURPOSE OF THE STUDY

Despite all efforts undertaken in the past there is no comprehensive and rigorous analytical framework to consider all of the relevant variables in gender equality issues, although there have been a number of European Commission projects such as PRAGES, GENDERA, GenSET, STAGES and GENOVATE, which have explored the gender equality (GE) dimension with different foci. While all these previous studies have illustrated numerous evaluation approaches, concepts, indicators etc. to provide examples of measuring different kinds of impacts, a clear understanding of the mechanisms between different gender equality-related policy initiatives and interventions (inputs) and outputs/results is still not available. In order to address these challenges, EFFORTI (Evaluation Framework for Promoting Gender Equality in Research & Innovation), an EU funded project, aims to clarify the mechanisms between gender equality inputs and the expected results not only on gender equality itself, but also on research and innovation (R&I). The evaluation framework provides the theory and tools for analysing how gender equality-related interventions contribute to the achievement of the three European Research Area's main objectives on gender equality and how those achievements affect the desired outcomes of (responsible) research and innovation. The uniqueness of the evaluation framework is that it goes beyond conventional research and innovation indicators, taking into account also evaluation dimensions like providing answers to the Grand Challenges and the promotion of Responsible Research and Innovation.

With the rise of the idea of evidence-based policy-making (e.g. Nutley et al. 2002; Solesbury 2001; Sanderson 2002), expectations have grown regarding the use of scientific evidence in policy-making. At the same time, establishing causal relationships between policy interventions and observed changes poses a theoretical challenge as well as empirical and methodological problems. One approach to address these challenges is the theory-based impact evaluation approach (TBIE): In theory-based impact evaluation (TBIE), causality is often defined as a problem of contribution, not attribution. "Why and how" questions are typically being asked instead of "how things would have been without" like counterfactual approaches do. The goal is to answer the "why it works" question by identifying the theory of change ("how things should logically work to produce the desired change") behind the program and assessing its success by comparing theory with actual implementation. The "theories" to be investigated on how gender equality and R&I outcomes interrelate (intervention logics), which in turn link the allocation of resources to the achievement of intended results and finally impacts are still to be developed. These might be complemented by academic theories about public interventions and already existing empirical evidence from former evaluations and impact assessments. The actual results of GE policies will depend both on policy effectiveness and on other context variables. Context factors are organizational structures and cultures, as well as national and regional structures, capabilities and policies. The application of a theory based impact evaluation approach will allow us to take these different levels of influences on policy effectiveness - mechanisms and context - systematically into account. Furthermore, it allows us to develop context sensitive and policy specific theories of change.

1 Acknowledgement: The project EFFORTI that constitutes the basis for this contribution receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 710470. The authors would like to acknowledge the contribution of all project participants and all project activities to the ideas that underpin this paper.

# METHODOLOGICAL APPROACH

Drawing on already developed and applied indicators in gender equality and R&I research (RIO Observatory, OECD STI Scoreboard etc.), but also on recent studies on RRI indicators (Ravn et al. 2015a, 2015b, European Commission 2015), we carried out a comprehensive desk research as a basis for the collection of a preliminary list of relevant indicators. Based on existing evidence, the project team first identified the most relevant indicators according to literature review; clustered these indicators into different categories, dimensions and sub-dimensions, which are based on GE-related literature and smart practice examples implemented in different organisations and contexts; and finally grouped these indicators according to an evaluation logic model. The indicators are differentiated between input, throughput, output, outcome and impact aspects. For each aspect, the indicators are illustrated at micro/ individual or team level, meso/organisational level and macro/policy or country level.

The indicators are based on the collection and review of "smart practices" implemented in Europe and beyond. The identification of smart practices was based on an assessment of the practices that are relevant, effective and efficient in the context that they operate in as to their quality of both evaluation and measurement (Kalpazidou Schmidt et al. 2017c). Smart practice examples evaluated measures of different nature and length: some constituted large national programmes with a long-term perspective, while others were of a more limited character. The selection of smart practices was based on the criteria of (1) the quality of the implemented measures, and (2) the impact of the measures. The quality of the measures was assessed based on the parameters of relevance, effectiveness, efficiency, and sustainability of the interventions, while the impact of the measures was assessed in relation to its subjective/objective dimension (Kalpazidou Schmidt & Cacace 2017). Synthesising the typologies developed by Kalpazidou Schmidt and Cacace (2017) and the fields of action identified by the GENERA project and building on further theoretical and empirical experiences, we developed an intervention typology. Examples of impact stories were developed for a broad spectrum of these intervention types in order to provide examples of the mechanisms regarding intervention intentions and to provide a common framework for understanding the multi-faceted interventions of the cases that will serve as a testing ground for the further development of the tentative evaluation model.

# CASE STUDY APPROACH FOR VALIDATION PURPOSES

The EFFORTI intervention logic model forms the conceptual basis for the case study work. The Intervention Logic Model considers inputs, throughputs, and outputs, as well as outcomes and impacts of the former two. The model also aims at showing how, once achieved, these objectives or effects can further affect desired R&I effects such as the number of patents and number of publications and citations, but also new R&I effects, such as providing answers to grand challenges and further promoting RRI. Additionally, the model includes three levels, i.e. team level (research quality, productivity, innovative outputs, and other RRI effects), organisational/ institutional level (workplace quality, recruitment capacity, efficiency, RRI orientation, competitiveness), and country/ system/ policy level (intensity, productivity, ERA orientation, etc.). However, some interventions will most likely overlap between different levels, which was taken into account in the development of the toolbox (EFFORTI Conceptual Evaluation Framework, D3.3, Kalpazidou et al. 2017.8). After having developed a first tentative evaluation framework, a series of case studies is foreseen to validate and further improve the model. Yin (1994.13) defines a case study inquiry as one that *"Investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident."* Therefore, the case study method lends itself to research where contextual factors are highly pertinent to the phenomenon of study (ibid). Case studies as a method have also been used extensively in evaluation research. We will use the case study method to inductively build on and validate the evaluation framework. The multiple case study work will shed light on those factors and mechanisms that shape and influence the effects of gender equality interventions in R&I on research and innovation outputs. It will attempt to explain what works (and what does not work) in what context and why. It will also explore whether the intervention is likely to work elsewhere and what is needed to make it work elsewhere. It will also attempt to explain how the national/ science system context influences the intervention in terms of the main contextual elements as well as the main agendas, strategies, and policies that shape the intervention. How the institutional context influences the intervention will also be taken into consideration – as will an assessment of whether the general conditions for effective gender equality policies are in place.

# CASE STUDY EXAMPLE

In order to illustrate what insights can be gained from this approach for the Evaluation Toolbox, we subsequently outline exemplary results of a case study that addresses the ERA goal: *integration of gender dimension in research and education:*

"FEMtech Research Projects" is a funding scheme of the Austrian Research Promotion Agency (FFG), which supports projects in applied research, technology / product and process development that integrate the gender dimension in research content. It can be classified as an international good practice example in the context of fostering gender in research content because it is one of the few funding programmes that comprehensively targets the implementation of the gender dimension in scientific and technological research projects. By means of content analysis of project descriptions, interviews with policy designers, program managers and representatives of three funded research projects, the effects of this funding instrument especially its contribution to research and innovation were analysed.

In the beginning of the Case Study an intervention logic model was developed:

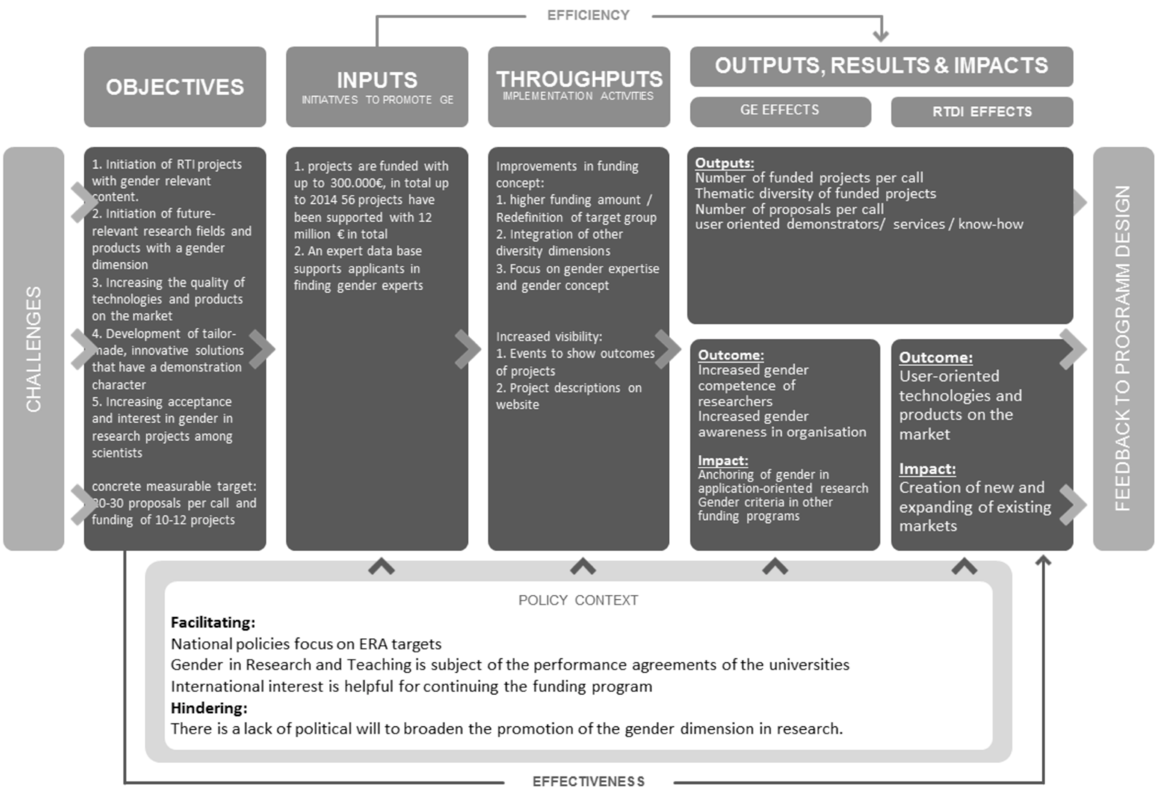



**Figure 1.** Intervention logic model "FEMtech Research Projects"

This log frame shows that the funding programme aims to initiate RTDI projects with gender relevant content in future-oriented research fields and in the long run wants to enhance the quality of technologies and products on the market. Funded projects shall develop tailor-made, innovative solutions that have a demonstration character. In addition, "FEMtech Research Projects" aims to increase acceptance and interest in the topic of integrating gender in research among scientists.

All in all 10-12 projects shall be funded per call with a maximum funding of 300.000€ per project. From 2008 until 2014, 7 calls have been launched. In total 56 projects were funded with a sum of 12 Mio. €.

Moreover, the logic model shows the expected outputs, outcomes and impacts of the funding scheme as they are stated in program documents[2] and formulated in interviews with policy makers and program managers (codes of interviewees: AU_CS2_02, AU_CS2_03, AU_CS2_04).

In the case study, it was then tried to measure possible outputs, outcomes and impacts of the funding program. The monitoring data for all seven calls between 2008 and 2014 shows a constantly rising number of submitted proposals with only one outlier in 2011.

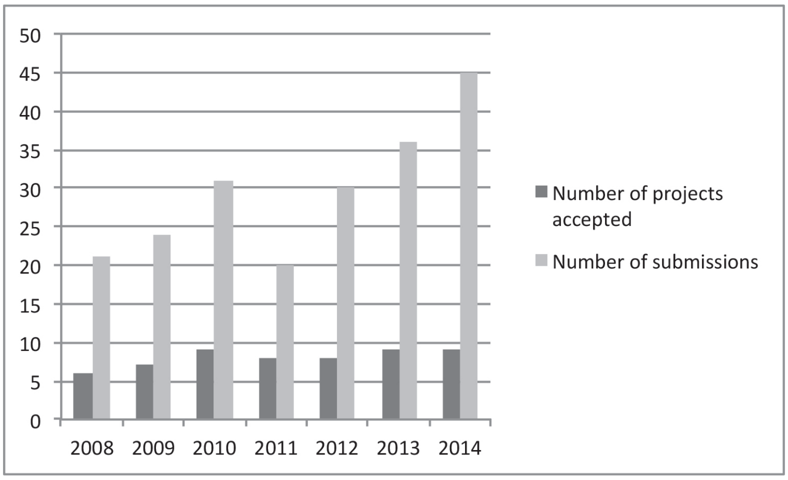



**Figure 2:** number of submitted and accepted projects per year
**Source:** FFG

1 See https://www.ffg.at/femtech-forschungsprojekte/5-ausschreibung

The number of funded projects stayed between six and nine per call because the amount of funding distributed was too low to fund more projects. The funded projects spread over six thematic categories from Energy/Ecology over Life Science to IT/communication, which indicates a thematic diversity of funded projects.

The measure's short-term output consists in the integration of gendered user involvement activities into technology development processes like gender divided test groups, gendered needs assessments, usability tests, participatory co-designing etc., ideally from the very beginning (see Nedopil/Schauber/Glende 2013; Rommes 2014). The result of this changed technology development process is information on gender-specific (and diversity-specific) user requirements for the product to be developed. But the analysis of the online-project descriptions shows, that not all of the projects focus on developing products, there were also studies funded to gain more gender-specific knowledge and projects that developed gender specific services. As the funded FEMtech research projects exhibit a broad range of research foci and project durations, the produced outputs vary widely as well. Also, most projects generated not one but several types of results. The most commonly produced result is the review of a product or service from a gender perspective. This can be explained with the considerably low funding of € 300,000 per project, which does not really make the development of a new product possible (AU_CS2_15). Less common are tutorials, didactic concepts / training concepts or manuals.

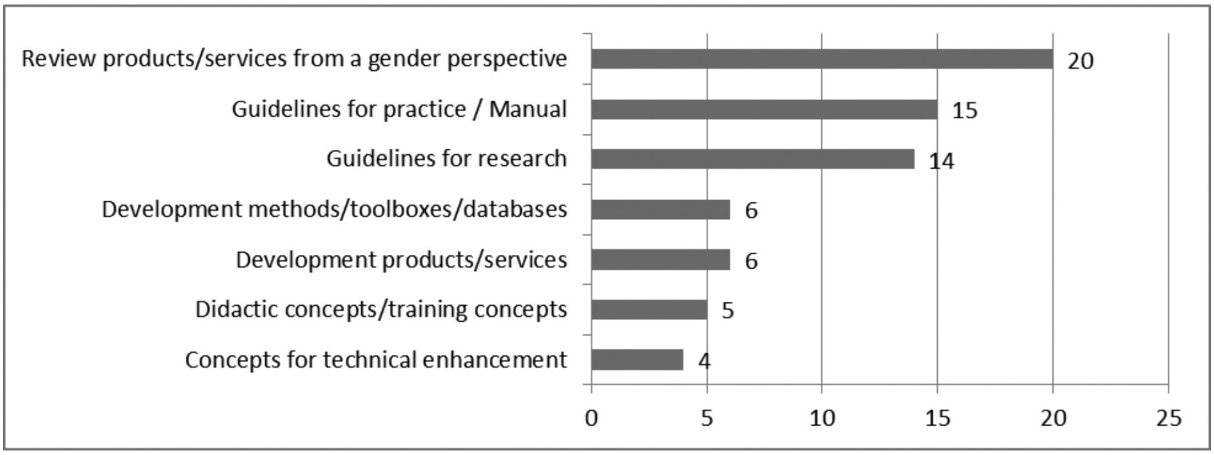



**Figure 3:** Type of results (number of projects)
**Source:** https://www.femtech.at/projekte, Analysis Joanneum Research

As many of the funded projects do not aim to develop new or improve existing products the outcome of the funding program cannot be measured only by counting user-oriented products and technologies on the market as it was expected in the intervention logic model. Another reason why this indicator is not useful is, that "FEMtech Research Projects" does not fund development processes until market entry. Therefore, information about the potential further development process after funding has ended is not available. Instead, in the case study the outcome of "FEMtech Research Projects" was measured by identifying different forms of further usage of project results. In 19 funded projects, starting points for further research were identified. 18 projects plan an application of project results in practice; another 12 are committed to apply the project results. Moreover, one interviewee reports of a market launch of a developed service in the upcoming months (AU_CS2_09).


**Figure 4:** Type of further use of results (number of projects)
**Source**: https://www.femtech.at/projekte, Analysis Joanneum Research

Follow-up projects are mentioned only in three project descriptions but are an important issue in the interviews with project representatives. Seven interviewees report of having already submitted another FEMtech project or a follow up project in another funding program.

In the interviews, also other outcomes of the projects were mentioned: the researchers gained gender competence in the course of the project and became self-confident regarding its practical use. For the research organisation the FEMtech Research Project also means a reference for further gender project applications. Some interviewees could also improve the scientific quality of the gender analysis in their research proposals for other funding schemes. The gained knowledge and competences regarding gender and/or new research methods were used in teaching, trainings and other research projects. Most interviewees also mentioned a sensitization of researchers regarding interdisciplinary and/or participative research through these projects.

This directly refers to the program target of increasing acceptance and interest in gender in research projects of scientists and the expected gender equality impact of anchoring gender in application-oriented research. The interviews showed an increasing awareness of the relevance of the gender dimension in research and also an increase of gender knowledge of researchers and representatives of companies who participated in the projects. This is also confirmed by the increasing quality of applications.

To investigate whether the funding program contributes to anchoring the gender dimension in application-oriented research, a social network analysis of funded organisations was conducted. It was examined whether the group of beneficiaries has grown from call to call. This analysis shows an expansion of research organisations and companies participating in "FEMtech Research Projects". This "spreading" is frequently happening via actors that submit regularly in the funding line, but with changing cooperation partners. It can be assumed that they have a multiplier function to involve other organisations in dealing with the gender dimension in research. But they can also take on a gatekeeper function in the future because they gain a substantial knowledge advantage, whereby other applicants with less experience are no longer competitive. All in all the community of organisations that already have conducted a FEMtech research project is still rather small compared to the number of research performing organizations in Austria. This can mainly be attributed to the fact that this funding scheme is rather small compared to other RTDI funding programmes. The lack of political will to broaden the promotion of the gender dimension in research, which was identified as a relevant context factor in the intervention logic model, reduces the expected impact of the "FEMtech Research Projects".

In the interviews, further possible effects of "FEMtech Research Projects" are addressed, which could not be investigated in the case study due to limited resources: FEMtech may have contributed to the implementation of gender criteria in other funding programmes. And interviewees report that organisations who conducted a FEMtech research project later on submit a proposal for a FEMtech Career project to start organizational change towards gender equality. This could not be investigated in this case study, as FFG monitoring data for other funding instruments could not be accessed.

We have presented some results of the "FEMtech Research Projects" case study and will now draw some conclusions from the case study for the development of the Evaluation Framework in EFFORTI: The case study shows that some expected effects cannot be detected due to lack of data. However, alternative ways of measuring outputs, outcomes and impact could also be demonstrated. The case study represents the first attempt to measure RTDI effects of FEMtech Research Projects as so far only concept and implementation evaluations have been carried out for this instrument. Indicators for impact assessment were outlined and will be included in the EFFORTI Evaluation Framework. With access to further funding data, impact measurement could be further developed.

# DISCUSSION AND CONCLUSIONS

Based on a thorough analysis of the relevant knowledge in gender equality, evaluation as well as science and innovation research and the structured analysis of smart practice examples, a first evaluation framework has been developed which was then used for the conduction of in total 19 case studies in seven EU countries (Austria, Denmark, France, Germany, Hungary, Spain, Sweden). The case studies cover a broad range of gender equality interventions, from mentoring instruments over structural change approaches up to incentives for integrating gender aspects into research and innovation projects. With this case study approach, we aim to validate and further develop the evaluation framework, a process of which the most recent results shall be shown and discussed at the Vienna Impact Conference.

Our approach of using a theory-based evaluation framework is appropriate even though it has hardly possible to measure concrete research and innovation outcomes and impacts of the GE programmes under consideration. One critique, however, can be that the theory of changes emphasizes differences between male and female researchers and might lead to the promotion of stereotypes. Furthermore, the work with log frames is rather linear and only partly suitable for complex environments, as we are fully aware.

The main and still unresolved problem is how to establish the link between the intervention and the research and innovation outcomes and impacts. Apart from subjective perceptions and anecdotal evidence, the interviewees could not contribute any confirmations.

The case studies underlined, however, the importance of the context yielding to the desired but also to some not desired effects. They also showed that the EFFORTI approach and the collection of indicators delivers a suitable background for programme evaluations.

# REFERENCES

**European Union** (2016). European Innovation Scoreboard 2016. Brussels.

**European Union** (Ed.). Research Innovation Observatory (https://rio.jrc.ec.europa.eu/en/stats/key-indicators).

**European Commission (Ed.)** (2015). Indicators for promoting and monitoring Responsible Research and Innovation. Brussels.

**Kalpazidou Schmidt, E. and Cacace, M.** (2017). Addressing gender inequality in science. the multifaceted challenge of assessing impact. Research Evaluation 2017, 1–13, doi. 10.1093/reseval/rvx003.

**Kalpazidou Schmidt, E., Bührer, S., Schraudner, M., Reidl, S., Müller, J., Palmen, R., Haase, S., Graversen, E. K., Holzinger, F., Striebing,**

**C., Groó, D., Klein, S., Rigler, D. and Høg Utoft, E.** (2017). A Conceptual Evaluation Framework for Promoting Gender Equality in Research and Innovation. Toolbox I - A synthesis report. EFFORTI - Deliverable 3.3.

**Nedopil, C., Schauber, C. and Glende, S.** (2013): The art and joy of user integration in AAL-projects. Brussels.

**Nutley, S., Huw, D. and Walter, I.** (2002). Evidence Based Policy and Practice: Cross Sector Lessons From the UK. Eds.. ESRC UK Centre for Evidence Based Policy and Practice. Working Paper 9. London.

**OECD** (2014). Science, Technology and Industry Outlook 2014, OECD Publishing. Paris.

**OECD** (2015). Science, Technology and Industry Scoreboard 2015. Innovation for growth and society, OECD Publishing. Paris.

**Ravn, T., Nielsen, M. W. and Mejlgaard, N.** (2015a). Synthesis report on existing indicators across RRI dimensions. Progress report D3.1. Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) (http.//www.technopolis-group.com/morri/).

**Ravn, T., Nielsen, M. W. and Mejlgaard, N.** (2015b). Metrics and indicators of Responsible Research and Innovation. Progress report D3.2. Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) (http.//www.technopolis-group.com/morri/).

**Rommes, E.** (2014): Feminist Interventions in the Design Process. In: Ernst, W. and Horwath, I. (Eds.): Gender in Science and Technology. Interdisciplinary Approaches. Transcript Verlag, Bielefeld. 41-55.

**Sanderson, I.** (2002). Evaluation, Policy Learning, and Evidence Based Policy Making, Public Administration 80(1), 1-22.

**Solesbury, W.** (2001). Evidence Based Policy: Whence it Came and Where it's Going. In: ESRC UK Centre for Evidence Based Policy and Practice (Ed.). Working Paper 1. London.

**Yin, R.** (1994). Case Study Research. Design and Methods, Second Edition, Sage Publications, London.

# AUTHORS

**SUSANNE BÜHRER**
*Fraunhofer ISI*
Breslauer Str. 48, 76139 Karlsruhe (Germany)
E: susanne.buehrer@isi.fraunhofer.de

**SYBILLE REIDL**
*Joanneum Research*
Sensengasse 1, 1090 Wien (Austria)
E: sybille.reidl@joanneum.at

**EVANTHIA K. SCHMIDT**
*Aarhus University*
Bartholins Allé 7, 8000 Aarhus C (Denmark)
E: eks@ps.au.dk

**RACHEL PALMEN**
*Universitat Oberta de Catalunya*
Av. Carl Friedrich Gauss 5, 08860 Castelldefels (Spain)
E: rpalmen@uoc.edu

**CLEMENS STRIEBING**
*Fraunhofer CeRRI*
Hardenbergstr. 20, 10623 Berlin (Germany)
E: clemens.striebing@iao.fraunhofer.de

**DORA GROO**
*Association of Hungarian Women in Science*
Napraforgo u 17, 1021 Budapest (Hungary)
E: dora.groo@nokatud.hu

# THE SHAPING A NEW UNDERSTANDING OF THE IMPACT OF HORIZON EUROPE: THE ROLES OF THE EUROPEAN COMMISSION AND MEMBER STATES

MATTHIAS WEBER, KATJA LAMPRECHT AND PETER BIEGELBAUER

DOI: 10.22163/fteval.2019.347

## ABSTRACT

This paper aims at tracing the process and the arguments that have led to the proposal of the 9th European framework programme for research and innovation 'Horizon Europe'. We are particularly interested in the role and influence of member states, especially in their interplay with the European Commission, on the actual shape of the FP9 proposal, next to important contributions to this still-ongoing debate, such as the interim evaluation of H2020, and various expert groups. The paper focuses on two of the novel elements of Horizon Europe, namely the concept of 'missions' as approach to help enhance the societal impact of 'Horizon Europe', and the proposal to establish a European Innovation Council.

## INTRODUCTION

Since the establishment of the European framework programmes for research and innovation (FPs) in the mid-1980s, they have been characterised by regular change in terms of underlying narratives and intervention instruments used (Biegelbauer and Weber 2018). Decision-making about a framework programme follows a formalised co-decision procedure involving European Commission, European Parliament and European Council (Pollak and Slominski 2006). However, the process leading to the formulation of the proposal for a framework programme is less clearly structured. It may benefit from political leadership, such as in the case of the current Horizon 2020 programme, which draws a lot of inspiration from the Lund declaration of the Swedish European Summit.[1] It also usually involves several elements of formal and informal consultations with Member States, stakeholders and experts. Overall, however, it is largely developed internally by the European Commission services. As formulated by one of the high-level expert groups looking into this issue, there is a great deal of strategic intelligence available to feed into the development of the framework programme, but the actual sense-making involved in the preparation of the specific proposal is opaque (EFFLA 2012).[2]

Given the importance of the FPs for R&I in Europe, both in financial and in symbolic terms, the governance question of how the FPs are actually "shaped" and by whom is of major importance for the legitimacy of the entire endeavour. This leads to the main research question of this paper: How has the process of preparing the initial proposal of a framework programme worked in the case of Horizon Europe?

Inspired by work on this matter, but also through involvement in the most recent process of this kind, we argue that the influence of external stakeholders, including the Member States, on the shaping of the framework programme proposal is rather limited, and that the internal processes, dynamics and conflicts within the European Commission are the key levers of change. As a consequence, if the ambition is to strengthen the legitimacy of the framework programme, providing more transparency to the internal processes and opening it up already in the preparatory phase would be more important than yet another external consultation or call for ideas.

In order to illustrate our argument, we will look at the two main changes that are most likely going to be introduced in the new Horizon Europe programme: missions and the European Innovation Council. In what remains, we will reconstruct the process of how these two new elements came into being. An emphasis is put on the role played by Member States in influencing the emergence and specification of these new concepts. The final section will draw some conclusions based on these tentative findings.

1 The Lund declaration was at the origin of the emphasis put in Horizon 2020 on addressing grand or societal challenges (see European Council 2009).

2 EFFLA was tasked to revisit the role of foresight in EU policy-making, and it stressed the importance of foresight for opening up the second phase of "sense-making" in particular in order to make it more transparent and hence strengthen the legitimacy of policy actions proposed by the EC. The subsequent phases of decision-making and implementation are much more formalised.

# CONCEPTUAL AND METHODOLOGICAL FRAMEWORK

In conceptual terms, the paper draws on a network perspective on governance, and how the internal governance networks within the European Commission interact with external stakeholders and networks in the course of the shaping of the framework programme (Torfing and Sørensen 2014). These interactions exhibit features of advocacy and interest representation. Member States are part of this game and use formal as well as informal channels to bring in their views on the future shape of the framework programme (Pernicka et al 2002). However, the actors are not driven by interests alone and the whole process features also elements of policy learning (Biegelbauer 2013) in the sense of learning from past experience with previous framework programs. These interactions are all taking place in a context of discourses that shape the policy field, and which rest on policy frames and narratives produced since the early days of the European unification process (Biegelbauer and Weber 2018). In the past such frames have recurrently stressed arguments of Europe being a "laggard in S&T" or the so-called "European Paradox".[3]

We therefore focus in the analysis of our two cases (missions and European Innovation Council) on the following aspects: next to tracing their patterns of emergence, we look into the underlying rationales and narratives, and also at the role of internal and external networks for shaping them. This is complemented by a final assessment of the two cases.

In methodological terms, this paper draws on a combination of document analysis, participation in some of the preparatory processes leading to the formulation of the Horizon Europe proposal, and a series of interviews with individuals involved in that process. The insights from the interviews are still preliminary, because the interview programme has not been finished yet. Moreover, as the preparation of Horizon Europe is still an ongoing process, the analysis presented in this paper must be regarded as preliminary, too, since relevant documents are published almost every month.

# THE PATTERNS OF EMERGENCE OF FP9

As foundation for the deeper analysis of the two examples of missions and EIC, it is instructive to briefly reconstruct the main phases of the shaping of Horizon Europe as a whole. Horizon 2020 started off with high ambitions regarding the role and contribution of European research and innovation funding to addressing the societal challenges that are at the core of the third pillar of Horizon 2020. However, it became soon clear that the envisaged impacts on societal challenges could at best arise in the long term and that the linkages between specific R&I projects and higher-order policy goals like addressing societal challenges were at best vague (see e.g. European Council 2015), also because governance practices did not really change compared to previous framework programmes. In addition, the overall research and innovation performance of Europe did not meet the ambitions initially formulated by the Innovation Union flagship initiative, as shown by corresponding output indicators. A growing pressure on EU (R&I) policy to demonstrate its value added to Member States and citizens raised the need to formulate convincing narratives in order to justify a rising, or at least stable, budgetary provision during the negotiation phase of the upcoming multi-annual financial framework (MFF). These observations may serve as backdrop for the four main phases of preparing the new framework programme.

## PHASE 1: EARLY THOUGHTS ABOUT A NEW FRAMEWORK PROGRAMME

First ideas about what a new framework programme should look like were launched already in the course of 2015/16 at the level of the research commissioner's cabinet. The RISE high-level expert group identified important issues to be considered for future R&I policy, such as i) difficulties of Europe in retaining fast-growing firms (partly due to a shortage of second-phase venture capital, partly as a result of an incomplete single market) (Soete et al. 2015), ii) the need for a more transformative and at the same time more open R&I policy (Andree et al. 2015), iii) the "double deficit" as compared to the US in terms of lagging behind not only in innovation, but also in key areas of science (Sachwald 2015), and iv) the growing divide between Member States in terms of innovation performance (Tsipouri 2017). At national level, first initiatives were taken as well to think ahead in terms of what the next framework programme should be about (e.g. the Austrian FP9 Think Tank). The "Lund revisited" conference (European Council 2015) and the Madelin Report further fueled the early debates about "Europe's mission to innovate" (Madelin and Ringrose 2016).

## PHASE 2: LEARNING FROM THE PAST, AND IDENTIFYING CHALLENGES FOR THE FUTURE

The culmination point of this second phase must be seen in the publication of the Lamy Report in mid-2017 (Lamy et al. 2017). Based on the interim-evaluation of Horizon 2020, the modelling of possible impacts of European research, and the foresight project BOHEMIA, the Lamy Report brought together eleven guiding principles for the next framework programme, among which also the introduction of a mission-oriented approach and support to the idea of a European Innovation Council were mentioned.

## PHASE 3: ELABORATING KEY ELEMENTS

Subsequent debates led to a further refinement of the ideas on the new framework programme, in particular to the missions and the European Innovation Council. Various EC expert groups, in particular RISE (Research, Innovation and Science Policy Experts High Level Group), ESIR (Expert Group on the Economic and Societal Impact of Research

3 The "laggard" argument has been with us since the first framework programmes that stressed the scientific backwardness of Europe in key technology areas, whereas the European paradox was first used in 1995 in the EC Green Paper on Innovation (EC 1995), which fed into the preparation of the 5th framework programme.

and Innovation) and the High-Level Group of Innovators, were involved in this process.[4] A policy paper by Mariana Mazzucato (Mazzucato 2018) and support studies on past experiences with missions (JIIP 2018) and future candidate themes for missions (Weber et al. 2018b) also fed into the debates.

### PHASE 4: FEEDBACK AND REFINEMENT

In the second half of 2018, Member States, European Parliament and stakeholders gave feedback and made suggestions for changes to the Commission proposal. An enhanced pilot of EIC is now foreseen to be launched in 2019, in order to pave the way for the new Horizon Europe concept of the EIC. However, there are still many open questions regarding the effectiveness and governance of the EIC, as well as regarding the enhanced role of the EC as financing agent. The identification, selection and governance of missions in between the different pillars and (within pillar) clusters of Horizon Europe are still unclear, in spite of launching an extensive EC-internal preparatory process during summer 2018. The process of refinement is likely to continue even after the legal decision on Horizon Europe will have been taken.

## THE SHAPING OF "MISSIONS"

### PATTERNS OF EMERGENCE

The notion of missions was not used explicitly by the Commission in the debate about the next framework programme until the publication of the Lamy Report in 2017. However, already before similar ideas had been raised, which were inspired by the gap between the high ambitions of the societal challenges pillar in Horizon 2020 and the largely supply- and S&T centric approach used to implement it (European Council 2015). Drawing on earlier work on demand-side innovation policy in the European Commission (e.g. in the Aho Report), Andree et al (2015), in a report of the RISE group advising the research commissioner, called for a demand-centric mission-oriented approach in future R&I policy in order to move beyond the technology-centric approach of Horizon 2020: *"While the move towards a challenge-driven approach in Horizon 2020 has been a good step forward, addressing now broader societal challenges, to have a real impact, such a programme would have to be truly "mission-oriented", fitting in as an integral part of larger policy objectives. To achieve this, R&I will have to be linked closer to the other EU policies, defining concrete missions in the realm of a broader EU energy policy, transport policy, environment policy, etc. In other words, what is lacking is coordination and synergies between supply and demand of R&I."* (p. 5).

The term 'mission' was explicitly used for the first time in a Commission report by Robert Madelin and David Ringrose (Madelin and Ringrose 2016), which was entitled "Opportunity now: Europe's mission to innovate". Here the notion of "mission" was used in a comprehensive sense, in order to promote a positive commitment to innovation in order *"to make society attentive to its future and resilient in face of crisis"* (p. 49).

After this first phase, the idea that a mission-oriented approach should be adopted in the EU framework programmes, however, was discussed in several circles, as reflected, for instance, in the first theses paper of the Austrian FP9 Think Tank (FP9 Think Tank 2016), which argues that *"contributing to the grand societal challenges of our times and bringing science closer to the people should be main objectives of FP9. With respect to the societal challenges element of the programme, a redesign is required to give full justice to the specificities of new mission-oriented programmes."* (p. 6). Other national papers outlining first ideas about the future framework programme, came up with similar suggestions.

In the follow-up to the Lamy report, policy papers by different expert groups (RISE 2018a; ESIR 2017) further contributed to the elaboration of a mission-oriented approach in FP9. With the Mazzucato paper (Mazzucato 2018), published in early 2018, the political legitimacy of missions was further consolidated. This was necessary because missions reassign a stronger role to the state in matters not only of research, but also of innovation and diffusion. The subsequent consultation on the mission-oriented approach lent a lot of support to the concept, but also raised quite some skepticism as to the governance capabilities and capacities of the European Commission to deliver on the high promises raised. This skepticism was underpinned by experiences from past mission-oriented initiatives, showing that missions require a highly developed governance and management system with strong leadership to succeed: Also a range of other critical success factors has to be taken into account, in particular when addressing 'transformative' and 'systemic' rather than just technological challenges (JIIP 2018). Moreover, the approach presented by Mazzucato (and largely adopted by the EC) seems to draw strongly on the rather technocratic experiences with "old" missions, which may well be suitable for technology-centric missions, but is less adapted to "new" missions geared towards societal challenges, which are more complex and wicked in nature.

Since the presentation of the Horizon Europe proposal, which was rather vague about the topics and the implementation modalities for missions, first steps have been made to develop a governance framework for missions. The debates about the identification and selection of priority themes for future missions to be addressed in Horizon Europe give evidence of the difficulties faced by the European Commission in bringing the missions concept to the ground, and which are reflected in a recent memorandum by the ESIR expert group (ESIR 2018). The challenges and uncertainties associated with the governance of missions in Horizon Europe are likely to be the reasons why for now only a comparative modest share of 10-15% of the budget are foreseen to be implemented under the umbrella of missions.

### RATIONALE: TURNING A VISION INTO PRACTICE

Although the strengthening of an orientation of European R&I towards societal challenges received a lot of support when Horizon 2020 was launched, it soon turned out that the gap between highly abstract challenges and the reality of specific projects was very wide. Even if explicit reference was made in project proposals to the relevance of the envisaged work for addressing societal challenges, the challenges often served only as umbrella to which lip service had to be paid. Neither was it possible to seriously assess or evaluate impacts of specific projects on

4 Between August 2017 and August 2018, the RISE group also launched a series of meetings with national think tanks ("Tour d'Europe") in order to discuss ideas for future European R&I policy with its peer expert groups in Member States (EC 2019).

the ability to address societal challenges, due to attribution problems resulting from the many other intervening factors and the long time horizon. What was missing was often an intermediate layer that would allow targeting ambitious but achievable goals with the help of a bundle or cluster of projects. Moreover, research and innovation activities at best promise an impact potential, but whether an actual impact will be achieved depends on demand-side conditions determining the uptake and diffusion of new solutions.

Overall, the mode of implementing Horizon 2020 did not change significantly as compared to earlier framework programmes. Guiding ideas or visions were translated in a systematic process of strategic programming and work programme development into individual topics to which consortia could apply, but limited room was given to non-conventional ideas.

In this light, missions represent an opportunity to introduce an intermediary level of orientation and guidance, in between the highly abstract societal challenges and the reality of specific projects, which should help overcome the fragmentation into a myriad of individual projects. Functionally, they describe a credible claim to make the change happen that is needed on the pathway towards successfully addressing societal challenges. A very important promise tied to this claim was the promise of impact. A clear goal, a clear timeline and a clear plan of how to bundle complementary projects into a package should contribute to achieving impact beyond the level of individual projects.

This technocratic vision of enhanced planning and implementation of research and innovation activities was meant to help overcome growing skepticism about the ability to achieve the ambitious goals tied to societal challenges under Horizon 2020. Such a convincing narrative was important to ensure support to an increase of the research budget within the multi-annual financial framework of the EU.

However, the appealing idea of missions also opened up Pandora's box. If taken seriously, a mission-oriented approach, in particular when applied to systemic and wicked challenges (e.g. circular economy, sustainable mobility, climate change) opens up many interfaces. These are primarily with policy areas and policy levels that would need to revisit the demand-side instruments and framework conditions key for the uptake of novel solutions: no uptake, no impact. This issue of who ultimately "owns" the missions has not yet been resolved. Seen from an impact perspective, sectoral policies "own" the issues to be addressed (e.g. secure and sustainable energy supply, sustainable mobility) and should therefore lead the definition of missions, but by restricting them to R&I missions the lead could be claimed to stay within R&I policy.

## EXTERNAL NETWORKS: SUPPORT FOR AN AMBITIOUS AND VAGUE GOVERNANCE APPROACH

The simple narrative behind missions was well received by the majority of external stakeholders. Several Member States had already before adopted similar ideas in their national policies, and adapted them to their specific conditions. The Challenge-Driven Innovation programme in Sweden may serve as an example. And most recently, the German government presented 12 missions as part of its revised High-Tech Strategy 2025 (BMBF 2018).

This is also reflected in the generally rather positive reactions to the mission-oriented approach as reflected in national position papers to the Horizon Europe proposal. The main points of criticism refer to the governance of the missions, and in particular to the question of respective influence of Member States and Commission on the definition, selection and subsequent implementation of missions, whereas other critical issues such as the coordination and alignment of R&I policy with sectoral policies received less attention.

Some other stakeholders, in particular in industry, remain more reserved about the mission-oriented approach. Not only do they fear the complexity of implementation, but also declining support to traditional key enabling technologies. Others, such as many RTOs, perceive missions as an opportunity to bring their inter-disciplinary competencies and their ability to manage large-scale projects involving different stakeholders to bear in the implementation of missions.

## INTERNAL NETWORKS: TRICKLING DOWN OF A POLITICAL IDEA

Internally to the European Commission, the mission-oriented approach gained support through a range of mechanisms. First of all, the societal challenges were largely supported as overarching frame, and further strengthened by the launch of the UN Sustainable Development Goals. As regards implementation, two different perspectives can be distinguished. On the one hand, the 'traditionalist stance' was in favour of the well-established approach to implementing framework programmes through thematic work programmes. On the other hand, the 'modernist stance' sought a revision of the implementation approach in order to truly deliver on the ambitions formulated with the societal challenges.

A second important concern was the autonomy of DG RTD in defining its policy agenda. The past years saw a transformation of DG RTD from a programme-implementing into a policy DG with a strong political agenda of its own. This political agenda was focused on matters like the European Research Area, but also the strengthening of the political and economic significance of R&I policy in general. Missions could be a means to give this significance higher visibility, but it implied tying missions to political goals that were largely defined in other policy areas. This tension was overcome by stressing the R&I-centric nature of the missions to be pursued, as an argument that the control over missions remains largely within DG RTD.

At the same time, the engagement with other Directorates General was intensified in the preparatory debates about possible themes for future missions. The Foresight Correspondents Network, for instance, brings together key strategic thinkers from the majority of DGs, and it was closely involved in the implementation of the EC's foresight project BOHEMIA that helped prepare the thematic orientation of Horizon Europe, and thus also of possible missions. In other words, the network served as a soft coordination and harmonization mechanism between R&I policy and various other EC policies already in the two years preceding the presentation of the Horizon Europe proposal.

## ASSESSMENT

The introduction of the mission-oriented approach in Horizon Europe is based on the widely shared recognition that the implementation model of Horizon 2020 is not sufficient to achieve the expected impacts on societal challenges that were promised at the outset of Horizon 2020. This view is also backed by many Member States.

While it is difficult to reconstruct precisely how and when the notion of 'missions' found its way into the Lamy Report, it was a concept under discussion in many different circles and tested in several Member States. In other words, the time was ripe for a new approach to implementing the framework programme; a necessity that was not particularly controversial in its general line of reasoning.

The situation is more complicated when it comes to the details of the mission-oriented approach. The overall appeal of the mission concept has led to an under-estimation of the governance challenges that a mission-oriented approach involves, from the selection of missions and the establishment of carrier organisations, to the coordination needs with demand-side sectoral policy and to coherent implementation of supply and demand side policies. The potential organizational interests within the Commission seem to have had an influence on the shaping of the more detailed specifications of the mission concept, but this process is still not finalised.

# THE SHAPING OF THE EUROPEAN INNOVATION COUNCIL

## PATTERNS OF EMERGENCE

The idea to establish a European Innovation Council (EIC) was announced for the first time by Carlos Moedas, Commissioner for Research, Science and Innovation, in his speech on 'Open Innovation, Open Science, Open to the World' in June 2015.[5] The concept of the EIC follows the perception of the European Commission of an ongoing deficit of the European innovation system, its innovation capacity to commercialize European high quality research and its ability to scale up innovative business, in particular in comparison with US ("European paradox").

Against this backdrop, the European Commission ran an open call for ideas in spring 2016 to develop further discussions, accompanied by numerous published stakeholder position papers (EC 2016a, EC 2016b). In that period, the spectrum of ideas varied from bundling innovation supporting instruments for reducing complexity ('one-stop-shop'), to allocating financial support for up-scaling or to concentrating on providing strategic intelligence and helping to reduce regulatory barriers in cooperation with other sectoral DGs.

Subsequently, the 'High Level Group on Maximising the Impact of EU Research and Innovation Programmes' (Lamy et al, 2017) and particularly the 'High Level Group of Innovators' (HLG Innovators, 2018) stressed the need to support and invest in European high-risk, market-creating breakthrough innovations, particularly in 'deep tech' innovation (relying on science and engineering advances) and to overcome hindering factors in Europe. Examples are missing large investments over a significant period (venture capital is too small, fragmented, short term, with lack of critical mass, bank lending is inherently risk-adverse, policy funding perceived too complex), national and local initiatives too small to compete on global level, an incomplete single market and regulatory barriers. The European Innovation Council was recommended to play a central role in implementing this focus and in providing a more simplified support scheme with bottom-up and multi-state approach of funding (grant-based at early stage for technology development and understanding pathways to commercialize, combination of grants and financial instruments when larger investment is needed). It should encourage collaboration and networking between innovators, firms, investors, etc. to stimulate scaling-up on EU-level, stimulate collaboration with national and regional agencies and help innovators overcome regulatory barriers.

With the1st phase of an EIC pilot (launched in October 2017) as part of the Horizon 2020 Work Programme 2018-2020, the European Commission bundled existing funding instruments: SME Instrument, Fast Track to Innovation, Future and Emerging Technologies (FET) Open and the EIC Horizon Prizes accompanied by opportunities for networking, mentoring and coaching (EC 2018a).

The third pillar ("Open Innovation") of the proposed Horizon Europe programme (EC 2018b) basically follows this approach with the idea to provide financial support along a linear innovation cycle and to overcome the growing lack of equity funding for risk-prone companies dealing especially with deep-tech products. In addition, the envisaged InvestEU Programme is meant to mobilise further public and private investment by a factor of about 14 (EC 2018d). Concerning the EU added value it is argued that the only possibility to provide large-scale venture capital is to act on the EU-level, with more effectiveness and comprehensiveness (e.g. common regulation, synergies with other EU programmes) and with increasing coherence of the overall innovation ecosystem.

## THE RE-DISCOVERY OF A FRAME

The main rationale used to underpin the call for a European Innovation Council is rooted in the "rediscovery" of the European paradox, a notion that was first coined in the mid-1990s, when the European Commission in its Green Paper on Innovation argued that 'one of Europe's major weaknesses lies in its inferiority in terms of transforming the results of technological research and skills into innovations and competitive advantages' (EC 1995, p. 5). The paradox, then, was suggesting that Europe was performing comparatively well in research, but was not successful in exploiting that potential economically.

A decade later, the existence of this paradox was increasingly questioned. Dosi et al (2006), for instance, argued that this paradox does not exist because Europe is behind the US also in scientific terms, for instance when looking at publication output per capita of population or of research personnel. Sachwald (2015), in a paper for the RISE group advising the European research commissioner, confirms this skepticism and speaks of a "double deficit", because although Europe produces more scientific publications than the US, these are less cited and less relevant to innovation. But also sectoral differences matter, because the US have their strongest scientific base in ICT, health and medicine, i.e. in areas where the mode of science-based innovation is particularly

5 'Europe does not yet have a world class scheme to support the very best innovations in the way that the European Research Council is the global reference for supporting excellent science. So I would like us to take stock of the various schemes to support innovation and SMEs under Horizon 2020, to look at best practice internationally, and to design a new European Innovation Council' (Moedas 2015)

pronounced. Other reasons for the comparatively poor innovation performance are seen in less developed entrepreneurship and start-up cultures in Europe (Henrekson and Sanandaji 2017), and in the limited capacities of many European firms to absorb new scientific knowledge (Czarnitzki et al., 2009). Also still remaining barriers to a truly single market in Europe hamper the incentives for firms to innovate.

However, in spite of these insights, the European paradox was adopted as the guiding narrative underpinning the call for the creation of a European Innovation Council. In the course of the publication of the proposal for Horizon Europe and its impact assessment (EC 2018c), the Commission calls for action on the EU-level as future breakthrough innovation will be science-based[6].

## EXTERNAL AND INTERNAL NETWORKS: HANDLING SKEPTICISM

The round of consultation launched in 2016 raised support for the intention, but also criticism of the concept of the proposed EIC. The position papers of Member States, as well as the joint position of European Research Area and Innovation Committee (ERAC 2016) and reflection papers of the RISE group are interesting in this regard (RISE 2017).

Member States came up with a diverse range of proposals regarding the focusing of the EIC, reflecting on the 'call for ideas' during 2016 and in preparation for the interim evaluation of Horizon 2020 in the beginning of 2017. Their statements, as synthesised by Weber et al (2018a), suggests models ranging from i) EIC as supporting instrument for start-ups with high potential to scale-up on European and global level with entrepreneurs as the main beneficiaries, ii) EIC as supporting instrument for 'excellence in innovation' model (partly described in BMBF 2016) for a wider target group, iii) EIC as driver for the integration of existing instruments enabling synergies up to iv) EIC as key towards an integrated research and innovation policy through coordinating and thus enhancing policy coherence between research policy, innovation policy and sectoral policy fields. In other words, this latter model aims at taking into account policies and framework conditions innovation on the demand side of innovation (FP9 Think Tank 2017), thus stretching out to both European and national policy levels (IPM 2017).

Furthermore, the RISE expert group stressed that a new narrative 'From Innovation to Innovators' shall be one of the guiding principles of the EIC, aiming to align innovation policy in Europe with the characteristics of emerging models of innovation. Moreover, the EIC was meant to become a one-stop shop for innovators of any nature, be they driven by technology, new business modes, new design, customer experience, or organizational development (RISE 2017).

Criticism first of all addressed aspects concerning potential duplication of national funding initiatives for SMEs and therefore an unclear division of labour with national and regional policies. Other points of critique referred to the limited European added value because the EIC addresses individuals or individual firms rather than collaborative innovation activities across borders. A risk to overlap with activities of the European Institute for Innovation and Technology (EIT) was also criticised, as was the exclusion of universities and research organizations as potential sources of disruptive innovations with major scale-up potential. Finally, the ability of a public institution like the EC to identify excellent innovations/innovators with a market-creating potential and to manage risk capital and entrepreneurship-centric initiatives was questioned (RISE 2018b).

In phase 3 of the elaboration of the next framework programme, the recommendation of the High Level Group of Innovators (HLG Innovators 2018) and the proposal of the Commission for Horizon Europe were published, the 1st phase of the EIC pilot had already started, the enhanced EIC pilot was not launched yet. In this period, the national position papers mainly followed the HLG of Innovators and its idea of supporting market creating 'deep tech' breakthrough innovations, by bringing together existing instruments under and EIC umbrella, bridging the 'valley of death' and combining funding and financial instruments to prepare innovators for large-scale private investment. Nevertheless, some items still remain unclear and are viewed with skepticism:

- **Narrowing down of Scope:** In contrast to the wide range of elements foreseen in the initial debates about the EIC, the proposed mission of the EIC has been narrowed down to science and technology-based market creating breakthrough innovation ('deep-tech') and on supporting entrepreneurs with potential to scale-up on European and global level. The coordination with Member State policies in order to complement national innovation initiatives without duplicating or even thwarting them, is essential. However, in referring to the proposed focus of the EIC (i.e. science and technology-based market creating breakthrough innovation), the RISE Group recommended in its recent paper (RISE 2018b) to carefully distinguish between 'deep-tech' and 'architectural' disruptive innovation. It further suggested the concept of 'Innovator Readiness Levels' instead of 'Technology Readiness Levels', in order to avoid the traditional linear mode of science-technology-market development.
- **Complementarity with other segments of Horizon Europe:** The proposed instruments of the EIC are based on a 'bottom up' approach and thus supporting innovators and innovations emerging within or at the crossroad of different sectors and disciplines. Interconnections and synergies with other pillars of Horizon Europe, in particular with the mission areas, and with the European Institute of Technology (EIT), while avoiding the creation of overlaps, have not yet been deepened in the discussions so far.
- **Governance:** The EIC portfolio is proposed to be managed following the ARPA-E approach (EC 2018c). The detailed concept and the requirements for dedicated programme managers and expert panels will be crucial. The participation of Member States for the implementation of the EIC and coordination with national agencies (co-funding partnerships are proposed by the Commission) will also be important. However, both aspects have not yet been developed in detail.
- **Appropriateness of budget allocation within the "Open Innovation" pillar:** A budget of EUR 10.5 billion for the European

6 'The EU innovation ecosystem generates as many start-ups as the US in number but only a few of them grow-up rapidly. This is even truer for start-ups carrying out breakthrough innovation and for the science-based ones ("deep tech"). The fact that the next wave of breakthrough innovation will be science-based calls for immediate action.' (EC 2018c).

Innovation Council is foreseen, including up to EUR 0.5 billion for European Innovation Ecosystems to 'boost the effectiveness of the European innovation system' (EC 2018b, p5 of Annexes). In view of the importance of coordination with several actors in Member States and the need to strengthen the single market and overcome the European fragmentation (RISE 2018b), the amount of EUR 0.5 billion seems rather low.

### ASSESSMENT

EU Member States have raised several concerns regarding the European Innovation Council, and as a result, the initial approach was considerably narrowed down to address a much more specific group of innovators than initially envisaged. Many other points of criticism have not been addressed, such as the extension of the single-beneficiary approach of the EIC, or the extensive role that the EC would assume in handling a variety of financial instruments which are well beyond the scope of funding instruments the EC is used to handle.

Possibly the most fundamental argument questioning the EIC concept as a whole concerns the main barriers to realizing market-creating breakthrough innovations in Europe, which some experts and Member States see in the remaining deficits of the European single market and the regulatory rigidities residing in sectoral policies, rather than in funding and advice to innovators. In other words, it is the wider ecosystem that hampers the success of market-creating breakthrough innovations. Interestingly, the ecosystem-oriented element of the EIC pillar of Horizon Europe is by far the smallest component in financial terms.

However, the strong support to the EIC concept from the Commissioner and his cabinet, backed largely by the high-level group installed, demonstrates that the EIC is a good example of rather limited influence of external voices, including those of the Member States, on the shaping of a key element of Horizon Europe.

## CONCLUSIONS

To come back to our research question and hypothesis regarding the role and influence of Member States on the shaping of Horizon Europe, the two examples show a more differentiated picture than suggested by our initial hypothesis.

The EIC is an element that has been driven top-down from the Cabinet. Some suggestions from Member States were taken up (e.g. regarding the creation of an umbrella approach rather than a strong institution, or the narrowing of the scope of the EIC), while other major concerns and criticisms that could have questioned the EIC in its entirety were left aside (e.g. regarding the role of the entire ecosystem for market-creating innovations).

The mission-oriented approach, while still being controversially discussed with regard to its governance and the selection of priorities, was generally received positively by Member States and several other stakeholders. The rationale behind missions is largely shared, but major controversies arise at the level of national interests in potential themes and the role of Member States in the governance of subsequent implementation. However, it is still too early to give an assessment of the influence of Member States on the final shape of governance modalities and priority-setting.

## REFERENCES


**Andree, D., Weber, M., Llerena, P.** (2015). A new role for EU Research and Innovation in the benefit of citizens: Towards an open and transformative R&I policy, Policy Paper by the Research, Innovation, and Science Policy Experts (RISE), European Commission, Brussels.

**Biegelbauer, P.** (2013). Wie lernt die Politik - Lernen aus Erfahrung in Politik und Verwaltung. Wiesbaden, VS Verlag für Sozialwissenschaften.

**Biegelbauer, P. and Weber, M.** (2018). EU research, technological development and innovation policy. Handbook of European Policies: Interpretive Approaches to the EU. Hubert Heinelt and Sybille Münch (eds). Cheltenham, Gloucestershire, UK, Edward Elgar: 241-259.

**BMBF** (2016). European Innovation Council (EIC) - Position paper of the Federal Ministry of Education and Research (BMBF); Retrieved October 2, 2018 from: https://www.bmbf.de/files/EIC_position_paper_eng.pdf

**BMBF** (2018). Forschung und Innovation für die Menschen. Die Hightech-Strategie 2025, Bundeministerium für Bildung und Forschung, Berlin.

**Czarnitzki, D. et al.** (2009). The Nexus Between Science and Industry; Evidence from Faculty Inventions. Leuven: Department of Managerial Economics, Strategy and Innovation Working Paper.

**Dosi, G. et al.** (2006). The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called 'European Paradox', Research Policy, 35(10), 1450-1464.

**EC** (1995). Green Paper on Innovation, European Commission, Brussels.

**EC** (2016a). Ideas for a European Innovation Council – Overview of Responses to the Call for Ideas, European Commission, Brussels, Retrieved October 2, 2018 from: https://ec.europa.eu/research/eic/pdf/eic_call_for_ideas-overview.pdf

**EC** (2016b). Ideas for a European Innovation Council – Summary of a Validation Workshop with Stakeholders held on 13 July 2016. European Commission, Brussels, Retrieved October 2, 2018 from: https://ec.europa.eu/research/eic/pdf/workshop/eic_workshop_summary_report_final.pdf

**EC** (2018a). H2020 Work-Programme 2018-2020, Towards the next Framework Programme for Research and Innovation: European Innovation Council (EIC) pilot (revised version), European Commission, Brussels, Retrieved October 2, 2018 from: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-eic_en.pdf

**EC** (2018b). Proposal for a Regulation of the European Parliament and of the Council establishing Horizon Europe – the Framework Programme for Research and Innovation. COM(2018) 435 final, European Commission, Brussels, Retrieved October 2, 2018 from: https://ec.europa.eu/commission/publications/research-and-innovation-including-horizon-europe-iter-and-euratom-legal-texts-and-factsheets_en

**EC** (2018c). Impact Assessment Accompanying the document establishing Horizon Europe, SWD(2018)307 final, European Commission, Brussels, Retrieved October 2, 2018 from: https://ec.europa.eu/commission/publications/research-and-innovation-including-horizon-europe-iter-and-euratom-legal-texts-and-factsheets_en

**EC** (2018d). Factsheets on InvestEU, European Commission, Brussels, Retrieved October 2, 2018 from: https://ec.europa.eu/commission/publications/investeu-programme_en

**EC** (2019). 101 Ideas on the future of Research and Innovation in Europe, European Commission, Brussels, Retrieved March 6, 2019 from: https://ec.europa.eu/info/publications/101-ideas-future-research-and-innovation-europe_en

**EFFLA** (2012). How to design a European foresight process that contributes to a European challenge driven R&I strategy process, Policy Brief No. 2, European Forum on Forward-Looking Activities, European Commission, Brussels.

**ERAC** (2016). ERAC Opinion on the idea of a European Innovation Council. ERAC 1209/16. Brussels, 14 July 2016.

**ESIR** (2017). Towards a mission-oriented research and innovation policy in the European Union. An ESIR memorandum, Report of the ESIR expert group, European Commission, Brussels, December 2017.

**ESIR** (2018). ESIR memorandum II. Implementing EU missions, Report of the ESIR expert group, European Commission, Brussels, October 2018.

**European Council** (2009). The Lund Declaration. Europe must focus on the Grand Challenges of our time, European Council, Brussels.

**European Council** (2015). Background paper to the Lund declaration 2015, European Council, Brussels.

**FP9 Think Tank** (2016). Theses Paper for shaping the next (9th) EU-RTD Framework Programme, Retrieved October 2, 2018 from: https://era.gv.at/object/document/2826/attach/FP9_Theses_Paper_endg_.pdf

**FP9 Think Tank** (2017). Fostering Impact and Sustainable Collaboration in FP9 within a new Common Research, Technology and Innovation Policy, Retrieved October 2, 2018 from: https://www.era.gv.at/object/document/3117/attach/FP9_Theses_Paper_2_final.pdf

**Henrekson, M. and Sanandaji, T.** (2017), Schumpeterian Entrepreneurship in Europe Compared to Other Industrialized Regions. Stockholm: IFN Working Paper No. 1170.

**HLG Innovators** (2018). Funding - Awareness - Scale - Talent (FAST) - Europe is back: Accelerating breakthrough innovation - Full set of recommendations from the Independent High-Level Group of Innovators on establishing a European Innovation Council. Retrieved October 2, 2018 from: https://ec.europa.eu/info/sites/info/files/eic_hlg_bz_web.pdf

**IPM** (2017). Draft Recommendations on the European Innovation Council, High Level Group on Innovation Policy Management, Brussels June 2017.

**JIIP** (2018). Mission-Oriented Research and Innovation: Inventory and characterisation of initiatives. Final report, Joint Institute for Innovation Research, European Commission, Brussels.

**Lamy, P** (2017). LAB – FAB – APP Investing in the European future we want. Report of the independent High Level Group on maximising the impact of EU Research & Innovation Programmes, Retrieved October 2, 2018 from: https://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/hlg_2017_report.pdf#view=fit&pagemode=none

**Madelin, R. and Ringrose, D** (eds.) (2016). Opportunity now: Europe's mission to innovate, European Commission, Brussels.

**Mazzucato, M.** (2018). Mission-oriented research & innovation in the European Union: A problem-solving approach to fuel innovation-led growth, European Commission, Brussels.

**Moedas, C.** (2015). Open Innovation, Open Science, Open to the World. Speech at 'A new start for Europe: Opening up to an ERA of Innovation' Conference, Brussels 22 June 2015.

**Pernicka, S., Feigl-Heihs, M. Gerstl, A., and Biegelbauer, P.** (2002). Wie demokratisch ist die europäische Forschungs- und Technologiepolitik? Der politische Entscheidungsprozess zum fünften Forschungsrahmenprogramm aus österreichischer Perspektive. Baden-Baden, Nomos.

**Pollak, J. and Slominski, P.** (2006). Das Politische System der EU. Wien, WUV (UTB).

**RISE** (2017). Europe's Future: Open Innovation, Open Science, Open to the World. Reflections of the Research, Innovation and Science Policy Experts High Level Group.

**RISE** (2018a). Mission-Oriented Research and Innovation Policy. A RISE Perspective, European Commission, Brussels.

**RISE** (2018b). The academic underpinnings of the European Innovation Council. Summary and recommendations of the RISE subgroup on EIC, European Commission, Brussels.

**Sachwald, F.** (2015). Europe's twin deficits: Excellence and innovation in new sectors, Policy Paper by the Research, Innovation, and Science Policy Experts (RISE), European Commission, Brussels.

**Soete, L., Veugeleurs, R., Sundgren, J.-E., and Glover, A.** (2015): R&I in times of crisis: dynamic interactions between R&I, public finances and economic growth, RISE presentation, 01/2015.

**Torfing, J. and Sørensen, E.** (2014). "The European debate on governance networks: Towards a new and viable paradigm?" Policy and Society 33(4): 329-344.

**Tsipouri, L.** (2017). Pockets of excellence as drivers of regional growth. Policy paper of the RISE high-level workshop on "The impact of smart specialisation strategies on pockets of excellence and regional growth" (Crete, 6-8 October 2015), European Commission, Brussels.

**Weber M., Rammer Ch., Dinges M., Dachs B., Hud M., and Steindl C.** (2018a). Erkenntnis- und Wissenstransfer im Kontext europäischer F&I-Politik. Studien zum deutschen Innovationssystem Nr. 12-2018.

**Weber, M., Andreescu, L., Cuhls, K., Dragomir, B., Gheorghiu, R., Giesecke, S., Ricci, A., Rosa, A., Schaper-Rinkel, P., Sessa, C., Baiocco, S., Curaj, A., Giuffrè, G., Zern, R.** (2018b): Transitions on the Horizon: Perspectives for the European Union's future research and innovation policies, Final Report of Project BOHEMIA (Beyond the Horizon. Foresight in Support of the EU's Future Policies on Research and Innovation), European Commission, Brussels.

*All position papers, that have been considered in the analyses, were retrieved on October 2, 2018 from https://era.gv.at/directory/238 and https://era.gv.at/directory/297*

# AUTHORS

**MATTHIAS WEBER**
*Center for Innovation Systems and Policy, Austrian Institute of Technology*
Giefinggasse 4, Vienna, 1210 (Austria)
E: matthias.weber@ait.ac.at (corresponding author)

**KATJA LAMPRECHT**
*Center for Innovation Systems and Policy, Austrian Institute of Technology*
Giefinggasse 4, Vienna, 1210 (Austria)
E: katja.lamprecht@ait.ac.at

**PETER BIEGELBAUER**
*Center for Innovation Systems and Policy, Austrian Institute of Technology*
Giefinggasse 4, Vienna, 1210 (Austria)
E: peter.biegelbauer@ait.ac.at

# POSTERS

## POSTER 1 / PAGE 156

**How to achieve, assess and advance impact and societal value of health research from a funders' perspective**

**Authors:** Gijbels, J. and Reijmerink, W.

**Contact:** ZonMw, The Netherlands Organisation for Health Research and Development (www.zonmw.eu), Joey Gijbels: gijbels@zonmw.nl, Wendy Reijmerink: reijmerink@zonmw.nl, DOI: 10.22163/fteval.2019.399

## POSTER 2 / PAGE 158

**Improving access to finance for young innovative firms with growth potential: evidence of impact of equity instruments on firms' output**

**Authors:** Stamenov, B. and Szkuta, K.,

**Contact:** European Commission DG JRC, Blagoy Stamenov: blagoy.stamenov@ec.europa.eu, Katarzyna Szkuta: katarzyna.szkuta@ec.europa.eu DOI: 10.22163/fteval.2019.400

## POSTER 3 / PAGE 160

**Understanding the policy pathway to research impact. Commonwealth Fisheries Harvest Strategy Policy – a pilot case study.**

**Authors:** Russell, T., Gash, A., Söderbaum, J., Keenan, T. & Dowd, A-M.,

**Contact:** Dr Anne-Maree Dowd, Executive Manager, Planning, Performance & Evaluation, anne-maree.dowd@csiro.au, www.csiro.au/en/about/our-impact, DOI: 10.22163/fteval.2019.401

## POSTER 4 / PAGE 162

**Improving access to finance for innovative firms with growth potential: evidence of impact of R&D grant schemes on firms' output**

**Authors:** Testa, G. and Szkuta, K.,

**Contact:** European Commission DG JRC, Giuseppina Testa: giuseppina.testa@ec.europa.eu, Katarzyna Szkuta: katarzyna.szkuta@ec.europa.eu, DOI: 10.22163/fteval.2019.402

## POSTER 5 / PAGE 164

**Preparing the future: developing research impact assessment capacity among stakeholders in Catalonia**

**Authors:** Adam Paula, Daban Montserrat, Solans-Domènech Maite, Radó-Trilla Núria, Arboix Esteve, Borràs Cristina, Cabezas Roger, Lasauca Xavier, Mas Jordi, Morales Alba, Nason Eddy, Puigdomènech Cantó Alba, Dedeu Toni, Barberà Albert, Subirada Francesco

**Contact:** RIACAT Group, Paula Adam: padam@gencat.ca, DOI: 10.22163/fteval.2019.403

## POSTER 6 / PAGE 166

**Evaluation of socio-economic impacts of the business R&D support in small economies. The case of the Czech Republic**

**Authors:** Tomas Ratinger, Vladislav Cadil, Sylvester Amoako Agyemang

**Contact:** Technology Centre CAS, Tomas Ratinger: ratinger@tc.cz, Vladislav Cadil: cadil@tc.cz, Sylvester Amoako Agyemang: amoakoagyemangs@gmail.com, DOI: 10.22163/fteval.2019.404

## POSTER 7 / PAGE 168

**Working in partnership to make innovation, entrepreneurship and business growth policy more impactful through experimentation and new evidence.**

**Authors:** Teo Firpo, James Phipps

**Contact:** Innovation Growth Lab (IGL), Nesta, 58 Victoria Embankment, London EC4Y 0DS (UK), innovationgrowthlab@nesta.org.uk www.innovationgrowthlab.org, teo.firpo@nesta.org.uk; james.phipps@nesta.org.uk, DOI: 10.22163/fteval.2019.405

## POSTER 8 / PAGE 170

**Preliminary remarks about science, technology, and innovation policy evaluation in latin america**

**Authors:** Adriana Bin, Rafaela M. de Andrade, Lissa Vasconcellos Pinheiro, Sergio Luiz Monteiro Salles-Filho

**Contact:** Rafaela M. de Andrade, Doctoral Candidate, Department of Science and Technology Policy, University of Campinas, rafamarcelly@gmail.com, DOI: 10.22163/fteval.2019.406

## POSTER 9 / PAGE 172

**The impact of the EXIST Business Start-up Grant on corporate growth: A group comparison for Dresden (GER)**

**Author:** Philip Dörr

**Contact:** Philip Dörr, FSU Jena, Chair of Economics/Microeconomics Carl-Zeiss-Straße 3, 07743 Jena, Germany, philip.doerr@uni-jena.de, DOI: 10.22163/fteval.2019.407

## POSTER 10 / PAGE 174

**What are the impacts of cluster policies? The german expirience**

**Authors:** Michael Rothgang (RWI), Bernhard Lageman, Anne-Marie Scholz (ISG)

**Contact:** Michael Rothgang, RWI-Leibniz-Institute for Economic Research, Hohenzollernstraße 1–3, 45128 Essen, Germany, michael.rothgang@rwi-essen.de, DOI: 10.22163/fteval.2019.408

## POSTER 11 / PAGE 176

**Making the invisible visible. Towards strategic measures of research and technology organisations (RTOs)**

**Authors:** Kirsi Hyytinen, Katri Kallio, Olli Kuusisto & Arho Suominen

**Contact:** VTT Technical Research Centre of Finland Ltd, Kirsi HyytinenTel. +358 40 5818495 Kirsi.Hyytinen@vtt.fi, DOI: 10.22163/fteval.2019.409

# How to achieve, assess and advance impact and societal value of health research from a funders' perspective

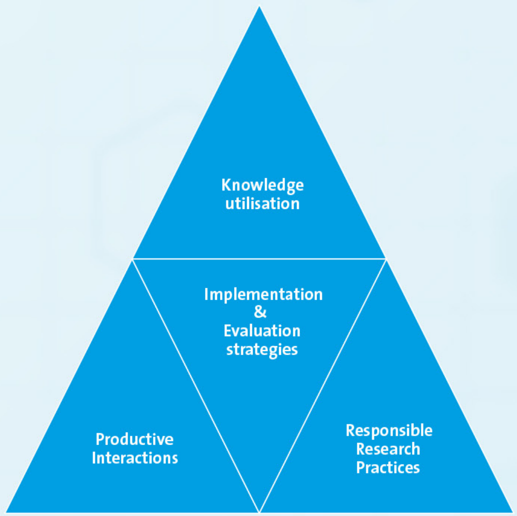
**Keywords:** impact assessment, societal impact, responsible research practices, productive interactions, funder, research programming

**Authors:** Joey Gijbels, gijbels@zonmw.nl & Wendy Reijmerink, reijmerink@zonmw.nl

ZonMw, The Netherlands Organisation for Health Research and Development (www.zonmw.eu)

## ZonMw's view on strengthening impact

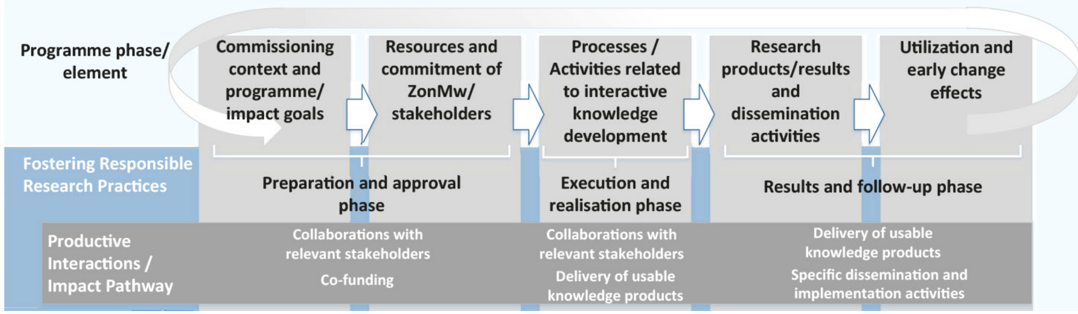
Responsible research programming generates valuable evidence which is optimally utilised in policy, practice and education by the stimulation of productive interactions




**Main elements of the ZonMw Impact Assessment Framework**

- Programming cycle
- Fostering Responsible Research Practices (fRRP)
- Productive Interactions
- Impact Pathway
- Basic information sources (programme bureau/project leader)

## ZonMw Impact Assessment Framework



| Cluster level | | | | |
|---|---|---|---|---|
| Programme level | Programme text<br>Grant/Funding call<br>Baseline measurement | Interim evaluation | Final evaluation<br>Study into societal knowledge use |
| Project level | Application form<br>Award letter | Progress report | Final report |

**ZonMw's definition of impact**

Impact defined in terms of demonstrable utilisation of health research instead of 'health and wealth of the nation'

## ZonMw Framework Fostering Responsible Research Practices: criteria and indicators for planning, monitoring and evaluation

| Societal relevance | Scientific quality | Integrity | Efficiency |
|---|---|---|---|
| Stakeholder participation* | Mixed methods designs | Transparancy (e.g. registration of research, open access, FAIR data) | Use of existing data/eResearch/ citizen science |
| Co-financing* | Practice-oriented research | Replication (research) | Stimulation of systematic reviews/ knowledge syntheses |
| Divers composition of steering committees | Pioneering/Innovative research | Prevention of publication bias (e.g. reporting guidelines) | Appropriate designs/ alternatives for RCT's |
| Holistic health concepts (e.g. positive health) | Interdisciplinary and international cooperation and knowledge development | Education and quality assurance | Handling of (potential) inclusion/ implementation problems |
| Participative knowledge infrastructure | Diversity of assessment process | Conflicting positions/ interests | Efficient arrangement of own programming processes |
| Added value of knowledge in policy, practice and education* | Variety of (transfer of) output* | | |

*Productive Interaction

**Productive Interactions** (http://www.siampi.eu/)

- Collaborations with relevant stakeholders
- Co-funding
- Delivery of usable knowledge products
- Specific dissemination and implementation activities

**Impact Pathway**

What do you want to achieve, why and when, how and with whom?

European Commission

# Improving access to finance for young innovative firms with growth potential: evidence of impact of equity instruments on firms' output

Stamenov, B. and Szkuta, K., European Commission DG JRC

## 1. Motivation / Objectives

High-growth enterprises are increasingly a target for government interventions, especially in Europe which lags behind the US in the number of fast growing innovative firms (so-called scale-ups). In response to this large "scale-up gap", the current policy debate has focused on new sources and forms of funding to enhance national and EU level support for scale-ups.

**Objective:** This paper examines the impact of public equity instruments on young innovative companies with growth potential. It focuses on the **output additionality**, i.e. the effects of equity instruments on the output of firms (measured by growth in terms of employment, turnover and innovative activities).

## 2. Main Research Questions

➢ What are the effects of public equity instruments on firms' output as measured by employment growth, turnover growth and innovation activities?

➢ What are major lessons learnt (in terms of design and implementation) from the evaluations of this type of policy measure?

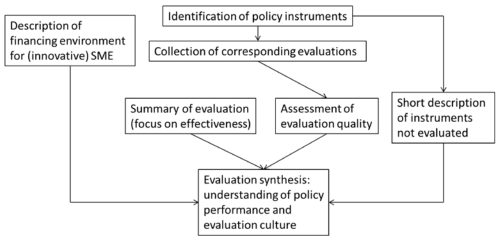
- The firm **profitability** declines in the beginning, with progressive growth after 3-4 years.
- The effects are **highly concentrated**, with the top 5-15% of supported firms absorbing the vast majority of generated returns.
- The majority of the national and regional equity measures achieve medium to high private investment **leverage levels** (2 to 3), with minimal or zero **crowding out** effects.
- Equity does not necessarily impact companies' **innovation performance** measured in terms of R&D spending. However, venture capital acts as a facilitator for the commercialisation of already developed innovations mainly through provision of finance, but also through partner networks and opening up of different channels of commercialisation.
- Most of the available evidence exists for **early stage investment** where public support is concentrated.

Output Additionality Summary Table

| **Effects on:** | **Equity instruments** |
|---|---|
| Employment | ●●● |
| Turnover | ●●● |
| Innovation | ●❍❍ |

●●● = major relevance ●●❍ = moderate relevance ●❍❍ = minor relevance

## 3. Methodology




Adapted from Edler, J., Ebserberger, B. and Lo, V. (2008). Improving Policy Understanding by means of Secondary Evaluation. *R&D Evaluation* 17(3)

Evidence sources

| All sources | Evaluation Reports | Academic articles |
|---|---|---|
| 31 | 19 | 12 |

**Types of equity instruments:**

- Public venture capital funds directly investing in companies (direct VC)
- Public funds invested in private VC funds (indirect or hybrid VC)
- Equity guarantees or government-backed loans to finance VC

**Countries covered:** Austria, Belgium, Denmark, Ireland, the Netherlands, Finland, Germany, the United Kingdom

## 4. Results

The main results arising from our review of the evaluation studies on the effectiveness of public equity instruments are:

- High-growth potential firms, which receive public funding in the form of equity, experience stronger increases in employment and turnover compared to the control groups.
- The employment growth rate oscillates from 50% to 145% and turnover from 125 to 800% (post-treatment), both are significantly higher compared to the untreated groups.

## 5. Relevance for Policy

The gathered evidence provides also more general lessons on the design and implementation of financial support for innovative ventures with growth potential.

- In favour of **syndication** of funds with a **leading role of the private sector** (as opposed to government funds only).
- Provision of **long-term investments** encompassing longer time spans and/or longer perspectives with less pressure on exits to achieve better outcomes in terms of employment and turnover growth.
- Intervening **both at the early and growth stages** while making the intervention more flexible at the margins (i.e. allowing for some investments just under/over the equity gap).
- Delivering **added value services** (e.g. networking and coaching) to the companies and ensuring that the fund is assisted by skilled professionals.
- **Flexible geographical boundaries** together with **larger size funds** to enable a diverse portfolio of investments and sufficient funding for follow up rounds of financing.
- **Extending the indicators of success**, beyond leverage effect, exits and fund profitability towards a closer examination of the effects on employment and turnover as well as the assessment of larger socio-economic benefits.
- The outcomes are heavily dependent on the number and quality of companies available to invest in and therefore on the general innovativeness of the regional or national economy (i.e. the **innovation ecosystem**).

**Blagoy Stamenov** (blagoy.stamenov@ec.europa.eu)
**Katarzyna Szkuta** (katarzyna.szkuta@ec.europa.eu)

© European Union, 2017

The European Commission's science and knowledge service
Joint Research Centre
EU Science Hub: *ec.europa.eu/jrc* @EU_ScienceHub Joint Research Centre
EU Science Hub - Joint Research Centre EU Science Hub
Joint Research Centre

# Understanding the policy pathway to research impact

## Commonwealth Fisheries Harvest Strategy Policy – a pilot case study

Russell, T., Gash, A., Söderbaum, J., Keenan, T. & Dowd, A-M.

CHIEF OF STAFF AND STRATEGY
www.csiro.au

The literature on public policy and related fields has been used to adapt elements of the CSIRO Impact Framework to better explain how science has influenced policy making and ultimately policy outcomes. Key to the adaptation are concepts of information exchange, cognitive reasoning, path dependency, institutional analysis, historical method and decision making.

## Background

The way in which CSIRO delivers impact in a public policy setting is not well understood. This pilot study seeks to improve this understanding by examining the role CSIRO played in the Commonwealth Fisheries Harvest Strategy Policy and Guidelines (HSP&G). The aim was to answer two key questions:

1. **How effective is CSIRO's engagement in policy processes?**
2. **To what extent did CSIRO contribute to the policy process?**

## Case study approach

The case study focuses on CSIRO's contribution to the development of a specific government policy: the Commonwealth Fisheries HSP&G. The HSP&G govern and guide the management settings for commercially-targeted species in Commonwealth fisheries, and are designed to ensure that key commercial fish species are managed for long–term biological sustainability and economic profitability.

A key reason for selecting the HSP&G for the pilot study is that it has clear

### INDICATORS OF IMPACT ON POLICY

A review of the literature suggests that the following indicators of policy impact are most suitable for application to the fisheries case study.

Table 1: Policy impact indicators.

| INDICATOR TYPE | DESCRIPTION | SUB-INDICATORS |
|---|---|---|
| ATTITUDINAL | This dimension considers the framing of debates, raising awareness of new issues, influencing political agendas, and affecting attitudes or perceptions of key stakeholders. | Key indicators include change in the awareness, values and beliefs of key stakeholders, changes in public opinion and increased interest in the issue (e.g. number of interest groups concerned) |
| DISCURSIVE COMMITMENTS | This dimension considers the rhetoric and the language used to discuss issues, as well as the endorsement of international declarations and agreements, as a source of policy impact | Key indicators include increased dialogue on key issues, a shift in the language surrounding debates, and evidence of changed rhetoric (such as major party policy positions) |
| PROCEDURAL CHANGE | This dimension considers changes in the processes through which policy decisions are made at the local, national and international levels | Key indicators include change in written government/organisational processes and procedures |
| BEHAVIOURAL CHANGE | Changes in behaviour of key actors and participants in the policy processes, including policy makers and government departments | Key indicators include changes in budgets and expenditure that align with new stated goals |
| CHANGE IN POLICY CONTENT | Changes in the content of legislation and other regulatory instruments | Key indicators include changes in the text of legislation or policy |

temporal bounds, beginning with a Ministerial Direction in 2005, and ending with the expected release of the revised HSP&G in 2018.

In addition, CSIRO, through its Oceans and Atmosphere Business Unit, has a long-standing relationship with the Australian Fisheries Management Authority (AFMA), which has been underpinned by a formal Memorandum of Understanding (MoU) since 2010. The key objectives of the MoU are to:

a. Facilitate the sharing of relevant strategic information with the aim of joint planning and closer alignment of strategies.

b. Provide a framework for identifying and prioritising key research initiatives and developing associated proposals for research and development and consultancy projects.

**MAPPING THE POLICY PROCESS**

The case study undertook an ‘actor-centric’ analysis and examined distinct policy phases to understand the HSP&G development and implementation, as well as the contributions of the various stakeholders (see Fig. 1).

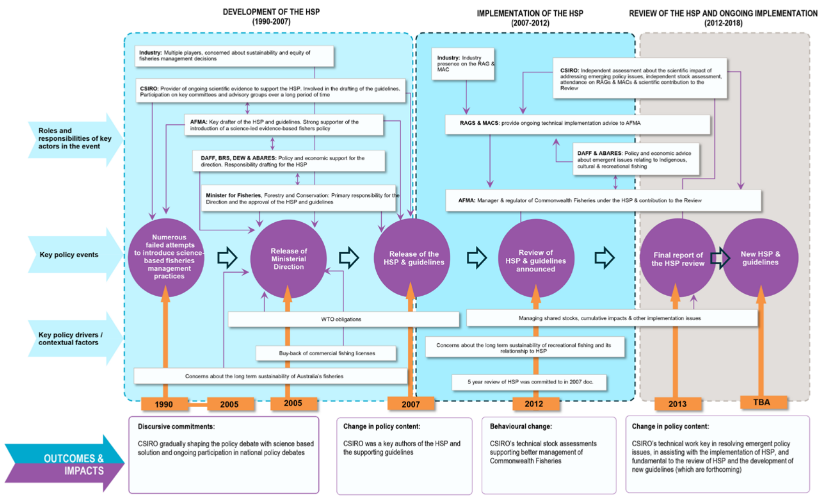



Figure 1: Timeline of events, actors, outcomes and impact.

# Key findings

- CSIRO played a leading role in providing the science that underpins the development of the policy and the guidelines for its implementation.
- The long-term, strategic nature of the engagement, spanning more than 25 years, was highly important and resulted in a high degree of trust between key actors.
- Multiple engagement mechanisms were employed across the continuum of the policy cycle to understand different perspectives and address fisheries management issues.
- CSIRO’s breadth of capability and reputation for scientific excellence was valued, not only in a traditional research-provider context but also as a trusted advisor both domestically and internationally.
- CSIRO was viewed as having made a seminal contribution to the policy cycle by the majority of stakeholders.
- Indicators of policy impact – including attitudinal and behavioural changes, discursive commitments, and changes to procedures and policies – were evident in each of the policy phases.

# Conclusion and next steps

The pilot study has provided a greater understanding of the activities, relationships and qualities that underpinned the successful translation of scientific research into a specific public policy. The study also demonstrated the importance of being able to engage across the continuum of the policy cycle, at different levels and through diverse mechanisms, which was integral to understanding issues and formulating appropriate, science-based responses. Finally, CSIRO’s multidisciplinary capability and international reputation for science excellence, were highlighted as valuable attributes in assisting partners to advance the fisheries management agenda.

However, it would be inappropriate to generalise findings from just one case study to broader lessons for science translation into policy. Further work is needed to gain deeper insights into some elements of this story and to confirm the findings in other areas where CSIRO is thought to have played a role in the policy process.


**FOR FURTHER INFORMATION**
Dr Anne-Maree Dowd
Executive Manager
Planning, Performance & Evaluation
**e** anne-maree.dowd@csiro.au
**w** www.csiro.au/en/about/our-impact



**ACKNOWLEDGEMENTS**
The authors would like to acknowledge the significant contribution made by Renate Hays and Jian Wang to the project’s activities.

European Commission

# Improving access to finance for innovative firms with growth potential: evidence of impact of R&D grant schemes on firms' output

Testa, G. and Szkuta, K., European Commission DG JRC

## 1. Motivation / Objectives

High-growth firms are increasingly a target for government interventions (European Commission, 2016). This is especially true for Europe which lags behind the US in the number of fast growing highly innovative enterprises (so-called scale-ups). In response to this large scale-up gap, the current policy debate has focussed on new sources and forms of R&I funding to enhance EU level support for scale-ups.

**Objective:** This paper examines the impact of R&D grant schemes on young innovative companies with growth potential. It focuses on the **output additionality**, i.e. the effects of R&D grants for scale-ups on the output of firms.

## 2. Main Research Questions

Our main research questions are:

1. What are the effects of R&D grants on scale-ups' output as measured by innovation activities, employment growth and firm performance (in terms of output, sales – including sales of new products and foreign sales/exports – value added and revenues)?
2. How do these R&D grants that specifically target young innovative firms with growth potential compare in term of employment, firm economic and innovative performance, and innovative activities with generic R&D grants and R&D subsidies commonly used as external funding to support both SMEs and large enterprises?

The results of the comparative analysis show:

- **The effects for R&D grants for young innovative firms** are **larger** than the effects of **generic R&D grants** and **R&D subsidies**.
- For generic R&D grants, the effects are higher when the grants induce changes in **firm behaviour** (collaboration and enhance firm human capital endowment) and when they target particular **technologies or sectors** (high-tech companies).
- The **combination of R&D grants and tax incentives** is more effective in increasing firm innovation than using only one instrument.

**Table 2**. Evidence sources

| | All sources | Academic articles | Evaluation reports |
|---|---|---|---|
| R&D grants for scale-ups | 20 | 13 | 7 |
| Employment | 11 | 5 | 6 |
| Economic and Innovation performance | 14 | 8 | 6 |
| Innovation | 7 | 3 | 4 |

## 3. Methodology

Our research approach draws on policy evaluation studies, and academic literature.

(a) **Sample papers: selection criteria**

- Direct R&D grants were selected only if the scope was to help young innovative companies grow faster (e.g. promote growth and exports, increase the commercialisation of innovation, enhance competitiveness).
- Young < = 10 years old.
- Examples of keywords used were "R&D" "grants", "SMEs", "young", "innovative firms", "high-growth firms", and "growth potential".

(b) **Selection criteria for generic R&D grants and R&D subsidies**

- Generic R&D grants were defined as R&D programmes grants targeting all companies (SMEs and larger enterprises) in all sectors.
- R&D subsidies include all R&D programmes (grants, loans and tax incentives), without distinguishing between instruments when reporting effects.

## 4. Results

The results on the effect of R&D grants for innovative enterprises with growth potential shows:

- Impact on employment: The average number of employees ranges from **7 to 16** per granted firms.
- Impact on both sales' growth, and share of innovative sales:
  - Strong and **positive** effect on total sales, and share of innovative sales.
  - **Time lag**. Effects on sales' growth take from **two** to **four** years to appear.
  - **Growth boosting effects**. Firms continue to grow for several years following the receipt of the subsidy (Autio and Ranniko, 2016; Soderblom et al., 2015).
    ➔ **Quality signal**
- Impact on innovation: Strong and **positive** effect on patent.

**Table 1**. Output Additionality

| **Effects** on: | R&D **grants** for **scale-ups** | **Generic** R&D **grants** | R&D **subsidies** |
|---|---|---|---|
| Employment | ●●● | ●●○ | ●○○ |
| Firm innovative and economic performance | ●●● | ●●○ | ●●○ |
| Innovation | ●●● | ●○○ | ●○○ |

●●● = major relevance, ●●○ = moderate relevance, ●○○ = minor relevance

## 5. Relevance for Policy

- **R&D grants** stimulate and prepare the companies **for the growth phase.**
- **Targeted funding** (technology focus) **delivers better results** for **disruptive innovations**, whereas generic grants for SMEs are better suited for knowledge diffusion as they mostly deliver new to the firm rather than new to the market results.
- **Selection mechanisms** built on milestones or subsequent phases of funding are **still rarely used** although their effects are very positive. This calls for a **greater use of this type of mechanisms**.
- The competitive and attractive R&D grants help companies to **attract follow up funding** (signalling effect especially for equity).
- Financial measures coupled with **complementary services** (e.g. networking, advice) **have a longer lasting effect**.
- **Tax incentives and grants** are **complementary** as regards to their impact on firm's growth and innovation activities given the evidence of higher impact of combined application (tax incentives and grants).

**Table 3**. Summary of R&D grants' design

| Description | Advantages | Challenges |
|---|---|---|
| | **R&D grants for scale-ups** | |
| • Phased approach, often linked to performance.<br>• Mostly delivered with additional services (training, mentoring, advice).<br>• Small cohorts of firms.<br>• Eligibility criteria more detailed and focused cf. generic grants (e.g. specific sectors, project managers' experience, company age). | • Phased approach allows distribution of funding based on results - not project proposals alone.<br>• Added-value services help entrepreneurs to deliver project to market. | • Phased approach requires clear milestones to enable monitoring of the process .<br>• Problems with picking winners if eligibility criteria very stringent. |
| | **Generic R&D grants** | |
| • Single grant.<br>• Financial support rarely linked with additional services.<br>• Larger cohorts of firms.<br>• Eligibility criteria more generic: R&D intensity, company's size, no age limits. | • Simple administrative rules<br>• Risk more equally distributed due to larger cohorts | • Risk of funding mostly new-to-the-firm innovation and/or issues with commercialisation given the lack of support during project development |

## 7. Sources/References


Acemoglu, D., U. Akcigit, N. Bloom, and W. Kerr (2013). Innovation, Reallocation and Growth, *National Bureau of Economic Research*, NBER Working Paper 18993.
Durufle, G, T. Hellmann and K. Wilson (2017). From start-up to scale-up: examining public policies for the financing of high-growth ventures. Working Paper, 04, Bruegel.
European Commission (2016), Europe's next leaders: the Start-up and Scale-up Initiative, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions, COM/2016/0733 final.
Edler, J., Ebserberger, B. and Lo, V. (2008). Improving Policy Understanding by means of Secondary Evaluation. *R&D Evaluation* 17(3): 175-186.
Edler, J., P. Cunningham, A. Gok, and P. Shapira (2016). Handbook of Innovation Policy Impact, EE Edlgar.
Jovanovic, B. (1982). Selection and the Evolution of Industry, *Econometrica*, 50, 649-670



**Giuseppina Testa** (giuseppina.testa@ec.europa.eu)
**Katarzyna Szkuta** (katarzyna.szkuta@ec.europa.eu)



The European Commission's science and knowledge service
Joint Research Centre
EU Science Hub: *ec.europa.eu/jrc* @EU_ScienceHub Joint Research Centre
EU Science Hub - Joint Research Centre EU Science Hub

Joint Research Centre

© European Union, 2017

Adam Paula[1], Daban Montserrat[2], Solans-Domènech Maite[1], Radó-Trilla Núria[1], Arboix Esteve[3], Borràs Cristina[4], Cabezas Roger[5], Lasauca Xavier[2], Mas Jordi[3], Morales Alba[6], Nason Eddy[7], Puigdomènech Cantó Alba[2], Dedeu Toni[1], Barberà Albert[8], Subirada Francesc[2] for the RIACAT Group

1 Agency for Health Quality and Assessment of Catalonia (AQuAS);
2 Secretary for Universities and Research Ministry of Business and Knowledge of the Government of Catalonia;
3 Catalan University Quality Assurance Agency;
4 Agency for Management of University and Research Grants;
5 I-CERCA (Agency for the Research Centres of Catalonia);
6 Catalan Association of Public Universities;
7 Ontario SPOR SUPPORT Unit
8 Catalan Ministry of Health. Government of Catalonia

# Preparing the future: developing research impact assessment capacity among stakeholders in Catalonia

The Department of Business and Knowledge and the Department of Health of the Catalan Government and the Agency for Health Quality and Assessment of Catalonia (AQuAS) created in 2018 the RIACat Group.

The mission of the RIACat Group is to put together a 3-year plan for the design,

## Introduction

- Over the last two decades, Catalonia has successfully implemented a research strategy to support outstanding excellence.
- Nevertheless, in current times, a scientific policy based only on excellence might turn outdated soon. Policy-makers are starting to realize that there is a need to set up a research impact agenda if wanting to keep successful rates in European competitions.

## 50 Participants

- Science policy-makers
- Other policy-makers (e.g. health, education, justice, agriculture)
- Assessment agencies (evaluators)
- University managers
- Research center managers
- Science dissemination
- Foundations

## Objectives of the RIACat Seminar

- To raise awareness about the importance of Research Impact Assessment (RIA)
- Familiarize participants with RIA concepts and options
- To understand the Catalan context and options for designing and implementing a RIA tool in Catalonia

**The International School** on Research Impact Assessment

The learning materials of the RIACat Seminar are based on an adaptation of the materials and delivery of the The International School on Research Impact Assessment (ISRIA)

More information is available at: www.theinternationalschoolonria.com

AQuAS was co-founder of ISRIA in 2013

development and implementation of a country-based cross-discipline research impact assessment agenda.

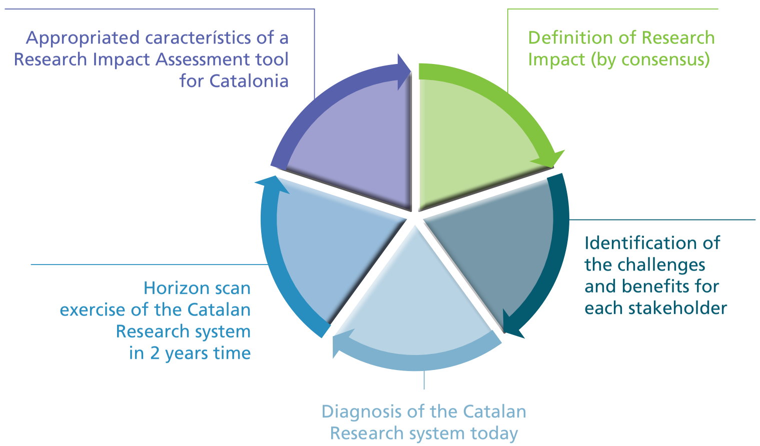
The first action was to organize a 3,5 day course for policy-makers and research stakeholders on research impact assessment to put every stakeholder 'in the same page' and learn from international experiences.

## Assessment agenda

The **program** combined general learnings about RIA, international lessons, discussions about applications to Catalonia, and sessions of mutual learning for the co-design of an impact agenda

| Day 1<br>Tuesday 22 | Day 2<br>Wednesday 23 | Day 3<br>Thursday 24 | Day 4<br>Friday 25 |
|---|---|---|---|
| **WHY?**<br>Setting the grounds of RIA<br>(Catalan) | **HOW AND WHO?**<br>Methods and stakeholders<br>(English) | **APPLICATIONS TO CATALONIA**<br>Needs and options<br>(English) | **APPLICATIONS TO CATALONIA**<br>Implementation<br>(Catalan) |
| What is impact and the foundations of RIA | Compare and contrast methods | **Panel:** Learning from International experiences | What tools do we have? |
| Purposes of RIA | The power and limitations of indicators | Understand our context, build an horizon scan | Planning for the future: what do we need? |
| Active learning: Defining impact | Impact narratives | **Panel:** What are the challenges and options for Catalonia | RIACat summary |
| | Engaging stakeholders with RIA | Active learning: what tool? | |
| | Active learning: purposes and stakeholders | | |

The **active learning** approach of the RIACat Seminar lead to the development of collective knowledge, useful for the design of an impact agenda for Catalonia




Paula Adam
padam@gencat.cat

# Evaluation of socio-economic impacts of the business R&D support in small economies. The case of the Czech Republic

*Tomas Ratinger, Vladislav Cadil, Sylvester Amoako Agyemang*
**Technology Centre CAS**

## Context

The counter-factual econometric analysis is becoming increasingly popular particularly for the assessment of input and output additionality. The use of methods based on comparison between strictly separated treated and control groups faces some obstacles in terms of data-availability and quality, interpretation and presentation of results to a non-specialist audience . A particular challenge is finding the right control group, particularly if country or investigated region is small and support measures plentiful. If an appropriate control group cannot be found among the unsupported companies, a so-called **generalized propensity score matching** (GPSM) can be used, provided that firms receive a different level of treatment (support). The control group is created from the closest similar firms, which received a different level of support.

## Research aims

- To assess the effects of the Czech programme TIP supporting R&D in private companies in the recent decade.
- To point out the methodological challenges of such an assessment, and to show how they can be addressed using the GPSM approach.

## Data

- Two data sources were used: monitoring data of the support programme TIP and economic data from the database Bisnode-MagnusWeb.
- The impact is measured at two hierarchical levels – the firm (micro) and the economy (macro) levels. At the micro level, the impact of R&D support on the firms' performance are measured by 5 indicators: 2 measuring net and gross R&D expenditure effects and 3 indicators relating to output – gross value added, profit and labour productivity. The macro level impact of R&D support is represented only by a fiscal indicator – average tax paid by the firms.
- The assessment of the R&D expenditure effects relates to the period of 2010 to 2016 while the impact of the support on the output indicators relates to 2013-2015 (the period after the economic recession 2009-2012). In the both cases, the year of 2009 was used as the base year for the analyses. For each of the output indicator, the difference-in-difference approach is adopted in order to remove the time invariant "fixed effects" ➔ the difference between the average figure for the "impact period" (2013 to 2015) and the base year (2009) is computed.

## Methodology

**Size categories of the beneficiary firms**
The probability of receiving a lower or higher support and treatments effects are more likely to be heterogeneous with respect to the firm size. Therefore the firms are divided into three categories (small, medium and large) using their annual turnover in the base year.

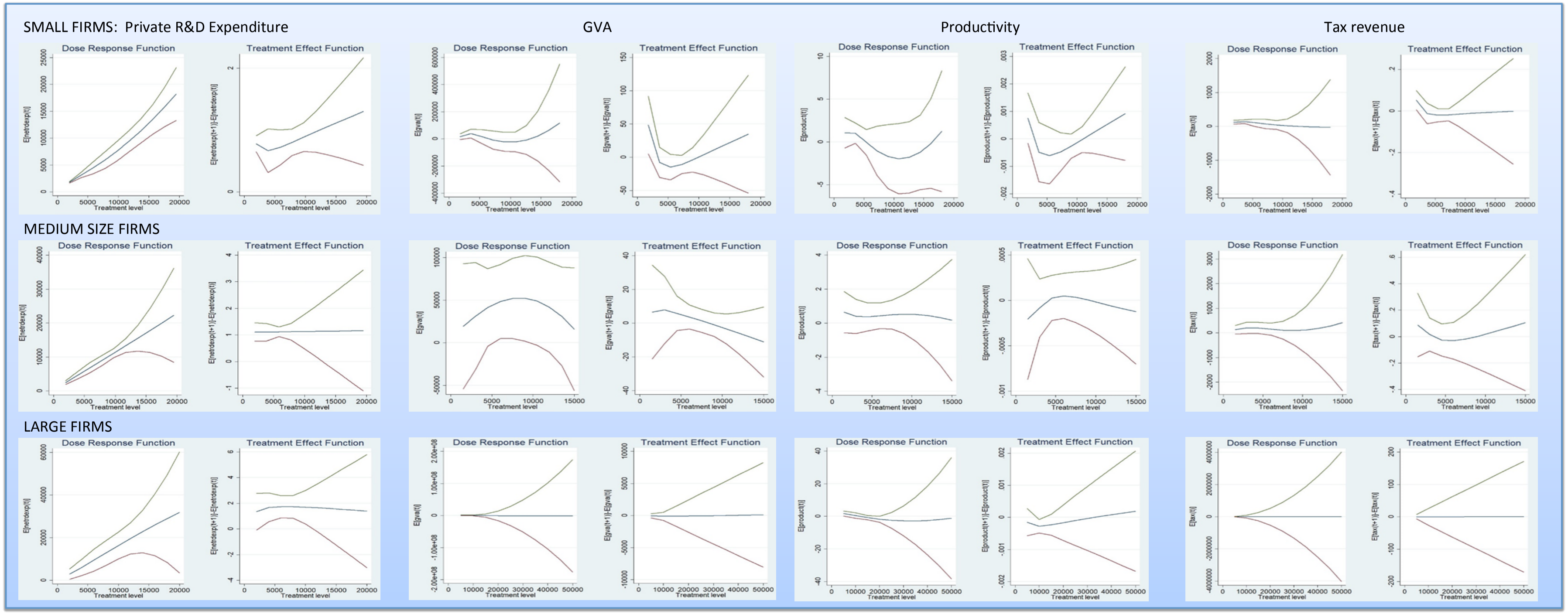
**Control Variables**
To deal with the pre-treatment biases in the impact estimation, 10 pre-treatment covariates (cash flow, fixed assets, current assets, equity, profit, long term debt/total assets, bank credit/total assets, intermediate consumption, depreciation, interest received by a firm) referring to the base year (2009) are selected and included in the model.

**A Linear Regression Analyses of the Outcome and Control Variables**
Firstly, the impact of R&D support on the defined outcome/performance variables is analysed by regressing the control variables and the treatment variable (R&D support) on the outcome variables with the aim to determine the relevance of the selection of the control (pre-treatment) variables in respect of the formation of the performance variables.

**The Generalised Propensity Score (GPS) Matching**
Afterwards, the conditional distribution of the treatment (support) given the covariates via Ordinary Least Squares (OLS) method is estimated. Then the normality of the disturbances using the Kolmogorov-Smirnov equality-of-distributions test is tested. ➔ GPS estimation
The next step is to test whether the propensity scores balance the sample. To achieve this, the treatment (support) is divided into 3 different intervals in our particular case..

**Estimation of the Dose-response Function**
Finally, each outcome variable (in all size categories of firms) is regressed on the conditional expected value of R&D support given by the pre-treatment variables ➔ dose-response functions.
The Dose-Response Function indicates the estimated impact of a given level of R&D support on the outcome indicator. In the way answering the counterfactual question, what would have happened to a given firm had that unit received a different level of treatment.
(Marginal) Treatment Effect Function shows the change of the outcome indicators in response to the change of the treatment level.

# Results (4 of 7 indicators)




**LEGEND**

-- Dose Response -- Treatment Effect
-- Low Bound -- Upper Bound
Confidence Bounds at 95 % level
Linear predictions

## Conclusions

- Treatment effects are heterogeneous depending on firm sizes. Generally, the R&D support effects are positive and more substantial in small and medium scale firms while there are no such effects confirmed for large-scale firms.
- The support could only impact positively on the output (profit, productivity and GVA) of small-scale firms only after reaching a certain minimum threshold of the support whereas output diminishes for medium scale firms after a point of saturation.
- The support has a complementary effect rather than a substitutability effect on the private R&D expenditures.
- The differences in impacts between the size categories of supported enterprises can be explained by the different economic performance of small firms compared to the medium and large companies and different strategic focus of R&D in each size category. Small companies are more likely to concentrate on activities with faster return on capital. Therefore R&D conducted by small firms is rather oriented on the direct support of their production activities or on the application of the new knowledge in the short-term horizon, while the strategic focus of R&D in large companies are more focused on long-term strategic research.

TECHNOLOGY CENTRE CAS
Ve Struhach 27, 160 00 Prague 6, Czech Republic
tel.: +420 234 006 100
fax: +420 234 006 251
e-mail: tc@tc.cz
www.tc.cz

Tomas Ratinger: ratinger@tc.cz , Vladislav Cadil: cadil@tc.cz, Sylvester Amoako Agyemang: amoakoagyemangs@gmail.com

Innovation Growth Lab

# Innovation Growth

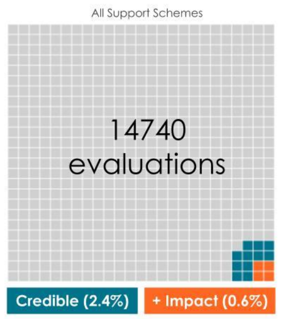
*Working in partnership to make innovation, entrepreneurship and business experimentation and new evidence.*

***€150 billion** spent every year in EU supporting business to start, innovate and grow. Yet we know little about what works, and what doesn't*

https://www.innovationgrowthlab.org/blog/much-%E2%82%AC152-billion-spent-across-eur

Good evaluations can change people's minds and lead to better decisions

But most impact evaluations are not credible




Source: Systematic reviews conducted by the What Works Centre for Local Economic Growth at the LSE (Credible: Level 3 Maryland Scale – Positive impact on employment)

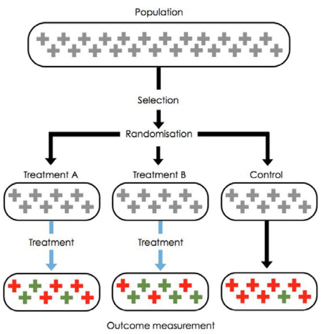
*There is also too little innovation within policy itself and when there is, be it incremental or radical, we cannot tell whether it is for better or worse.* ***How can an organisation develop and test new ideas systematically?***

*A really effective way is using a Randomised Controlled Trial (RCT). While currently underused and not always feasible, these are often seen as the most robust approach to demonstrating causality*

*RCTs can be used as part of a broader experimental approach to test anything from small tweaks to programmes to the overall impact of business or innovation schemes*

## What is a randomised controlled trial (RCT)?

An RCT is an experiment where participants are randomly allocated to receiving an intervention. The randomisation enables you to compare the effectiveness of the new intervention against what would have happened if you had changed nothing




## IGL partners

nesta · Ewing Marion Kauffman Foundation · argidius · Australian Government Department of Industry, Innovation and Science · FFG Austrian Research Promotion Agency · designsingapore · Innovation Norway · Innovate UK · Scottish Enterprise · Department for Business, Energy & Industrial Strategy

IGL is a global partnership bringing together governments, foundations and researchers to scope, develop and test different approaches to increase innovation, support high-growth entrepreneurship and accelerate business growth.

Over the last few years we are seeing a growing number of RCTs. Many of these are (co)funded by the IGL Grants programme, which has supported over 30 trials with close to $3 million from the Kauffman Foundation, Nesta and the Argidius Foundation. We are assisting a number of government agencies in their own journey to experimentation.

We are starting to learn valuable lessons about how we can encourage innovative ideas and support businesses, and many more lessons will emerge as the trials now in the field start to deliver results.

https://www.innovationgrowthlab.org/blog/what-are-we-learning-policy-experiments-increase-innovation-and-entrepreneurship

## IGL in numbers

**15** agencies we've worked with

**>1,100** attendees at our conferences and workshops

**55** trials supported

**>85** researchers in our network

Amount invested in trials through IGL **$2.8 million**

**26** countries with IGL partners or projects

# Lab

@IGLglobal

*growth policy more impactful through*

What IGL does

1. **Support** partners to identify policy solutions and run trials
2. **Fund** trials with the IGL Grants Programme
3. Build and **connect** communities
4. Promote wider **adoption** of trials
5. Create useful **resources**
6. **Disseminate** lessons


Creating a global community around experimentation

Our **three annual conferences** have been attended by >650 senior policymakers, practitioners and researchers from over 40 countries

**85%** of IGL2018 participants were very likely to recommend the IGL Conference to their colleagues

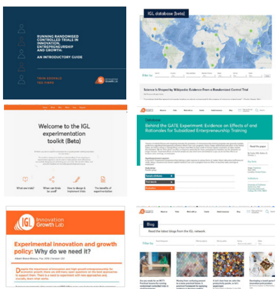
Our **five research meetings** brought together top researchers to present and discuss new experimental research in innovation, entrepreneurship and growth

Register your interest for the IGL2019 Conference in Berlin:

https://www.innovationgrowthlab.org/igl2019-global-conference


Some IGL outputs and resources

1. A guide on how to conduct trials in the field of innovation, entrepreneurship and growth
2. An online database of trials from around the world, ongoing and completed
3. An online toolkit to help those wanting to know why and how to become experimental
4. A policy brief on why we need more experimental policies and how to become more experimental
5. A regular series of blogs covering advice and results from trials and new policy ideas



***Visit our website or get in touch with us to find out more***

innovationgrowthlab@nesta.org.uk

www.innovationgrowthlab.org

# PRELIMINARY REMARKS ABOUT SCIENCE, TECHNOLOGY AND INNOVATION POLICY EVALUATION IN LATIN AMERICA

Adriana Bin*, Rafaela M. de Andrade**, Lissa Vasconcellos Pinheiro*, Sergio Luiz Monteiro Salles-Filho**
*School of Applied Sciences **Department of Science and Technology Policy, Geosciences Institute (University of Campinas/UNICAMP)

**Research funded by: CNPq and FAPESP**

## INTRODUCTION

The focus of this work is to present an on-going experience of collecting, coding and analyzing Science, Technology and Innovation (STI) Policy Evaluations in Latin America (LA), with emphasis in the evaluation design and methods. The research is part of a broader initiative named Science and Innovation Policy Evaluations Repository (SIPER), coordinated by Manchester Institute of Innovation Research (MIoIR). SIPER is a central source of knowledge on science and innovation policy evaluations. Its aim is twofold: (i) to provide on-line access to a unique collection of policy evaluations, located at a single location; and (ii) to allow policy learning by providing an informed analysis of the database contents that is both searchable by policy makers and other stakeholders and which provides the basis for additional academic analysis.

## METHODOLOGY

The research follows a three-phase methodology: collect, code and analyze. Collection phase refers to identifying evaluation studies of STI policies in LA countries. As defined by SIPER project methodology, qualified evaluations to be included in the study are those: (i) on science and innovation policy; (ii) evaluating a clearly identifiable, specific program or group of programs; (iii) having a distinguishable methodology; and (iv) providing some sort of evidence.

**Table 1 – Evaluations collected and inserted on SIPER**

| Country | Collected | Characterized |
|---|---|---|
| **Argentina** | 26 | 24 |
| **Brazil** | 33 | 17 |
| **Chile** | 35 | 20 |
| **Colombia** | 18 | 16 |
| **Mexico** | 18 | 18 |
| **Uruguay** | 16 | 16 |
| **Total** | 146 | 111 |

Coding phase was dedicated to characterization of collected documents following SIPER requirements based on a survey, which includes: (1) Related policy measure characteristics; (2) Evaluation characteristics: (2.1) Basic; (2.2) Topics covered; (2.3) Design; (2.4) Data collection methods; (2.5) Data analysis methods; (2.6) Quality Issues; and (3) Document properties. Finally, the last phase comprehends the use of codified information in order to discuss state-of-art of STI evaluation practice in LA.

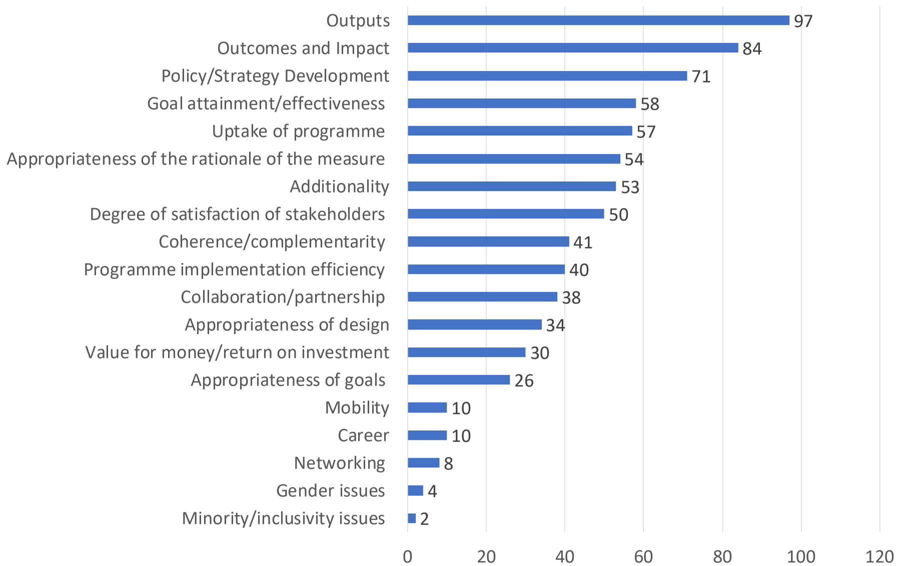
## RESULTS (1)

Preliminary results show that STI policy evaluation activity in LA is recent and

## RESULTS (2)

Quasi-experimental and non-experimental designs represent the majority of evaluation reports, corresponding to 52 and 46 evaluations, respectively. 84 evaluations measure the long-term impacts of policy measures, with emphasis on the use of scientific and technological, social and economic indicators. Among the evaluations that measured economic impact, 26 analyzed cost-benefit or returns on investment. A set of 53 evaluations measures additionality, 25 of which analyze only additionality of output, 8 additionality of input and 1 behavioral additionality. There is also a combination of input and output additionality (11 documents) and of the three types (5 documents).

**Chart 2 – Aspects of the program examined by the evaluations**




Regarding data collection, the use of existing databases (101 evaluations) stands out; 51 evaluations employ surveys and 44 employ interviews; analysis of scientific publications appear in 18 documents and focus group /workshop/meetings in 17.

heterogeneous across countries. In general, the purposes of these evaluations are both formative and summative (62 documents), followed by those that are only summative (36) and only formative (13) evaluations.

**Chart 1 – Types of policy**

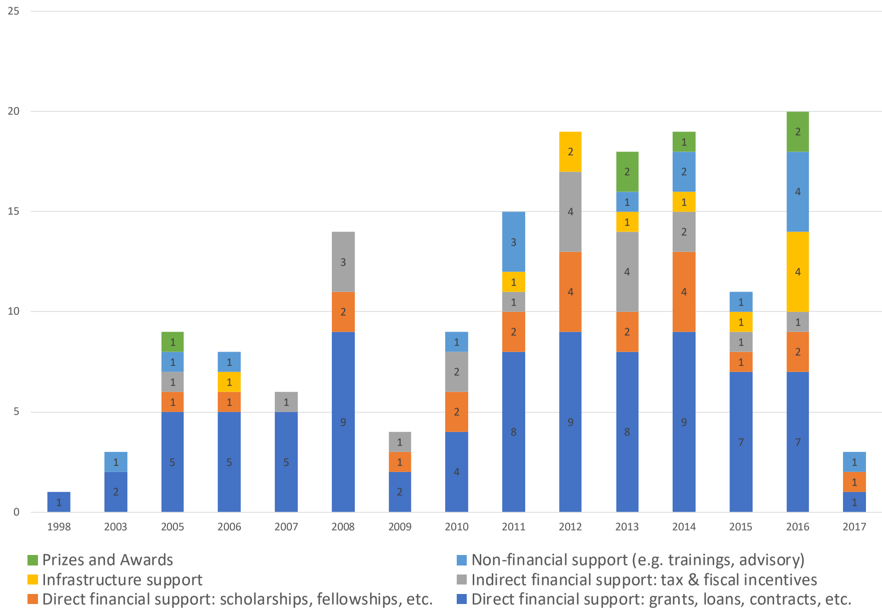


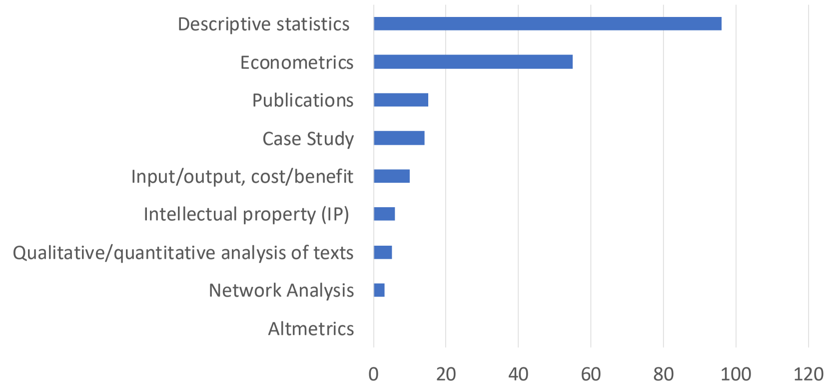

**Table 2 – Who conducted**

| Who | Total |
|---|---|
| External (independent) | 66,7% |
| Internal | 24,3% |
| External (independent) and internal | 4,5% |
| External (within government) | 3,6% |
| External (within government) and External (independent) | 0,9% |

**Table 3 – Timing**

| Timing | Total |
|---|---|
| Interim | 94 |
| Ex-post | 11 |
| Ex-ante | 6 |

**Chart 3 – Data Analysis**




## CONCLUSIONS

- The analyzes carried out so far indicate a growing movement towards the institutionalization of STI policy evaluation practices in Latin America, in line with the growing importance of these policies and the perception of their contribution to countries' economic and social development.
- However, there are few variations on the methodological designs and indicators used, evidencing the need for substantive advances in this field.
- Complementary analyzes should be performed after the completion of the collection and characterization phases of the evaluations, seeking to identify the occurrence of a relationship between the variables analyzed, as well as the countries' profile regarding STI policy evaluation.
- Visit the SIPER repository: **http://si-per.eu**

## REFERENCES

EDLER, J. et al. (2012). The practice of evaluation in innovation policy in Europe. Research Evaluation, London, v.21, n.3, p. 167-182.

SCRIVEN, M. (2009). Meta-Evaluation Revisited. *Journal of MultiDisciplinary Evaluation*, 6(11), iii-viii.

STUFFLEBEAM, D.L.; CORYN, C.L.S. (2014). *Evaluation theory, models, and applications*. Second edition. San Francisco (EUA): Jossey-Bass.

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

# The impact of the EXIST Business Start-up Grant on corporate growth: A group comparison for Dresden (GER)

## Introduction

The EXIST Business Start-up Grant (BSG) is one of the most important governmental programs in Germany to support founders by turning their business idea into action. This paper investigates the start-ups' corporate development using a peer group comparison on longitudinal data.

**Research Question**

Do BSG-funded start-ups outperform non-funded industry peers in terms of:

(i) Risk of cessation
(ii) Survival time
(iii) Employment (FTE) development
(iv) Revenue development

**Keywords**

- corporate growth, governmental start-up assistance, innovation policy, longitudinal analysis, peer group comparison

**Conference pillars mainly addressed**

- Leading edge concepts, tools and methods to assess impact of R&I policy
- Effects of and policy learning from impact evaluation

**Acknowledgement**

- Cornelia Ernst, Dresden University of Technology
- Marco Rösler, dresden|exists

## Methods

**Data set**

The self-collected dataset was created by conducting desk research and field research (online survey) due to a lack of reliable and publicly accessible longitudinal micro-level data at the time (2017).

Cross-sectional

- Treatment group (TBSG): limited liability corporations (Ltds.) in Dresden (GER) funded by the BSG (n=21) [provided by dresden|exists, the local authority responsible for the BSG]
- Control group (CBSG): non-funded peers (n=18) which would have been eligible for a BSG funding. The eligibility criteria were assessed on information related to the Ltds.' time of incorporation, e.g. its registered object, in retrospect by two independent experts (four-eyes principle). Two independent datasets were merged, one provided by Dresden Chamber of Commerce and one retrieved from the database Amadeus.

Longitudinal

- Period of incorporation: 2008 - 2011
- Observation period: First five post-incorporation years [data on corporate development obtained from the two independent datasets and an online survey]

| Measurement | (i) Risk of cessation | (ii) Survival time | (iii) Employment development | (iv) Revenue development |
|---|---|---|---|---|
| Method | | Analysis of Differences | | |
| | Cox proportional hazards resgression | Survival time analysis | Two-way Analysis of variance | Cumulative odds ordinal logistic regression |
| Variables | Survival time by Treatment | Survival time by Treatment | Treatment and Year on FTE, Treatment and Industry on FTE | Revenue by Treatment, Industry, and Year |
| Specification | • Log-rank test<br>• Scaled Schoenfeld residuals | • Kaplan-Meier survival curves | • Type III sums of squares<br>• Bonferroni adjustment | • Proportional odds<br>• Post-estimation: discrete and marginal change |

## Results

**(i) Risk of cessation**

- Visual inspection of proportional risk of cessation (Figure 1) is inconclusive because the curves cross each other but lie in the same range overall. However, the result of a performed log-rank test ($p = .820$) provides evidence that the survival distributions of the two groups are not statistically different. This is supported by a test based on Schoenfeld residuals ($p = .397$).

**(i) Risk of cessation**

- Visual inspection of proportional risk of cessation (Figure 1) is inconclusive because the curves cross each other but lie in the same range overall. However, the result of a performed log-rank test ($p$ = .820) provides evidence that the survival distributions of the two groups are not statistically different. This is supported by a test based on Schoenfeld residuals ($p$ = .397).

**(ii) Survival time**

- Is almost equal (Table 1) between the two groups which blends into the result of almost equal risk of cessation.

**(iii) Employment development**

- Related to employment overall, Wilcoxon's two-sample rank sum test (Table 1) provides evidence to negate an overperformance by funded Ltds., which is displayed in Figure 2. On the contrary, groups' means indicate higher employment for the funded Ltds. (Table 1).
- For employment development, the results of Analyses of Variance are ambiguous and highly effected by outliers which is depicted in Figure 2. Over or underperformance of funded Ltds. depends on the industry sector.

**(iv) Revenue development**

- Related to revenue overall, independent t-test (Table 1) provides evidence to negate an overperformance by funded Ltds. Groups' medians indicate the same (Figure 3).
- Moreover, there are higher probabilities of generating higher revenue for control peers.

Table 1: Results analysis of differences for Survival time, Employment and Revenue by Treatment.

| | | n | M | Mdn | t-test | Wilcox. | $r^2$ | M-W stat. |
|---|---|---|---|---|---|---|---|---|
| (ii) Survival time (5 years) | Total | 39 | 52.8 | 60 | | .901 | .000 | .509 |
| | Treatment | 21 | 52.3 | 60 | | | | |
| | Control | 18 | 53.3 | 60 | | | | |
| (iii) Employment | Total | 72 | 8.9 | 3 | | .032** | .072° | .654 |
| | Treatment | 47 | 9.6 | 2 | | | | |
| | Control | 25 | 7.3 | 6 | | | | |
| (iv) Revenue (e.v.n.a.) | Total | 76 | 4.2 | 4 | .082* | | .040° | |
| | Treatment | 46 | 3.8 | 3.5 | | | | |
| | Control | 30 | 4.7 | 4 | | | | |

n represents the amount of observations for Ltd. i in point of time j.
*, **, ***. Denote significance at 10%, 5%, and 1%, respectively (two-tailed test).
°, °°, and °°°. Denote effect sizes: small (°): $.01 \leq r^2 < .09$, medium (°°): $.09 \leq r^2 < .25$, and large (°°°): $.25 \leq r^2$.

Figure 1: Kaplan-Meier survival curves.

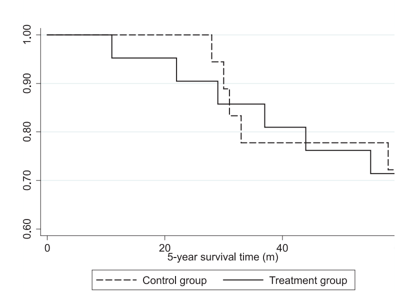



Figure 2: Median values of Employment by Year, n = 72.

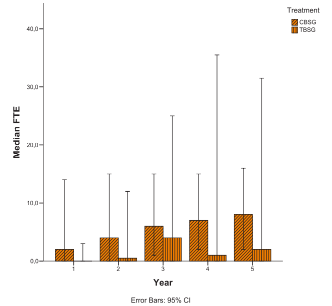
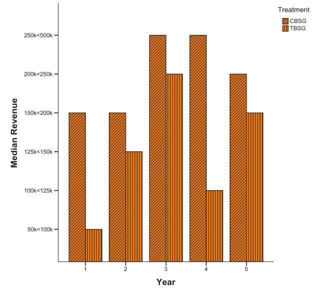



Figure 3: Medians values of Revenue by Year, n = 76.




## Conclusion and Policy recommendation

- Funded Ltds. do not outperform industry peers in terms of (i) risk of cessation and (ii) survival time.
- Rather non-funded industry peers perform better than funded Ltds in terms of (iii) employment and (iv) revenue development.
- According to the online survey, for 8 out of 11 respondents it would have been unlikely or even very unlikely having founded without the BSG funding. So, treatment and control groups' Ltds. might differ in their pre-treatment willingness to incorporate.

→ The BSG in Dresden might not fund the founders ready to fly high, as intended, but helps start-up seeds to see the light of the day, which tally with the "theory of external assistance as the support option of last resort" (Juita-Elena (Wie) Yusuf 2017).

**Policy recommendation**

1. The start-up agents responsible for the BSG funding should increase active sourcing in order to not rely on the people who come in and apply for a BSG funding.
2. BSG funding might provide even more guided preparation during the one year funding period with regard to the "theory of outside assistance as a knowledge resource" by Chrisman and McMullan (2004).
3. Funded Ltds. should be encouraged to self-reliance by loosening university-related ties since proximity to university is not necessarily related to better performance of the start-ups (Doutriaux 1987).

## Limitation and Further research

- Group sizes of 21 and 18 companies do not meet the self-proclaimed sample size threshold of $n > 30$, with regard to Student's t-distribution and the associated t-test.
- The results are not representative for Germany and must be interpreted even for the case of Dresden with caution.

**Further research**

1. Replicating on a larger sample size to meet the threshold of n=30, at least.
2. Detecting peers with another method, e.g. text mining the companies' object, to certify judges' matching (investigator triangulation).
3. Verifying the same pre-treatment conditions for the two groups to justify the matching.
4. Conducting qualitative analysis to investigate the reasons for the corporations' development on an individual basis.
5. Taking advantage of variables from the self-collected dataset, which are not considered in this paper already, e.g. Year of incorporation.

References

Chrisman, J. J., & McMullan, W. E. (2004). Outsider Assistance as a Knowledge Re-source for New Venture Survival. Journal of Small Business Management, 42(3), 229–244.
Doutriaux, J. (1987). Growth pattern of academic entrepreneurial firms. Journal of busi-ness venturing, 2(4), 285–297.
Juita-Elena (Wie) Yusuf. (2017). The effectiveness of entrepreneurial start-up assistance programs: Evidence from the U.S. Panel study of entrepreneurial dynamics. In Preventing Ageing Unequally (pp. 1–9). OECD Publishing.

IHK Industrie- und Handelskammer Dresden

dresden | exists
WISSEN. GRÜNDEN. UNTERNEHMEN.

Philip Dörr
FSU Jena, Chair of Economics/Microeconomics
Carl-Zeiss-Straße 3, 07743 Jena, Germany
philip.doerr@uni-jena.de

INSTITUT FÜR SOZIALFORSCHUNG UND GESELLSCHAFTSPOLITIK
ISG
rwi Leibniz-Institut für Wirtschaftsforschung

# WHAT ARE THE IMPACTS OF CLUSTER POLICIES? THE GERMAN EXPERIENCE

Michael Rothgang (RWI), Bernhard Lageman, Anne-Marie Scholz (ISG)*

## Starting point

### Cluster policies in Germany

Cluster policies have changed structural policies in Germany in a substantial and sustainable manner in the past decades:

1. All State Governments use cluster promotion to promote regional growth and innovation, while the Federal Government implemented several own cluster programmes.[2]
2. Cluster policies are rather popular due to (i) a sound theoretical basis in cluster literature, (ii) flexibility of the instrument with cluster organizations as intermediaries that take up firm needs.
3. There are substantial differences between programs in respect to funding and targets.

To improve cluster policies it is important to understand the mechanisms of their policy success and failure.

### Theory-based expectations

Cluster policies should

- promote existing clusters and function as midwife of emerging clusters,
- encourage cooperation between firms, academia and state authorities,
- increase the developmental potential of regional economies and promote technology evolution.

### However...

- Evidence on the actual effects of cluster policy is limited,[1]
- The experiences with practical cluster policy in Germany have not been assessed in a systematic manner yet.

## Empirical Approach

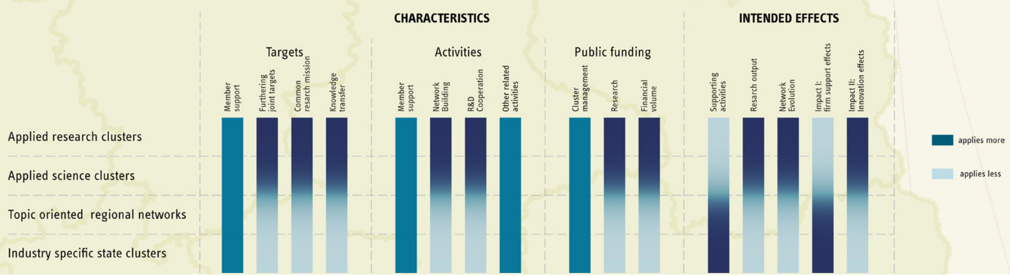
Qualitative analysis of the evidence on characteristics of German cluster policies, their intended effects, and the available evaluation studies; three components are combined:

- Cluster initiatives: Internet-based survey of 600 cluster initiatives from which 370 focus on innovation and knowledge transfer, deeper analysis of 30 randomly chosen technology-oriented cluster initiatives,
- Cluster policies: Survey of about 25 cluster programmes of the Federal Government and the 16 State Governments,
- Evaluation studies: Analysis of the available cluster literature (formal evaluations, empirical studies) on German cluster policy.

## Findings

1. Cluster policies (and the related concept of network promotion) embrace a large range of different instruments from financial support of common activities to R&D funding of projects related to a common cluster mission.
2. Cluster initiatives are the intermediate organizations needed to promote cluster development; based on different characteristics, four types of cluster initiatives could be identified which relate to different intended policy effects.

### Types of Cluster Initiatives, Characteristics, and Effects



## Research Questions

1. How does the diversity of practical cluster policies relate to their intended effects?
2. To what extent and on what methodological basis were the Federal and State cluster programs evaluated?
3. What effects were identified and which questions remain open in respect to (a) the impact of cluster policies and (b) the evaluation methods applied?

## Evaluation studies

### Availability and methods

- Many official evaluation studies, however mostly ex ante or accompanying evaluations (some ex post evaluations mainly of Federal Programmes),
- In many cases simple qualitative assessments, based on a limited set of indicators,
- We find a small number of more in-depth cluster policy impact analyses in the research literature (some qualitative, some quantitative).

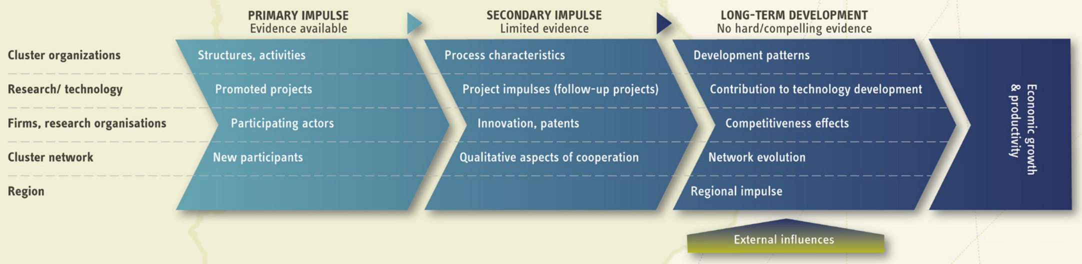
### Challenges

- Most programme effects are the result of complex effect mechanisms with many external influences.
- Politicians' interest in short-term results contradicts basic structural features of these programmes.
- Major policy effects can only become visible after longer time-lags and based on a resource-consuming thorough analysis.

3. Several evaluation studies have been conducted both at the Länder level and for Federal programmes. These evaluations were rather different in respect to which policy effects could be assessed.[3] Major results are:
   - Qualitative studies show that the policies were successful in shaping supportive environments for cluster actors and in offering assistance for existing and emerging R&D cooperation networks.[4]
   - These activities were successful only in a part of the cases; there is more relevant knowledge on effects on the cluster than on the programme level.
   - Success on cluster level was determined by multiple factors, including program design, financial endowment, abilities of the cluster management, but also the individual constellation of the cluster/cluster initiative.
   - Quantitative analyses identified effects for individual firms' input and output indicators. However, doubts arise with respect to the causal interpretation of some of the output-related results because of the probable time structure of impact patterns.[5]

4. While primary effects (induced activities and network development) could be clearly identified, much less is known about secondary policy effects (such as resulting inventions and innovations on the firm level, and impact on cluster population and performance on the cluster level).[6]

### Evidence on Cluster Policy Impact from an Evolutionary Perspective




## Implications

1. To gain more relevant insights into the long-term impact of cluster policy, impact research should reconstruct chains of effects of policy-induced cluster development and confront the findings with the original narratives of policy makers on expected cluster evolution.
2. New developments in the evaluation methodology of cluster policies are needed: improved qualitative approaches to understand the sequential effects of policy impulses (e.g. process tracing), the thoughtful application of advanced quantitative approaches based on theoretical impact models which take account of the complexity of impact patterns, innovative combination of qualitative and quantitative approaches.


Selected References
[1] Uyarra, E., Ramlogan R. (2016), see also the methodological discussion in Schmiedeberg, C. (2010).
[2] Kiese, M. (2012).
[3] Wessels, J. (ed.) (2008).
[4] E.g. Dohse, D. (2007) and Töpfer, S., U. Cantner, H. Graf (2017).
[5] E.g. Falck, O., S. Heblich, S. Kipar (2010), Crass, D., C. Rammer, B. Aschhoff (2016) and Engel, D., V. Eckl, M. Rothgang (2017).
[6] Rothgang, M., J. Dehio, B. Lageman (2017).

*We thank Martina Böhmel for her contribution to the research project.

Contact: Michael Rothgang, RWI-Leibniz-Institute for Economic Research, Hohenzollernstraße 1-3, 45128 Essen, Germany, michael.rothgang@rwi-essen.de

# Making the invisible visible

## Towards strategic measures of research and technology organisations (RTOs)

Kirsi Hyytinen, Katri Kallio, Olli Kuusisto & Arho Suominen
VTT Technical Research Centre of Finland Ltd

### RTOs are tackling with systemic and complex challenges

There are no simple solutions to complex societal problems. Modern innovation theories emphasizes that tackling system level problems requires considering three perspectives:

1) Solutions require deep integration of technologies and service-based novelties.
2) Collaboration between multiple actors from different sectors of society is required.
3) Developing and disseminating solutions requires understanding of customer and citizen needs and the policy-making context.

### Current innovation measures do not make the impact visible

Target of RTOs is to solve broad societal challenges. However, current innovation indicators do not capture RTOs' multiple roles and impact in systemic environment.

There are two main reasons why traditional evaluation and measures fail. The focus is on:

- techno-economic aspects of innovation and impact, which are not able to capture the reality of innovations (e.g. system, social and service innovations) and "hidden performance".
- input and direct outputs. Therefore we lack data about impact.

### VTT's strategic evaluation framework

VTT's new strategic evaluation framework and related measures of success (KPIs) integrate multi-criteria approach of innovation to the traditional balanced scorecard (BSC). The framework expands traditional innovation indicators towards broader societal transformations in accordance with VTT's strategy. They make VTT's long-term impact visible in four categories:

**1) Benefit for society**
**2) Benefit for customers**
**3) Excellence in everything we do**
**4) Sufficiency of financial resources**

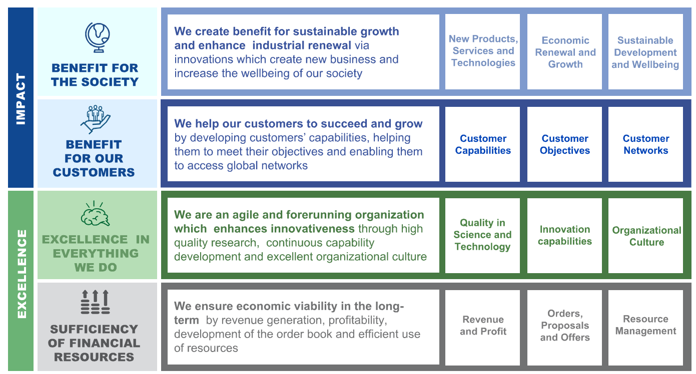



***Figure 2.** Framework to measure success at VTT.*

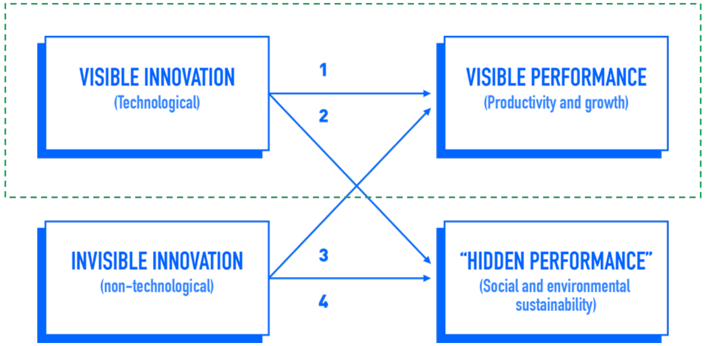



***Figure 1.*** *The focus in the impact analysis is typically in the linkages between the visible innovation and visible performance, which do not capture the non-technological innovation and hidden performance of innovations. (Djellal & Gallouj, 2010).*

Biased information may cause inaccurate analysis and interpretations and lead to inappropriate decisions. In order to paint **a truthful and comprehensive picture of RTOs' impact**, and to provide better information for managing and learning, we need diversified evaluation approach and measures.

The long-term societal impact is based on VTT's ability to create customer impact in shorter term. To ensure this, we need to support customers' success and growth and actively create new partner networks. A prerequisite for creating impact is high-level competence in research and innovation. Finally, operative excellence and financial profitability should secure the balance of VTT's finance in the long run.

## Conclusions

- **To make the invisible visible we need more diversified evaluation approach and indicators which enable to analyse RTOs impact in the systemic and complex environment.**
- **RTOs' impact cannot be demonstrated with one – or even a few single indicators: a comprehensive picture of impact requires multi-criteria approach, a variety of indicators and data gathering from multiple sources.**
- **Policy measures and evaluation practices could be updated to perceive the significance and 'hidden performance' of innovations in systemic context: innovations do not emerge without evaluation mechanisms that support their creation.**

VTT – beyond the obvious

www.vttresearch.com

**Contact**
**Kirsi Hyytinen**
Tel. +358 40 5818495
Kirsi.Hyytinen@vtt.fi

# SCIENTIFIC- & ORGANISING COMMITTEE OF THE CONFERENCE: IMPACT OF RESEARCH AND INNOVATION POLICY AT THE CROSSROADS OF POLICY DESIGN, IMPLEMENTATION AND EVALUATION

## SCIENTIFIC COMMITTEE

1. **ERIK ARNOLD**
   Chairman of the Technopolis Group and Adjunct Professor in Research Policy at the Royal Institute of Technology (KTH), Stockholm
2. **GEMMA. E. DERRICK**
   Co-director of Centre for Higher Education Research and Evaluation, Lancaster University
3. **LEONIE VAN DROOGE**
   Senior Researcher at the Rathenau Instituut, The Hague
4. **JAKOB EDLER**
   Executive Director at Manchester Institute of Innovation Research, University of Manchester
5. **JORDI MOLAS-GALLART**
   Research Professor at Spanish Council for Scientific Research (CSIC)
6. **SYBILLE HINZE**
   German Centre for Higher Education Research and Science Studies GmbH
7. **PHILIPPE LAREDO**
   Director of Research at Institut Francilien, Recherche, Innovation et Société/ Université Paris-Est
8. **MIREILLE MATT**
   Director of Research, INRA, Laboratoire GAEL, UMR, Université Pierre Mendès France, Grenoble
9. **WOLFGANG POLT**
   Director POLICIES – the Institut for Economic and Innovation Research of JOANNEUM RESEARCH Ltd., Austria
10. **ANKE REINHARDT**
    Deputy Head of Information Management and Director of Evaluation and Monitoring at the German Research Foundation (DFG)
11. **KLAUS SCHUCH**
    Managing Director of fteval – Austrian Platform for Research and Technology Policy Evaluation / Scientific Director of ZSI – Centre for Social Innovation
12. **MICHAEL STAMPFER**
    Managing Director at WWTF – Vienna Science and Technology Fund
13. **LIVIU STIRBAT**
    Acting Head of Unit for Better Regulation in the European Commission's Directorate-General for Research and Innovation
14. **KATHARINA WARTA**
    Managing Director of technopolis group Austria
15. **MATTHIAS WEBER**
    Head of Competence Unit at Austrian Institute of Technology (AIT), Innovation Systems & Policy

## ORGANISING COMMITTEE

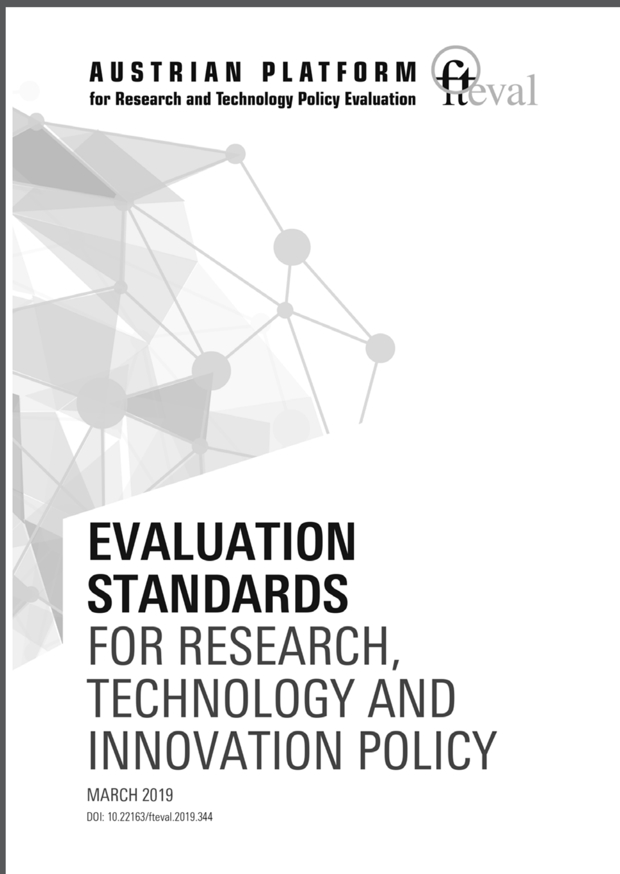
1. **MICHAEL DINGES**
   Senior Expert Advisor at Austrian Institute of Technology (AIT), Innovation Systems and Policy Unit
2. **BRIGITTE ECKER**
   Managing Director of WPZ Research GmbH
3. **Doris KAISERREINER**
   Executive Assistant at the Austrian Platform for Research & Technology Policy Evaluation
4. **PATRICK LEHNER**
   Head of Impact and Knowledge Exchange, Data Protection and Open Innovation in Science at the Ludwig Boltzmann Society
5. **SABINE MAYER**
   Analyst at FFG – The Austrian Research Promotion Agency
6. **FALK RECKLING**
   Head of Department at the FWF – Austrian Science Fund
7. **KLAUS SCHUCH**
   Managing Director of fteval – Austrian Platform for Research and Technology Policy Evaluation / Scientific Director of ZSI – Centre for Social Innovation
8. **MARIO STEYER**
   Policy Advisor, Austrian Federal Ministry for Transport, Innovation and Technology
9. **MICHAEL STRASSNIG**
   Programme Manager of Vienna Science and Technology Fund WWTF
10. **JÜRGEN STREICHER**
    Researcher at Joanneum Research
11. **KATHARINA WARTA**
    Managing Director at Technopolis Austria

# NOW AVAILABLE!

## LATEST EDITION OF THE

## 'EVALUATION STANDARDS FOR RESEARCH, TECHNOLOGY AND INNOVATION POLICY'

OF THE AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION


**Download in German:**
https://fteval.at/content/home/plattform/standards

**Download in English:**
https://fteval.at/content/home/plattform/standards/index.jsp?langId=2

**You want print copies?**
Please send an eMail to office@fteval.at

# FTEVAL JOURNAL
# FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION




## ... RTI POLICY EVALUATION AT THE INTERFACE BETWEEN PRACTICE AND ACADEMIA ...

The **'fteval Journal for Research and Technology Policy Evaluation'** contributes in a quality assured way to exchange between various stakeholder groups in the area of science, research, technology and innovation policy evaluation. It addresses policy-makers, practitioners, evaluators and the academic community. Thematic editions alternate with thematically open issues. By now 47 editions were published, which can be accessed and downloaded from the fteval-repository and the fteval homepage. The fteval Journal is open access. The journal and each published paper receive a separate DOI. All paper contributions are quality controlled, but not subject to peer-review in the strict academic sense.

Researchers and practitioners from the RTI policy domain are invited to send contributions to the editor of the **'fteval Journal for Research and Technology Policy Evaluation'** (office@fteval.at). The contributions can be submitted either in German or English.

If you want to receive the **'fteval Journal for Research and Technology Policy Evaluation'** free of charge, please fill out this form and send it to **office@fteval.at** or drop it into the infobox at one of our conferences and events.

first name ..........................................................................

last name ..........................................................................

institution ..........................................................................

address ..........................................................................

postal code ..........................................................................

city/place ..........................................................................

country ..........................................................................

e-mail ..........................................................................

By filling in this information you agree to* receive the fteval Journal by mail (free of charge)
* to receive the electronic newsletter of fteval, which informs about pertinent issues of RTI evaluation in Austria and the EU via email (such as calls for contributions for the next fteval Journal) and thus you entitle the fteval to process your indicated personal data for the purpose of transmission of information to you.
You can revoke consent at any time by informing us via eMail (office@fteval.at).

**EDITORIAL BOARD**

Rebecca Allinson, **Technopolis UK**; Balázs Borsi, **Eszterházy Károly College**; Elke Dall, **Centre for Social Innovation**; Michael Dinges, **Austrian Institute of Technology**; Leonid Gokhberg, **National Research University Higher School of Economics**; Wolgang Polt, **Joanneum Research**; Andreas Reinstaller, **WIFO**; Klaus Schuch, **fteval** (chief editor); Michael Stampfer, **WWTF**; Lena Tsipouri, **University of Athens**

**AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION (fteval)**

c/o ZSI – Centre for Social Innovation GmbH

Linke Wienzeile 246, A-1150 Vienna

**T** +43 1 495 04 42 - 79

**F** +43 1 495 04 42 - 40

**E** office@fteval.at

**W** www.fteval.at

**DESIGN**

**W** carotte.at

**E** caroline@carotte.at

**ILLUSTRATIONS:**

Kinga Jakabffy

**W** https://kingasgrapes.com

**PRINT**

agenskettterl Druckerei, Bad Vöslau

Gedruckt auf PEFC™-zertifiziertem Papier.

**AUSTRIAN PLATFORM FOR RESEARCH AND TECHNOLOGY POLICY EVALUATION**

**Platform fteval members:**

Austrian Federal Ministry of Education, Science and Research (bmbwf), Austrian Federal Ministry for Digital and Economic Affairs (bmdw), Austrian Federal Ministry for Transport, Innovation and Technology (bmvit), Austrian Cooperative Research (ACR), Austrian Council for Research and Technology Development, Austrian Institute of Technology (AIT), AQ Austria – Agency for Quality Assurance and Accreditation Austria, Austria Wirtschaftsservice (AWS), Christian Doppler Research Association (CDG), convelop Cooperative Knowledge Design GmbH, Austrian Research Promotion Agency (FFG), Austrian Science Fund (FWF), Institute for Advanced Studies (IHS), Industriewissenschaftliches Institut (IWI), Joanneum Research Forschungsgesellschaft mbH, Austrian Institute for SME Research (KMFA), Ludwig Boltzmann Society (LBG), Austrian Academy of Science (OEAW), Technopolis Group Austria, Vienna Business Agency – A service offered by the City of Vienna, Austrian Institute of Economic Research (WIFO), WPZ Research GmbH, Vienna Science and Technology Fund (WWTF), Centre for Social Innovation (ZSI)

**AUSTRIAN PLATFORM**
**for Research and Technology Policy Evaluation**

fteval

c/o ZSI – Centre for Social Innovation GmbH
Linke Wienzeile 246, A-1150 Vienna

**T** +43 1 495 04 42 - 79
**F** +43 1 495 04 42 - 40
**E** office@fteval.at
**W** www.fteval.at

DOI: 10.22163/fteval.2019.414

The fteval Journal for Research and Technology Policy Evaluation is a forum to discuss current evaluation practices in the field of research, technology and innovation policy.

**ZVR-Zahl:** 937261837
**ISSN-Nr.** 1726-6629
**© Vienna 2019**