

IMPACT PATHWAYS: TRACKING AND COMMUNICATING THE IMPACT OF THE EUROPEAN FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION¹

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

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, Ra-

dauer, 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 benchmarks:** 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 re-design” (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 Beselaar, 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):

- **Proximity** - 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;
- **Attribution** - 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;
- **Traceability** – 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;
- **Holism** - 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;
- **Stability** - 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

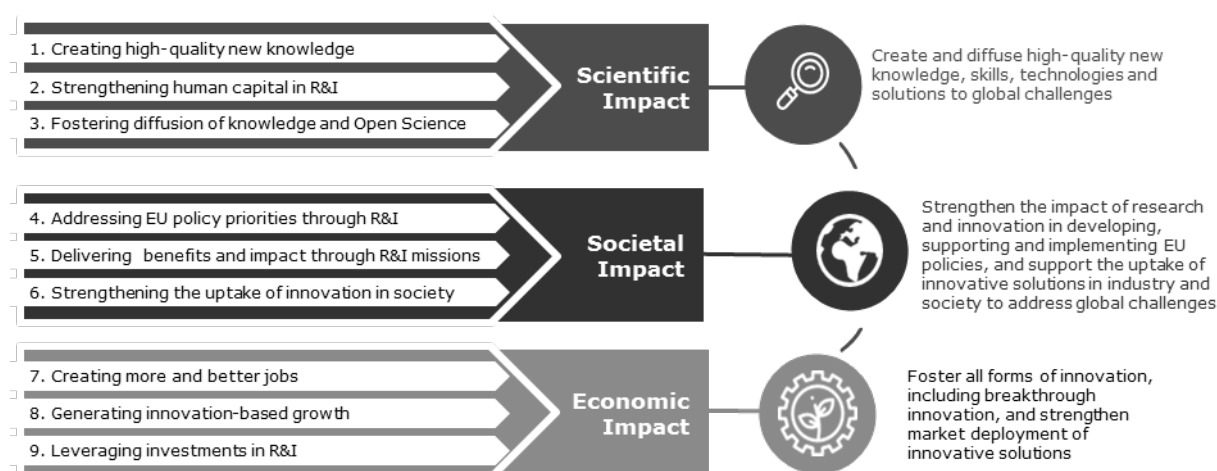


Fig. 1: Proposed Key Impact Pathways of Horizon Europe

Source: European Commission (2018), Impact Assessment accompanying the Commission proposal for Horizon Europe, the Framework Programme for Research and Innovation, SWD (2018) 307 final

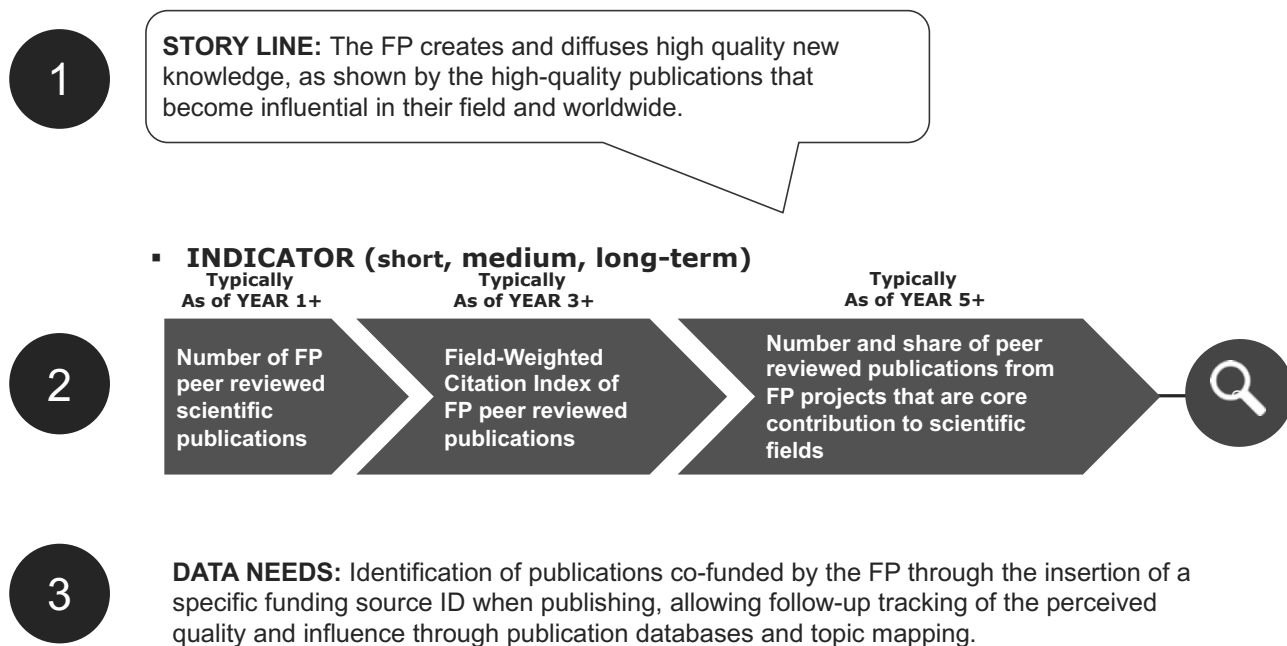


Fig. 2: Example of a Key Impact Pathway – Creating high-quality new knowledge

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.
- Pawson, 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



PATHWAYS to Impact of R&I POLICIES

